Modelling Patterns in Construction Labour Productivity

Asitha Rathnayake (corresponding author), (agbrr2@cam.ac.uk)
Laing O'Rourke Centre for Construction Engineering and Technology, University of Cambridge, United Kingdom

Danny Murguia, (dem52@cam.ac.uk)
Laing O'Rourke Centre for Construction Engineering and Technology, University of Cambridge, United Kingdom

Ashan Senel Asmone, (asa79@cam.ac.uk)
Laing O'Rourke Centre for Construction Engineering and Technology, University of Cambridge, United Kingdom

Campbell Middleton, (crm11@cam.ac.uk)
Laing O'Rourke Centre for Construction Engineering and Technology, University of Cambridge, United Kingdom

Keywords: construction, productivity, measurement, learning

Abstract

The construction industry has abundant data but lacks sufficient actionable information. Deriving such information is crucial for addressing one of its primary challenges: insufficient productivity. This paper presents a workflow to extract, analyse and present available productivity data on sites in a consistent manner. It was demonstrated using data from structural frames of 11 buildings, totalling 116 floor levels. Seventy-five percent of the levels were found to spend over 3 worker-hours per m^2 of floor area on average. Additionally, most buildings exhibited cubic learning curves, showing performance improvements in mid-levels due to worker familiarisation. However, unlike previous studies, all the buildings demonstrated significantly poorer labour productivity towards the end. This might be due to not minimising the workforce proportionally to reduced work, coupled with complexities in logistics and design on higher levels. While having a learning curve does not signify good performance, its shape was found to reflect labour management practices on projects.

1. Introduction

Productivity is defined as the ratio of output to input or input to output. The most common metric of productivity used in construction research is labour productivity, where the input is measured in worker-hours (Rathnayake & Middleton, 2023). Construction has historically suffered from low levels of productivity, and one of the reasons for this is the lack of consistent productivity measurement methods (Hwang & Soh, 2013). This inconsistency has made it difficult to compare and combine productivity data from different projects. Additionally, the studies developing productivity measurement frameworks for projects (e.g. Ayele and Fayek (2019)) typically focus more on distinguishing between good and bad performance rather than understanding how productivity varies in projects over time. On the other hand, studies that explore the variation of project productivity over time (e.g. Nguyen and Nguyen (2013) and Pellegrino et al. (2012)) do not analyse the quality of the absolute productivity values. This highlights a literature gap between these two types of studies. The aim of this study is to bridge this gap by developing a consistent productivity measurement framework that can be used to establish benchmarks for productivity and understand how productivity varies over time.

2. Learning Curve Theory

The learning curve theory is a theory that discusses how construction productivity varies over time (Thomas et al., 1986). It states that the unit inputs (i.e. time, worker-hours and cost) required to complete repetitive activities decrease as the number of repetitions increases. The observed learning effect in construction activities results from multiple factors, such as increased worker familiarisation, better coordination of equipment and crews and the development of more efficient techniques (Thomas et al., 1986). The learning effect is visually represented using learning curves, which illustrate the unit input (i.e. the input for the X^{th} unit) or cumulative average unit input (i.e. the total input up to and including the X^{th} unit divided by X) against the production quantity. These curves are usually presented on a log-log plot. The learning effect is quantified using the *learning rate*, which is related to the slope of the learning curve on a log-log plot according to Equation 1 (Lutz et al., 1994).

$$\text{Learning Rate} = 2^{\text{Slope of the learning curve}} \tag{1}$$

2.1. Straight-line Learning Curve Model

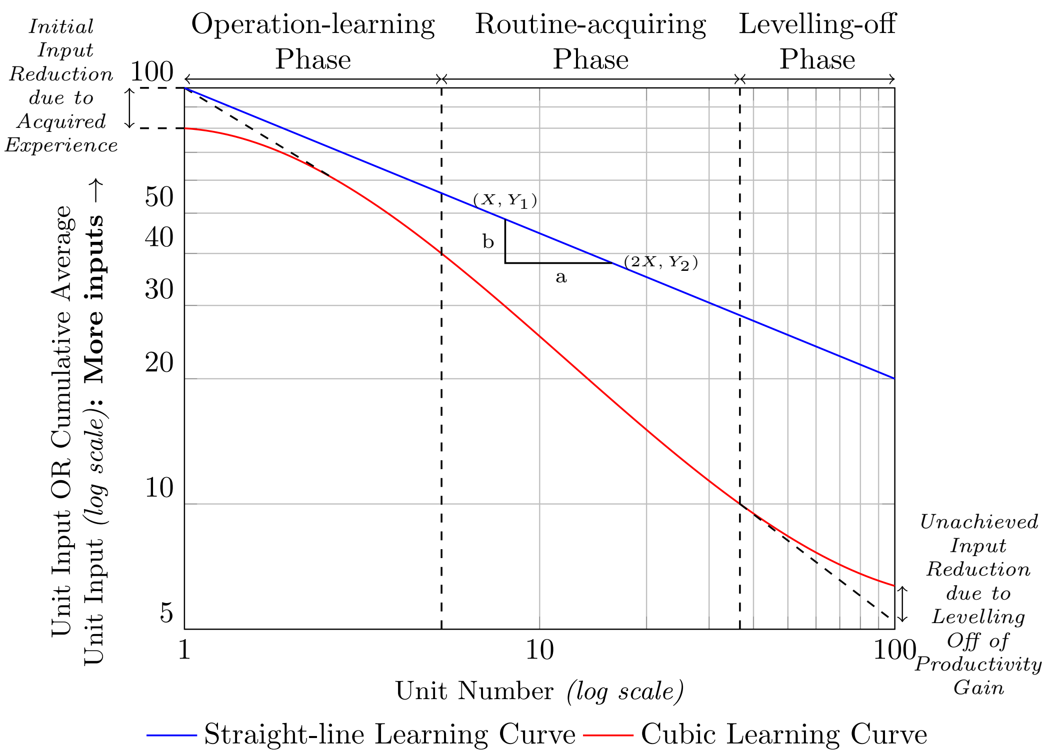
The most commonly used learning curve model is the straight-line model, which assumes a constant *learning rate* throughout an activity, resulting in a straight-line curve on a log-log plot (Thomas et al., 1986). Figure 1 illustrates this. Any learning curve has three phases: (1) the operation-learning phase, where input per unit rapidly decreases as workers learn the tasks (note that inverse relationships on log-log plots show a large variation on the Y-axis when there is a small variation on the X-axis at the start); (2) the routine-acquiring phase, where input per unit decreases more gradually as workers become familiar with the job and (3) the leveling-off phase, where input per unit reduction is minimal as the work approaches completion (Thomas et al., 1986).

Equation 1 for the *learning rate* can be simplified for a straight-line learning curve by calculating the slope between two points where production quantity doubles, such as from (X, Y_1) to $(2X, Y_2)$ in Figure 1 (see Equation 2). According to Equation 2, a lower *learning rate* indicates more learning, while a 100% *learning rate* indicates no learning.

$$\begin{aligned} \text{Learning Rate} &= 2^{\text{Slope of the learning curve}} \\ &= 2^{\frac{b}{a}} = 2^{\frac{\log_{10}(Y_2)-\log_{10}(Y_1)}{\log_{10}(2X)-\log_{10}(X)}} = 2^{\frac{\log_{10}(Y_2)/(Y_1)}{\log_{10}(2X)/(X)}} = 2^{\frac{\log_{10}(Y_2)/(Y_1)}{\log_{10}(2)}} = 2^{\log_2\left(\frac{Y_2}{Y_1}\right)} = \frac{Y_2}{Y_1} \\ &= \frac{\text{Unit input at production quantity 2X}}{\text{Unit input at production quantity X}} \quad \text{OR} \\ &\quad \frac{\text{Cumulative average unit input at production quantity 2X}}{\text{Cumulative average unit input at production quantity X}} \end{aligned} \tag{2}$$

2.2. Cubic Learning Curve Model

The cubic learning curve model features a varying *learning rate* and follows a cubic function (see Figure 1). Thomas et al. (1986) attribute its shape to two main factors: (1) workers having "acquired experience" due to prior experience with similar tasks in the



Note: This graph shows two hypothetical learning curves illustrating how inputs (i.e. time, worker-hours and cost) per unit or cumulative average input per unit decrease with the number of repetitive units.

Figure 1: Straight-line and Cubic Learning Curves

immediate past, leading to less input needed initially and (2) fewer opportunities for improvement in the leveling-off phase.

Everett and Farghal (1994) evaluated 12 different models using a dataset of 60 construction activities and found that cubic models best captured worker-hour variations. However, when only initial data points were available, the straight-line model provided better future performance estimates. In summary, while the cubic model generally offers the best approximations for construction activities, the straight-line model is also widely used for its simplicity and adequate accuracy (Couto & Teixeira, 2005; Nguyen & Nguyen, 2013; Pellegrino et al., 2012).

3. Research Method

3.1. Step 1: Developing the Workflow

First, meetings were conducted with the project teams of three case study projects from four Tier 1 contractors in the UK to identify the sources of productivity data. During these meetings, we also discussed the limitations of the available data and the assumptions that would need to be made. Next, the required data was extracted from various platforms, productivity calculations were conducted and the results were presented to the teams. The entire process of data extraction, analysis and visualisation was refined multiple times based on the input of project personnel. Finally, we used these findings to develop a standard workflow applicable to different projects across multiple companies.

3.2. Step 2: Application of the Workflow

Data from 11 commercial and residential buildings in London (Buildings 1-11), belonging to 8 projects, were used to demonstrate the developed productivity measurement workflow. The superstructure stage of these buildings was completed between 2020 and 2023. The sample size included 116 building levels, and we calculated the labour productivity of each of them, in terms of worker-hours per m^2 of floor area. Then, we calculated quartiles to understand different levels of labour productivity.

Next, we developed learning curves for each building. Superstructure worker-hours were considered as inputs (Y-axis), and building levels were taken as units (X-axis). Cumulative average inputs were used to smooth the variations between levels and identify underlying trends. Worker-hours were divided by the corresponding floor areas to adjust for differences among levels. This results in the labour productivity definition used in this study (i.e. worker-hours per m^2). Note that Level 1 of Buildings 1, 2, 4, 6 and 7 was excluded due to the use of different structural systems with non-repeated activities and different crew arrangements compared to the rest of the levels.

Linear and cubic regression models were tested for each building to identify the best fit. R^2 (coefficient of determination) values of 2%, 13% and 26% signify small, medium and large effect sizes between independent and dependent variables (Cohen, 1988). If two variables have a linear relationship, it is possible to find a quadratic or cubic model with a higher R^2 due to overfitting. Therefore, the shape of the data points was considered if both linear and cubic models showed high R^2 ($\geq$26%) values. Due to the log-log nature of the plots, more points are concentrated towards the ends, potentially skewing regression results. Therefore, a straight-line learning curve was assumed in cases where the curve demonstrated a sufficiently linear relationship from the start. Conversely, a cubic learning curve was assumed if only the cubic model showed high R^2 values.

In addition to the meetings with the project teams of the three case study projects (Buildings 1-4), we conducted eight interviews with participants from the other five projects (Buildings 5-11) to validate the proposed model. Their inputs were also used to explain the learning curve results.

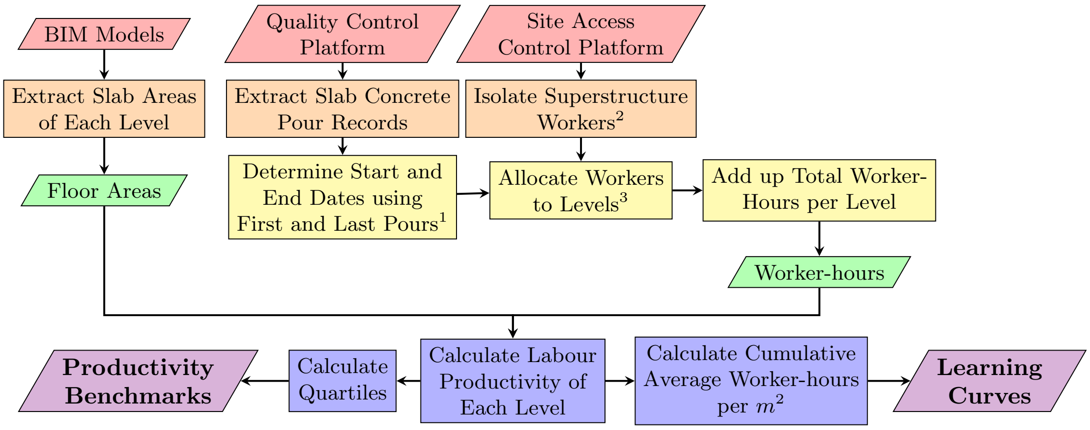
4. Workflow to Measure Construction Labour Productivity

Table 1 presents the data needed to calculate labour productivity and where they could be found on typical UK construction projects. One of the early efforts to determine the durations for each level was to use the master programmes updated by planners. However, it was found that planners would typically adjust only the end dates of activities on a weekly basis without changing start dates or durations. As an alternative, we decided to use installation records collected in the quality control platforms. Figure 2 presents the proposed workflow that shows the steps taken from extracting these data to presenting and analysis, as applicable for superstructure work in buildings. This workflow was validated using qualitative inputs from project participants.

Table 1: Sources of Productivity Data

Productivity Term	Data Needed per Level	Source
Output	Floor area	BIM models
Input	Start and end dates	Quality control platforms
Input	Worker-hours spent	Site access control platforms

Two practical considerations when using the workflow are as follows. First, periods of non-continuous work should be excluded when determining the start and end dates of levels to avoid overallocation of worker-hours. For example, there were instances



Note: This figure shows the steps to convert data from various sources (red) into productivity outputs (violet).
[1] Level X would typically start on the day after the first slab pour on Level X-1.
[2] Identify the trade descriptions of superstructure workers beforehand.
[3] Workers are typically equally divided among levels with concurrent work.

Figure 2: Proposed Workflow to Measure Superstructure Productivity

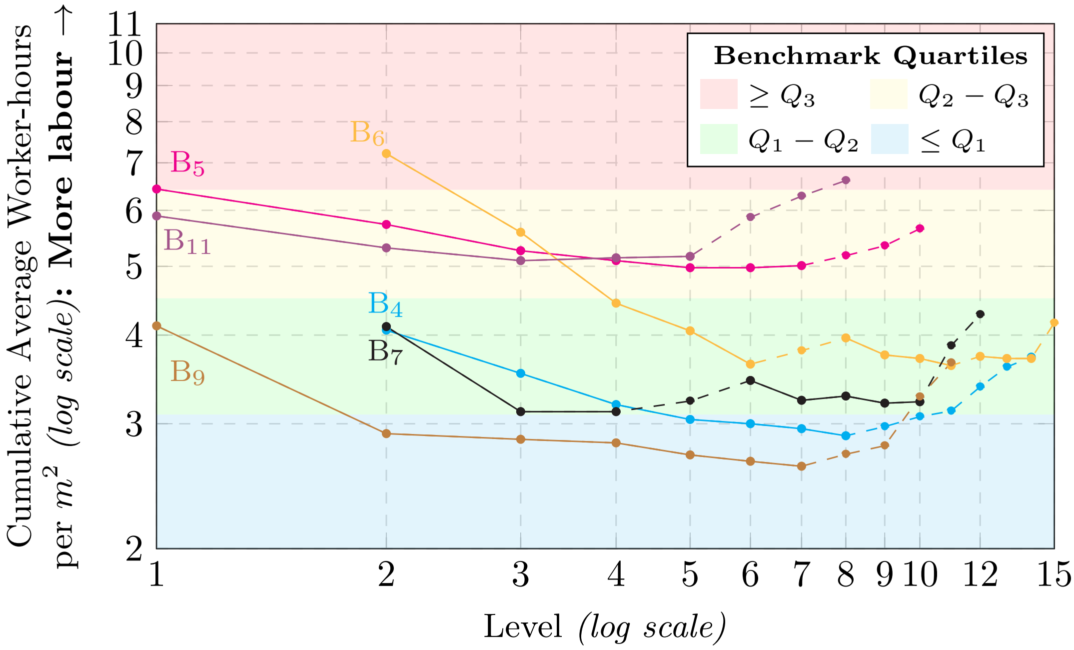
of work on a level being almost complete, but workers returning after a few months to pour concrete on a small slab portion. Such pours were excluded by checking the site cameras. Secondly, the recorded number of on-site hours from site access control platforms was found to be unreliable for estimating the actual working hours of each worker. The access control points (i.e. turnstiles) were sometimes located outside the work site, allowing workers to be in areas without work. However, the site access control platforms provided reliable information on how many workers were on-site each day. These values were multiplied by the number of paid hours to estimate the total number of working hours.

5. Productivity Benchmarks and Variations in Labour Productivity

The first, second and third quartiles of labour productivity were 3.1, 4.5 and 6.4 worker-hours per m^2, respectively. This indicates that 25%, 50% and 75% of the levels spent less than 3.1, 4.5 and 6.4 worker-hours per m^2 on average, respectively. The quartile ranges are visually represented in Figures 3, 4 and 5. For example, a building level falling into the red region has worse performance compared to at least 75% of the levels in the dataset. These values can be used as benchmarks for other similar projects.

Figures 3, 4 and 5 present the plots of $\log_{10}$(cumulative average worker-hours per m^2) versus $\log_{10}$(level) for the 11 buildings. For example, consider Building 5 in Figure 3. Its Levels 1 and 2 had floor areas of 1,614 and 1,663 m^2 and used 10,365 and 8,408 worker-hours, respectively. For Level 1, the cumulative average worker-hours per m^2 was $10365/1614 \approx 6.42$. For Level 2, it was $(10365 + 8408)/(1614 + 1663) \approx 5.73$. The remaining levels and buildings were plotted in a similar manner. The eleven buildings can be divided into three categories based on the shapes of the plots: (1) buildings that exhibit a cubic learning curve; (2) buildings that exhibit a straight-line learning curve and (3) buildings that exhibit no learning effect. Note that these graphs represent the "apparent learning effect" of the superstructure work package based on overall building performance and may not always reflect the improvement of individual workers over time. Thus, even in Figures 3 and 4, there are regions that exhibit no learning effect, indicated by dashed lines. During these periods, although workers may have increased the speed of performing tasks, overall performance may have suffered due to factors beyond their control (discussed in Section 5.1.2).

5.1. Buildings that Exhibit a Cubic Learning Curve



Note: Each coloured line represents the learning curve for one building. Dashed lines correspond to periods with no "apparent learning effect" for the superstructure work package as a whole.

Figure 3: Buildings with Cubic Learning Curves

Buildings 4, 5, 6, 7, 9 and 11 exhibited cubic learning curves (see Figure 3) with R^2 values of 98%, 99%, 96%, 67%, 86% and 96%, respectively. However, two main differences are noticeable between the cubic learning curves in Figure 3 and the 'ideal' cubic learning curve in Figure 1: (1) the initial input reduction due to acquired experience is not visible in Buildings 7, 9 and 11 and (2) the cumulative average labour productivity tends to worsen during the levelling-off phase (i.e. the cumulative average worker-hours per m^2 increases towards the end) in all six buildings.

5.1.1. Lack of Initial Input Reduction in Some Buildings

The negative slope of the learning curves for Buildings 4, 5 and 6 increased when transitioning from the first two levels to the second two (e.g. Level 2-3 compared to Level 3-4 in Building D). However, in Buildings 7, 9 and 11, there was a general reduction in negative slope during the initial levels. This indicates that Buildings 4, 5 and 6 experienced reductions in worker-hours at the outset due to workers having performed similar tasks in the immediate past (see Section 2.2). The common feature among Buildings 7, 9 and 11 was that they belonged to projects with multiple buildings employing different structural systems. Buildings 7, 9 and 11 began construction after the other buildings, which could explain the lack of experience with similar tasks in the immediate past.

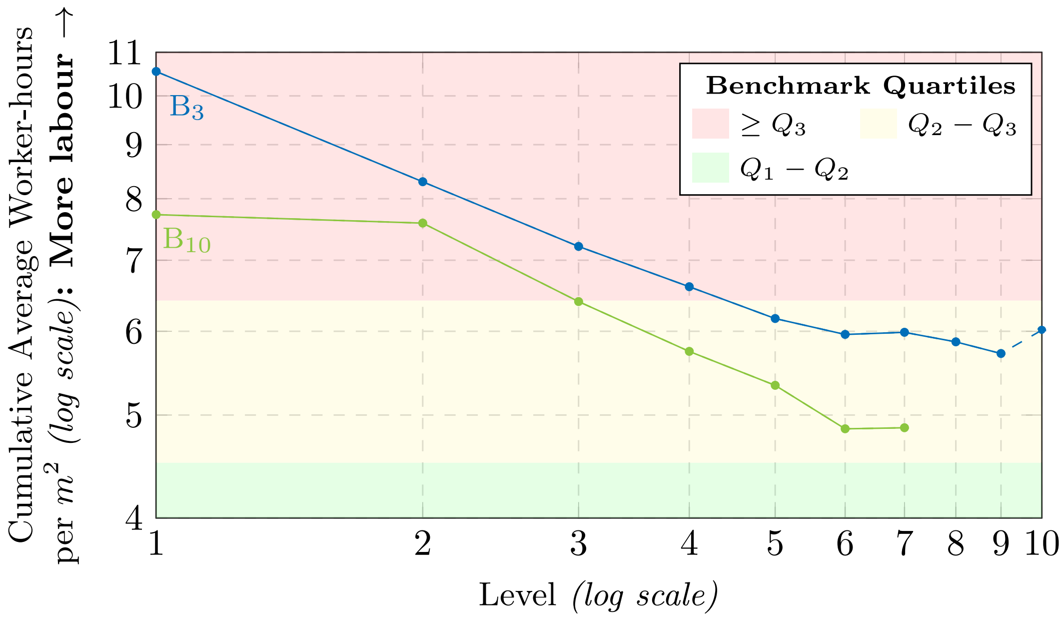
5.1.2. Worsening of Cumulative Average Labour Productivity During the Levelling-off Phase

Higher building levels take longer durations (i.e. more worker-hours) due to the extended time taken by cranes and workers to reach them, and design complexities. One planner mentioned incorporating a 'height factor,' adding an extra minute per lift for every 5 floors when planning high-rise projects (i.e. 0 minutes from the ground floor to Level 5, 1 minute from Level 6 to 10 and so on). Workers also have to cover longer distances from the site entrance or welfare areas to access higher levels. Additionally, roof levels typically have different structural designs to accommodate mechanical systems, higher ceiling heights and other architectural considerations.

Another reason for low labour productivity is the failure to reduce the number of workers proportionally to the decreasing size of the floor plates. Buildings 4, 5, 7, 9 and 11 had considerably smaller floor areas (compared to previous floors) starting from Levels 10, 8, 5, 9 and 6, respectively. These periods correspond to points of upward slope in the learning curves (see Figure 3). The change in floor area happened relatively early on in Building 7, which seems to have allowed for another learning curve to occur from Levels 6 to 10. According to project personnel, it is common for workers to be retained on-site even when there is not much work, to prevent them from joining other companies.

Work disruptions could also reduce labour productivity. For example, the upward slope in Building 6 after Level 6 (see Figure 3) was due to an 'unlearning effect' (Barrie & Paulson, 1992) caused by an eighteen-day superstructure work stoppage to commence facade installation.

5.2. Buildings that Exhibit a Straight-line Learning Curve



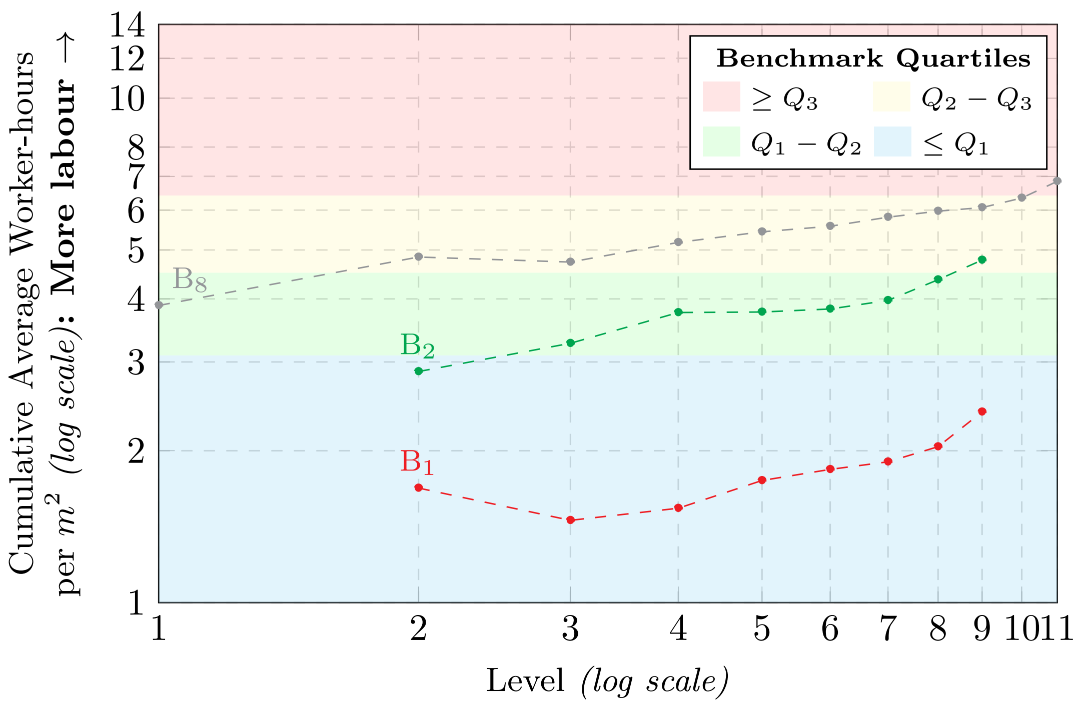
Note: Each coloured line represents the learning curve for one building. The dashed line in B_3 corresponds to a period with no "apparent learning effect" for the superstructure work package as a whole.

Figure 4: Buildings with Straight-line Learning Curves

Buildings 3 and 10 exhibited straight-line learning curves (see Figure 4) with R^2 values of 94% and 92%, respectively. The slope of Building 3's learning curve is -0.26. According to Equation 1, this corresponds to a *learning rate* of $2^{-0.26} \approx 84\%$. As per Equation 2, this indicates an approximate 16% reduction (i.e. 100% - 84%) in the cumulative average worker-hours per m^2 when transitioning from Level 1 to Level 2, Level 2 to Level 4 and then from Level 4 to Level 8. Similarly, Building 10's learning curve has a slope of -0.27, suggesting a *learning rate* of approximately 83%.

The main difference between Buildings 3 and 10 and the six buildings depicted in Figure 3 is that Buildings 3 and 10 maintained almost constant *learning rates* until the end (except for the last level in each building). Therefore, unlike the previous six buildings, which demonstrated lesser labour productivity improvements in the latter stages due to the levelling off of potential improvements, the project teams for Buildings 3 and 10 were able to continuously enhance their performance. A common feature of Buildings 3 and 10 is that they were part of projects involving multiple buildings, with other superstructure work continuing after the completion of superstructure work in Buildings 3 and 10. Consequently, it is possible that the project teams directed workers to other buildings as the output to be delivered in Buildings 3 and 10 decreased.

5.3. Buildings that Exhibit No Learning Effect



Note: Each coloured line represents the learning curve for one building. The three buildings in this graph showed no "apparent learning effect."

Figure 5: Buildings with No Learning Effect

Buildings 1, 2 and 8 did not exhibit the learning effect (see Figure 5). This means that the cumulative average worker-hours per m^2 in these buildings generally increased over time, indicating more labour being used at higher levels to achieve the same output as at lower levels. A common feature noticeable in all these buildings is that, similar to the six buildings in Figure 3, the number of workers on-site has not decreased proportionally to the floor area. In fact, in some cases, the number of workers has increased over time. According to construction personnel, increasing the number of workers on-site is a common solution employed to accelerate progress. As a result, Building 8 was able to achieve and maintain high levels of progress. However, Buildings 1 and 2 were not as successful. The reason was the unavailability of work onsite for the increased workforce (due to a lack of good production system designs). This is beyond the scope of this paper.

5.4. Effect of Initial Productivity on the Learning Effect

Buildings 3 and 10, with straight-line learning curves (see Figure 4), spent more worker-hours per m^2 for their first levels compared to Buildings 4, 5, 6, 7, 9 and 11, which have cubic learning curves (see Figure 3). Therefore, Buildings 3 and 10 might have had better opportunities to improve labour productivity over time while maintaining almost constant learning rates. Similarly, Buildings 4, 5, 6, 7, 9 and 11, with cubic learning curves, spent more worker-hours per m^2 for their first levels compared to Buildings 1, 2 and 8, which did not exhibit a learning effect (see Figure 5). Consequently, Buildings 1, 2 and 8 might not have had many opportunities to improve labour productivity and instead may have had to focus more on increasing progress. This aligns with one of the criticisms of learning curve theory by Thomas (2009), who mentions that learning curves can appear in cumulative productivity data of projects that began poorly.

Labour is one of the most challenging resources to track on-site (Rathnayake et al., 2024). Therefore, it may not always be evident when labour is being wasted. The presence or absence of a learning effect is not necessarily an indicator of good performance in a project, as it can be affected by initial performance. Nevertheless, learning curves can help understand how labour is managed on projects. Periods of upward slope in learning curves should be analysed to determine whether the additional labour-hours are due to changes in the output to be delivered, conscious management decisions or poor worker management practices.

6. Conclusions

This study presented a workflow that could be used to extract, analyse and present available productivity data on sites in a consistent manner. Most of the buildings demonstrated cubic learning curves with significantly poor labour productivity towards the end. This is due to the extended time taken by cranes to lift items to higher levels, work complexities in roofs and the failure to reduce the number of workers proportionally to the size of the floor plate.

One of the unique contributions of this study was the combination of research on developing productivity measurement frameworks with research on exploring the variation of productivity over time (i.e. learning curves). This showed that the presence or absence of a learning effect is not an indicator of good performance in a project, as it can be affected by initial performance. However, the developed workflow can be used to understand how labour is managed on projects and to take steps for improvement.

There were several limitations in this study. First, in labour productivity calculations, workers would not be allocated to a particular building level beyond the date of the final slab concrete pour. However, work such as stripping formwork and performing remedial tasks might still be occurring at that level. The worker-hours spent on such tasks would be counted for higher levels. It is worth exploring whether this misallocation is one of the reasons for the apparent poor productivity at higher levels. Secondly, labour productivity is no longer a suitable indicator of overall productivity due to the use of offsite construction methods and equipment-intensive activities. The variations in other resources should also be analysed to better quantify performance. Finally, the application of the workflow was limited to the superstructure work package.

In the future, the use of more accurate labour tracking technologies will be explored to overcome the limitations in labour allocation. Additionally, more advanced definitions of productivity, such as multifactor productivity (i.e. output per labour, material and equipment), will be considered. Finally, the productivity measurement workflow will be extended to cover other work packages such as substructure work, facade installation and MEP installation.

Acknowledgements

We would like to thank Laing O'Rourke for funding this research. Additionally, we thank our industry partners for their continued engagement with this research.

References

Ayele, S., & Fayek, A. R. (2019). A framework for total productivity measurement of industrial construction projects. *Canadian Journal of Civil Engineering*, *46*(3), 195–206. https://doi.org/10.1139/cjce-2018-0020

Barrie, D. S., & Paulson, B. C. (1992). *Professional construction management: Including CM, design-construct, and general contracting* (3rd). McGraw-Hill.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd). Routledge. https://doi.org/10.4324/9780203771587

Couto, J. P., & Teixeira, J. C. (2005). Using linear model for learning curve effect on highrise floor construction. *Construction Management and Economics*, *23*(4), 355–364. https://doi.org/10.1080/01446190500040505

Everett, J. G., & Farghal, S. (1994). Learning curve predictors for construction field operations. *Journal of Construction Engineering and Management*, *120*(3), 603–616. https://doi.org/10.1061/(ASCE)0733-9364(1994)120:3(603)

Hwang, B.-G., & Soh, C. K. (2013). Trade-level productivity measurement: Critical challenges and solutions. *Journal of Construction Engineering and Management*, *139*(11), 04013013. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000761

Lutz, J. D., Halpin, D. W., & Wilson, J. R. (1994). Simulation of learning development in repetitive construction. *Journal of Construction Engineering and Management*, *120*(4), 753–773. https://doi.org/10.1061/(ASCE)0733-9364(1994)120:4(753)

Nguyen, L. D., & Nguyen, H. T. (2013). Relationship between building floor and construction labor productivity: A case of structural work. *Engineering, Construction and Architectural Management*, *20*(6), 563–575. https://doi.org/10.1108/ECAM-03-2012-0034

Pellegrino, R., Costantino, N., Pietroforte, R., & Sancilio, S. (2012). Construction of multistorey concrete structures in italy: Patterns of productivity and learning curves. *Construction Management and Economics*, *30*(2), 103–115. https://doi.org/10.1080/01446193.2012.660776

Rathnayake, A., & Middleton, C. (2023). Systematic review of the literature on construction productivity. *Journal of Construction Engineering and Management*, *149*(6), 03123005. https://doi.org/10.1061/JCEMD4.COENG-13045

Rathnayake, A., Murguia, D., & Middleton, C. (2024). Measurement of construction productivity: State of the practice in the UK. *Construction Research Congress 2024*, 445–454. https://doi.org/10.1061/9780784485286.045

Thomas, H. R. (2009). Construction learning curves. *Practice Periodical on Structural Design and Construction*, *14*(1), 14–20. https://doi.org/10.1061/(ASCE)1084-0680(2009)14:1(14)

Thomas, H. R., Mathews, C. T., & Ward, J. G. (1986). Learning curve models of construction productivity. *Journal of Construction Engineering and Management*, *112*(2), 245–258. https://doi.org/10.1061/(ASCE)0733-9364(1986)112:2(245)