Hutchins Center Working Paper #65

July 2020

Quality-Adjusted Price Indices to Improve Productivity Measures in Highway Construction

Omar Swei

David Gillen

Anuarbek Onayev

University of British Columbia

ABSTRACT

In this paper we propose an approach to generate a quality-adjusted producer price index for highway construction. We use price data for highway construction projects across the contiguous United States from 2005 through 2017. The data set includes approximately 5,000 unique items for each project. These items are distilled to form 60 item baskets. We redefine our output from being lane-miles to 'lane-miles of service.' The indicator of quality is the deterioration rate of a roadway, which is measured using data on pavement roughness which links to deterioration and the time between required servicing; an improvement in quality reduces deterioration and increases the time span between maintenance and reconstruction, something which adds significant value to state budgets. We use a chained-Fisher price index and find that our proposed, quality-adjusted producer price index exhibits lower annual growth by 2.0 percent than the unadjusted price index. Given price inflation has been overestimated in the past by failing to account for quality changes, our finding suggests the lack of productivity growth in construction, specifically highways, bridges and infrastructure, may have been significantly underestimated.

The authors are indebted to participants at the Brookings Productivity Measurement Initiative Authors Conference on April 23, 2020 for helpful discussion and comments. In particular, they thank Dan Sichel and Chad Syverson for their detailed comments and helpful suggestions. They also want to thank Emily Wong for her help in processing some of our original data. The authors did not receive financial support from any firm or person with a financial or political interest in this article. None is currently an officer, director, or board member of any organization with an interest in this article.

1. Introduction

Productivity is simply a measure of what has been accomplished with a collection of factor inputs when they are mixed together in a certain way under the guidance of operating managers and in the presence of institutional, governance and regulatory controls. Many would say productivity is 'efficiency in production', i.e., how much output from a set of inputs (Syverson, 2011). This statement is clean and accurate but we prefer to recognize the environment of productivity and 'accomplishments' or 'outcomes' is a different notion from outputs because it raises the question of what is the output, what is being produced and how do you measure it; a question which is central in driving the research discussed in this paper.

The U.S. Bureau of Economic Analysis data show that annual labor productivity (private non-farm business) in the U.S. had averaged 2.0 percent from 1987 through 2018; capital productivity had averaged -0.6 percent and multifactor productivity 0.8 percent. Real GDP growth over this same period had averaged 2.8 percent. These aggregate, economy-wide numbers mask a substantial amount of variability among economic sectors and across and within industries. As Syverson (2011) has noted, the wide ranging, large and persistent productivity differences within industries have raised the question of why; in other words, what leads some firms in an industry to be nearly twice as productive as the marginal firms? There are ongoing research efforts around improving productivity growth measures generally but we focus on one specific sector of the economy: construction. Because the construction industry supports more than 7.6 million jobs in the United States and contributes more than 4 percent to national GDP, its productivity growth will contribute substantially in raising overall living standards.[1] Interestingly, the latest data made available by the U.S. Department of Commerce suggest that construction productivity has *decreased* more than 30 percent since 1997 while, over that same time period, productivity across the wider economy has *increased* by that same margin, as shown in **Figure 1**. This reality has renewed discussions around stagnant productivity growth in construction, a point of discussion that has resurfaced every 20-30 years in the academic literature.[2]

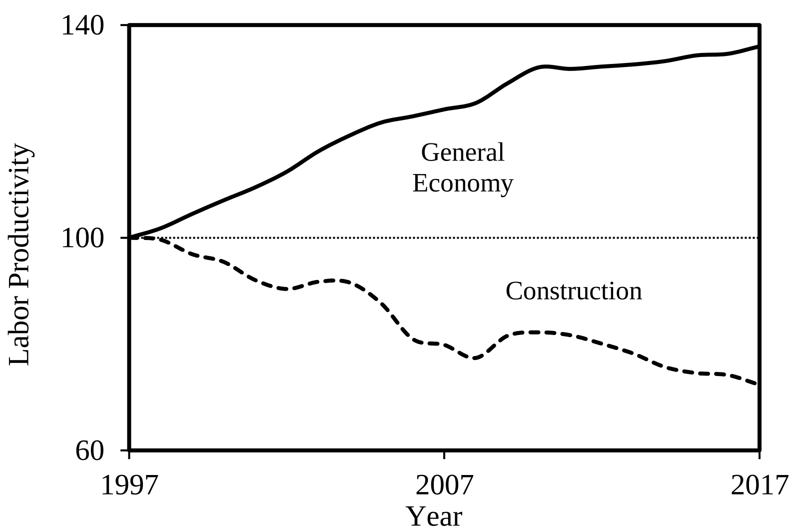


Figure 1. Labor productivity growth since 1997 across the general economy and construction. Output is measured as real value added per data from U.S. Department of Commerce.

. . .

1. Construction spending in the U.S. in 2018 was $302 billion in the public sector and $992 billion in the private sector (nominal dollars). It employed 8.83 million workers. Value added of construction was 4.4 percent of U.S. GDP in 2018. In 2018 there were 1,184,900 single family residents completed (1,256,200 in 2019).
2. Examples in the literature that comment on stagnant productivity growth in construction include Dacy (1965) and Allen (1985).

More recent attention around this subject has been led by the U.S. Bureau of Labor Statistics (BLS) (see Sveikauskas et al. 2014, 2016 and 2018). This research has sought to generate improved measures of labor productivity growth in construction across four important subsectors: (1) single-family housing, (2) multi-family residential, (3) highways, streets and bridges and (4) industrial construction.[3] The development of the productivity measures focuses on labour productivity rather than total factor productivity (TFP) or multi-factor productivity (MFP) because, as the authors say, it is difficult to measure capital materials requirement inputs. The view was it was difficult to obtain accurate measures of productivity growth because of the myriad of price deflators used against revenues to create an output index over time.[4] The central emphasis of the papers by Sveikauskas et al. was the use of accurate, relevant output price deflators. The deflators for these four areas came from the Census (single family deflator), BEA (multi-family deflator), FHWA (construction cost index for highways) and BLS (deflator for industrial construction). The output measures for each of these subsectors were based on the value of construction put in place provided by the U.S. Census.[5]

The findings of Sveikauskas et al. indicate that, among the four major subsectors of construction, only one has failed to exhibit positive productivity growth: highway, road, and bridge construction. More specifically, their results indicate direct and subcontractor labor productivity trends of 1.1, 1.7, -2.4 and 7.5 percent for single-family residential; multi-family residential; highways, streets and bridges; and industrial construction, respectively. These findings strongly motivate further study of highway, road, and bridge construction, particularly given the industry's size and importance. Public spending for capital, operations, and maintenance on highway infrastructure in 2017, for example, exceeded $177 billion, representing nearly 1 percent of U.S. GDP.

To support this goal, our study aims to generate further improved producer price deflators for the highway construction sector. Studies such as those of Sveikauskas et al. rely on FHWA's National Highway Construction Cost Index (NHCCI), which tracks quarterly changes in input prices across all highway construction activities. Quality improvements in both material and capital inputs for highway construction have been noted by several researchers (Goodrum et al. 2009, Goodrum and Haas 2004). Furthermore, according to the U.S. Congressional Budget Office, the share of public expenditures on highway construction that go for operations and maintenance (O&M) activities has also drastically shifted from 37 to 46 percent over the time frame of the analysis by Sveikauskas et al. (**Figure 2**). We, therefore, aim to complement this important initiative led by FHWA and BLS by proposing a methodology to create price deflators that can adequately capture quality improvements over time and can be designed for specific technologies and activities.

. . .

3. The time period covered for highways, streets, and bridges is 2002 through 2016 while the measures for the other sectors covers 1987-2016.

4. Construction is much too heterogeneous to have a single output and even the four major subsectors of constructions are too heterogeneous to yield an easily aggregated output. An alternative measure of real output for construction can be obtained by dividing the total expenditure on construction by a price index. However, this solution is not straightforward since which price index to choose becomes the problem. In addition, the productivity measure still suffers from the same general problem of productivity measurement which include managing changes in quality.

5. [5] Syverson (2011) has noted that physically-based rather than revenue-based output measures exhibit more variation. He argues there is a negative correlation between quantity-based TFP measure and price. The result is revenue-based productivity measures understate the differences in producers' physical efficiencies (Syverson, 2011, footnote 1).

Once quality changes are taken account of in the price index, we find the quality-adjusted producer price index for asphalt concrete highway construction exhibits lower annual growth by 2.0 percent. This suggests a strong upward bias in existing, non-quality adjusted producer price deflators. Since most productivity studies examining changes in productivity in construction use expenditures or revenues as the bases for calculating output, an upward biased price deflator will result in a lower estimation of productivity growth.

The following section briefly examines four different but relevant literatures that influence our work: the measurement of highway capital stock and costs, the measure of productivity in construction, quality adjustments for productivity measures and the highway construction literature. Following this we describe our method for introducing quality adjustment into a price index and explain how we measure quality improvement. We finally report our quality-adjusted price index and compare it with an unadjusted one, motivating our discussion of future research directions in this area.

Figure 2. Spending on operations and maintenance (O&M) as percent of highway construction spending since 2002, per data provided by U.S. Congressional Budget Office.

Literature on Measuring Highway Capital Stock and Costs of Infrastructure

A broad literature has examined the issue of what constitutes the highway capital stock and, more importantly, how it should be measured. This literature grew out of the interest in public capital's contribution to economic growth stemming from Aschauer (1989). The issue of how to measure highway capital was examined intensively by Fraumeni (1999, 2007, 2008 and 2009) whose focus was using a national income accounts approach.[6]

. . .

6. The three national income perspectives are (1) highway capital outlays, or how much was spent in order to build a highway unit; (2) highway capital stock, which is an input measure which captures the designed service level of a highway unit or its effective productivity; and (3) highway gross output, which signifies the amount spent to build and maintain roadways.

Fraumeni's 1999 paper was concerned with measuring the productive capital stock principally because it was the relationship between highway capital and productivity that was of interest. Fraumeni argues that in constructing a quality-adjusted public capital stock series there is a need to use measures of productive capital stock adjusted for deterioration and quality change. Choosing this method as distinct from wealth capital stock measures has a significant impact on the estimates of the public capital stock which will, in turn, affect measures of productivity.[7] The distinction between productive and wealth measures for capital stocks is subtle. Productive capital stock is adjusted for current and past decreases in efficiency whereas wealth capital stock measures are also adjusted for future declines in efficiency.[8]

Fraumeni (2007) presents measures of the contribution of highways to economic growth using national income accounting measures of productive highway capital stock; productive highway capital stock measures potential productive capacity. The measures do not include spillovers and only consider the use of highways for businesses and government.[9] The capital outlays for highways are distinguished by Interstate, non-interstate and local; new construction and reconstruction; and pavement, grading and structures. Fraumeni (2008, 2009) update the highway capital stock measures produced earlier in Fraumeni (1999). A key issue in this paper is whether the pavement curves used in the productive capital stock measures were still valid.[10]

More recently, a paper by Bennett et al. (2020) provides measures of the value of infrastructure, which they divide into basic (transportation & utilities), social (schools, hospitals) and digital (communications and cloud related). Among other measures, they provide what they term prototype estimates of investment in highways, at the state level, and they note significant Interstate variability. They estimate that as of 2020 maintenance and repair is approximately 15 percent of gross highway investment. It is important to note that this estimate differs from the earlier cited number from the Congressional Budget Office of 46 percent for operations and maintenance, highlighting the challenge in classifying highway investments. Their final contribution is to examine trends in price deflators and quality change in infrastructure assets.

Bennett et al. note that prices for highways and streets are volatile, with annual price increases of 10 percent from 1970 through the early 1980s and stable prices occurring during the latter 1980s through late 2000s, after which they noticeably increase. A key measurement issue is whether costs are expressed

. . .

7. Wealth capital measures the market value of capital. As an example: BEA would include cars and appliances in its measure of wealth capital, whereas these two items would normally be included in consumption in the national accounts.

8. Productive capital stock is an appropriate concept for measuring the contribution of capital to economic growth while wealth capital stock is appropriate for measuring the market value of capital. Productive capital stock measures are adjusted for past and current decreases in efficiency while wealth stock measures are adjusted for past, current and future decreases in efficiency.

9. For example, the value consumers place on travelling to visit friends and relatives would not be captured in these measures.

10. The pavement curves relate to pavement serviceability, the ability of highways to move traffic at a designed speed. If the speed declines or the operating cost to users increase, the "serviceability" declines and the efficiency of the highway capital stock declines. Pavement serviceability decreases with time and use, with serviceability decreasing least on Interstates, and the rate of decline increases for non-interstate and local roads respectively in part reflecting the initial standards to which they are built.

per mile or per lane-mile; if a road is widened by one lane in either direction, for example, total miles would not change but lane-miles would.[11]

Brooks and Liscow (2019) examine infrastructure costs for highways (specifically the Highway Interstate System) across states and observe considerable variation. Their choice of unit of measure, miles of Interstate, does not taken account of construction such as lane widening, increasing shoulder width or additional amenities associated with safety measures or amenities such as noise berms or walls. Expenditures on all of these items would not vary with length of the system but would be sensitive, in part, to a measure such a lane-miles. They note labor and materials price variation does not explain the large cost differences.

In Brooks and Liscow (2020) the authors focused on explaining the rise in expenditures for infrastructure and use the U.S. Interstate highway system as their empirical example. The metric used is total miles and data are at the state level. Also, the costs are related only to new construction while resurfacing and maintenance are excluded.[12]

Amekudzi-Kennedy et al. (2019) provide an extensive review of traditional asset valuation methods. The authors cover traditional valuation methods and introduce new approaches to how value is interpreted, particularly given the growing use of smart infrastructure and performance-based decision-making. Their discussion of value in establishing the 'worth' of assets is essentially distinguishing value in 'use' versus value in 'exchange.' They make a distinction, like Fraumeni (1999), between wealth-based value and productivity-based value, although they do not use those terms explicitly. They argue that valuation of infrastructure should go beyond use and exchange to include value based on existence, sustainability and user preferences.

The work of Fraumeni, Bennett et al., Brooks and Liscow and Amekudzi-Kennedy et al. illustrate the challenges of creating a price index and properly measuring the capital stock. What should be included for separating maintenance and expenditures and additions to the capital stock can be an issue. The literature also questions what constitutes output and value and how the two should enter the calculation of what we have termed as outcomes. These measurement and conceptual issues figure prominently in developing accurate price indices and in designing a means to measure quality change. Redefining the nature of what is produced in terms of its impact facilitates the development of a hedonic index to capture quality improvements.

Literature on Measuring Productivity in Construction

Measuring output in highway construction productivity has had a history of using a physical measure of output and has transitioned to revenue or expenditure measures which generate the output measure by dividing by a price index. Allen (1985) explores explanations for the productivity decline in the

. . .

11. The authors state their estimates of price increases 'lineup with' the results of Brooks and Liscow (2019) in terms of cost per mile for interstate highway construction. From Brooks and Liscow one can infer an implied annual rate increase of 5.3% while the National Accounts for highways and streets shows an annual increase of 6%. The other distinction is Brooks and Liscow are looking at interstate highway construction whereas Bennet et al. figures are for highway and street construction which would include urban, near urban and rural; of course, the construction figures exclude land.

12. They argue that a large portion of the increase in costs and variation over states is the increase in average incomes with higher income citizens demanding higher quality highways and a shift to a rise in 'citizen voice' whereby citizens could more easily challenge decisions by government.

construction industry between 1968 and 1978 in the United States. He uses a production function approach rather than a growth accounting one to better capture changes in capital-labor ratio and institutional and industrial features of the construction sector. He initially found there were real factors explaining the observed productivity slowdown: the shift from high productivity multiple family units to low productivity single family homes, labor quality, decreasing establishment size, the capital-labor ratio and unionization. He also claimed that the metric of output growth had been underestimated since the denominator, price deflators, were upward biased. Pieper (1989) criticizes Allen, claiming Allen's measures of some real factors (e.g., capital-labor ratio) were incorrect. Allen (1989) replies with arguments that his original approach was valid and after correction, real factors and mismeasurement of the price deflators account for 56.5% of the decline in observed decrease in construction industry productivity over the 1968-78 period.

Pieper (1991) provides a comprehensive review and assessment of historical construction deflation methods used by the Bureau of Economic Analysis (BEA) to prepare constant-dollar construction components for U.S. GDP. He also provides suggestions on how current deflators (current meaning 1991) could be improved. He points out the early deflation indices were cost indices made up of factor input prices such as materials prices and labor wage rates. But the problem is cost indices overstate price increases because they fail to account for changes in productivity or changes in prices due to competitive conditions. The preference is to use price indices for construction output rather than use input costs to deflate the value of construction output. However, price indices have their weaknesses as they assume homogeneity of the output for physical measures.[13] Pieper is optimistic on the use of hedonic price indices as a way of taking into account heterogeneity and possibly quality differences.

Parker (1991) in a 'Comment' on Pieper's work argues the lack of progress in developing new price deflators results from a number of sources: the lack of support from the private sector in providing data and the fact the Office of Management & Budget (OMB) hamstrings the BEA in affecting their ability to develop and pay for new surveys. He notes that academic interest in the area was meager at best. Parker is favourably disposed to the use of cost indices but recognizes their weakness, therefore suggesting that they serve as an upper bound on price indices.

A number of authors have examined productivity trends in the construction engineering and management literature (Allmon at al. 2000; Goodrum and Haas, 2001; Goodrum and Haas, 2002; Zhai, 2009). The interesting feature of this literature is the disaggregation of the aggregate output of the construction industry into construction activities; this is in a similar spirit to the shift from service-based expenditures to disease-based expenditures in measuring output in health care services (Sheiner and Malinovskaya 2016).[14] Allmon et al. (2000) examined labor productivity using activities such as residential framing, commercial web joists, compaction, hand trenching and ceiling tile installation. They measure unit labor costs, output and direct work rates and report that productivity increased for all tasks studied in the paper; a total of 8 out of 20 activities for which data exists. However, it appears decreasing real wages were driving much of the measured productivity increase. Goodrum and Haas (2002) also

. . .

13. Price per square foot of floor space is seen as weak since higher priced structures are not necessarily larger but have more amenities. Using price per square foot assumes residential housing stock is homogeneous. The same criticism was levelled at the FHWA price index which treats highway projects as homogeneous.

14. The traditional approach to measuring output in health care was spending on health goods and services which was deflated by some price index. The newer approach is to use expenditures on disease treatments. This change in approach affects the price indices used to deflate expenditures.

examined activities as a basis for measuring productivity. They estimated productivity changes in 200 activities associated with different types of equipment technology. Their principle findings are productivity increased for most all activities and that much of the increase came from technological advances and from increases in the capital labor ratio. This result was reaffirmed in Goodrum, Haas and Glover (2002). These findings are in contradiction to the productivity results found with aggregate measures, perhaps due to output measurement problems resulting from the increasing complexity of projects. Goodrum, Zhai and Yasin (2009) examined productivity change in construction using activities with an emphasis on changes in materials technology. Based on 100 activities, they find that material weight, installation and modularity all contributed to significant improvements in labour productivity.

Review of Productivity and Quality Adjustment Literature

The two primary methods used to adjust for long run price changes and account for any quality changes are the matching model and hedonic models. The matching model, as the name implies, compares the price of an item in period *t* with the same item in *t+1*. The fundamental purpose of this model is to hold constant the characteristics of a transaction (Triplett 2006). There are a number of possible sources of error with matching, including missing items (discontinuance or specification change), some adjustment in an items size, makeup or effectiveness etc. and a completely new item. The second method for adjusting for quality is hedonic models whereby changes in quality are estimated from data that characterize all of the features of the product as well as the product price. A regression of quantity of each characteristic on the price for the bundle of characteristics will yield a set of coefficients.[15] The coefficients will reflect both demand (user valuation) and supply influences (resource costs). A distinction is made between a hedonic function and hedonic price index. A hedonic function is an estimated relationship between prices of similar products and the characteristics they have whereas a hedonic price index makes use of a hedonic function (Triplett 2006, Chapter 3).

Griliches (1961) was one of the first to investigate the distortion in measured price indices in the face of quality change. His investigation of quality changes in automobiles framed the question as, 'what would this item cost with the new combination of features (i.e., qualities) relative to some base period when such features were not available'? By applying his approach in the automobile sector, Griliches is able to show that, after considering quality changes, a 1960 model car is less expensive than a 1954 model. Hall (1971) tackles a somewhat different problem, still linked to quality differences, in looking at input markets. The issue is determining the relative qualities in different vintages of capital by observing relative prices in second hand markets.[16] In a 1995 investigation of quality changes for personal computers, Berndt, Griliches and Rappaport (1984) found that when improvements in quality were included in the price index, real quality-adjusted prices for personnel computers had fallen by 30 percent. Taking these results in the context of measured productivity for construction, in our case highway construction, failure to account for quality changes may result in an overestimate of the price index which will in turn lead to an underestimate of real output and a lower measured change in productivity.

. . .

15. The coefficients are often referred to as implicit prices for the characteristics and are influenced by both the demand for and supply of similar characteristics.

16. Hall points out the knotty problem of disentangling the effect of age (depreciation) from quality difference when observing data in secondhand markets.

Hedonic regressions are generally represented in a generic way but there are three methods of implementation; the time dummy approach, the characteristics/repricing approach and the imputation approach (Silver, 2018). The first regresses product prices on product characteristics and time dummies which reflect the date of the product transaction. Parameters on the time dummies provide measures of the proportionate change in prices relative to a reference point in time (year, quarter, month).[17] The hedonic characteristics or repricing approach regresses product price on product characteristics for a given period and uses the parameters as weights. The average product is a tied bundle of average characteristics. Holding the average bundle constant, the parameters estimated in period 0 and again in period 1 can answer the question, 'what is the price change of the average product or average characteristics (quality adjusted) valued in period 0 relative to period 1'? A ratio of the values from period 0 weights and period 1 weights (parameters) is a constant quality product price index. The third method, the hedonic imputation approach, is essentially product matching. For each product compared the quantities are held constant and only characteristic prices change.[18]

A Brief Overview of Highways

Transportation agencies are under increased pressure to maintain and preserve their roadway assets with limited available resources. The latest *Highway Statistics* report from the Federal Highway Administration (FHWA) estimates that 20 percent of roadways across the United States are either in a fair or worse condition (U.S. Department of Transportation 2018). In addition, the projected annual capital investment required to improve the condition of the existing system is almost $30 billion real dollars higher than current levels (U.S. Department of Transportation 2020). In short, transportation agencies are being forced to do more with less. The challenge of maintaining an aging roadway system has motivated federal and state authorities to devote significant resources towards the creation of pavement management systems, which are used to collect and process highway condition information and forecast future performance. To achieve these objectives, an area of considerable research over the last several decades has been the development of analytical techniques to model pavement performance (i.e., deterioration).

Pavement performance broadly refers to the deterioration of an individual facility (i.e., asset) over time. For pavements, common distress mechanisms include rutting, cracking, faulting and surface disintegration. The majority of state and municipal agencies as well as FHWA, whose *Highway Statistics* track aggregate, highway-system condition measures, rely on pavement roughness to measure pavement performance. This fact can be attributed to three important factors: (1) the collection of this data is relatively inexpensive (Abulizi et al. 2016); (2) although measurement errors persist, roughness data are significantly more reliable than are measurements for other distress mechanisms (Schwartz 2007); and (3) pavement roughness correlates strongly with other performance measures of interest for practitioners (Garg et al. 1988; Li et al. 2011).

The evolution of pavement roughness across time for a facility will generally follow the saw-tooth diagram shown in **Figure 3** where, for a given set of capital, material, and labor inputs, an asset will

. . .

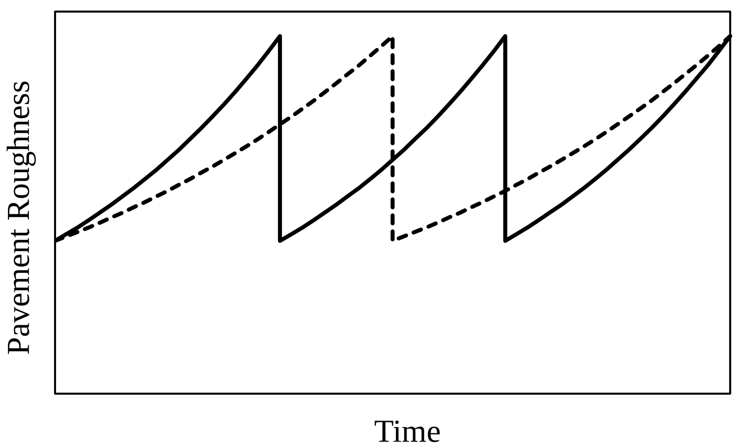
17. The parameters on the product characteristics are assumed invariant over time and space.

18. Detailed implementation descriptions can be found in Triplett (2006) and Silver (2018).

deteriorate at some exponential rate (Ouyang and Madanat 2006). Should a facility's roughness reach some critical, unacceptable level of performance, a construction activity (often referred to as an intervention) will typically take place to renew the condition of the pavement segment, leading to the idealized representation in **Figure 3**. Per **Figure 3**, a higher performing pavement, which could arise due to an increase in the *quantity* (e.g., a thicker pavement) or *quality* of inputs, will impact the total cost of ownership for a planning agency. Although the majority of research in highway construction has emphasized the effect of material inputs on performance[19], several technological and policy advances have played a critical role in improving the overall quality of highway construction. While we do not exhaustively list these innovations in this paper, we do provide a few examples to clarify these advances for the reader.

Highway construction requires large amounts of capital inputs to produce a high-quality product.[20] In addition, policy changes by state DOTs in process specifications and contractual arrangements for highway construction have also influenced the quality of roadway production. An intuitive example is the promotion of quality check/quality assurance (QC/QA) programs by state agencies over the last two decades. Patel et al. (1997) and Fernando (1997) describe the introduction of these QC/QA initiatives around material handling and initial pavement condition in Illinois and Texas and their anticipated impact on highway performance (few studies have quantified the actual effectiveness of these programs). The integration of QC/QA programs by many state DOTs has also led to the restructuring of contracts for many highway projects; it is now commonplace in many states to use incentive-based contracts to motivate high-quality construction by contractors (D'Apuzzo and Nicolosi 2010).

Figure 3. Illustration of the Evolution of Pavement Roughness over Time.



Note: For two highways subject to similar traffic and climate conditions, the illustration shows the evolution of roughness for a low-performance (solid line) and high-performance (dashed line) constructed facility.

. . .

19. Prozzi and Madanat (2003) provide an excellent overview of research initiatives aimed at capturing the effects of material inputs on performance. The authors classify this literature as empirical (i.e., data-driven), mechanistic (i.e., scientifically-based response functions to capture the relationship between performance and inputs) and mechanistic-empirical (i.e., mechanistic models that integrate the statistical rigor of empirical models).

20. One example of these capital inputs is compaction equipment, which ensures that material layers are of both proper and uniform in density. FHWA has played an important role in promoting the development of intelligent compaction technologies for soils, aggregate bases, and pavement surfaces to improve the life-cycle performance of highway maintenance and construction (Chang et al. 2011).

These technological, policy and structural changes in the highway sector have, theoretically, led to a higher quality product for a similar set of inputs. To our best awareness, there are no existing studies that have numerically quantified quality improvements over time nor embedded them within measures of productivity growth for highway construction. We aim to address this gap via the use of extensive longitudinal data around historical infrastructure performance and bid data within the contiguous United States. Specifically, the upcoming section details the creation of four new, quarterly producer price indices of highway construction disaggregated by material and activity. We subsequently evaluate the sensitivity of one of our indices (asphalt concrete construction), which constitutes 94 percent of highways in the United States, to the inclusion of quality improvement measures. These findings will support future planned studies on productivity growth in the highway construction sector in the coming years.

Methodology: Building a Quality-Adjusted Price Index

A Baseline Highway Producer Price Index

A producer price index summarizes average changes in prices for a set of inputs across time. Its importance for productivity studies cannot be understated; if output is measured via revenues, and the monetary value of those revenues change across time due to inflation, then it is necessary to have available a price deflator that can transform those revenues into some set of common, comparable values. For the highway sector, there are three well-known forms of price indices that have been previously implemented: Laspeyres, Paasche, and Fisher. A detailed discussion around the use of these approaches and their implementation by state departments of transportation (DOTs) for highway construction can be found in Nassereddine, Whited and Hanna (2016) and Shrestha, Jeong and Gransberg (2017).

Both the Laspeyres and Paasche approaches are fixed-base indices with weights based on quantities in a given reference year. Because both techniques neglect the substitution effect and declines in the usage of certain commodities due to possible price increases, both state DOTs and FHWA have adopted a chained-Fisher approach for developing their own highway construction producer price indices. FHWA's National Highway Construction Cost Index (NHCCI) is, in fact, the reference price deflator for the previously cited productivity analyses initiated by the U.S. Bureau of Labor Statistics around road, highway, and bridge construction. The general form of the chained-Fisher price index is shown equation (1) and equation (2). R_t captures the relative shift in average producer prices, C_t, between consecutive time periods (i.e., t and t-1). R_t is computed as the geometric mean of the Laspeyres and Paasche index using relative quantities, q, and unit-prices, p, for each individual item (i) across n item baskets used to consolidate bid items into categories.

$$R_t = \sqrt{\frac{\sum_{i=1}^{n} p_i^t q_i^{t-1}}{\sum_{i=1}^{n} p_i^{t-1} q_i^{t-1}} \times \frac{\sum_{i=1}^{n} p_i^t q_i^t}{\sum_{i=1}^{n} p_i^{t-1} q_i^t}} \tag{1}$$

$$C_t = C_{t-1} \times R_t \tag{2}$$

For highway construction, a frequently cited source of data is those made available by Oman BidTabs, which since 2005 has tracked price data for highway construction projects across the contiguous United States. The financial value of pavement-related activities used in our study (once possible outliers are removed) exceeds over $170 billion (nominal) dollars, making it a reliable source of data that is leveraged by FHWA in the generation of its own NHCCI. For each project, Oman BidTabs reports quantities and unit-prices for all required tasks, the date of the auction, the project's geographic location at the county-level, the type of activity pursued (e.g., new construction), and other variables (e.g., the identification of the winning contractor) that are of value for future policy research.[21]

Two important challenges underly the processing of our available data: the identification of appropriate "item baskets" and the detection of possible outliers. State DOTs frequently vary in their conventional units (e.g., cubic yards vs. tonnage) and terminology (e.g., hot-mixed asphalt vs. asphalt paving) for similar activities. Further complicating this matter is that multiple agencies have altered their internal naming convention of activities over the timeframe (i.e., post-2005) of our study. To deal with these challenges, we rely on available text classifiers to identify reasonable item baskets for our indices. While our dataset includes more than 5,000 unique item descriptions, this process reduces them into 60 unique item baskets.[22] We also detect outliers by inspecting difference between unit-prices for winning bids and competing contractors based on criteria reflecting conversations with stakeholders from state DOTs, private contractors, and trade associations.

Given that the goal of our work is to complement future productivity studies in highway construction, we compute four baseline producer price indices for two common technologies (Portland cement concrete vs. asphalt concrete pavements) and broad classifications of activities (general construction vs. maintenance and rehabilitation). Since the dominant material input for asphalt concrete highways, bitumen, is a by-product of crude oil production, its price growth has been impacted by recent price volatility in energy markets, causing the cost of its production to differ substantially from Portland cement concrete. Specifically, while bitumen constitutes only 4-5 percent of material inputs for asphalt concrete, it currently makes up around 40 percent of the average cost of its production. Producer price indices disaggregated by technology will additionally complement our intended focus of analyzing and commenting on differences in productivity growth across firms that utilize these two prevalent materials, which we further discuss towards the end of this paper. We have also created producer price indices for construction maintenance and rehabilitation given the increasing shift in public spending towards these activities for an aging transportation system. **Table 1** highlights differences in the available sample size used to generate each index; since asphalt concrete is the predominant technology used for highway construction, its producer price index relies on 500 percent more bid data (in terms of nominal dollars) than the Portland cement concrete index. As can be noted below, approximately 16 percent of spending in our dataset is tied to maintenance activities, which is consistent with the results of Bennet et al. (2020).

Table 1. Nominal value of pavement-related highway spending used for each producer price index from January 1, 2005 to December 31, 2017.

Producer Price Index	Nominal Value 2005-2017
(1) Asphalt Concrete Highway Construction	$124 billion
(2) Portland Cement Concrete Highway Construction	$24.5 billion
(3) Asphalt Concrete Highway Maintenance & Rehabilitation	$18.3 billion
(4) Portland Cement Concrete Highway Maintenance & Rehabilitation	$5.57 billion

. . .

21. While this dataset is rich in terms of its breath and scope, there are significant challenges involved in processing its information; we discuss these issues given their importance in generating our own reliable indices but keep this discussion brief given the larger focus and emphasis of our work.
22. Examples of relevant item baskets include Concrete Pavement, Bituminous Pavement, Joint Sealing, and other common construction activities.

Our baseline producer price indices are designed by tracking changes in highway construction input quantities and prices across time. Since the production of highways serves as an intermediate good for the broader economy, we should also be concerned with tracking potential shifts in the resulting outcomes. A better performing road for a consistent set of capital, labor, and material inputs will require fewer future maintenances, improve user travel times, and facilitate a more efficient economy. This means that only tracking changes in production inputs to form a price index is insufficient; we also need to "redefine" our output via a measure that reflects its utility to society and be able to track its evolution temporally. In the upcoming section, we define a measure of quality for highway construction based on significant research among the highway engineering community and propose a strategy to embed it within our own producer price indices.

Measuring Quality Improvements

To measure quality improvements in highway construction, we rely on data collected and monitored by FHWA as part of its Long-Term Infrastructure Performance (LTIP) program. The program, which was initiated by the National Research Council and Transportation Research Board in the 1980s before transitioning to FHWA in 1991, includes more than 150,000 field measurements of highway performance (measured in terms of roughness) from over 2,500 roads across North America. These roads have been built over a range of years and differ in their structural designs, exposed traffic volumes, and climactic regions. We leverage these data to generate an infrastructure performance model that adequately extracts improvement in construction quality over time and use that information to appropriately adjust our highway producer price index.

We have decided to measure quality improvements in highway construction via this approach for a couple of important reasons. Changes in methods to procure highway projects are, more so than most other sectors, motivated by knowledgeable, informed consumers. While quality improvements for telecommunications, for example, are driven by consumer preferences, consumers are frequently uninformed around the underlying technology used to achieve these enhancements. State DOTs, on the other hand, not only have a sense of their intended performance goals for a new infrastructure project but also rely on their own internal engineers to develop explicit specifications (e.g., pavement thickness and geometric design) for suppliers (i.e., contractors) to achieve those objectives. Although we do not have access to the complete specifications for each construction project within our database nor its ex-ante forecasted performance, we are able to access its measured quality directly thru ex-post performance observations. Improvements in construction quality should improve the overall performance (i.e., lower deterioration) of highway infrastructure and reduce the rate of depreciation for these assets over their life-cycle. We use the term performance as synonymous with deterioration given the convention used by researchers in this domain.

The performance model we propose builds off of an important body of research that has emerged over the last 30 years, which has identified the relevance of important topics such as stationarity, endogeneity, and measurement errors in determining an appropriate model specification[23]. These studies have generally assumed that variation in pavement performance across time can be adequately captured by incorporating a set of regressors that incorporate relevant structural and/or environmental (e.g., traffic

. . .

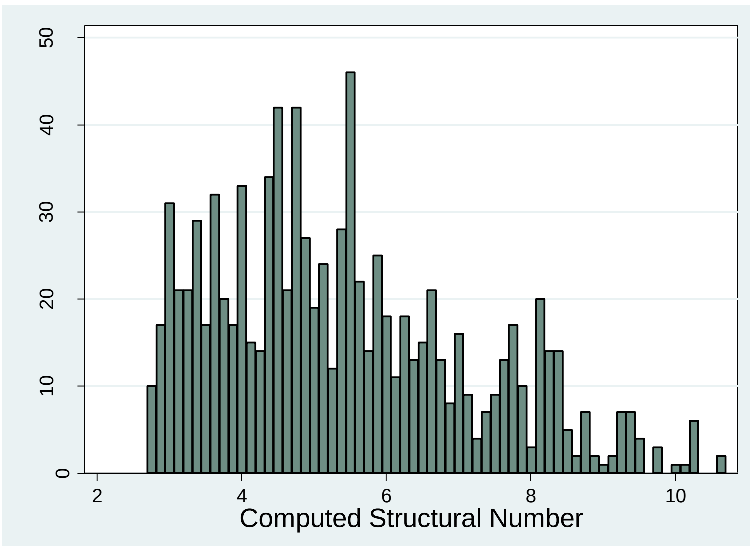
23. For a discussion around each of these issues, we encourage readers to review the work of Ben-Akiva and Ramaswamy (1993), Chu and Durango-Cohen (2008), Hong and Prozzi (2015), Swei, Gregory, and Kirchain (2018), and Yehia and Swei (2020).

conditions) factors that affect a highway's deterioration. Two particularly important explanatory variables per previous research are: (1) the design of the facility and (2) the exposed truck traffic volume of the highway. We therefore begin our analysis by estimating the below model similar to traditional research:

$$\Delta \ln D_{it} = \alpha_1 + \alpha_2 \ln SN + \alpha_3 \ln AADTT \quad (3)$$

D_{it} is the performance of the i^{th} facility in year t, α_1, α_2 and α_3 are parameter estimates that quantify the effect of a facility's structural capacity per its structural number (SN) and its average annual daily truck traffic (AADTT) loading. The SN for a constructed facility is a continuous variable that has been used in pavement design to characterize the *anticipated* structural capacity of a new facility. Two key characteristics affect its value: the thickness of the individual layers that make up a pavement and their material properties based off of a layer coefficient. A pavement's SN does not, however, comment on its true structural capacity, which will depend on the quality of construction conducted by the contractor. **Figure 4** presents a truncated histogram of the computed SN for each observation in our dataset.

Figure 4. Distribution of computed structural number (SN) for 5th through 95th percentile values.



The deterioration for each of our indexed panels is based on an average of a series of measurements. Because modern inspection technologies are prone to measurement errors, state DOTs will typically conduct at least five measurements of their asset performance during each site visit and subsequently report an average condition rating. In other words, the reported performance of a highway facility across our dataset is an uncertain, latent measure.[24] A consequential challenge in estimating Equation 3 is that, due to the latent nature of pavement performance, three sources of error have been shown to be present in the estimation of equation (3) per our previously cited references:

$$\Delta \ln D_{it} = \alpha_1 + \alpha_2 \ln SN + \alpha_3 \ln AADTT + \varepsilon_{it} + u_{it} + u_{i,t-1} \quad (4)$$

. . .

24. The challenge in modeling infrastructure performance due to its latent nature can be found in Humplick (1992), Ben-Akiva and Ramaswamy (1993), Madanat and Ibrahim (1995), Madanat, Karlaftis, and McCarthy (1997).

where ε_{it} is a random error term underlying the deterioration of facility i in time period t and u_{it} and $u_{i,t-1}$ are measurement error terms in time period t and t-1. Because the accuracy of the measured condition of a highway facility varies across observations, one successful approach employed previously to deal with this issue is to estimate an iterative, reweighted least squares model in which the weight for each observation, w_{it}, is set equal to the inverse of the observation's theoretical variance. If one assumes that each error term is independent then the weight for each observation would be:

$$w_{it} = \frac{1}{\mathrm{Var}[\Delta \ln D_{it}]} = \frac{1}{\mathrm{Var}[\varepsilon_{it}] + \mathrm{Var}[u_{it}] + \mathrm{Var}\left[u_{i,t-1}\right]} \tag{5}$$

The error term capturing uncertainty in pavement deterioration, ε_{it}, has been shown to have consistent variance, σ^2, across a dataset. The variance for u_{it} and $u_{i,t-1}$ can be approximated via first order methods. Specifically, if an arbitrary random variable, x, with finite variance is transformed by some function $g(.)$ that is differentiable, as is our case, the first-order approximation of the underlying variance for u_{it} and $u_{i,t-1}$ follows:

$$\mathrm{Var}[g(x)] = \left(g'(\mathrm{E}[x])\right)^2 \mathrm{Var}(x) \tag{6}$$

$$\mathrm{Var}[u_{it}] = \frac{\sigma^2_{m_{it}}}{\mu_{it}{}^2 n_{it}} \tag{7}$$

$$\mathrm{Var}\left[u_{i,t-1}\right] = \frac{\sigma^2_{m_{i,t-1}}}{\mu_{i,t-1}{}^2 n_{i,t-1}} \tag{8}$$

where σ^2_m is the variance in the measured condition of a facility, μ is the average measured condition of a facility, and n is the number of measurements during a site visit (typically five). Should the weights be appropriately specified, then one would anticipate the weighted residuals would follow a standard normal distribution.

The major point of departure in our work and those of previous pavement researchers is the incorporation of a another fixed-effect across time: year of construction, which is referred to as *YEAR* in equation (9) and equation (10). By incorporating year of construction, we are able to easily extract an intuitive estimate around the annual rate of improvement in the life-cycle performance of a facility derived from quality gains in its production beyond those related to its planned structural capacity (measured via SN). We estimate equation (9) and equation (10) to test the sensitivity of our *Year* variable to the inclusion of AADTT which, at an aggregate level, has remained fairly constant across the temporal horizon of our PPIs.

$$\Delta \ln D_{it} = \alpha_1 + \alpha_2 \ln SN + \alpha_3 \ln AADTT + \alpha_4 \ln YEAR \tag{9}$$

$$\Delta \ln D_{it} = \alpha_1 + \alpha_2 \ln SN + \alpha_4 \ln YEAR \tag{10}$$

Of course, the rate of performance improvement may vary across time periods, which would motivate an approach similar to the hedonic regression with time dummies as described in the literature review section. However, there are two important challenges in our context that have motivated us to model quality improvements per this approach: (1) the significant time lag (i.e., 5-10 years) to receive performance updates on a recently constructed highway and (2) the limited number of constructed highways available to us post-2005. The proposed linear model allows us to generate a first order, year-over-year estimate of annual quality improvements relying on a large portion of data generated prior to 2005. Future research efforts on our part will include the collection and analysis of more recently constructed roads that can be used to employ the time dummy approach.

With our two datasets, we are able to monitor across time (1) changes in highway production inputs and (2) improvements in the performance of the resulting outputs. We can use these two insights to update our baseline producer price index, which we do so in our study by redefining our output.

A Quality-Adjusted Highway Producer Price Index

Roads and highways form a valuable commodity that support the transport of goods, services, and people across the economy. Federal, state, and municipal governmental agencies not only care about the total quantity of paved surfaces (typically measured in units of lane-miles) across their networks but also their value to their citizens. One possible mechanism to embed measured quality improvements in a producer price index is to 'redefine the good.' Sheiner and Malinovskaya (2016) provide an excellent synthesis on the use of this approach for previous health care studies, highlighting that we can better measure output in medical treatments by monitoring their success (e.g., added life-years) rather than their total number.[25]

Similar to the medical sector, a key challenge in redefining our output for highway construction is reaching a consensus on a preferred metric. While there is no formal agreement between researchers and agencies on such a measure, the topic has been a point of discussion among practitioners and governmental officials following the enactment of the Moving Ahead for Progress in the 21st Century Act (MAP-21) in 2012. The program, which has funded federal surface transportation projects in the United States, includes requirements for state DOTs to establish performance and outcome-based programs in the management of their bridge and pavement assets. One particular performance-based metric that both FHWA and the infrastructure management community has broadly supported is the service interval for a facility (i.e., time until a major construction activity will be required) (Elkins et al. 2013). The longer the time between construction activities, the lower the cost of the good over its lifetime for agencies with limited fiscal resources. While an alternative definition of highway construction could be one that not only adjusts for a facility's serviceable life but also its supported traffic, there has been little change over our time period of focus (2005 to 2017) in the utilization of arterial and interstate roads, which primarily fall under the jurisdiction of state DOTs, in the United States (**Figure 5**).

Having redefined output of our good by its service interval (SI), we are able to adjust our original producer price index, C_t, by normalizing it by its SI[26]. We refer to our adjusted producer price index time period t and $t-1$ as C^*_t and C^*_{t-1}:

$$C_t^* = {C_t}/{SI_t} \qquad (11)$$

. . .

25. Sheiner and Malinovskaya (2016) discuss three approaches for incorporation; cost of living, redefining the good and cost of quality improvements. Among the three the 'redefine the good' approach is intuitively appealing for our purpose. The product being purchased is not an output like 'miles of road' but an outcome, such as additional years of service for connectivity. Connectivity can be seen as a function of lane miles, roadway management, time, governance, environment, etc.).

26. See page 23 of Sheiner and Malinovskaya (2016) for a similar derivation but for quality-gains in health care.

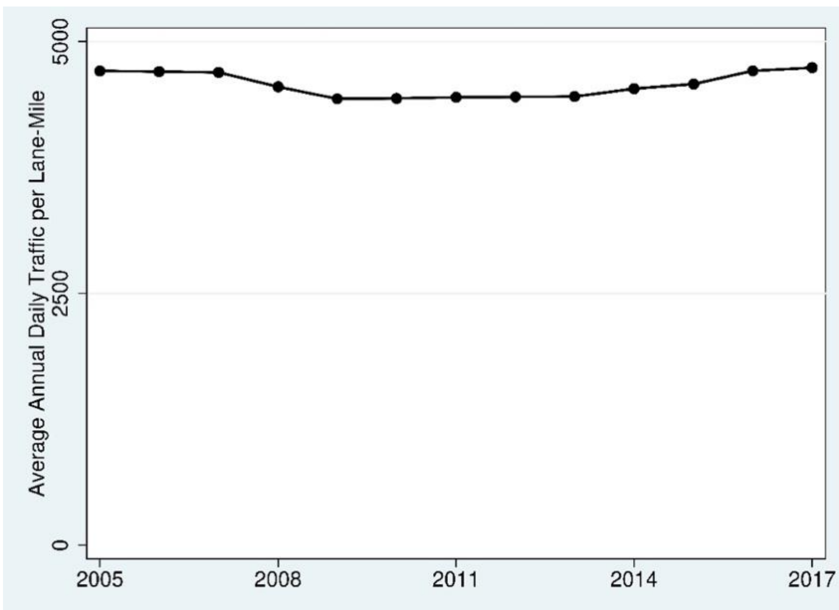
$$C^*_{t-1} = {C_{t-1}}/{SI_{t-1}} \quad (12)$$

$$C^*_t = \frac{C_t}{SI_t} = \frac{C_{t-1} \times R_t}{SI_t} = \frac{C^*_{t-1} \times SI_{t-1} \times R_t}{SI_t} \quad (13)$$

$$C^*_t = C^*_{t-1} \times R_t \times \frac{SI_{t-1}}{SI_t} \quad (14)$$

where SI_{t-1} and SI_t is the average service interval for a newly constructed facility in time periods t-1 and t.

Figure 5. Average annual daily traffic per lane on arterial roads and interstates from 2005-2017 per FHWA's *Highway Statistics 2018*.



We can easily compute the service interval for a typical highway facility constructed in a given year via our estimated performance model. For sake of simplicity, we can assume that the average rate of deterioration for a facility constructed in a given year can be approximated by r, which accounts for the previously discussed factors (e.g., structural capacity) that impact the performance of a highway. State DOTs will have a general expectation around the initial condition of any constructed facility, D_0, and a critical, threshold value in which the roadway will be viewed as unacceptable and require an intervention in the future, $D_{critical}$. With these three variables, the service interval for a newly constructed highway follows:

$$D_{critical} = D_0 e^{r*SI} \quad (15)$$

$$SI = \frac{\ln D_{critical} - \ln D_0}{r} \quad (16)$$

With average performance rates, r_t and r_{t-1}, in time periods t and t-1, we can further simplify our ratio between SI_{t-1} and SI_t as part of our updated producer price index via:

$$\frac{SI_{t-1}}{SI_t} = \frac{\ln D_{critical} - \ln D_0 / r_{t-1}}{\ln D_{critical} - \ln D_0 / r_t} = r_t / r_{t-1} \tag{17}$$

$$C_t^* = C_{t-1}^* \times R_t \times \frac{r_t}{r_{t-1}} \tag{18}$$

Beyond its simplicity, there are at least two other advantages with our proposed producer price index quality adjustment. First, for many highway projects completed over the last 20-30 years, we do not know the true serviceable life of the highway given that an intervention has still yet to take place. Our performance model allows us to estimate a serviceable life in the absence of such information. Second, state DOTs will frequently vary in their definition of $D_{critical}$ (e.g., see Chen, Hildreth, and Mastin 2019); our proposed producer price index, however, is unaffected by its value. In fact, it is fully possible that our approach actually still *underestimates* quality improvements in highway construction given that QC/QA programs have sought to improve the initial condition, D_0, for highway facilities. A better initial condition should play a role in increasing the serviceable lifetime of a facility. Unfortunately, due to the nature of our dataset, we have limited access to initial condition information and are therefore unable to comment on its evolution over time.

In the following section, we demonstrate the impact of measured quality gains in highway construction on our producer price indices by applying this approach to our asphalt concrete series. We have decided to only apply it to this index given that nearly 94 percent of highways across the United States are constructed with this technology and, more importantly, the available sample size for concrete-based activities is limited. We still present our three other unadjusted indices as part of our results given that they will be of utility for future productivity studies in the absence of comparable indices in the existing literature. The results for the asphalt concrete series should at least provide an order of magnitude estimate of possible adjustment rates for the other three indices.

Empirical Results

Figure 6 and **Figure 7** show our baseline, chained-Fisher producer price indices across technologies and activities. In general, these producer price indices for highway construction have grown faster than the consumer price index. Although we do not show FHWA's NHCCI in these figures, its pattern and overall shape is similar to that of the asphalt concrete producer price index. This result makes intuitive sense given that it is the predominant technology used in highway construction. Having said that, average producer price growth is faster for our asphalt concrete producer price index that FHWA's NHCCI. This could be due to the removal of other technologies and activities exhibiting lower price growth, differences in criteria rules used to identify outliers, or dissimilar item basket categories and coverage (FHWA uses approximately half as many item baskets). While future engagements with FHWA could help explain possible differences, we are generally pleased that these indices are, overall, quite similar to the benchmark, NHCCI.

Unsurprisingly, the average price to produce asphalt concrete pavements has outpaced Portland cement concrete highways. The rapid rise in crude oil prices in the late 2000s naturally affected prices for bitumen, which was already the costliest material input involved in its production prior to this event.

While falling outside the scope of our work, this fact helps to partially explain the lower growth in prices for asphalt maintenance and rehabilitation activities, which tend to require higher labor and lower material inputs than new construction projects. We can also note that our Portland cement concrete indices have exhibited higher volatility across time; this is largely due to the low sample size mentioned previously, in which the general construction and maintenance and rehabilitation indices for asphalt concrete rely on 500 percent and 350 percent more data, respectively, than their Portland cement concrete counterparts.

Figure 6. Chained-Fisher Producer Price indices for Asphalt Concrete (AC) and Portland Cement Concrete (PCC) Highway Construction.

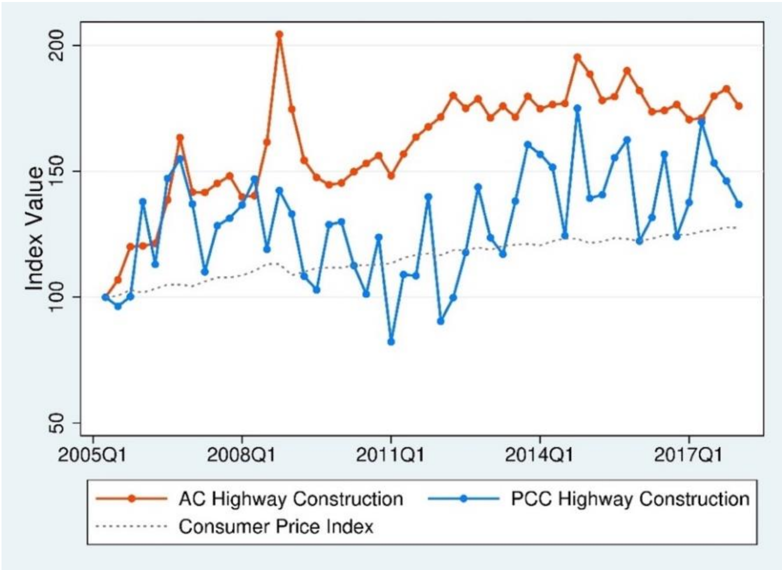


Figure 7. Chained-Fisher Producer Price Indices for Asphalt Concrete (AC) and Portland Cement Concrete (PCC) Maintenance and Rehabilitation (M&R).

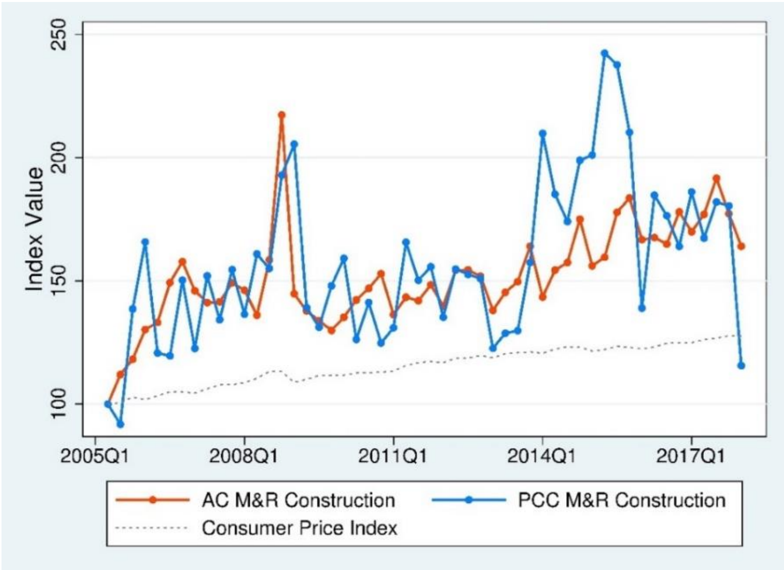
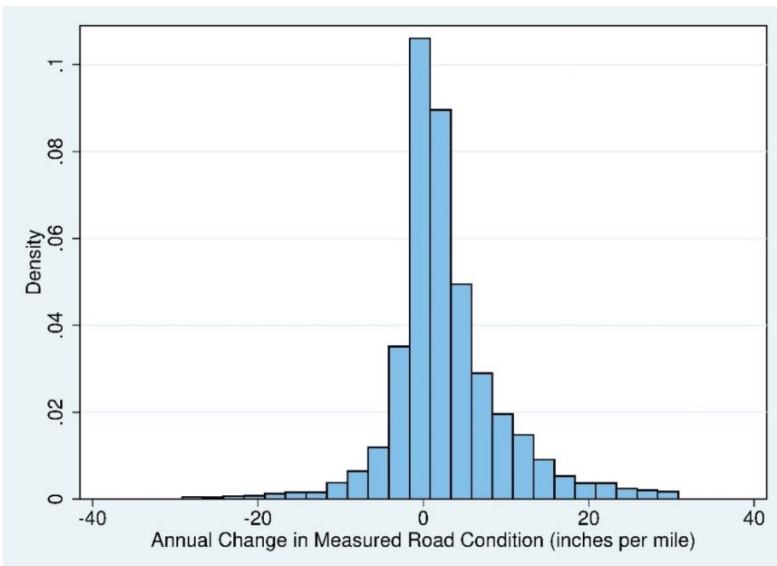


Figure 8 presents the average measured change in deterioration (inches per mile) for each observation in our longitudinal dataset. The average condition in time period *t* and *t*-1 is computed using five measurements with commonly available inspection technologies. This figure highlights a tremendous challenge in estimating the true performance of a constructed highway facility; that is, our samples are subject to considerable uncertainty due to the persistent errors underlying conventional inspection methods. As can be noted in **Figure 8**, for a large number of instances, a constructed facility has experienced a negative change in its average measured roughness between time periods, which is counter to the idealized saw-tooth diagram shown previously in **Figure 3**. While the underlying distribution is skewed to the right, if inspection technologies were perfectly accurate, we would anticipate no observations with a negative value. Despite this challenge, researchers have demonstrated that a high-fidelity performance model can be estimated despite this highlighted issue if specified properly.

Figure 8. Measured average change in pavement roughness (i.e., performance) across consecutive time periods for each facility.



Note: The histogram removes instances in which the measured deterioration improves between time periods due to a possible intervention activity.

Table 2 distills the estimated performance model for asphalt concrete highway construction. In general, all fixed-effects reject the null hypothesis that they are equal to zero at the 5 percent level and are directionally as anticipated. A highway (1) constructed with a higher planned structural capacity (captured via its SN); (2) subject to lower truck traffic; and (3) more recently produced has, on average, exhibited a lower rate of deterioration and a higher serviceable life per Model 2. The impact of anticipated structural capacity, furthermore, does not alter considerably if we remove construction year from the analysis per Model 1.

The parameter estimates of -0.81 and -0.82 for α_4 can be used to characterize an annual rate of improvement in performance for highways due to advances in its production when holding constant the planned material inputs used in its design. Since our study emphasizes the implication of these findings on productivity measures, we focus our attention on the resulting producer price indices based on these findings.

Table 2. Specified pavement performance model. Parameters with a **/* reject the null hypothesis at the 5%/10% level.

Parameter	Model 1 Estimate	Model 2 Estimate	Model 3 Estimate
α_1	0.0240**	6.1673**	6.2924**
α_2	-0.0144**	-0.0129**	-0.0090*
α_3	0.0057**	0.0056**	
α_4		-0.8094**	-0.8222**
Sample Size:	1,039	1,039	1,039

Specified Model:

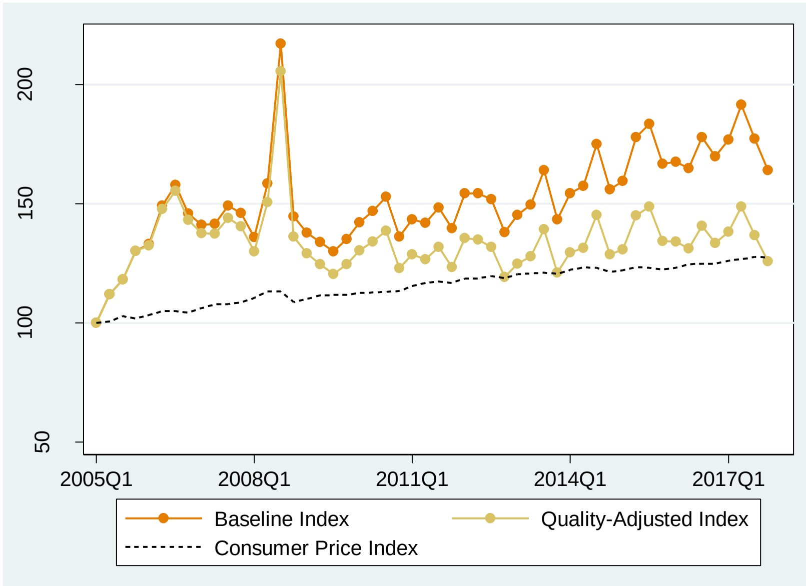
$$\Delta \ln D_{it} = \alpha_1 + \alpha_2 \ln SN + \alpha_3 \ln AADTT + \alpha_4 \ln YEAR$$

Figure 9 plots our quality-adjusted producer price index for asphalt concrete highway construction based on our specified deterioration models. This has been accomplished by integrating (1) our original producer price index with (2) our deterioration models that track both temporal quality improvements in construction and typical structural designs over time via equation (18) as part of our "redefine the good" approach. The proposed producer price index exhibits lower annual growth by 2.0 percent, a non-trivial amount in the context of productivity research. The divergence in these indices is particularly striking from 2010 onwards as the effects of upward bias in the original price index compound.

While this finding is significant and considerable, we recognize important shortcomings in this analysis. For example, our modeling approach computes an average quality improvement rate across time rather than considering possible shifts across time. We have primarily done this because we have access to data that do not necessarily overlap with our bid dataset temporally. While not used in this study, our previous initiatives with a select few state DOTs have provided us with more recent information. However, these data are not necessarily representative of the national average nor do they fully address the significant time lag (i.e., 5-10 years) in our context to receive performance updates for recently constructed highways.

Despite these shortcomings, our initial findings suggest a potentially strong, upward bias in existing producer price deflators for highway construction. Given that most productivity studies in our context use revenue as an output measure, an upwardly-biased price deflator will lead to the estimation of lower productivity growth over time. We view this as a valuable contribution to the productivity literature with further reflections and discussions presented in the following section.

Figure 9. Our original chained-Fisher producer price indices (referred to as baseline) and quality-adjusted producer price indices for asphalt concrete highway construction.



Summary and Conclusions

From 1968 onwards, reported productivity growth in the U.S. construction sector has continually lagged productivity in the rest of the economy. In fact, U.S. Department of Commerce data indicate that between 1997 and 2017 construction productivity decreased by nearly 30 percent. Early authors have attempted to explain these low productivity growth measures as a consequence of shifts in output design, changes in market structure and firm size, reduced labor quality, reductions in the capital/labor ratio and institutional changes such as unionization. In addition to these claimed real factors, there have also been claims that measurement errors in the underlying price deflators, which are used to develop real output measures, are upward biased, resulting in a downward bias in measured productivity.

While early studies of construction showed zero to negative productivity growth, these measures have come into question following sub-sector and subsequent activity-specific studies. These latter studies have shown positive productivity increases stemming from technological advances, falling real wages and increases in the capital/labor ratio. The most recent attention around this subject has been led by the U.S. Bureau of Labor Statistics. This effort sought to generate improved measures of labor productivity growth across four subsectors of construction; single and multiple family residential construction, industrial construction and highway, streets and bridge construction. The emphasis was on the use of accurate, relevant price deflators to use against revenues to create a real output index over time.

Our research has aimed to further improve the price deflators for the highway construction sector by creating new price deflators that can capture quality improvements and can be disaggregated across technologies and activities. A number of authors have pointed out that quality improvements have taken place in both materials and capital inputs. In addition, expenditures on operations and maintenance have shifted from 37 to 46 percent in the time period considered by the BLS research.

The starting point for our work was with the considerable research work that has been carried out by engineers and state DOTs in developing analytical models of pavement performance. This has been driven by the challenge of maintaining an aging roadway system. This required collecting extensive information on highway condition information and forecast future performance. The majority of states rely on a roadway roughness index to measure performance. The value of the roughness index stems from its ease of collection, reliability relative to other measures and its strong correlation with other measures of pavement performance. The measure of quality changes is based on roughness measurements for over 2,500 roadway sections in North America.

Initially, we calculated baseline chained-Fisher producer price indices developed from price data from individual highway construction projects from 2005-2017 in the 48 contiguous U.S. states. Within this dataset are over 5,000 unique expenditure items which we distilled to 60 unique item baskets. A key feature of our index was our redefinition of 'output' from simply miles of roadway to include the 'service life of the miles.' This yields a modified price index which accounts for average performance rates between time periods. The producer price index was developed for two technologies of Portland cement concrete (PCC) and asphalt concrete (AC) which were each separated by general construction maintenance and rehabilitation. The quality parameter was determined using hedonic regression of the anticipated structural capacity of a facility and year of construction on the change in roadway deterioration (roughness). From the two datasets we were able to distinguish changes in inputs and changes in quality improvements.

Our results show the chained-Fischer producer price indices across technologies (PCC and AC) and activities have grown faster than the consumer price index and AC technology has outpaced the PCC technology. The baseline chained-Fischer produce price index has grown faster than the quality-adjusted price index by 2.0 percent. Since the price index appears in the denominator in calculating real output, our results imply productivity growth has been underestimated.

A reasonable question that one may ask is 'so what'? Does this new result change anything? Is productivity growth at an acceptable level? The construction industry supports over 7.6 million jobs in the United States and contributes more than 4 percent to national GDP. Having an accurate measure of productivity change is clearly important since productivity growth will contribute substantially to overall living standards. While we may want to understand how to improve the industry's productivity, we must first know where we presently stand: in that regard, getting the measurement right is essential.

While 2.0 percent is significant, natural follow-up questions are 'Why does productivity change in construction?' and 'How do we invest to improve the well-being of people?' Certainly, the shift in spending from new construction to more operations and maintenance will mean lower productivity with conventional measures. Such activities are labor intensive and limited state budgets will mean it is unlikely we would see a shift to more capital-intensive technologies. The collection of firm-level data would likely help us in this regard to better understand how market structure, management structure and technical and materials adoption affect productivity. For example, Syverson (2004) describes the effects of product substitutability on the selection of firms and the equilibrium dispersion of firm productivity. When products become more substitutable production within an industry relocates. Less productive firms disappear and output shifts toward more productive firms. There is strong evidence that a higher degree of substitutability leads to narrow productivity dispersion and a higher median productivity. There is also evidence that different technologies, asphalt concrete and Portland cement concrete for example, operate differently.

The research also identifies gaps and challenges. We chose to redefine our product as an outcome rather than an output. Are there alternative definitions we could use that would improve the results? How sensitive would the results be to these definitions? One of the limitations of our data is there is a large time gap in knowing how well a facility actually performs; could other data sources from certain state DOTs allow us to further enhance our producer price index? While we have cited many sources of quality improvement in construction, we do not know which ones are driving quality gains in the sector. For example, there is a shift to greater capital intensity but the usability of such capital and robustness against breakdown will affect the variation in productivity measures. Such improvements naturally occur at the firm level before being aggregated to subsectors. Having firm-level data would address the difficulty of dealing with heterogeneous outputs and the large and diverse set of activities, both of which plague the creation of an accurate price index. Additionally, the efficient use of material and capital inputs by firms likely not only depends on their individual actions but also the context in which they operate. What is the impact, for example, of constructing in urban versus rural locations on the productivity of a firm? Firm-level data would also allow us to address questions posed by Syverson (2004) and understand the distribution of differences in multi-factor productivity (MFP) across firms. It is, after all, not just labor productivity but also MFP that is of interest.

REFERENCES

Abulizi, N., Akiri Kawamura, Kazuya Tomiyama, and Shun Fujita. 2016. "Measuring and evaluating of road roughness conditions with a compact road profiler and ArcGIS." *Journal of Traffic and Transportation Engineering*, 3(5): 398-411.

Allen, Steven. 1989. "Why Construction Industry Productivity is Declining: Reply." *The Review of Economics and Statistics*, 71(3): 547-549

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

Allmon, Eric, Carl Haas, John Borcherding, and Paul Goodrum. 2000. "U.S. Construction Labor Productivity Trends, 1970-1998." *Journal of Construction Engineering and Management,* 126(2): 97-104

Amekudzi-Kennedy, Adjo, Samuel Labi and Prerna Singh. 2019. "Transportation Asset Valuation: Pre-, Peri- and Post Fourth Industrial Revolution." *Transportation Research Record: Journal of the Transportation Research Board*, 2673(12): 163-172.

Aschauer, D.A. 1989. "Is public expenditure productive?" *Journal of Monetary Economics*, 23(2): 177-200

Ben-Akiva, Moshe, and Rohit Ramaswamy. 1993. "An approach for predicting latent infrastructure facility deterioration." *Transportation Science*, 27(2): 174-193.

Bennett, Jennifer, Robert Kornfeld, Daniel Sichel and David Wasshausen. 2020. "Measuring Infrastructure in BEAs National Economic Accounts", Research paper part of the *NBER Economics of Infrastructure Project*, February 2020.

Brooks, Leah and Zachary Liscow. 2019. "*Is Infrastructure Spending Like Other Spending?*" Working paper presented at *NBER Economics of Infrastructure Conference*, November 2019, Cambridge, MA.

Brooks, Leah and Zachary Liscow. 2020. "*Infrastructure Costs*". SSRN-ID 3428675, March 2020.

Chang, George, Xu Qinwu, Jennifer Rutledge, Larry Michael, David White, and Pavana Vennapusa. 2011. *Accelerated Implementation of Intelligent Compaction Technology for Embankment Sub grade Soils, Aggregate Base, and Asphalt Pavement Materials* Washington, D.C.: U.S. Department of Transportation, Federal Highway Administration.

Chen, Don, John Hildreth, and Neil Mastin. 2019. "Determination of IRI Limits and Thresholds for Flexible Pavements." *Journal of Transportation Engineering Part B: Pavements*, 145(2): 04019013.

Chu, Chih-Yuan, and Pablo Durango-Cohen. 2008. "Estimation of infrastructure performance models for transportation infrastructure using panel data." *Transportation Research Part B: Methodological*, 21(1): 57-81.

Congressional Budget Office. 2018. *Public Spending on Transportation and Water Infrastructure, 1956 to 2017*. Washington, D.C.: Congressional Budget Office.

D'Apuzzo, Mauro, and Vittorio Nicolosi. 2010. "A New Methodology for Stochastic Modelling of Pay Factors in Hot-Mix Asphalt Pavements." *Road Materials and Pavement Design*, 11: 559-585.

Elkins, Gary E., Travis M. Thompson, Jonathan L. Groeger, Beth Visintine, and Gonazlo Rada. 2013. *Reformulated Pavement Remaining Service Life Framework*. Washington, D.C.: U.S. Department of Transportation, Federal Highway Administration.

Fernando, Emmanuel. 1997. "Applicability of New Flexible Pavement Smoothness Specification for Asphalt Overlays." *Transportation Research Record: Journal of the Transportation Research Board*, 1575: 18-24.

Fraumeni, Barbara. 1999. "Productive Highway Capital Stock Measures", Washington, D.C.: U.S. Department of Transportation, Federal Highway Administration, Contract DTFH61-97-C-00010, BAT-98-006.

Fraumeni, Barbara. 2007. "Productive Highway Capital Stocks and the Contribution of Highways to Growth of GDP, Volume I", Washington, D.C.: U.S. Department of Transportation, Federal Highway Administration.

Fraumeni, Barbara. 2008. "Productive Highway Capital Stocks and the Contribution of Highways to Growth." Paper presented at *NBER Summer Institute: Conference on Research in Income and Wealth* (July 15)

Fraumeni, Barbara. 2009. "The Contribution of Highways to GDP Growth." *National Bureau of Economic Research Working Paper No. 14736*.

Garg, Arun, Alan Horowitz, and Fred Ross. 1988. "Establishing Relationships Between Pavement Roughness and Perceptions of Acceptability." *Transportation Research Record: Journal of the Transportation Research Board*, 1196: 276-285.

Goodrum, Paul, Dong Zhai, and Mohammed Yasin. 2009. "Relationship between Changes in Material technology and Construction Productivity." *Journal of Construction Engineering and Management,* 135(4): 278-287.

Goodrum, Paul, and Carl Haas. 2004. "Long-Term Impact of Equipment Technology on Labor Productivity in the U.S. Construction Industry at the Activity Level." *Journal of Construction Engineering and Management*, 130(1): 124-133.

Goodrum, Paul, and Carl Haas. 2002. "Partial Factor Productivity and Equipment Technology Change at Activity Level in US Construction Industry." *Journal of Construction Engineering Management*, 128(6): 463-472.

Griliches, Zvi. 1961. "Hedonic price indexes for automobiles: An econometric analysis of quality change." *The Price Statistics of the Federal Government*, New York, USA: Columbia University Press.

Hall. Robert. 1971. "The measurement of quality change from vintage price data." in: Z. Griliches, ed., Price indexes and Quality Change, Cambridge, MA, USA: Harvard University Press, 240-271.

Hong, Feng, and Jorge Prozzi. 2015. "Pavement Deterioration Model Incorporating Unobserved Heterogeneity for Optimal Life-Cycle Rehabilitation Policy." *Journal of Infrastructure Systems*, 21(1): 0414027.

Li, Qiang, Danny Xiao, Kelvin Wang, Kevin Hall, and Yanjun Qiu. 2011. "Mechanistic-empirical pavement design guide (MEPDG): a bird's-eye view." *Journal of Modern Transportation*, 19(2): 114-133.

Madanat, Samer, and Wan Hashim Wan Ibrahim. 1995. "Poisson Regression Models of Infrastructure Transition Probabilities." *Journal of Transportation Engineering*, 121(3): 267-272.

Madanat, Samer, Matthew Karlaftis, and Patrick McCarthy. 1997. "Probabilistic infrastructure deterioration models with panel data." *Journal of Infrastructure Systems*, 3(1): 4-9.

Nassereddine, Hala, Gary C. Whited, and Awad S. Hanna. 2016. "Developing a Chained Fisher Construction Cost Index for a State Highway Agency." *Transportation Research Record: Journal of the Transportation Research Board*, 2573: 149–156.

OECD (2001), *Measuring Productivity: Measurement of Aggregate and Industry Level Productivity Growth*: OECD Manual, Paris, France.

Ouyang, Yanfeng, and Samer Madanat. 2006. "An analytical solution for the finite-horizon pavement resurfacing planning problem." *Transportation Research Part B: Methodological*, 40(9): 767-778.

Parker, Robert. 1991. "Comment: The Measurement of Construction Prices: Retrospect and Prospect", Chapter 8 in Ernst Berndt and Jack Triplett (eds), *Fifty years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth*, University of Chicago Press, 268-272.

Prozzi, Jorge, and Samer Madanat. 2003. "Incremental Nonlinear Model for Predicting Pavement Serviceability." *Journal of Transportation Engineering*, 129(6): 635-641.

Schwartz, Charles. 2007. "Implications of Uncertainty in Distress Measurement for Calibration of Mechanistic-Empirical Performance Models." *Transportation Research Record: Journal of the Transportation Research Board*, 2037: 136-142.

Sheiner, Louise, and Anna Malinoskaya. 2016. *Measuring Productivity in Health Care: An Analysis of the Literature*, Washington, D.C.: Hutchins Center on Fiscal and Monetary Policy, Brookings Institution.

Shrestha, K. Joseph, H. David Jeong, and Douglas Gransberg. 2017. "Multidimensional Highway Construction Cost Indexes Using Dynamic Item Basket." *Journal of Construction Engineering and Management*, 143(8): 04017036

Silver, Mick. 2018. "How to Measure Hedonic Property Price Indexes Better (European Review on National Accounts and Macroeconomic Indicators." *Eurostat Review on National Accounts and Macroeconomic Indicators*.

Sveikauskas, Leo, Samuel Rowe, and James Mildenberger. 2018. "Measuring Productivity Growth in Construction." *Monthly Labor Review*, 141(1): 1-33.

Sveikauskas, Leo, Samuel Rowe, James Mildenberger, Jennifer Price, and Arthur Young. 2016. "Productivity Growth in Construction." *Journal of Construction Engineering and Management*, 142(10): 1-10.

Sveikauskas, Leo, Samuel Rowe, James Mildenberger, Jennifer Price and Arthur Young. 2014. "Productivity growth in Construction." Working Paper 478, Washington, D.C.: Bureau of Labor Statistics, U.S. Department of Labor.

Swei, Omar, Jeremy Gregory, and Randolph Kirchain. 2018. "Does pavement degradation follow a random walk with drift? Evidence from variance ratio tests for pavement roughness." *Journal of Infrastructure Systems*, 24(4): 04018027.

Syverson, Chad. 2011. "What Determines Productivity." *Journal of Economic Literature*, 49(2): 326-365.

Syverson, Chad. 2008. "Markets: Ready-Mixed Concrete.", *Journal of Economic Perspectives*, 22(1): 217-233

Syverson, Chad. 2004. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics*, 86(2): 534-550.

Triplett, Jack E. 2011. "Health System Productivity." *Oxford Handbook of Health Economics*. New York, USA: Oxford University Press, 707-32.

Triplett, Jack E. 2006. *Handbook of Hedonic Indexes and Quality Adjustments in Price Indexes, Special Application to Information Technology*, OECD, Paris

U.S. Department of Transportation. 2018. *Highway Statistics 2018*. Washington, D.C.: U.S. Department of Transportation.

U.S. Department of Transportation. 2020. Status of the Nation's Highways, Bridges, and Transit. Washington, D.C.: U.S. Department of Transportation.

Yehia, Ayatollah, and Omar Swei. 2020. "Probabilistic infrastructure performance models: an iterative-methods approach." *Transportation Research Part C: Emerging Technologies*, 111: 245-254.

Zhai, Dong, Paul Goodrum, Carl Haas and Carlos Caldas. 2009. "Relationship between Automation and Integration of Construction Information Systems and Labor Productivity." *Journal of Construction Engineering and Management,* 135(8): 746-753.

Hutchins Center
on Fiscal & Monetary Policy
at BROOKINGS

The mission of the Hutchins Center on Fiscal and Monetary Policy is to improve the quality and efficacy of fiscal and monetary policies and public understanding of them.

Questions about the research? Email communications@brookings.edu.
Be sure to include the title of this paper in your inquiry.