Reexamining Lackluster Productivity Growth in Construction

Daniel Garcia, Raven Molloy

2023-052

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

Reexamining Lackluster Productivity Growth in Construction

Daniel Garcia
Federal Reserve Board of Governors

Raven Molloy
Federal Reserve Board of Governors

April 2025

Of all major industries, construction is the only one to have registered negative average productivity growth since 1987. Mechanically, this lackluster performance owes to the fact that indexes measuring the cost of building a constant-quality structure have risen much faster than those measuring the cost of producing other goods. We assess the extent to which growth in construction costs could be biased upward by improvements in unobserved structure quality. Even under generous assumptions, our estimates of the magnitude of this bias are not large enough to alter the view that construction-sector productivity growth has been weak. Next, we calculate new estimates of single-family residential construction productivity growth by state and metropolitan area from 1980 to 2019. These estimates reveal that productivity has declined the most in areas with a larger fraction of construction in the urban core and with tighter housing supply constraints, especially in locations with long permitting times.

1. Introduction

According to data published by the Bureau of Labor Statistics (BLS), productivity growth in the construction industry has been the slowest among all of the major industry categories since at least 1987, when the current estimates begin.[1] Table 1 shows that productivity growth was actually *negative* on average for the construction industry over the 1987 to 2019 period, whereas it averaged at least 1 percent per year in all other major industries.[2] Moreover, Figure 1 shows construction productivity growth was consistently low during this whole period. While the chronic issue of low productivity growth in construction is not new (Stokes 1981, Allen 1985), this topic has recently been the focus of new research (Goolsbee and Syverson 2022, D'Amico et al. 2024). One reason for this renewed interest is that low productivity growth has implications for housing affordability, as increases in productivity could have allowed for the construction of more, higher-quality structures at a lower cost, helping to mitigate the growing imbalance between housing supply and housing demand over this period.

This paper has two main contributions to the literature on construction productivity. First, we question whether productivity growth has actually been so low for the past three decades.   The absence of any productivity growth sounds implausible given the variety of labor-saving innovations in the industry such as nail guns (Sichel 2022), more pre-fabricated inputs (Haas et. al. 1999, Teicholz 2013), and the vast improvements in information technology (e.g. architectural design software) since the 1980s. The accuracy of the official statistics is questionable because the measurement of construction productivity requires a number of assumptions, some of which could understate actual productivity growth.   One important example is the possibility that unobserved improvements in structure quality (e.g. energy efficiency, quality of interior finishes, etc.) have biased growth in the construction deflator upward, thereby biasing growth in the estimates of real output downward. We perform a detailed analysis of many possible sources of bias—mainly but not exclusively related to unobserved improvements in structure quality—and find that measurement error is unlikely to be large enough to overturn the conclusion that productivity growth in the construction sector has indeed been quite low. These results suggest that labor-saving innovations have been either modest or largely offset by other factors that have increased the costs of construction.

Our second contribution is to produce new state and metro-level estimates of productivity growth in new single-family residential construction from 1980 to 2019. These estimates are new in the literature, partly because producing them is not trivial, due to both data limitations and volatility in regional data. These estimates suggest that single-family productivity growth has likely been weak in much of the U.S. and has contracted sharply in some locations. Using these estimates we document new stylized facts about differences in productivity growth across locations and that shed light on some possible reasons why construction productivity growth has been so low.

---

[1] Estimates of productivity growth by industry can be found at https://www.bls.gov/productivity/
[2] We do not discuss data post-2019 because of concerns about measurement issues 2020-2022. The response rate for the construction-put-in-place survey fell in 2020, creating an increase in imputed data used to construct nonresidential construction spending. Moreover, real output growth in nonresidential construction was likely biased down in 2021 and 2022 due to measurement issues related to high cost pressures for inputs used in construction (Brandsaas et al. 2023). Productivity data are currently available through 2023. The published estimates for the construction sector fell by 0.3pp per year from 2019 to 2023, so the recent data continue the trend of low productivity growth.

The paper begins with a detailed description of how construction productivity is measured in the official statistics: as the quantity of structures produced divided by labor input, where the quantity of structures is calculated as the total nominal value of structures built in a variety of subsectors divided by price deflators specific to those subsectors. Figure 2 shows that on average, the deflators for the construction industry have risen at a much faster pace than those used to deflate nominal output in other industries. Hence, either the price of producing structures really has increased at a much higher rate than the price of producing other goods and services over the past 30 years, or the construction-sector price deflators have been biased upward by a growing amount over time.

There is reason to suspect a role for deflator mismeasurement because a proper deflator should measure the change in the price of structures holding quality constant, but quality is a function of many characteristics, some which are harder to measure than others. For example, the analysis in Goolsbee and Syverson (2023) controls for quality changes by accounting for changes in housing size, but this adjustment alone could be insufficient, as it does not account for other aspects of quality such as number of bedrooms and bathrooms, energy efficiency, and quality of craftmanship that may be hard to record as data but are appreciable to owners and property inspectors. As we discuss below, the price deflator methodology employed by the Census does account for characteristics like size and number of bedrooms, but it does not account for other qualities that have likely improved over time.[3]

Given the potential importance of unobserved quality, we first quantify the bias from this source using three different approaches. Our first approach uses detailed industry construction cost data from R.S. Means to estimate the change in construction costs for specific housing types. This method is much less susceptible to unobserved quality bias than the Census Bureau's method because we can hold many more features of a housing unit fixed, such as the type of roof and the material used for kitchen countertops. The construction costs that we generate under this approach rise by about the same amount as the deflator used for new single-family construction, suggesting that the influence of unobserved quality increases on this deflator has been negligible. Our second approach measures aspects of structure quality that are observable to us but not used in the Census Bureau's calculations. We obtain three such measures: an assessment of structure quality from property tax assessors, a rating of structure quality from the resident of the home, and an estimate of energy efficiency. We estimate that improvements in energy efficiency have boosted structure values by only about 0.2 percentage point per year from the late 1980s to the late 2010s. The tax assessors' and residents' quality ratings have not changed much at all over time, likely because these measures are better suited for cross-sectional comparisons of quality rather than for changes over long periods of time. Using the cross-sectional correlation between quality and house value and the generous assumption that all homes built in the 1980s were low quality and all homes built in the 2010s were high quality, we estimate that unmeasured quality improvements could have boosted structure prices by about 0.6 percentage point per year. Our final approach is the application of an econometric technique for assessing the magnitude of unobserved variable bias (Oster 2019). It is based on how observed structure characteristics like unit size and number of bathrooms are correlated with the change in structure values over time, as well as on an assumption about how unobserved structure characteristics might be correlated with changes in structure value. This technique suggests that unobserved quality

---

[3] For example, structures now are likely much more energy efficient, using better insulation and more efficient heating and cooling systems. They are also more fire-resistant as they are more likely to use grounded electrical outlets. And they use higher-quality materials, such as reinforced concrete for foundations.

improvements have biased the growth rate of the single-family deflator upward by no more than 0.8 percentage point per year from 1987 to 2019.

In sum, our three different approaches suggest that unobserved structure quality has biased up the single-family deflator by an amount ranging from zero to 0.8 percentage point per year. After accounting for the fraction of nominal construction output that is deflated by this price index and making assumptions about the effect of unobserved quality on deflators used for other sectors of construction, we calculate that the resulting bias to productivity growth in the aggregate construction sector is no more than 0.5 percentage point per year. We also discuss other potential sources of bias to measured productivity growth and conclude that the magnitude of these other sources is probably small as well. In conclusion, it does seem that that productivity growth has been quite low in the construction industry, even if it has not been as low as implied by the official statistics.

Having established that construction productivity growth has truly been quite low, the second goal of this paper is to provide new stylized facts about how productivity growth has varied across the country. Reliable regional estimates of construction productivity growth are generally not available, and so we construct our own measures.[4] We focus on the new single-family construction sector since data on output for this sector is available, unlike other sectors of construction. Specifically, we create new estimates of productivity growth by state and metropolitan area from 1980 to 2019 using data on single-family permit issuance, the average size of new homes, and the number of construction workers.

The regional estimates suggest productivity growth has been quite low through much of the United States, but still with a fair amount of geographic heterogeneity. For example, at the low end some states experienced productivity declines of about -4 percent per year, while at the high end some states experienced productivity growth of about 1 percent per year. Some of the states with the lowest productivity growth are small and relatively densely-populated like Connecticut, Rhode Island and Vermont. Other states with relatively strict regulatory constraints on new construction, like Massachusetts, New York and California also had fairly low productivity growth. Meanwhile states with relatively high productivity growth include West Virginia, South Carolina and Montana.

We next turn to metropolitan area estimates and estimate regressions of productivity growth on local attributes that could be related to structure costs and hence productivity growth. One of the main findings is of a negative association between productivity growth and measures of housing supply regulations. To measure the latter, we mainly rely on the Wharton Residential Land Use Regulatory Indexes, but we also find a negative association with other measures of housing supply constraints. This finding complements finadings in in D'Amico et al (2024), who find a correlation between regulation, firm size, and the *level* of productivity. The correlation between productivity growth and regulation that we find is generally stronger than that documented in Sveikaukas et al (2016), likely because our results are based on a long-run estimate of productivity growth while their estimates are based on higher-frequency correlations, which are likely to be noisier given the high cyclicality and volatility of productivity growth at the regional level. An additional contribution of our analysis is that we dig into the types of regulation that matter for productivity growth. We find that delays in permit approval are most strongly and robustly correlated with productivity growth. Other aspects of the regulatory

---

[4] In particular, the BEA publishes state-level nominal output indexes that are deflated using a *national* construction price index, which thus assumes away differences across areas in construction costs.

environment matter as well, including restrictions on the number of permits allowed and impact fees. Beyond regulation, we also find a positive association between productivity growth and the fraction of construction taking place outside of the urban core (further away from the city hall), perhaps as construction projects in areas that are less built-up tend to be larger and hence exploit economies of scale. By contrast we find that productivity growth is not related to initial metro size or initial metro density, suggesting that there is scope for productivity growth even in large, dense cities.

Our research is not the first attempt to examine slow productivity growth in the construction sector. Some previous research suggests that productivity growth may have been higher than the official statistics suggest. For example, Goodrum, Haas and Glover (2002) analyze data on the labor hours needed to complete 200 different construction-sector work activities in 1976 and 1998 and find that productivity increased materially for most activities, with an average increase of 31 percent. Also, Sveikauskas et al (2016) and Sveikauskas, Rowe, Mildenberger (2018) argue that some construction sectors like multifamily and industrial have experienced more robust productivity gains. And Allen (1985) found that slow productivity growth from 1968 to 1978 could be explained by mismeasurement of the nonresidential construction deflators and a shift towards single-family construction.[5] By contrast, Goolsbee and Syverson (2023) argue that productivity growth in the single-family sector has likely been minimal since the 1970s, given the roughly flat trajectory of aggregate square footage of new single-family homes per employee. D'Amico et al. (2024) show that small firms are more common in residential construction than in manufacturing, and argue this difference could partly explain stagnant productivity in construction since small firms are less able to exploit economies of scale. Our paper complements this earlier work by showing that productivity growth may have been restrained by regulation and shifts in the location of construction, even as some labor-saving improvements may have been pushing in the opposite direction.

The remainder of the paper proceeds as follows. Section 2 provides an overview of how productivity in the construction sector is measured. The price index used as the deflator for new single-family construction has a very large influence on aggregate construction, and so this section provides details on how this price index is calculated. Section 3 provides evidence on the potential role for measurement error in the single-family deflator and also discusses other measurement issues. Section 4 presents the estimates of construction productivity growth by state and metropolitan area. Section 5 concludes and discusses other possible reasons why construction productivity growth has been so low for so long.

Section 2. Measurement of Productivity in the Construction Sector

2.1 Measurement of Nominal Output and Real Output

The BLS measures productivity in the construction sector by aggregating real output of 22 subsectors and dividing by an estimate of labor input for the entire industry.[6] The residential subsectors are new single-family construction, new multifamily construction and improvements. The nonresidential subsectors span a wide range of structures such as offices, warehouses, manufacturing structures like factories, power and communication infrastructure, and highways. The nominal shares of the 15 largest subsectors are reported in Appendix Table 1. Following the methodology used in the National Income

---

[5] His findings are less relevant today, however, due to changes in nonresidential price index methodology and since single-family construction has not continued to be a growing share of total construction.
[6] Measuring productivity for each subsector separately is complicated by classification issues with labor input, as some workers in the construction industry may operate in more than one subsector.

and Product Accounts (NIPA), real output for each subsector is calculated by dividing nominal output by a deflator that is specific to that subsector. Nominal output for each subsector is based on construction spending from the Census Bureau's Value of Construction Put In Place program.[7] For new single-family residential construction, spending is estimated from the sales prices of newly-built single-family homes and assumptions about how the construction of a unit is spread over time from start to completion. For residential improvements, nominal spending is estimated from homeowner expenditures in the Consumer Expenditure Survey. For multifamily and nonresidential construction, nominal spending is from a survey that asks builders to estimate the nominal value of structures put in place each month.

The price deflators used to convert nominal output to real output are drawn from a variety of sources. The price deflator for new single-family construction is the price index for new single-family homes under construction produced by the Census Bureau, which we will refer to henceforth as the "single-family price index". As we will describe in more detail in section 2.2, this index estimates the constant-quality price of new single-family structures based on the sales prices and characteristics of new single-family homes. For data since 2005, the deflator for new multifamily construction is the price index for new multifamily units under construction produced by the Census Bureau, calculated using a similar method as the single-family price index. From the late 1970s to 2004, the multifamily deflator was a price index developed by the BEA for the purpose of deflating nominal construction spending (de Leeuw 1993). The deflator for residential improvements is an average of the single-family price index, the PPI for inputs to residential maintenance and repair, and the Employment Cost Index for the construction industry. Meanwhile the deflators for the nonresidential sectors differ by sector and time period. Some nonresidential sectors, such as office and health care, use a Producer Price Index (PPI) for new buildings in that specific sector. These sector-specific PPIs were developed in the 2000s and the starting dates differ a bit for each sector. For years between 1997 and the start data of each PPI, the BEA uses a sector-specific cost index that it developed from a construction cost estimator (Grimm 2003). Prior to 1997, the deflators used for all nonresidential sectors are an unweighted average of the single-family price index and the Building Cost Index produced by the Turner Construction Company. For some nonresidential sectors, like lodging, there is no sector-specific PPI and an unweighted average of the single-family price index and the Turner Building Cost Index is used for the entire time period from 1987 to the present.

Table 2 lists all of the price indexes that are used as inputs to the deflators and reports the average share of nominal construction activity for which each is used over our 1987-2019 sample period. The single-family price index has the largest influence on aggregate construction, both because of the large share of new single-family construction and because this price index is used to deflate other sectors as well. In total, any bias in the single-family price index will affect nearly half of aggregate output in the construction industry. In Section 3 we will discuss how each of the other price indexes might also be affected by unobserved quality improvements. We will also discuss other potential sources of bias that will be relevant to various subsets of the price indexes listed in Table 2.

Section 2.2. Methodology for the price deflator for new single-family construction

---

[7] The productivity statistics use a concept of output called "sectoral output," which is defined as the total amount of goods and services produced in an industry for sale either to consumers or to businesses outside that industry. Because the value of inputs is not subtracted from output, accurate measurement does not require accurate measurement of the sector's inputs.

Since much of our analysis will focus on measurement issues pertaining to the price index for new single-family homes under construction, it is helpful to describe how it is computed in more detail. Full details of the methodology can be found on the Census website.[8] Musgrave (1969) describes the development of this method.

The Census Bureau computes the single-family price index using sales prices and characteristics of new homes sold. The first step is a set of hedonic regressions, modeling structure value as a function of various housing unit characteristics. [9] These characteristics include structure square footage, number of bedrooms, number of bathrooms, presence of a basement deck, patio, or garage, type of exterior wall material, and type of heating/air conditioning. These characteristics are often included in hedonic pricing models (Sirmans et al 2006). Regressions are run separately for each time period and for five separate market strata: single-family attached homes and single-family detached homes in each of the four Census regions. The next step is to calculate two price indexes using the coefficient estimates from these regressions: a Laspeyres index and a Paasche index. The Laspeyres index is a weighted average of the estimated coefficients for each housing unit characteristic, with the weights based on the housing unit characteristics in 2005. The Paasche index is similar but uses the current period housing characteristics as the fixed weights. Finally, the price of single-family homes under construction is the geometric mean of the Laspeyres and Paasche indexes (i.e. a Fisher Ideal index).

The dependent variable in the hedonic regressions is an estimate of structure value. For homes that are built by contractors, structure value is computed as the amount paid to the contractor. However, for homes that are built for sale, structure value is not easily observable. The Census Bureau multiplies the home's sales price by a fixed factor (0.84) to subtract out the value of the project attributable to land as well as some other non-structure costs like the value of moveable appliances. [10]

The Census Bureau's methodology will generate an unbiased estimate of changes in structure costs if it controls for all aspects of structure quality that are correlated with house value and that have changed over time. However, not all aspects of structure quality are included in their analysis. Some examples of omitted structure characteristics include the energy efficiency of windows and doors, the types of interior finishes such as flooring and kitchen countertops, and the durability of the materials used. In the analysis below we will refer to "unobserved quality" as all features of the structure that are correlated with structure value but not included in the Census Bureau's methodology for calculating the price index for single-family homes under construction.

Section 3. Bias from Unobserved Quality and Other Sources of Bias

This section starts by examining the potential for changes in unobserved structure quality to bias estimates of growth in the Census Bureau's price index for new single-family homes under construction. We take three approaches: the creation of an alternate price index that holds many more housing characteristics fixed than the Census Bureau's methodology; an analysis of aspects of quality that we can observe in other data sources but that are not included in the Census Bureau's methodology; and an econometric method for estimating the magnitude of unobserved variable bias. Next, we discuss the

---

[8] https://www.census.gov/construction/cpi/pdf/descpi_uc.pdf
[9] See Sirmans, Macpherson and Zietz (2005) for a review of studies using hedonic models of house prices.
[10] See more information here: https://www.census.gov/construction/c30/methodology.html. For contractor-built homes, the Census Bureau inflates sale amounts by a factor of 1.1 to account for other expenses related to lot development. Contractor-built houses are weighted to also represent owner-built houses.

scope for unobserved quality to bias the price deflators used for sectors of construction other than new single-family homes, as well as potential sources of bias other than unobserved structure quality. Finally, this section concludes by summarizing the maximum possible and likely bias owing to all of the issues discussed in this section.

3.1 An Alternate Price Index to Measure Construction Costs

One way to assess the potential bias of the price index for new single-family homes under construction is to create an alternate measure of construction costs that holds many more aspects of structure quality fixed over time than the Census Bureau's price index does. If increases in this alternate measure of construction costs were much smaller than the increases in the Census Bureau's single-family price index, we might conclude that unobserved quality has biased up the Census Bureau's price index. To create this alternate price index, we use information from a company named R.S. Means that estimates the cost of building various types of residential structures. These estimates are used by builders and contractors to develop cost estimates for their construction projects, and therefore should be quite reliable. The cost estimates are created by adding up the cost of materials and installation for all of the individual components of a structure and then adding in costs for overhead, architectural fees, and other general costs.

The advantage of using R.S Means estimates to study changes in construction costs is that R.S. Means allows the user to specify many detailed attributes of the structure. For example, one can specify that a home has a laminate kitchen countertop, and therefore one can compare the cost of homes with laminate kitchen countertops at different points in time. By contrast, a general shift from laminate kitchen countertops to granite countertops would increase the average sales price of new homes and bias the Census Bureau's single-family price index upward because type of countertop is not included in the Census Bureau's regression. Nevertheless, the RS Means estimate is not entirely free from bias. Continuing with the example of countertop quality, the RS Means estimate would be biased if the quality of laminate countertops has changed over time.

R.S. Means provides cost estimates for a variety of home types (1-story, split level, 2 story, etc.) and four quality levels of each type: economy, average, custom and luxury. We calculate the construction costs for 1-story homes and 2-story homes at each of these quality levels, yielding a total of 8 cost estimates at each point in time. As shown in Table 3, we allow unit size, number of bathrooms, type of exterior and roof, type and length of kitchen countertops, and many other unit characteristics to differ by level of quality. The costs of the characteristics also vary by quality. For example, the cost per linear foot of a laminate countertop is higher for an average quality home than for an economy quality home, presumably reflecting the use of a higher quality material.

R.S. Means also contains information about the amount of time required to complete various tasks and the total cost to complete these tasks. Goodrum, Haas and Glover (2002) analyze data from R.S. Means and other cost-estimating firms on the labor hours needed to complete 200 different construction-sector work activities in 1976 and 1998 and find that productivity increased materially for many activities, with an average increase of 31 percent. Though this task-level analysis is interesting, the set of tasks undertaken to build a structure has changed over time. If new tasks tend to be lower productivity than existing tasks, then these new tasks would mute the productivity gains from the existing tasks. RS Means does not describe exactly which sets of tasks are required to build a specific type of structure, so it is not possible to examine how the set of tasks has changed. Moreover, some

aspects of the construction process, such as measures that increase worker safety, may have changed in a way that increases costs even for a given set of tasks. Hence, in our analysis we prefer to look at cost changes over time for a completed structure, as these cost estimates are more comprehensive than task-based estimates.

Table 3 reports the estimated construction costs for each unit type in 1987 and 2019. The cost increases for all 8 housing types range from 2.7 to 3.6 percent per year, with the unweighted average equal to 3.2 percent. Meanwhile, the Census Bureau's single-family price index rose by 3.2 percent per year over this period. The result that the R.S. Means cost estimates do not show substantially smaller increases than the Census Bureau's price index suggests that the Census Bureau's omission of the many housing unit characteristics included in R.S. Means has not led to a material bias. While it is true that the R.S. Means estimates do not hold all housing characteristics fixed, the fact that holding many important characteristics fixed does not lead to a much lower estimate of cost increase suggests that the role of unobserved quality change is small.

3.2 Observed Measures of Quality

3.2.1 Tax Assessors' and Residents' Ratings of Structure Quality

For the purposes of assessing the value of residential property, tax assessors in many jurisdictions report the quality of the structure. The excerpt below, taken from the real estate assessment website of Fairfax County, Virginia, provides an example of the factors that affect the assessor's quality evaluations:[11]

> *The Average category covers many standard tract-built houses. These are built to at least minimum building code standards and the quality of materials and workmanship is acceptable. Good category houses are typically found in better quality tract developments or can be designed for an individual owner. The shape of the structure is generally somewhat more complex than the Average category and good quality standard materials are used throughout. The Excellent category covers properties in higher end subdivisions or standard custom houses. Excellent properties have a higher level of design and materials when compared to Good. Luxury properties are typically individually designed custom houses and exhibit very high standards of design, materials, finish, and workmanship.*

Thus, these quality ratings will capture many elements of structure quality that are not otherwise recorded in the data. These quality assessments are included in CoreLogic's Residential Real Estate database, which contains the property characteristics from tax assessment records for 99% of the US housing stock. The categories of ratings vary across jurisdictions. Appendix Table 2 reports the frequency of all the ratings that appear in the CoreLogic dataset. We group the responses "excellent", "luxury", "above average" and "good" into an indicator for high quality and the remaining non-missing responses into an indicator for medium/low quality.

We examine the correlation of this measure of structure quality with house prices by regressing the sales prices of new homes on an indicator for high structure quality and the housing unit characteristics included in the Census Bureau's methodology. To this end, we use the sales prices of new homes from

---

[11] https://icare.fairfaxcounty.gov/ffxcare/content/desc.htm

the deeds transactions in CoreLogic's Residential Real Estate database, which includes transactions from 2000 to 2019. New homes are identified using a new construction indicator calculated by CoreLogic based on owner transfer records where CoreLogic has identified the seller as a builder.[12] We also require that new homes have a year of first sale no more than three years after the year built. In our sample, 91 percent of the new homes were first sold within a year of the year they were built. Because structure quality is only available for 39 percent of the sample, we also include an indicator for homes where quality is missing.[13]

As shown in the first column of Table 4, the price of homes with high structure quality is 0.16 log points, or about 18 percent higher than the price of lower-quality homes conditional on the other characteristics that the Census Bureau uses to calculate constant-quality new home prices.[14] Therefore this measure of structure quality does indeed seem to capture an important housing attribute that is missing from the Census Bureau's methodology. What matters for our purpose is how this aspect of quality has changed over time. However, the CoreLogic data cannot speak directly to changes in structure quality from 1987 to 2019, both because we only have data starting in 2000 and because we suspect that the assessor's measure of quality may be relative to other homes in the same year rather than an absolute measure of quality.   46 percent of the homes built in 2000-2004, the first five years of this sample, were high quality, compared with 52 percent in 2014-2019.

Nevertheless, we can use a back-of-the envelope calculation to estimate the largest possible effect that these results imply for bias in the single-family price index. Specifically, if we assume that no new homes were high-quality in 1987 and all new homes were high quality in 2019, the cumulative change in the single-family price index would be biased upward by almost 20 percent, which translates to an annualized growth rate of about 0.5 percentage point per year.

Next we examine a similar measure of quality: a resident's rating of the quality of their home. This rating is reported in the American Housing Survey (AHS), which is a nationally representative survey of housing units with a primary goal of measuring the size, composition and quality of the US housing stock.   For this analysis, we use AHS data on newly-built single-family detached homes covering two time periods: an "early" period, which includes data on homes built between 1970 and 1989 as observed in the 1985, 1987 and 1989 National samples, and a "recent" period, which includes data on homes built between 2000 and 2019 as observed in the 2015, 2017 and 2019 National samples.

The AHS asks the resident to rate the quality of their home as a place to live on a scale from 1 to 10. Since the AHS asks a separate question about neighborhood quality, we are reasonably confident that

---

[12] This measurement of new construction based on information identifying the seller as a builder is consistent with the recommendations in Coulson, Morris, and Neil (2019), who show that estimates of the new home premium can vary meaningfully when new homes are identified solely based on age since some recently built homes could include "flips." Our main estimates are robust to excluding properties first sold one or more years after the property was built.

[13] Quality ratings appear to be missing in many cases because many counties do not record structure quality. Specifically, quality tends to be missing for all housing units in a county or available for most housing units in a county. Results are robust to dropping observations with missing quality, limiting the sample to counties where less than 25 percent of the observations are missing quality, and including county fixed effects.

[14] We find that the high-quality premium is similar in high-cost and low-cost areas, as well as in the first five and last five years of our sample period. The high-quality premium is also robust to including state and metro area fixed effects in the regression. Results available upon request."

the home quality rating reflects structure quality and not local amenities. In this sample the resident's rating of housing quality is generally in the top third of the range and did not change much between the two sample periods (see Appendix Table 3). Just like the tax assessor measure, we suspect that this rating reflects an assessment of the quality of the home relative to other homes in the same time period rather than relative to homes in an earlier time period. Even so, we can use the data to estimate the cross-sectional correlation between quality and home value conditional on other housing unit characteristics.

We regress the natural logarithm of house value (as reported by the survey respondent) on a set of housing unit characteristics, indicators for Census region, an indicator for homes built in the "recent" period, and indicators for different quality ratings. Although the set of housing unit characteristics is not as complete as the set used by the Census Bureau for calculating the single-family price index, we still obtain a good approximation of the cumulative price increase from the early period to the recent period. Specifically, controlling for housing characteristics the value of homes in the recent period is 169 percent higher than the value of homes in the early period (not shown). The single-family price index rose by 153 percent between these two periods, a very similar amount. Column 2 of Table 4 shows that homes with the highest quality rating are about 0.16 log points (17 percent) higher value than those with a rating of 7 or below.   Therefore, this analysis supplies supporting evidence that conditional on the housing characteristics used by the Census Bureau, high-quality new homes are roughly 20 percent higher value than low-quality new homes. As with the CoreLogic data, this analysis cannot speak directly to changes in quality over time. But a back-of-the-envelope calculation similar to the one using the CoreLogic estimate would generate a similar result.

3.2.2 Energy efficiency

Another aspect of housing quality that we examine in the AHS data is energy efficiency. Many improvements in housing quality over the past 40 years are intended to improve energy efficiency. Some examples include double-paned windows, better insulation, and more efficient heating and cooling systems. Although these improvements are difficult to measure individually, we can get a sense of the cumulative changes in energy efficiency of new homes by comparing the total energy use of new homes built in the 1970s and 1980s to that of new homes built in the 2000s and 2010s.

For this exercise we calculate total expenditures on utilities as the sum of annual expenditures on electricity, natural gas, heating oil, water and other fuels. We deflate these nominal expenditures by the Consumer Price Index for utilities in order to obtain an estimate of the quantity of energy used for each home. Then we regress the energy use for each house on indicators for unit square footage, indicators for Census region, and an indicator for homes built in the recent period. The coefficient on the indicator for homes built in the recent period shows how energy use has changed over time after conditioning on changes in the size and geographic location of housing units.[15]

As reported in Table 5, the energy use of homes built in the 2000s and 2010s was almost 25 percent lower (column 1) or $740 (column 2) per year lower than that of homes built 1970s and 1980s. While

---

[15] It is important to control for unit size because new homes have become larger over time and larger homes use more energy. Ideally it would be nice to control for more detailed geographic information since weather patterns, and therefore the need to heat and cool homes, can vary materially within Census region. However, this information is not available in the public-use data.

this dollar amount is not insignificant, it is only about 4 percent of the annual rental expenditures of the homes in the recent sample. To compare the energy savings with average value of the structure, we estimate cumulative savings over the life of the home by dividing the annual energy savings by a cap rate of 5.3 percent, which is the ratio of rental income to property value reported in Jorda et al. (2019) for all U.S. residential property from 1870 to 2015. Next we calculate average structure value of the homes in our sample by multiplying the average home value in the recent sample by (1-0.41), since Davis, Larsen, Oliner and Shui (2021) estimate that the share of house value attributable to land is 0.41.[16] These calculations suggest that energy savings over the life of a home are about 7 percent of average structure value. Given that this improvement in energy efficiency occurred over a 30-year period, this aspect of quality boosted structure value by only 0.23 percent per year. Estimates are even smaller if we use a cap rate derived from the 2019 annual reports of large single-family rental corporations, for which we calculate cap rates above 9 percent (calculations available upon request).

3.2.3 Other measures of quality

A third aspect of structure quality that we can observe in the AHS data is whether the home has various types of appliances: a dishwasher, a washing machine, and a clothes dryer. In principle, moveable appliances like these should not be included in structure value. In fact, part of the Census Bureau's time-invariant adjustment to sales prices is to subtract the value of appliances. However, since the adjustment is time-invariant, the price index for new single-family homes will be biased if the ratio of total value of appliances to total structure value has changed over time. Moreover, the presence of these appliances could be correlated with other aspects of structure quality. For example, homes with a dishwasher could be more likely to have higher quality kitchen countertops and cabinets. In the AHS, the fraction of new homes with dishwashers increased from 0.73 in the 1980s to 0.93 in the 2010s, while the fraction of homes with dryers rose from 0.88 to 0.96. These increases, while not very large, could signal that moveable appliances, or possibly other unobserved housing attributes that are correlated with these appliances, have become a larger fraction of home value. We assess this possibility by including indicators for each of these appliances in the regression described above. As shown in column 3 of Table 4, the coefficient estimate on the indicator for homes built in the recent period barely changes, suggesting that the contribution of such appliances to total home value has not changed over time. In support of this conclusion, a survey of homebuilders found that appliances were only a small share of total structure cost and that this share did not increase from 1998 to 2019.[17] We conclude that appliances have not increased as a share of total home value from the 1980s to the 2010s, and therefore have not led to a material bias in the single-family price index.

We can also assess changes in structure quality over time using supplemental information provided by the R.S. Means company. In conjunction with providing estimates of the cost of building specific types of structures, they describe the general characteristics of the structures whose costs they assess. In Appendix Table 4 we summarize the descriptions of average-quality new homes in 1987 and 2019. Many elements of new single-family homes have remained the same over this 32-year period. The

---

[16] https://www.fhfa.gov/PolicyProgramsResearch/Research/Pages/wp1901.aspx. As we will discuss below, this estimate is for all homes less than 10 years old, not only newly-built homes. The land share is probably lower for new homes since lot sizes have fallen over time. A lower land share would raise our estimated structure value and therefore lead to an even smaller estimate of the improvements in energy efficiency relative to structure value.
[17] https://www.nahb.org/-/media/8F04D7F6EAA34DBF8867D7C3385D2977.ashx

average new home is still built with a concrete foundation and framed with 2x4 studs and ½" plywood sheathing. It has asphalt shingles on the roof, ½" drywall for the interior walls, and similar flooring. That said, some elements of homes built in 2019 are higher quality. Foundations in 2019 were made of reinforced concrete and insulated, whereas foundations in 1987 were not. The average quality new home in 2019 included a 40-gallon electric water heater, whereas the typical water heater in 1987 was only 30 gallons and gas-fired. Electric water heaters tend to be cheaper and more energy efficient than gas, so this shift reflects a clear quality improvement. Overall, this evidence suggests that building quality has increased a bit over time, but the changes do not seem dramatic. We find similar results for luxury-quality homes (not reported).

Section 3.3 Econometric bounds on the contribution of unobserved quality

As a final way to assess the magnitude of measurement error attributable to unobserved quality, we turn to an econometric technique developed by Oster (2019). This technique is useful for placing bounds on the magnitude of coefficient bias for scenarios in which observed controls are an incomplete proxy for omitted variables. The Oster (2019) estimator uses as inputs observables (how the coefficient of interest and model R-squared change when the observed controls are included) and two assumptions about unobservables. These assumptions are: 1) the maximum R-squared if all relevant explanatory variables were observed and 2) the influence of remaining unobservables relative to the influence of the controls we do observe. We adapt this method to our case, where the coefficient of interest is a time period indicator (i.e. the change in structure price conditional on observed characteristics).

Oster (2019) shows a consistent estimate of the coefficient of interest (β*) can be approximated using the following formula:

$$\beta^* \approx \tilde{\beta} - \delta \left[\mathring{\beta} - \tilde{\beta} \right] \frac{R_{\max} - \tilde{R}}{\tilde{R} - \mathring{R}}.$$

where $\tilde{\beta}$ and $\tilde{R}$ are the coefficient estimate and R-squared from the model with full controls, βo and Ro coefficient and R-squared from the baseline model, δ relates the importance of unobservables relative to the importance of observables, and Rmax is the maximum R-squared when all possible controls (observed and unobserved) are included.

As a rule of thumb, Oster (2019) suggests bounding values of Rmax=1.3*$\tilde{R}$ and of δ=1. These values are calibrated by re-analyzing estimates from randomized experiments, which provide unbiased coefficient estimates by design. The second assumption implies that the remaining unobserved characteristics are as important as the observables.   In our case, we think δ=1 is a reasonable bounding assumption, as it implies that various aspects of unobserved quality like interior finishes are as important to home values as the observables like square footage and number of bathrooms.

We begin with the AHS data since the data cover the full time period of interest. Table 6 shows that in a regression with only region indicators, the coefficient on the indicator for homes built in the recent period is 1.19. When the full set of Census variables are included, this coefficient decreases to 0.99 and the R-squared increases by 0.18. Thus, the observed measures of quality reduce the estimated increase in home value over this 30-year period by 0.2 log point. With Rmax=1.3*$\tilde{R}$ and δ=1, the lower bound for the unbiased coefficient on the indicator for homes built in the recent period would be 0.77. Converting

the coefficient estimates to annualized growth rates, we find the constant-quality price of structures would have risen at an annual rate of 2.6 percent, rather than the estimated 3.4 percent when only the Census controls are included. In other words, this calculation suggests that unobserved increases in quality have biased the rate of increase of structure prices by up to 0.8 percentage points per year.

Next we conduct the same econometric exercise using the CoreLogic data described in section 3.2. The results in Table 6 show that the unobserved quality improvements may have biased the rate of increase by up to 0.2 percentage points this year. This estimate is smaller than in the AHS data because the time period coefficient falls by less when the observed measures of quality are included and because the R-squared increases by more.

Although we cannot test the appropriateness of the bounding assumptions directly, two types of evidence suggest that δ is unlikely to be larger than 1. First, we can look at the correlation of the resident's or tax assessor's assessment of structure quality with house value, since these variables are observable measures of quality that are excluded from the Census Bureau's analysis. These correlations are smaller than the correlation of unit size with house value, and either the same size as or small than the correlations of many other housing attributes with house value (see Appendix Table 5). Therefore, assuming that the correlation between unobserved characteristics with house prices is as large as the correlation between observed characteristics and house prices seems like a reasonable upper bound. Second, we note that the only way that δ could be larger than one, or even equal to one, is if the unobserved measures of quality increased by much more than the observed measures of quality. This seems unlikely to us given the large increases in observed quality: in the AHS the fraction of new single-family homes larger than 2500 square feet increased by more than 50 percent from 27 percent in the 1980s to 44 percent in the 2010s. And the fraction with at least 3 bathrooms tripled from 11 percent to 35 percent.

To summarize our results on bias to the single-family price index from unobserved structure quality, our estimates range from very small (when comparing the single-family price index to alternate cost estimates from RS Means) to 0.8 percentage point per year (when using the econometric method). We will use the estimate of 0.8pp in the spirit of calculating the largest possible amount of measurement error.

Section 3.4 Quality bias in nonresidential sectors and other sources of bias

So far, we have focused on the potential for unobserved quality to bias growth of the price index for new single-family homes under construction. What about other deflators for sectors of construction? Since the price index for new multifamily units under construction is calculated using a very similar methodology as the single-family price index, the bias in this deflator could be similar.[18] We suspect

---

[18] Specifically, the Census Bureau also creates this index from a sample of property sales prices and the characteristics of the buildings. Eriksen and Orlando (2022) use the RS Means cost estimator to calculate the construction cost of two multifamily building types from 2012 to 2020. They find much smaller increases in construction costs (less than 2 percent per year) than the increase in the Census Bureau's multifamily price index (5 percent per year). This result suggests that increases in building quality have biased up the multifamily price index. That said, Eriksen and Orlando's calculations assume that management and design overhead are a fixed percentage of building cost; increases in these costs might also have caused the multifamily price index to increase by more than their estimates.

that unobserved structure quality is likely to have a negligible influence on the PPIs for new nonresidential buildings because they are based on changes in the costs of very specific inputs. For a similar reason we suspect that the PPI for inputs to residential maintenance and repair will not be influenced by changes in the quality of construction materials. We also think that the ECI for construction workers should not be influenced by changes in structure quality since it measures only labor costs. The influence of unobserved structure quality on the BEA's cost indexes—the nonresidential indexes used between 1997 and the introduction of the PPIs and used for the multifamily index used before 2005—is probably smaller than that for single-family price index because these indexes were created based on the estimated cost of labor and materials for specific structure types, not based on building sales prices. However, since the inputs used for these indexes may not have been as detailed as the inputs used in the PPIs, there could be some scope for increases in input quality to boost these indexes. Therefore, we will assume that the bias related to unobserved structure quality for these indexes is half of the bias that we assume for the single-family price index. The quality bias in the other price indexes used as deflators—the Handy Whitman index, the AUS telephone index, and the Turner Building Cost Index—is unclear, as we do not have much information on their methodologies. Since they are also based on input costs rather than property sales prices, we will also assume that the bias from unobserved quality in these indexes is half as large as for the single-family price index.

Next we assess the potential for sources of bias beyond unobserved structure quality. One issue is that structure value is not observed directly for most homes under construction, but rather is assumed to be a constant fraction of total house value. This assumption would lead to an upward bias in the single-family deflator if the share of land had, in fact, risen over time. Prior research has found that the share of house value attributable to land has risen since the 1980s as land prices have risen more than structure prices (Case 2007; Davis and Heathcote 2007; Davis and Palumbo 2008; Davis, Larsen, Oliner and Shui 2021). However, most of this research has measured the average land share for all existing residential structures in the US. Buyers of new homes may react to higher land prices by substituting towards smaller lots (Molloy, Nathanson, and Paciorek 2022) or to areas where land prices are lower, reducing the land share for newly built homes. Davis, Larsen, Oliner and Shui (2021) measure the land share for homes that are less than 10 years old and find that the average land share only increased from 38 percent in 2012 to 41 percent in 2019. Moreover, this increase was concentrated in a small number of counties where land prices are high and new construction is less common. If we take their estimates of land shares by county and calculate a weighted average of the change in land share using single-family construction as weights, we find that the land share did not increase at all from 2012 to 2019. Although this analysis covers only a short sample period, a survey conducted by the NAHB found that the ratio of finished lot costs to sales price for new single-family homes was actually lower in 2019 than it was in 1998 (the first available year).[19] Not only do we suspect that land shares for new single-family homes may not have risen that much from 1987 to 2019, but the bias to measurement of real construction output is mitigated by the fact that the estimate of nominal construction expenditures for new single-family homes uses the same assumption of a constant land share. Therefore any bias would be present in the numerator and denominator of the calculation for real single-family output and would cancel out. The bias from rising land shares in the single-family price index would only matter for other sectors of construction that use this price index in their deflator, since these other sectors measure nominal construction spending from structure values and do not use any assumptions about land shares.

---

[19] https://www.nahb.org/-/media/8F04D7F6EAA34DBF8867D7C3385D2977.ashx

One final measurement issue is that the deflators for some sectors are based on input prices rather than output prices. This will overstate the increase in the final cost of the structure because any productivity improvement should allow a structure to be produced at a lower cost, even if all of the input costs have not changed. Pieper (1991) finds little bias from this issue based on comparing productivity estimates for the period 1963-1982 using three different methodologies and finding similar growth rates. Another reason to suspect that this bias is small is that the growth rates of the ECI for construction workers and the PPI for residential maintenance and repair—two price indexes that measure input costs and are used as deflators—were only 0.5pp and 0.2pp higher, respectively, than the growth rate of the single-family price index from 1987 to 2019 (after adjusting the single-family price index downward for bias owing to unobserved structure quality). Since these indexes combined deflate 13 percent of total nominal construction expenditures, the bias to aggregate productivity growth would only be 0.05pp per year. We cannot conduct a similar analysis for the sectors of nonresidential construction that use the Handy Whitman or AUS telephone cost indexes because we do not have access to these price indexes, nor do we have any alternative measures of output prices to compare them to. If we assume that the bias for these sectors is similar to the bias that we calculated based on the ECI for construction workers and the PPI for residential maintenance and repair, then we would find an additional bias of 0.03pp per year.

Section 3.5 Implications for aggregate construction sector productivity

The implication for productivity growth in the aggregate construction sector depends on what portion of the construction sector is affected by each type of bias discussed above. Based on the fraction of nominal output associated with each deflator, our calculation that omitted quality could bias up the single-family price index by 0.8pp per year at most, and our assumptions about the role of omitted quality in other construction sector deflators, we estimate that omitted quality could have biased downward total construction productivity growth by 0.5pp per year. The other sources of measurement error discussed above may have contributed an additional 0.1pp per year (see Table 7). Cumulatively these factors add up to less than ¾ percentage point per year, even though the calculations are based on fairly generous assumptions. We have based our calculations on generous assumptions in order to determine the largest possible role for measurement error in explaining low productivity growth. More modest assumptions would, of course, reduce the magnitude of our estimates and make the case for measurement error even weaker

Adding the cumulative bias to reported productivity growth, we estimate that productivity was essentially flat in the construction sector from 1987 to 2019 (see Figure 3). From one perspective, this could be considered a material difference from the published data because the level of bias-adjusted productivity in 2019 was 21 percent higher than the published level. From another perspective, the bias-adjusted estimate dos not change the qualitative result that productivity growth in this sector has been quite low. Figure 3 illustrates that our bias-adjusted estimate of productivity growth in the construction sector remains much lower than in other industries.

Section 4. Regional evidence on productivity growth

In this section, we first describe how we calculate new estimates of productivity growth in the new single-family construction sector from 1980 to 2019 for states and metropolitan areas. Next, we explore what local characteristics are associated with productivity growth, such as initial housing costs, the proximity of new construction to the downtown area, and physical barriers to construction.

Section 4.1. Productivity growth across states and metropolitan areas

Because the Bureau of Labor Statistics does not publish estimates of construction productivity by geography, we calculate our own estimates. For the numerator of our productivity estimates, we calculate the total quantity of housing structure produced in a year as the number of single-family housing units permitted in that year multiplied by the average square footage of homes built in that year. We focus on the production of new single-family homes because we do not have data on real quantities for other types of structures. The permit data are from the Census Bureau's Residential Construction branch and the square footage data are from CoreLogic's tax assessor data.

For the denominator of our productivity estimates, we calculate labor input as the total number of employees in the construction industry because the available data on employment by industry do not allow for clear estimates of the number of people working specifically on new residential construction.[20] For the state-level data, estimates are similar when based on the number of workers in "construction of buildings" (NAICS 236) or the number of residential construction workers calculated from a set of 6-digit NAICS sectors related to residential construction.[21] Specifically, the correlations of our baseline measure of productivity growth with measures that use these two alternate employment definitions are 0.89 and 0.87, respectively. These employment measures are based on establishment-level data, which seems appropriate for state-level estimates.[22] However, for smaller areas like counties or metropolitan areas, it seems likely that a non-trivial amount of construction work could take place outside the location of the establishment. Therefore, for the metro-level employment estimates we use the number of construction workers in the Decennial Census and American Community Survey, which record the number of workers living in a given location. It seems more likely that construction workers work on projects in the metropolitan area where they reside.

The permit data and state-level employment data are annual so we are able to calculate annual productivity estimates by state. Our metro-level estimates cover 1980, 1990, 2000, 2010 and 2019. The sample period for both types of geographies start in 1980 because that is the first available year of the permit data.[23]

Just as with the national productivity growth estimates, our estimates of local productivity growth are likely biased downward because they do not account for improvements in the quality of homes beyond

---

[20] Identifying all workers in new single-family construction is difficult because many construction firms employ specialty trade contractors. In the NAICS classification, specialty trade contractors can be separated into residential and nonresidential sectors. But the residential specialty trade contractors will include people working on remodeling as well as new construction. In the SIC classification, specialty trade contractors cannot be disaggregated into residential and nonresidential.

[21] Data on employment by industry and state are from the QCEW, which has annual data by NAICS starting in 1990. For each state, we extend the estimate of workers in "construction of buildings" back to 1980 using estimates from SIC category 1521 (single-family housing construction) and SIC category 1531 (operative builders). For the 10 years of overlap between the SIC and NAICS data, the correlation of state-level employment growth rates is 0.9.

[22] We exclude the District of Columbia because establishments located in DC likely work on projects in Virginia and Maryland.

[23] When we average the metropolitan area estimates by state (using housing unit weights for metropolitan areas in multiple states), the correlation with state-level productivity growth is 0.87. We find this strong correlation reassuring given that the metropolitan area estimates use a different source for construction employment and are based on decadal data rather than annual data.

unit square footage. Therefore, we multiply our estimates by a trend that increases by 0.8 percent per year, the largest-possible magnitude of the bias for single-family homes found in section 3. Because this adjustment is the same for all locations, it does not affect the regression results reported below. Ideally, we would adjust the productivity estimates based on the magnitude of changes in non-size related housing characteristics for each location. But we do not have reliable data on changes in these other structure characteristics by state or metropolitan area. If changes in these other structure characteristics are not correlated with structure size but are correlated with location characteristics like housing supply regulation, the correlations that we report in this section could be biased. Appendix Table 6 shows that some measures like number of bedrooms and number of bathrooms are strongly correlated with size, while other characteristics like presence of central air conditioning or a porch are weakly correlated with size. We leave this issue as a possible limitation of our productivity growth estimates and hope that further research can develop more comprehensive measures of structure output at the local level.

Because productivity is cyclical and noisy, calculating the average growth rate from the first year of the sample to the last year of the sample could be an imperfect measure of the long-run trend in local productivity. Instead, we regress the natural logarithm of annual productivity in each location on a time trend. The coefficient on the time trend provides an estimate of the average productivity growth rate in the state that is more robust to the start and endpoints of the time series.

The average permit-weighted estimate of productivity growth in our state-level and metro-level samples are -0.6 and -0.7 percent per year, respectively. These estimates are somewhat lower than average growth of aggregate (bias-adjusted) productivity from 1987 to 2019, which we estimate to have been about 0.1 percent per year.[24] It is plausible that productivity growth in the single-family sector has been lower than other types of construction because the projects and firms tend to be smaller. Indeed, Sveikauskas, Rowe and Mildenberger (2018) find that single-family productivity growth was lower than multifamily productivity growth from 1987 to 2016. Another possibility is our estimates of productivity growth may be biased down by more than the aggregate estimates because we do not account for improvements in the quality of homes beyond unit square footage, whereas the aggregate estimates account for changes in some other aspects of structure quality.

Figure 4 show the estimates of productivity growth for each state. The estimates vary notably by state, with productivity growth around -3 percent per year at the low end and around 2 percent per year at the high end. About 80 percent of the states have an estimate of productivity growth near or below zero. In that sense, new single-family productivity growth appears to have been low throughout much of the United States.

The figure shows that the states with the lowest productivity growth are Connecticut, Rhode Island and Vermont, small states that tend to be relatively densely populated. Massachusetts, New York and California—states with relatively strict regulatory constraints on new construction—also have fairly low productivity growth. Meanwhile states with relatively high productivity growth include West Virginia, South Carolina and Montana.

---

[24] We find a similar estimate of aggregate productivity growth when we use the coefficient on a linear trend rather than calculating average annualized growth from 1987 to 2019.

Section 4.2. Correlation of productivity growth with local characteristics

To examine local attributes that are related to productivity growth, we turn to the metropolitan area estimates. Table 8 reports results from a regression of productivity growth on various local characteristics. We focus on characteristics that seem like they could plausibly be related to construction costs, and therefore productivity growth. All attributes are standardized to have a mean equal to zero and standard deviation equal to 1. In addition, some specifications include region fixed effects.

The first column of Table 8 shows results for about 300 metro areas, which is the full sample of metro areas for which we were able to estimate productivity growth. Productivity growth is negatively associated with median housing values in 1980, perhaps because more expensive areas tend to have higher housing supply constraints. Relatedly, productivity growth also tends to be weakly lower in areas where buildable area is constrained by a higher share of water to total land and water area.

We also examine whether the density of new construction is related to productivity growth in two different ways. First, we measure average density in 1980 as housing units per square kilometer. Second, we measure the fraction of single-family construction that took place in suburban or exurban locations, defined as the fraction of single-family units built 1980 to 2019 that are far from the city hall, where "far" is defined as more than the median distance between the city hall and all single-family homes built 1930 to 1979.[25] This fraction is positively related to productivity growth, indicating that productivity growth has been higher on average in areas where more construction has taken place outside the urban core. Meanwhile, average density in the metro area is unrelated to productivity growth, perhaps because density varies widely within metro areas.   We also do not find a correlation between productivity growth and initial metro size, measured as the log of housing units in 1980. The standard error is small enough that we can reject that the coefficient on initial city size is less than -0.3 with a p-value of 0.06. We find a similar result when we measure city size using land area. Finally, productivity growth is positively associated with growth in the housing stock.[26] This relationship could reflect the fact that locations with higher productivity growth have been able to grow more. It could also reflect the possibility that locations with more teardowns have a lower net increase in the housing stock, and the cost of teardowns could have risen faster than the cost of building on vacant land. Column 2 of the table shows these results are robust to controlling for Census region indicators.[27] Interestingly, even conditional on these attributes, productivity growth has been notably higher in the South.

The models reported in columns 3 and 4 include the Wharton Residential Land Use Regulation Index. We average the results from two waves of the survey that were conducted in 2006 and 2018 because

---

[25] The location of city hall is mainly from Holian (2019). We add a few locations that are missing from that dataset by searching for the city hall location on the internet. We measure housing unit location by year built using tax assessor data from CoreLogic.

[26] This association is not mechanical for at least two reasons. Areas with stronger gains in housing units may also have faster construction employment growth, and so might not have stronger productivity growth on average. Also, depreciation including teardowns will lead to differences between the number of permits issued and changes in the net housing stock.

each survey measures the true amount of regulation with noise. We find productivity growth is negatively associated with the Wharton regulation index, as shown in columns 3 and 4. Moreover, including this variable reduces the correlation between house value and productivity growth, suggesting that the correlation with house value mainly reflected regulatory constraints. These results complement the findings in D'Amico et al. (2024), who find a correlation between regulation, firm size, and the *level* of productivity.   Similarly, Sveikaukas et al. (2016) find a small negative correlation between state-level changes in regulation and construction productivity growth. We suspect that we find a larger correlation with regulation with regulation than they do because annual changes in their measure of regulation are noisy.

Table 9 shows that a range of measures of regulation tend to be negatively correlated with productivity growth at the metro and state level. The correlations tend to be stronger at the state level than at the metro level, perhaps because regulation and productivity growth are both measured with more noise at the metro level. The 2006 wave of the Wharton survey, which is the one used by D'Amico et al. (2024), has the strongest correlation with productivity growth among the various regulatory measures.

To dig a little deeper into how regulation may reduce productivity growth, Table 10 shows the correlation of productivity growth at the metropolitan level with each of the components of the Wharton survey. The strongest and most robust correlation is with approval delays, which capture average review times for a range of types of residential projects including by-right projects (permitted under current rules), not-by right projects (requiring exemptions to current rules) and subdivision approvals. According to the version of the Wharton survey conducted in the late 1980s, permit approval times were low across the US (Linneman et al. 1990). Hence, it seems that increases in regulatory delays have reduced productivity growth. Supply restrictions—limits on the number of permits issued or on the size of multifamily buildings—also tend to be negatively correlated with productivity growth, perhaps because these limits prevent builders from taking advantage of economies of scale. Impact fees, which raise the cost of construction projects, also tend to be negatively correlated with productivity growth. Moreover, areas where a larger number of local entities (such as zoning boards, planning commissions, and environmental review boards) are needed to approve a by-right building project tend to have lower productivity growth. In addition to lengthening approval timelines, the need for a large number of approvals can reduce the likelihood that a project is approved, thereby reducing the amount of work that a construction firm can accomplish. Finally, locations where the state legislature is more involved in influencing residential building activity or growth management also tend to have lower productivity growth. This result may be surprising because states should have some incentive to encourage development in at least some parts of the state. However, it could be that state involvement further complicates the approval process or ends up requiring higher-cost designs.

Returning to Table 8, the final column shows that dropping the variables that are unrelated to productivity growth, the 7 remaining variables explain 24 percent of the heterogeneity in productivity growth across metro areas. While this percentage is not immaterial, much of the geographic variation remains unexplained.

Section 5. Discussion and conclusions

In sum, our findings suggest that the mismeasurement of construction-sector deflators has likely biased down estimates of construction productivity growth by ¾ percentage points per year, at most. Though this bias is not negligible, it is modest enough that we do not overturn the conclusion that construction

productivity growth has indeed been quite low—still near zero—and much lower than productivity growth in other sectors. All things considered, it seems unlikely that the mismeasurement of construction productivity growth itself has had an effect on housing policy or on the construction-sector labor market. Moreover, because the construction sector's share of nominal aggregate output averaged only 11 percent over our sample period, the implications of this bias for aggregate productivity growth are negligible.

Beyond the deflators, we doubt that mismeasurement of other components of construction productivity has led to material downward bias. Specifically, we do not have any particular reason to think that nominal output growth would be biased down by a large amount. And mismeasurement of labor input may have been biasing productivity growth *upward,* since one large source of measurement error in labor input is an undercount of undocumented workers.[28] This undercount may easily have become larger from the 1980s to the 2010s as the unauthorized immigrant population expanded.[29] With labor input having grown by a larger amount than measured, measured growth in labor input would be biased downward, leading to an upward bias in productivity growth. Goolsbee and Syverson (2023) also find little role for labor input in explaining low growth in construction-sector productivity.

Further evidence supporting low productivity growth can be found in statistics measuring the average length of time from start to completion of single-family homes.[30] This timeline increased from 6.2 months in the mid-1980s to 7.0 months in 2019, suggesting that any time-saving productivity improvements have been more than offset by delays elsewhere in the construction process.[31]

Low productivity growth in the construction sector is not unique to the United States. A study by McKinsey shows that construction productivity growth from 1995 to 2015 was less than 1 percent per year in 22 out of the 38 countries that they examined and it was negative in 12 of these countries (McKinsey 2017, Exhibit E2). Countries with low productivity growth include developed countries like the US, France and Spain as well as less-developed countries like Malaysia and Columbia. The common international experience of low productivity growth bolsters the conclusion that low growth is not due to measurement error.

The fact that construction productivity growth has truly been low for at least the past three decades has many important implications. For one, the rising cost of housing has reduced housing affordability, which has likely influenced household formation decisions as well as other household spending decisions. And because housing cost increases have not been uniform across the country, differential changes in housing affordability may have caused workers to make different location choices (see e.g. Hsieh and Moretti 2019 and Ganong and Shoag 2017). Stagnant construction productivity has also likely had material effects on wages, labor supply and labor demand in the construction industry.

---

[28] Labor input is defined as the total number of annual hours worked of all people in the industry. Any possible mismeasurement of labor quality would result in mismeasurement of total factor productivity, not labor productivity.

[29] The Pew Research Center estimates that the unauthorized immigrant population tripled from 1990 to 2017. https://www.pewresearch.org/fact-tank/2021/04/13/key-facts-about-the-changing-u-s-unauthorized-immigrant-population/

[30] https://www.census.gov/construction/nrc/data/time.html

[31] Using the Census Survey of Construction microdata files, we still find a lengthening of construction timelines from 1999 to 2019 after controlling for structure size and location.

Due to the extensive implications of low productivity growth in the construction sector, it is important for researchers to explore why productivity growth has been so low. The state-level and metro-level estimates of residential productivity growth that we create and present in this paper provide suggestive evidence in this direction. We show that productivity growth has been lower in areas with more land use regulation. D'Amico et al. (2024) show that regulation of construction projects reduces developers' investment in technology, average revenues and average revenues per employee. More generally, these regulations can increase the cost of construction by creating delays in the construction process, increasing overhead costs, and requiring higher-cost designs. Indeed, we find that permit approval times are the component of the regulation index that is most strongly correlated with productivity growth. Millar, Oliner and Sichel (2016) find that time-to-plan for nonresidential construction projects has lengthened by more in metropolitan areas with more restrictive land use regulation. And Brooks and Liscow (2023) find that increases in Interstate infrastructure costs have been associated with an increase in land use litigation.

Beyond regulation, our results show that productivity growth has been lower in areas with more construction in the urban core rather than suburban or exurban communities. Building in the urban core can be more costly because the density of existing structures is higher. It is difficult to take advantage of gains to scale on small parcels of land where only one or two homes can fit compared with large developments of hundreds of new homes. Moreover, construction in the urban core is more likely to be a teardown, which adds to the cost relative to building on vacant land. To illustrate that more construction today takes place in denser areas than in the past, we calculate the pre-existing population density where new homes were being built in the early 1990s and compare with the population density in areas where new homes were built in the late 2010s. Figure 5 shows the distribution of population density for each cohort of homes. Indeed, new homes built between 2016 and 2019 were more likely constructed in tracts with a population density above 3000 persons per square mile, while new homes built between 1991 and 1994 were more likely to be built in tracts with less than 100 persons per square mile.

Finally, it is worth highlighting one of the null results of our analysis: productivity growth has not been lower in larger metropolitan areas than smaller metropolitan areas. This result is important because it suggests that city size itself does not become a drag on future productivity growth as cities grow. In other words, there is no evidence that productivity growth is low in larger cities because they have become fully built-out.

While the cross-sectional analysis in our paper is informative, about ¾ of the heterogeneity in productivity growth across locations remains unexplained by the geographic characteristics that we examine. Other factors are clearly important for explaining the variation in productivity growth across space, and these factors also might shed light on why aggregate productivity growth has been so low. For example, it would be valuable for future research to directly examine the construction costs of building teardowns one at a time relative to building multiple homes in larger subdivisions. Relatedly, future research should also examine the potential drag on productivity growth from the rising share of home improvements in total construction spending, as it seems plausible that renovation is more costly than new construction.

Another important area for future research is the potential role that modular or manufactured homes could play in boosting construction productivity. These types of homes allow for much more output per

worker because they take advantage of factory-production methods and returns to scale. Even though the technology to produce this type of housing has existed for many decades, it is still not common, perhaps owing to building codes and other regulations (Schmitz 2020). More generally, construction is still a very labor-intensive industry with low investment in intellectual property. Productivity growth in the education services industry—another sector that has a very low capital-to-labor ratio—has also been very low over the past 3½ decades. Future work should explore these and other possible explanations for why productivity growth in the construction sector has been so low.

References

Adelino, M. and Robinson, D. (2022) The Environmental Cost of Easy Credit: The Housing Channel. Working paper.

Allen, S. G. (1985). Why Construction Industry Productivity Is Declining. The Review of Economics and Statistics, 67(4):661-669.

American Institute of Planners. 1976. Survey of State Land Use Planning Activity. Washington, DC: Department of Housing and Urban Development.

DC: Department of Housing and Urban Development.Brandsaas, E., Garcia, D., Nichols, J. and Sadovi, K. (2023). "Nonresidential construction sending is likely not as weak as it seems." FEDS Notes. https://doi.org/10.17016/2380-7172.3283

Bartik, Alexander and Gupta, Arpit and Milo, Daniel, The Costs of Housing Regulation: Evidence from Generative Regulatory Measurement (September 14, 2024). Available at SSRN: https://ssrn.com/abstract=4627587 or http://dx.doi.org/10.2139/ssrn.4627587

Brooks, L. and Liscow, Z. 2023. Infrastructure Costs. American Economic Journal: Applied Economics 15(2): 1-30.

Case, K. E. (2007). The value of land in the United States 1975-2005. In Ingram, G. K. and Hong, Y.-H., editors, Land Policies and Their Outcomes, pages 127-147. Lincoln Institute of Land Policy.

Coulson, N.E., Morris, A.C. and Neill, H.R. (2019), Are New Homes Special?. Real Estate Economics, 47: 784-806. https://doi.org/10.1111/1540-6229.12165

D'Amico, Leonardo and Glaeser, Edward L. and Gyourko, Joseph and Kerr, William R. and Ponzetto, Giacomo A. M. (2023). Why Has Construction Productivity Stagnated? The Role of Land-Use Regulation. https://ssrn.com/abstract=4679195 or http://dx.doi.org/10.2139/ssrn.4679195

Davis, M. and Heathcote, J. (2007). The price and quantity of residential land in the United States. Journal of Monetary Economics, 54(8):2595-2620.

Davis, M. and Palumbo, M. (2008). The price of residential land in large us cities. Journal of Urban Economics, 63(1):352-384.

Davis, M. A., Larson, W. D. , Oliner, S. D. , and Shui, J. (2021). The price of residential land for counties, ZIP codes, and census tracts in the United States. Journal of Monetary Economics, 118(C):413-431

Diewert, E., Heravi, S., and Silver M. (2008). Hedonic Imputation versus Time Dummy Hedonic Indexes. NBER Working Papers 14018, National Bureau of Economic Research, Inc.

de Leeuw, F. (1993). A Price Index for New Multifamily Housing. Survey of Current Business 1993-02, Bureau of Economic Analysis.

Eriksen, M. and Orlando, A. (2022). "The Causes and Consequences of Development Costs: Evidence from Multifamily Housing." *SSRN https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4225004.*

Ganong, P. and Shoag, D. (2017). Why has regional income convergence in the U.S. declined? Journal of Urban Economics, 102(C):76 90.

Glaeser, E. L. and Ward, B. A. (2009). The causes and consequences of land use regulation: Evidence from Greater Boston. Journal of Urban Economics, 65(3):265-278.

Goodrum, P., Haas, C., and Glover, R. (2002). The divergence in aggregate and activity estimates of US construction productivity. Construction Management and Economics, 20(5):415-423.

Goolsbee, A.D. and Syverson, C. (2023). The strange and awful path of productivity in the U.S. construction sector. NBER Working Paper 30845.

Grimm, B. T. (2003). New Quality Adjusted Price Index for Nonresidential Structures. Working paper 2003-03, Bureau of Economic Analysis

Gyourko, J., Hartley, J. and Krimmel, J. (2021). "The Local Residential Land Use Regulatory Environment Across U.S. Housing Markets: Evidence from a New Wharton Index" *Journal of Urban Economics 124.*

Gyourko, J., Saiz, A., and Summers, A. (2008). "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index" *Urban Studies.*

Haas, C., Allmon, E., Borcherding, J. D., and Goodrum, P. M. (2000). U.S. Construction Labor Productivity Trends, 1970-1998. Journal of Construction Engineering and Management, 126(2).

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

Jackson, K. (2016). Do land use regulations stifle residential development? Evidence from California cities. Journal of Urban Economics, 91(C):45-56.

Linneman, Peter, Anita Summers, Nancy Brooks, and Henry Buist. 1990. "The State of Local Growth Management." Wharton Real Estate Center Working Paper 81.

Steven Manson, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. IPUMS National Historical Geographic Information System: Version 17.0 [dataset]. Minneapolis, MN: IPUMS. 2022. **http://doi.org/10.18128/D050.V17.0**

McKinsey Global Institute. (2017). *Reinventing Construction: A Route to Higher Productivity*. Retrieved from
https://www.mckinsey.com/~/media/mckinsey/business%20functions/operations/our%20insights/reinventing%20construction%20through%20a%20productivity%20revolution/mgi-reinventing-construction-a-route-to-higher-productivity-full-report.pdf

Millar, J. & Oliner, S. & Sichel, D. (2016). Time-to-plan lags for commercial construction projects. Regional Science and Urban Economics, Elsevier, vol. 59(C), pages 75-89.

Molloy, R. and Nathanson, C. G. and Paciorek, A. (2022). Housing supply and affordability: Evidence from rents, housing consumption and household location. Journal of Urban Economics, Elsevier, vol. 129(C).

Musgrave, J. C. (1969). The Measurement of Price Changes in Construction. Journal of the American Statistical Association, vol. 64, pages 771-786.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence, Journal of Business & Economic Statistics, 37:2, 187-204, DOI: 10.1080/07350015.2016.1227711

Paul P.E. (1991). The Measurement of Construction Prices: Retrospect and Prospect. NBER Chapters, in: Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth, pages 239-272. National Bureau of Economic Research, Inc.

RSMeans (1987). 1987 Square Foot Costs with RSMeans data.

RSMeans (2018). 2018 Square Foot Costs with RSMeans data. Gordian.

Saks, Raven E. (2008). "Job Creation and Housing Construction: Constraints on Metropolitan Area Employment Growth" *Journal of Urban Economics*.

Schmitz, J. A. (2020). Solving the Housing Crisis will Require Fighting Monopolies in Construction. Working Papers 773, Federal Reserve Bank of Minneapolis.

Sichel, D. E. (2022). The Price of Nails since 1695: A Window into Economic Change. Journal of Economic Perspectives, 36(1):125-150.

Stokes, H Kemble, J. (1981). An Examination of the Productivity Decline in the Construction Industry. The Review of Economics and Statistics, 63(4):459-502.

Sirmans, G., Macpherson, D., and Zietz, E. (2005). The Composition of Hedoing Pricing Models. Journal of Real Estate Literature, vol. 13, pages 3-43

Sirmans, G., MacDonald L., Price, J., Macpherson, D., and Zietz, E. (2006). The Value of Housing Characteristics: A Meta Analysis. The Journal of Real Estate Finance and Economics, vol. 33, pages 215-240.

Sveikauskas, L., Rowe, S., Mildenberger, J., Price, J., and Young, A. (2016). Productivity Growth in Construction. Journal of Construction Engineering and Management, 142(10).

Sveikauskas, L., Rowe, S., Mildenberger, J., Price, J., and Young, A. (2018). Measuring productivity growth in construction. Bureau of Labor Statistics Monthly Labor Review.

Teicholz, P. (2013). Labor-productivity declines in the construction industry: Causes and remedies (another look). Working paper, AECbytes.

Table 1
Average Productivity Growth by Major Industry, 1987 to 2019

	Percent Change Annual Rate
Construction	-0.4
Services	1.2
Transportation	1.3
Agriculture	1.3
Mining	1.6
Nondurable Manufacturing	1.7
Finance, Insurance and Real Estate	2.0
Utilities	2.2
Durable Manufacturing	2.9
Trade	3.2
Information	4.7

Source. Bureau of Labor Statistics, Office of Productivity and Technology.

Table 2
Deflators Used in the Construction Sector

	Nominal Output Share 1987-2019
Price index for new single-family homes under construction	46.5
Used for the new single-family sector	31.6
Used for residential improvements	6.6
Used for nonresidential sectors	8.3
Price index for new multifamily units under construction	2.3
BEA multifamily price index	2.5
Nonresidential Producer Price Indexes	
Warehouse PPI	3.1
Office PPI	2.5
Industrial PPI	2.7
New school PPI	0.9
Health care	0.8
Home maintenance PPI (input cost index)	6.6
ECI for construction workers (input cost index)	6.6
Turner building cost index	8.3
BEA nonresidential price indexes	8.1
Handy-Whitman cost indexes (input cost index)	6.9
AUS telephone cost index (input cost index)	2.5

Source. Authors' calculations based on nominal output data from the Bureau of Economic Analysis and summaries of NIPA methodology from various years.

Table 3
Construction Cost Estimates from R.S. Means

	Cost in 1987	Cost in 2019	Percent Change (annual rate)
1-story home			
Economy wood siding (800sf, 1 bath)	51,010	157,089	3.6
Average wood siding (1200 sf, 1 bath)	77,908	220,808	3.3
Custom brick veneer (1800sf, 2½ baths)	162,899	443,288	3.2
Luxury solid brick (2400sf, 2½ baths)	267,022	622,518	2.7
2-story home			
Economy wood siding (18000sf, 2 baths)	90,653	254,738	3.3
Average wood siding (2200sf, 2 baths)	114,930	321,208	3.3
Custom brick veneer (2800sf, 3½ baths)	202,876	548,146	3.2
Luxury solid brick (3600sf, 3½ baths)	327,282	790,914	2.8

Note. Economy and average homes are assumed to have an asphalt roof, a 1-car garage, an open porch and breezeway, and laminate kitchen countertops. Custom and luxury homes have a cedar shake roof, a 2-car garage, an enclosed porch and breezeway, and marble kitchen countertops. All 1-story homes have a 30-gallon gas water heater. For 2-story homes, the economy and average homes have a 30-gallon gas water heater while the custom and luxury homes have a 50-gallon gas water heater. Economy homes have 2 kitchen cabinets and 6 linear feet of countertops. Average, custom and luxury homes have 3, 4 and 5 cabinets and 14, 20 and 25 linear feet of countertops, respectively. The custom and luxury homes have a burglar alarm. All homes have air conditioning as well as a broom closet, smoke detector, dishwasher, garbage disposal, refrigerator, range, oven, microwave, washing machine, and dryer.

Table 4
Effect of Structure Quality on Ln(Home Value)

	CoreLogic	American Housing Survey	
	(1)	(2)	(3)
Built in 2019 (relative to 2000)	0.569	--	--
	(0.018)		
Tax assessor rating = high quality	0.164	--	--
	(0.026)		
Tax assessor quality rating missing	0.064	--	--
	(0.024)		
Built 2000-19 (relative to 1970-89)		0.956	0.938
		(0.006)	(0.006)
Resident quality rating = 10	--	0.159	0.153
(relative to 7 or less)		(0.009)	(0.009)
Dishwasher			0.237
			(0.009)
Washing machine			0.130
			(0.027)
Dryer			0.015
			(0.019)
Years	2000-2019	1985-2019	1985-2019
Housing characteristics	Yes	Yes	Yes
Division FE	Yes	No	No
Region FE	No	Yes	Yes
Number observations	3,398,815	35,298	35,298
Adjusted R2	0.59	0.68	0.68

Source. CoreLogic Residential Real Estate database and American Housing Survey National Samples 1985, 1987, 1989, 2015, 2017, 2019. CoreLogic sample includes newly-built single-family detached homes and the dependent variable is the home's sales price "High quality" are homes designated by the property tax assessor as "luxury", "excellent", "above average" or "good." Standard errors in column 1 are clustered by county. AHS sample restricted to single-family detached homes built 1970-1989 and appearing in the 1985-89 samples or built 2000-19 and appearing in the 2015-19 samples. The dependent variable in AHS is the resident's estimate of home value. The resident's quality rating is on a scale from 1 to 10, but few respondents rate their home quality below 7 (see Appendix Table 3). In CoreLogic the housing characteristics are unit square footage, number of bedrooms, indicators for fireplace, garage, basement, and various types of exteriors. In the AHS the housing characteristics are indicators for unit square footage, number of bedrooms, number of bathrooms and presence of central air conditioning, a fireplace, a garage and a basement. Indicators for appliances are three separate indicators for the presence of a clothes washer, a dryer and a dishwasher.

Table 5
Energy Use of Newly-Built Homes

	Ln(Energy Use)	Energy Use
Built 2000-2019	-0.223	-742
	(0.004)	(14)
Unit square footage		
1000 to 1499 sf	0.223	477
	(0.012)	(38)
1500 to 1999 sf	0.332	788
	(0.012)	(37)
2000 to 2499 sf	0.419	1062
	(0.012)	(38)
2500 to 2999 sf	0.498	1344
	(0.012)	(40)
3000 to 3999 sf	0.599	1735
	(0.012)	(40)
>=4000 sf	0.685	2103
	(0.013)	(43)
Region		
Midwest	-0.150	-622
	(0.008)	(27)
South	-0.129	-520
	(0.008)	(24)
West	-0.226	-795
	(0.008)	(26)
Constant	7.877	3245
	(0.012)	(40)
Number of observations	40,722	40,722
Adj. R2	0.158	0.167

Source. American Housing Survey National Samples 1985, 1987, 1989, 2015, 2017, 2019. Sample restricted to single-family detached homes built 1970-1989 and appearing in the 1985-89 samples, or built 2000-19 and appearing in the 2015-19 samples. Energy use defined as total annual expenditures on electricity, natural gas, heating oil, water and other fuels, deflated by the Consumer Price Index for Utilities.

Table 6
Estimated Bias from Unobserved Quality Following Oster (2019)

	Coefficient on Time Period Indicator	R^2	Implied Structure Price Growth Rate (annual rate)
Regression estimates in AHS data			
Baseline	1.191	0.48	4.05
Census controls	0.989	0.66	3.35
Implied unbiased coefficient	0.767	0.86	2.59
Regression estimates in CoreLogic data			
Baseline	0.637	0.27	3.41
Census controls	0.563	0.59	3.01
Implied unbiased coefficient	0.521	0.77	2.79

Note. Baseline regression includes U.S. Census region indicator variables as controls. The 1985-89 AHS sample includes homes built from 1970 to 1989 and the 2015-19 AHS sample includes homes built from 2000 to 2019; the time period indicator is equal to one for properties built after the 2000s and zero otherwise. The CoreLogic dataset covers the period from 2000 to 2019; the time period coefficient reported is for the year 2019 (2000 is the omitted year). The implied unbiased coefficient assumes that δ=1 and that the R^2 of the regression including all unobserved variables would equal 1.3 times the R^2 of the regression with Census controls. See text for more details.

Table 7
Contributions to Bias in Aggregate Construction Sector Productivity Growth

	Percentage Points Annual Rate
Unobserved structure quality	
Reduces SF and MF price indexes by 0.8pp per year	0.39
Reduces some other nonresidential price indexes by 0.4pp per year	0.11
SF and MF price indexes include land prices	0.00
Price indexes for some sectors based on input prices	0.08
Total	0.58
Published productivity growth	-0.45
Productivity growth adjusted for total bias	0.13

Source. Author calculations described in text.

Table 8
Correlation of Metro-Level Productivity Growth 1980-2019 with Various Local Characteristics

	(1)	(2)	(3)	(4)
Ln(housing units in 1980)	-0.109	-0.082	-0.049	--
	(0.119)	(0.122)	(0.123)	
Ln(housing units per square mile in 1980)	0.195	0.057	-0.068	--
	(0.136)	(0.159)	(0.164)	
Ln(housing units 2019/housing units 1980)	0.434	0.295	0.278	0.311
	(0.107)	(0.125)	(0.143)	(0.123)
Fraction of suburban and exurban construction	0.195	0.179	0.268	0.243
	(0.103)	(0.104)	(0.110)	(0.107)
Fraction water area	-0.149	-0.150	-0.171	-0.200
	(0.112)	(0.116)	(0.116)	(0.113)
Ln(median house value 1980)	-0.544	-0.413	-0.107	--
	(0.121)	(0.162)	(0.187)	
Wharton regulation index	--	--	-0.337	-0.422
			(0.175)	(0.151)
Midwest region	--	0.402	0.155	0.026
		(0.364)	(0.396)	(0.361)
South region	--	0.716	0.786	0.706
		(0.350)	(0.370)	(0.346)
West region	--	0.093	-0.492	-0.614
		(0.431)	(0.446)	(0.359)
Adjusted R2	0.14	0.14	0.23	0.24
Observations	289	289	212	212

Note. The dependent variable is the coefficient from regressing ln(productivity) on a time trend, multiplied by 100. All independent variables except the region fixed effects are standardized to have a mean equal to zero and standard deviation equal to one. Fraction of suburban and exurban construction is defined as the fraction of single-family units built 1980 to 2019 that are far from the city hall, where "far" is defined as more than the median distance between the city hall and all single-family homes built 1930 to 1979.

Table 9
Correlations of Various Measures of Housing Supply Regulation
with Construction Productivity Growth

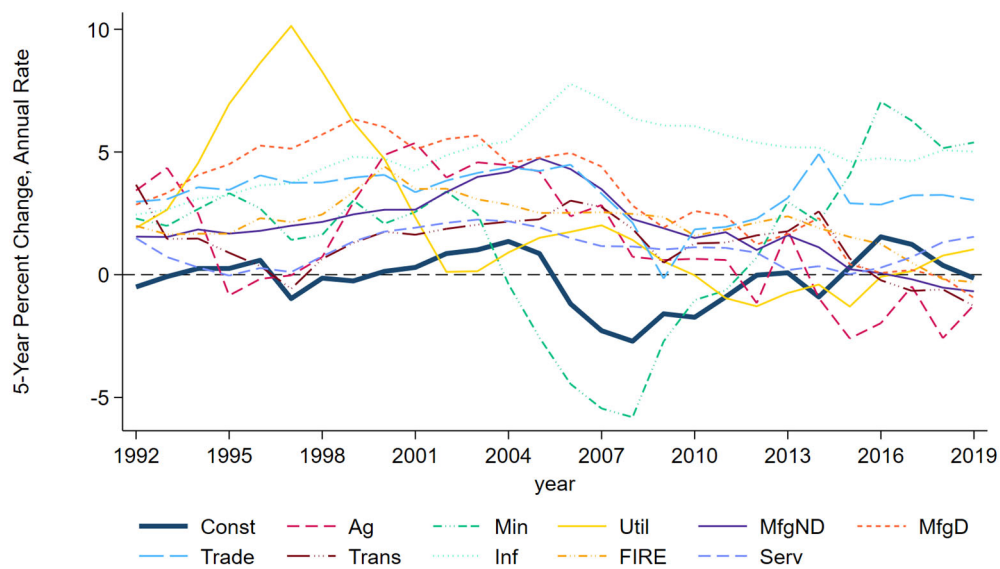|	Metro-Level Regressions						State-Level Regressions					
Wharton (1988)	-0.25 (0.14)											
Wharton (2006)		-0.56 (0.09)					-1.18 (0.16)					
Wharton (2018)			-0.27 (0.12)					-0.49 (0.22)				
Saks regulation index (1970s and 1980s)				-0.48 (0.17)								
Bartik et al. 1st Principal Component (2024)					-0.35 (0.11)				-0.94 (0.18)			
Bartik et al. 2nd Principal Component (2024)						-0.17 (0.10)				-0.91 (0.18)		
AIP (1976)											-0.76 (0.20)	
Zoning court cases per capita (1980-2010)												-0.58 (0.21)
R^2	0.02	0.12	0.02	0.10	0.03	0.01	0.54	0.08	0.35	0.33	0.22	0.12
Number observations	142	255	238	66	266	266	48	49	50	50	50	50

Note. Each column reports coefficient estimates from a bivariate regression of productivity growth on the measure of regulation named in the row. The 1988 Wharton survey is described in Linneman et al. (1990). The 2006 Wharton survey is described in Gyourko, Saiz and Summers (2008). The 2018 Wharton survey is described in Gyourko, Hartley and Krimmel (2021). The Saks regulation index combines 6 different sources of regulatory constraints from the 1970s and 1980s (Saks 2008). The AIP measure is from a survey conducted by the American Institute of Planners (1976). Zoning court cases per capita are from Ganong and Shoag (2017). The Bartik et al. (2024) measures are the first and second principal components from regulations measured using large language models to machine learning algorithms and recent local statutes and administrative documents. Estimates are provided for individual municipalities and townships; we aggregate to metropolitan areas and state using the average number of housing units 2019-2023 as weights. They interpret the first principal component as reflecting the complexity of the regulatory environment and the second as reflecting exclusionary zoning.

Table 10
Correlations of Various Types of Housing Supply Regulation
with Construction Productivity Growth

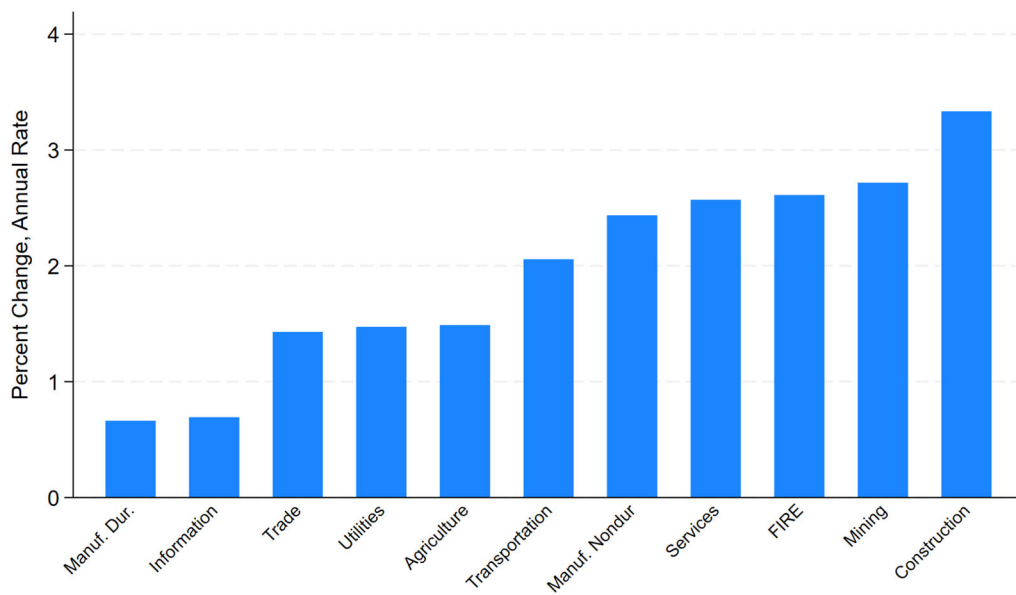|	2006	2018	Average of 2006 and 2018
Local political pressure	-0.12	0.08	-0.13
	(0.10)	(0.11)	(0.14)
State political involvement	-0.53	-0.13	-0.55
	(0.10)	(0.12)	(0.14)
Court involvement	0.07	-0.06	0.04
	(0.10)	(0.12)	(0.17)
Local zoning approval	0.10	0.16	0.33
	(0.10)	(0.11)	(0.17)
Local project approval	-0.25	-0.10	-0.36
	(0.10)	(0.11)	(0.15)
Local assembly	-0.16	0.17	-0.06
	(0.09)	(0.11)	(0.14)
Supply restrictions	-0.41	-0.26	-0.40
	(0.11)	(0.09)	(0.12)
Density restriction	-0.05	-0.19	-0.20
	(0.10)	(0.11)	(0.14)
Open space	-0.26	-0.06	-0.29
	(0.10)	(0.11)	(0.14)
Impact fees	-0.19	-0.29	-0.33
	(0.10)	(0.10)	(0.14)
Approval delay	-0.57	-0.45	-0.68
	(0.10)	(0.10)	(0.12)
Affordable housing	--	-0.41	--
		(0.11)	

Note. Each row shows the results of a separate regression of metro-level productivity growth on the type of regulation named in the row. See Gyourko, Hartley and Krimmel (2021) for a full description of each type of regulation. The first column reports results using the 2006 Wharton survey (sample size is 255), the second column reports results using the 2018 Wharton survey (sample size ranges from 242 to 245) and the third column reports results averaging the results from the two surveys (sample size ranges from 221 to 224). The affordable housing component was not asked in the 2006 survey. All types of regulation are standardized to have a mean of zero and standard deviation of one.

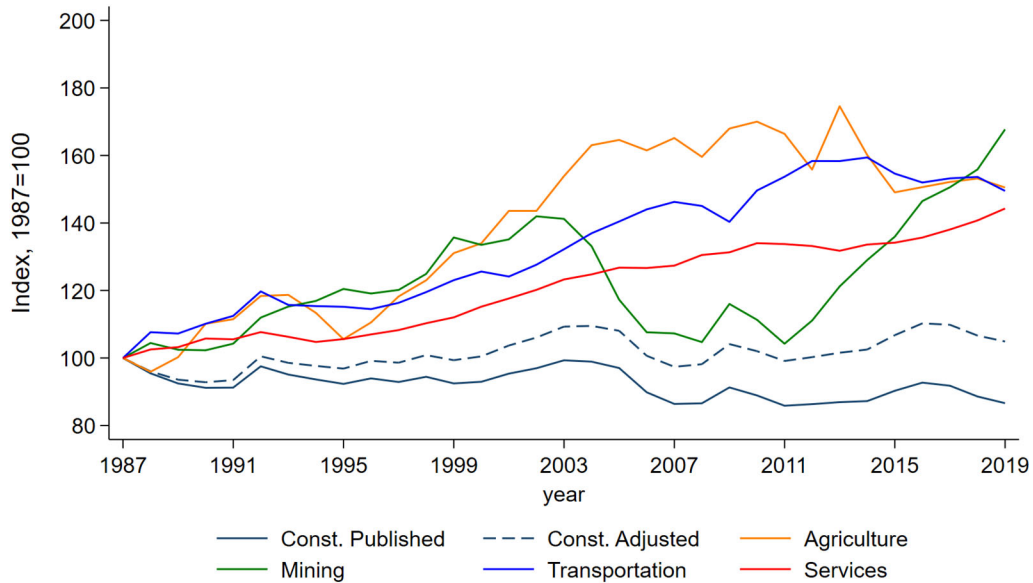**Figure 1**
Productivity Growth by Major Industry

Source. Bureau of Labor Statistics, Office of Productivity and Technology.

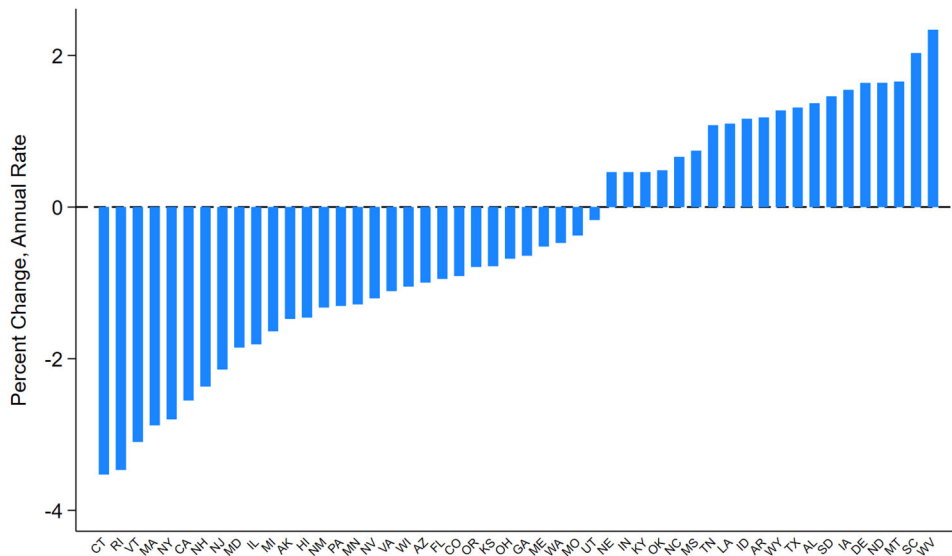**Figure 2**
Changes in Deflators for Major Industries

Source. Bureau of Labor Statistics, Office of Productivity and Technology.

Figure 3
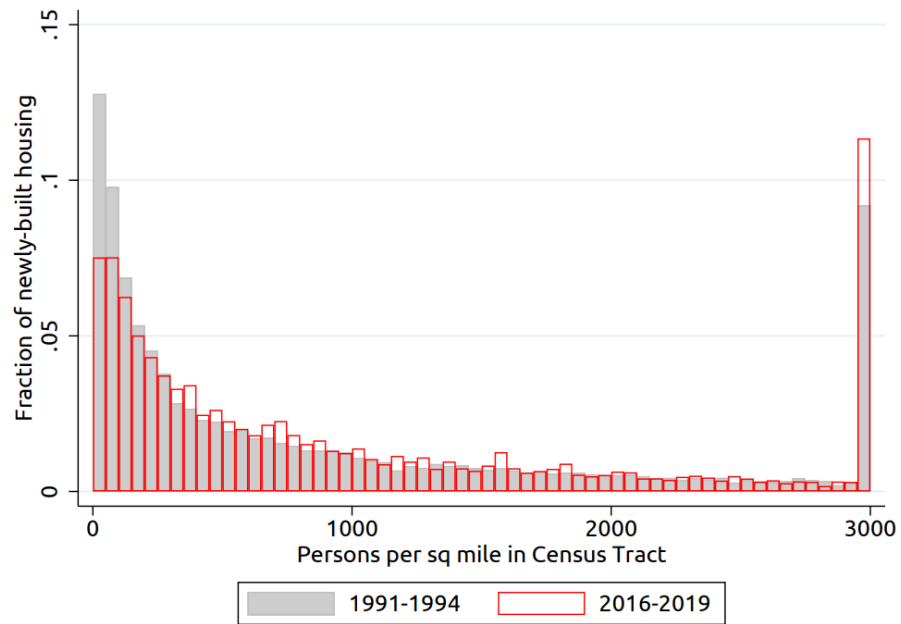**Productivity in Construction and Selected Major Industries**

Note. Source of the published productivity statistics is the Bureau of Labor Statistics, Office of Productivity and Technology. The adjusted construction estimates multiply the published estimates by a time trend that increases by 0.6 percent per year, our estimate of total bias shown in Table 7.

Figure 4
New Single-Family Residential Productivity Growth 1980-2019 by State

Note. Productivity is defined as new single-family permits multiplied by average structure size of units built in that year divided by the total number of construction workers. The numerator is multiplied by a trend that increases by 0.8 pecent per year to account for the maximum bias of single-family construction found in section 3. Growth is measured by regressing ln(productivity) on a linear time trend.

Figure 5
Density of new single-family housing

Source. IPUMS NHGIS time series tables, CoreLogic RRE, and authors' calculations. Sample of newly built housing restricted to single-family detached. Population density is measured in 1990 for 1991-1994 new housing and in 2010 for 2016-2019 new housing. Census tracts with more than 3000 persons per square mile are topcoded to 3000.

Appendix Table 1
Construction Subsector Output Shares

	Share of Nominal Construction Output 1987-2019
Residential	
New single-family	31.4
Improvements	19.7
New multifamily	4.8
Nonresidential	
Power	6.9
Office	6.7
Industrial	6.5
Health care	4.1
Lodging	2.6
Shopping malls	2.6
Telephone	2.5
Warehouse	1.8
Education	1.8
Amusement	1.6
Food establishments	1.2
Land transportation	1.0
Other	4.9

Source. Authors' calculations based on nominal output data from the Bureau of Economic Analysis.

Appendix Table 2

Property Tax Assessor's Designation of Structure Quality

Quality	Percent of Observations
Poor	0.01
Below Average	0.58
Low	0.22
Economical	0.03
Average	17.44
Fair	0.73
Good	11.69
Above Average	5.26
Excellent	2.19
Luxury	0.35
Missing	61.51

Source. CoreLogic Residential Real Estate database.
Sample includes newly-built single-family detached
homes from 2000 to 2019.

Appendix Table 3

Distribution of Resident's Rating of Home Quality

	Homes built 1970-1989	Homes built 2000-2019
Rating = 1 to 6	6.6	3.5
Rating = 7	7.6	6.9
Rating = 8	21.2	22.0
Rating = 9	16.7	18.7
Rating = 10	48.0	49.0

Source. American Housing Survey National Samples 1985,
1987, 1989, 2015, 2017, 2019. Sample restricted to single-
family detached homes built 1970-1989 and appearing in
the 1985-89 samples, or built 2000-19 and appearing in the
2015-19 samples.

Appendix Table 4
Housing Quality Descriptions from RS Means

	Average Quality 1987	Average Quality 2019
Foundations	• Concrete footing 8" deep x 18" wide • 8" cast in place concrete 4' deep • 4" concrete slab on 4" crushed stone base	• Concrete footing 8" deep x 18" wide • 8" reinforced concrete foundation wall 4' deep, dampproofed and insulated • 4" concrete slab on 4" crushed stone base and polyethylene vapor barrier
Framing	• 2x4 studs 16" O.C. • ½" plywood sheathing • 2x6 rafters 16" O.C • 2x6 ceiling joints 16" O.C • ½" wafer board subfloor on 1x2 wood sleepers 16" O.C.	• 2x4 studs 16" O.C. • ½" plywood sheathing • 2x6 rafters 16" O.C • 2x6 ceiling joists • ½" plywood subfloor on 1x2 wood sleepers 16" O.C.
Exterior walls	• Beveled wood siding, #15 felt building paper • Brick veneer on wood frame with 4" average quality brick • Stucco on wood frame with 1" stucco finish • Solid masonry 6" concrete block load bearing wall with insulation and brick/stone exterior	• Beveled wood siding, #15 felt building paper • Brick veneer on wood frame with 4" average quality brick • Stucco on wood frame with 1" stucco finish • Solid masonry 6" concrete block load bearing wall with insulation and brick/stone exterior
Roofing	• 240# asphalt shingles • #15 felt building paper • Aluminum gutters, downspouts, and flashings	• 25-year asphalt shingles • #15 felt building paper • Aluminum gutters, downspouts, drip edge and flashings
Windows	Wood double hung	Double hung
Exterior doors	3 flush solid core wood exterior doors, storms and screens	3 flush solid core wood exterior doors with storms
Interior walls	½" taped and finished drywall Primed and 1 coat paint	½" taped and finished drywall Primed and 2 coats paint
Flooring	• 40% finished hardwood • 40% carpet with underlayment • 15% vinyl tile with underlayment • 5% ceramic tile with underlayment	• 40% finished hardwood • 40% carpet with ½" underlayment • 15% vinyl tile with ½" underlayment • 5% ceramic tile with ½" underlayment
Interior doors	23 hollow core doors	Hollow core and louvered
Heating	Gas or oil-fired warm air furnace	Gas fired warm air heat
Electrical	• 200 AMP service • Romex wiring • Incandescent lighting fixtures, switches receptacles	• 100 AMP service • Romex wiring • Incandescent lighting fixtures, switches, receptacles
Kitchen cabinets	14 LF wall and base with plastic laminate countertop and sink	14 LF wall and base with plastic laminate countertop and sink
Water heater	30-gallon gas fired	40-gallon electric

Source. R.S. Means Company "Square Foot Costs", volumes 1987 and 2019.

Appendix Table 5
Correlation of Housing Unit Characteristics with Ln(House Value)

	AHS		CoreLogic	
	1985-89	2015-19	2001-05	2015-19
Characteristics included in Census regression:				
Unit size	0.45	0.52	0.59	0.41
Number bedrooms	0.36	0.35	0.33	0.23
Number bathrooms	0.46	0.48	0.51	0.37
Fireplace	0.40	0.27	0.13	0.08
Garage	0.29	0.21	0.03	-0.03
Porch	0.18	0.06	--	--
Basement	0.11	0.14	0.01	0.08
Central air conditioning	0.17	0.05	-0.06	-0.20
Characteristics not included in Census regression:				
Resident rating of unit quality	0.20	0.13	--	--
Tax assessor rating of unit quality	--	--	0.42	0.31

Note. Each row shows the bivariate correlation between home values and the variable named in the row. The 1985-89 AHS sample includes homes built from 1970 to 1989 and the 2015-119 AHS sample includes homes built from 2000 to 2019. In the AHS data unit size has nine discrete values and resident rating has 5 discrete values. The tax assessor rating equals 1 for ratings of "good", "above average", "excellent" or luxury." In the CoreLogic data the correlations reported are with ln(square footage) and the tax assessor rating equals 1 for ratings of "good", "above average", "excellent" or "luxury."

Appendix Table 6
Correlation of Housing Unit Characteristics with Unit Square Footage

	AHS		CoreLogic	
	1985-89	2015-19	2001-05	2015-19
Number bedrooms	0.46	0.59	0.58	0.59
Number bathrooms	0.46	0.67	0.68	0.70
Fireplace	0.36	0.36	0.23	0.08
Garage	0.24	0.25	0.07	0.06
Porch	0.16	0.08	--	--
Basement	0.36	0.28	0.01	0.01
Central air conditioning	0.18	0.18	0.12	0.05
Resident rating of unit quality	0.20	0.16	--	--
Tax assessor rating of unit quality	--	--	0.39	0.26

Note. Each row shows the bivariate correlation between square footage and the variable named in the row. The 1985-89 AHS sample includes homes built from 1970 to 1989 and the 2015-119 AHS sample includes homes built from 2000 to 2019. In the AHS data unit square footage has 9 discrete values and resident rating has 5 discrete values. In the CoreLogic data the correlations reported are with ln(square footage) and the tax assessor rating equals 1 for ratings of "good", "above average", "excellent" or "luxury."