Determining the Relationship between Direct Work and Construction Labor Productivity in North America
Four Decades of Insights

Neve, Hasse H.; Wandahl, Søren; Lindhard, Søren; Teizer, Jochen; Lerche, Jon

Published in:
Journal of Construction Engineering and Management

DOI (link to publication from Publisher):
10.1061/(ASCE)CO.1943-7862.0001887

Creative Commons License
Unspecified

Publication date:
2020

Document Version
Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

Citation for published version (APA):
Determining the relationship between direct work and construction labor productivity in North America: Four decades of insights

Hasse H. Neve¹, Søren Wandahl², Søren Lindhard³, Jochen Teizer⁴, Jon Lerche⁵

¹ Ph.D. Student, Dept. of Engineering, Aarhus Univ, Inge Lehmanns Gade 10, Aarhus 8000, Denmark, ORCID: https://orcid.org/0000-0003-2311-3529 (Corresponding Auhtor). Email: hn@eng.au.dk

² Professor, Dept. of Engineering, Aarhus Univ, Inge Lehmanns Gade 10, Aarhus 8000, Denmark, ORCID: https://orcid.org/0000-0001-8708-6035. Email: swa@eng.au.dk

³ Associate Professor, Dept. of Civil Engineering, Aalborg Univ., Thomas Manns Vej 23, Aalborg 9220, Denmark, ORCID: https://orcid.org/0000-0002-0807-5352. Email: sml@civil.aau.dk

⁴ Associate Professor, Dept. of the Built Environment, Aarhus Univ, Inge Lehmanns Gade 10, Aarhus 8000, Denmark, ORCID: https://orcid.org/0000-0001-8071-895X. Email: teizer@eng.au.dk

⁵ Ph.D. Student, Dept. of Business Development and Technology, Aarhus Univ, Birk Centerpark 15, Herning 7400, Denmark, ORCID: https://orcid.org/0000-0001-7076-9630. Email: jon.lerche@btech.au.dk

Abstract

Labor productivity in construction has fallen behind other industries in most of the world and has declined continuously for decades in the US. To change this, the construction industry needs to know where to focus. This research aims to show how important craftsmen efficiency is for national construction labor productivity (CLP) development. Statistical analysis was used to compare craftsmen efficiency and CLP data from North America (NA) in the period 1972-2010. Craftsmen efficiency data were extracted from published work that measured direct work (DW) through work sampling, and CLP data were extracted from national databases. A statistically significant relationship between DW and CLP was established. This revealed that adding 36 seconds of DW to every work hour could yield a yearly return of 5.4 billion USD to the NA gross domestic product (GDP). Results show that more focus on activity and project level efficiency is crucial for changing the trends of national CLP. Industry leaders and policy makers now have a solid foundation for taking corrective actions for an industry plagued by low productivity.

Keywords:

Construction labor productivity, efficiency, effectiveness, direct work, work sampling
Introduction

Previous work in comparing labor productivity in construction with other industries has pointed out a significant difference. Examples are Canada (Harrison 2007; Nasir et al. 2014), the US (WEF and BCG 2016), and the rest of the world (MPSC 2015). To make it worse, it is reported that construction labor productivity (CLP) has been declining continuously in the US (Arditi and Mochtar 2000; Nasir et al. 2014; Teicholz et al. 2001; Teicholz 2013; WEF and BCG 2016) and internationally (Abdel-Wahab and Vogl 2011; MPSC 2015). The development in both US and Canadian CLP remains a big challenge, bearing in mind that labor cost accounts for between 20% and 50% of a construction project’s cost (Buchan et al. 1991; Forsberg and Saukkoriipi 2007; Gilleard 1992; Harmon and Cole 2006; Laufer 1980; Laufer and Jenkins 1982). The big share of labor cost of the total project cost makes it a deciding factor when trying to maximize project profits (Thomas and Mathews 1986). Horner and Duff (2001) quantify the problem with a concrete example in the UK, noting that an increase of just 10% in CLP is equivalent to a saving of £1.5bn (2.16bn USD 2001 value).

Changing the declining CLP in the construction industry requires more than insights from macroeconomic CLP data. It requires deeper insights into how construction crafts spent their work time and what factors increase and decrease their efficiency. A reason is that the denominator in single-factor CLP is ‘hours,’ with the numerator being construction’s gross domestic product (GDP). One could argue that the focus should be on increasing construction GDP instead of optimizing time usage, but as Teicholz (2013) work shows, this does not solve the problem of declining CLP, at least not in the US.

This research’s objective is to explore the relationship between craftsmen efficiency on activity and project level, and CLP on national level. The purpose is to investigate how important craftsmen efficiency is and to highlight the impact of potential change. The data stems from previously published work and national CLP databases. It is anticipated that awareness of the economic potential will foster an increased construction industry focus on both efficiency and effectiveness.

Background

The background review focuses on literature and research that investigates how craftsmen work time is spent on construction sites and how the value-adding part of the work time affects CLP. The value-adding part of the work time is known as direct work (DW). It is of interest because it is a measure of efficiency and indicator of productivity. The work sampling (WS) method has been used for decades to measure direct work in the construction industry and thereby the understanding of how time is utilized (Gong et al. 2011). The WS method is a quantitative method based on direct observations, which are categorized in appropriate categories defined to match the work of interest (Terp et al. 1987).

The WS method has been refined since its first use in the construction industry, where only two categories were used: 1) DW and 2) non-productive work. Today, seven categories (Gong et al. 2011) or more (Kalsaas 2010) are used. They are integrated into both continues improvement processes (CII 2010; Gouett et al. 2011; Hwang et al. 2018) and feedback loops designed to reveal productivity potentials (Neve et al. 2019). The only category that has not changed when using the WS method in construction is the category of DW (Gong et al. 2011). Despite the fact that the WS method has proven its ability to increase craftsmen time spent on DW in individual projects (Gouett et al. 2011; Hwang et al. 2018), WS studies conducted in the US have revealed both stagnation (Allmon et al. 2000) and decline (Gong et al. 2011) in DW levels over time.

However, for policy makers and construction industry leaders to act and recommend the use and thereby investment in, for example, work sampling, they need to trust that an increase in DW increases CLP. Table 1 summarizes previous literature and research views on the relationship between DW and CLP.

### Table 1. Views on the relationship between DW and CLP.

<table>
<thead>
<tr>
<th>DW and CLP are related</th>
<th>DW and CLP are not related</th>
<th>Do not consider</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=62</td>
<td>N=5</td>
<td>N=17</td>
</tr>
</tbody>
</table>

Table 1 shows that most of the reviewed work views DW and CLP as a related measure, many do not address the topic, and only a few argue that DW and CLP are not related. Some research have though investigated the relationship by applying statistical methods (Table 2). Most of these studies concluded that DW and CLP are related.

Table 2. Statistical conclusions on relationships between DW and CLP

<table>
<thead>
<tr>
<th>DW and CLP are related</th>
<th>DW is not related to CLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=7</td>
<td>N=1</td>
</tr>
</tbody>
</table>

Research objective and scope definition

The challenge with the above findings is that all of the previous research addresses only DW’s relationship with CLP on activity and project levels. This makes it very difficult for both policy makers and industry leaders to provide broad recommendations across the whole construction industry on using the WS method and other similar process-optimization tools. Further, not knowing whether DW measured on the activity and project levels effects CLP measured on the national level just adds to the challenge of giving industry-wide recommendations. This constitutes a gap in the current body of knowledge, and this research sets out to close this by answering the research question (RQ):

RQ: Does a statistically significant relationship exist between DW measured on the activity and project levels and CLP measured on the national level?

The RQ and its implications will be answered and exemplified based on data from selected countries in North America (NA: USA and Canada). NA is selected because the majority of identified DW values (322 out of 466, whereas one publication often contains several DW values) stems from WS studies conducted in countries on this specific continent. Of the 322 NA DW values, 313 provide data that represent NA DW development very well by giving continuous data from 1972 to 2010 from various location, project, and activity types. The remaining 144 identified DW values were obtained from WS studies conducted on other continents, including Europe (54), Asia (45), Africa (37), and South America (8), thereby making a comparative analysis very difficult.

Even though comparing national level CLP between Canada and the US is challenging due to differences in both data collection methods and available price indexes, previous research has shown that comparisons at the macro level can present general trends between countries and create valuable insights (Abdel-Wahab and Vogl 2011; Harrison 2007; Nasir et al. 2014). Comparing results of WS studies across borders is possible because the only category of interest in this research is the DW category, which for all WS studies has had a narrow definition relating only to the value-adding part of work.

Methods

The method section will firstly outline how national CLP data for the US and Canada were collected and compiled into representative CLP for NA. Secondly, the collection and quality check of data from NA WS studies is outlined, and finally, the used statistical analysis of CLP and WS data is described.

CLP data collection

CLP data for the US were collected from multiple sources since no single US entity collects and manages all data necessary for calculating CLP. Current GDP for the US construction industry came from the US Census
Two additional studies were identified with data from 2013 (Tsehayae and Fayek 2016) and 2017 (Siriwardana et al. 2017) but were excluded due to the big gaps between years.

Bureau and was extracted firstly from 1972 to 1993 (CB-B 2019) and secondly from 1993 to 2010 (CB-A 2019). The US Bureau of Economic Analysis (BEA-A 2019) provided the price index used to deflate the current dollar GDP into constant 2012 USD. The price index was chosen because it covers the period of interest (1972-2010) and focuses on structures, though only from private investments. US Bureau of Labor Statistics-Current Employment Statistics (BLS 2019) was the source of total yearly hours for production and non-supervisory employees for the construction industry.

CLP data for Canada were extracted from the single source of Statistics Canada (SC 2019). Yearly 2012 constant local currency GDP 1997-2010 for the Canadian construction industry was firstly extracted (SC-A 2019). Secondly, yearly 1992 constant local currency GDP 1972-1997 was extracted (SC-B 2019). It was necessary to use the 1972-1997 data in constant 1992 local currency since neither the price index used for this table nor the original current dollar table were accessible. The relative yearly changes in construction GDP 1972-1997 was used to backtrack changes for the SC-A (2019) table enabling the creation of data for constant 2012 local currency construction GDP 1972-2010. The constant 2012 local currency Canadian GDP was converted to constant 2012 USD by using the Organisation for Economic Co-operation and Development (OECD) conversion rates of Purchasing Power Parities (OECD-A 2019). Total yearly hours for all employees in the Canadian construction industry were extracted from the tables SC-C (2019); SC-D (2019).

Continuous data from the USA and Canada were available for both countries in the period of 1972-2010, thus aggregated CLP development for North America was calculated for this period. This was done by firstly adding the constant 2012 USD construction GDP for the US and Canada together, secondly adding the yearly total hour for the US and Canada construction industry together, and thirdly taking the two added datasets of construction GDP and hours and dividing them by each other to obtain the aggregated NA CLP in constant 2012 USD per hour.

Compiling the two data sets presents challenges because the two countries use different standards for measuring CLP, which creates uncertainty that must be addressed. The challenges are firstly that Canada and the USA use different methodologies for collecting CLP (construction GDP and hours) data since Canada is part of OECD, and the US is not. Secondly, Canada has a dedicated price index for construction, which the US does not. Thirdly, when extracting data from Statistics Canada, one can extract only total hours for all employees in the whole construction industry, whereas in the US Bureau of Labor Statistics, one can choose to extract only hours used by production and non-supervisory employees. Since this research aims to understand general macro trends for NA, the above uncertainties are recognized and will be included in the discussion of results.

Collection of work sampling studies

An extensive systematic literature review addressing or applying the concept of WS in the context of construction was conducted. The purpose was, firstly, to understand how previous literature from the whole world have addressed the relationship between DW and CLP (Table 1 and 2 in introduction). Secondly, the review aimed to identify which studies from the US and Canada contained DW measurements from WS studies, which could be used for further analysis. The result of the reviews second part was 313 usable DW values, from WS studies, providing continuous data from 1972-2010. These studies are presented in results. Two additional studies were identified with data from 2013 (Tsehayae and Fayek 2016) and 2017 (Siriwardana et al. 2017) but were excluded due to the big gaps between years.
To gain insights on the unmanipulated dataset from Canada and the US, databases from 1972 until 2010 were analyzed. The gathered data come from published literature and national databases from Canada and the US in the period from 1972 to 2010.

After having collected the data, the quality of the data was scrutinized in two steps to secure a solid foundation for further analysis. First, the 313 DW values (sample) were checked for outliers through a standard deviation based strategy described by Field (2018). The test revealed that no data from the sample had to be removed. Secondly, a normality test was performed to test whether the sample represented the population. Testing the data’s representation of the population is done by using the central limit theorem.

To test the normality of the 313 DW values, the following tests were done:

1) visual inspection of histogram, normal P-P plot, normal Q-Q plots and box plots,
2) skewness and kurtosis test with DW skewness of 0.216 (SE=0.138), and kurtosis of -0.587 (SE=0.275) (Cramer 2003; Cramer and Howitt 2004; Doane and Seward 2011).

The tests found the sample normally distributed, thus this research moves forward assuming that the sample represents normality and thereby the population.


Data analysis

Data analysis will be performed to test if a relationship exists between DW measured on activity and project levels and CLP measured on a national level. The gathered data come from published literature and national databases from Canada and the US in the period from 1972 to 2010.

As outlined above, the WS data set has gaps in some years. Therefore, three different manipulations of the dataset were carried out to gain further insights into the relationship. The four datasets included one unmanipulated (raw mean values) and three manipulated:

1) raw mean values,
2) filled gaps – 3-year mean: the gaps in the dataset are filled with the mean value of the DW value on both sides of the gap, and
3) 2-year average for both CLP and DW. A 2-year average is the mean of, for example, 1972 and 1973, which gives one data point and so forth,
4) 3-year moving average: as an example, the 1975 values would be the average of 1974, 1975, and 1976 and then moving forward.

To gain insights on the four different datasets, four statistical analyses were performed in SPSS:

1) curve estimation with 11 equations done firstly to understand if a statistically significant relationship can be established for the four different datasets and secondly to understand which equation provides the best predictive capabilities. The 11 equations: linear, logarithmic, inverse, quadratic, cubic, compound, power, S, growth, exponential, and logistic (the linear is found best fitting, thus the introduction of the last three tests),
2) linear regression analysis providing a linear equation,
3) ANOVA analysis providing a p-value which reveals the statistical significance of the linear regression models’ predictive capabilities, and
4) a t-Test enabling the calculation of 95% confidence intervals for the linear regression models coefficients.

Common for the two first tests are that they rely on the interpretation of the correlation coefficient (R). Previous recommendations (Cohen 1988; Cohen 1992) outline that R=0.3 reflect a medium effect size and research (Liu et al. 2011) have previously used R=0.318 as an acceptable level in the same context and with similar goals thus this research goes forward with R=0.3 as the minimum limit for accepting any relationship established through the statistical analysis. The R-value can be squared (R^2) to instead reflect the predictive capabilities of the independent variable in the analysis. The R^2 value corresponding to R=0.3 is 0.09, and hereof R^2=0.09 will be the lower limit acceptance reference. Further, all established relationships must have a statistical significance level above 95% (p≤0.05) to be accepted as valid.

**Results**

In total, 313 DW values in 23 papers from NA were identified. Of these, 9 came from Canada (Agbulos and AbouRizk 2003; Choy and Ruwanpura 2006; Christian and Hachey 1995; Da Silva 2006; Handa and Abdalla 1989; Heinz 1984; Hewage and Ruwanpura 2006; Shahtaheri 2012; Shahtaheri et al. 2015) and 14 from the US (Allmon et al. 2000; Gong et al. 2011; Gouett et al. 2011; Jenkins and Orth 2004; Liou and Borcherding 2016; Logcher and Collins 1978; Oglesby et al. 1989; Picard 2002; Salim and Bernold 1994; Thomas 1991; Thomas and Daily 1983; Thomas et al. 1984; Thomas and Holland 1980; Thomas Jr 1981).

To reach this research’s objective, curve estimation was done for the 4 datasets. The result of the curve estimation is the predictive capability (R^2) with the lower limit at R^2=0.09 and statistical significance level at 95% (p≤0.05) of each established relationship for the 11 equations. Table 3 below reveals that all analyzed relationships have an R^2-value that is above the lower limit of R^2=0.09. The statistical significance test for each relationship reveals that all but four relationships are statistically significant above the 95% limit (p≤0.05). The 4 equations failing that mark were the equations inverse and S when used on the filled gaps – 3-year mean data set and quadratic and cubic when used on the 2-year average dataset.

**Table 3.** Curve estimations of the five datasets with 11 equations.

<table>
<thead>
<tr>
<th>DataSets</th>
<th>Linear</th>
<th>Logarithmic</th>
<th>Inverse</th>
<th>Quadratic</th>
<th>Cubic</th>
<th>Compound</th>
<th>Power</th>
<th>S</th>
<th>Growth</th>
<th>Exponential</th>
<th>Logistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw mean values</td>
<td>R^2</td>
<td>.198</td>
<td>.186</td>
<td>.196</td>
<td>.203</td>
<td>.203</td>
<td>.189</td>
<td>.179</td>
<td>.163</td>
<td>.189</td>
<td>.189</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.011</td>
<td>.014</td>
<td>.019</td>
<td>.037</td>
<td>.037</td>
<td>.013</td>
<td>.016</td>
<td>.022</td>
<td>.013</td>
<td>.013</td>
</tr>
<tr>
<td>Filled gaps - 3-year</td>
<td>R^2</td>
<td>.135</td>
<td>.116</td>
<td>.096</td>
<td>.167</td>
<td>.164</td>
<td>.124</td>
<td>.107</td>
<td>.088</td>
<td>.124</td>
<td>.124</td>
</tr>
<tr>
<td>mean</td>
<td>p</td>
<td>.022</td>
<td>.034</td>
<td>.055</td>
<td>.038</td>
<td>.040</td>
<td>.028</td>
<td>.042</td>
<td>.066</td>
<td>.028</td>
<td>.028</td>
</tr>
<tr>
<td>2-year average</td>
<td>R^2</td>
<td>.238</td>
<td>.240</td>
<td>.236</td>
<td>.241</td>
<td>.241</td>
<td>.228</td>
<td>.231</td>
<td>.229</td>
<td>.228</td>
<td>.228</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.029</td>
<td>.028</td>
<td>.030</td>
<td>.096</td>
<td>.095</td>
<td>.033</td>
<td>.032</td>
<td>.033</td>
<td>.033</td>
<td>.033</td>
</tr>
<tr>
<td>3-year moving average</td>
<td>R^2</td>
<td>.199</td>
<td>.197</td>
<td>.193</td>
<td>.200</td>
<td>.200</td>
<td>.189</td>
<td>.187</td>
<td>.183</td>
<td>.189</td>
<td>.189</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.004</td>
<td>.005</td>
<td>.005</td>
<td>.018</td>
<td>.018</td>
<td>.006</td>
<td>.007</td>
<td>.006</td>
<td>.006</td>
<td>.006</td>
</tr>
</tbody>
</table>
With the above Table 3 revealing that 7 out of 11 equations can establish statistically significant relations (P) with predictive capabilities (R²) above the lower limit for all four datasets the necessary knowledge now exists to answer the RQ:

Does a statistically significant relationship exist between DW measured on the activity and project levels and CLP measured on the national level?

The unambiguous answer is, yes, a statistically significant relation does exist. To illuminate the inherent implications of the above finding, this research moves forward firstly by identifying the most appropriate equation to describe the relationship in question by analyzing each equation’s capacity. The equation’s capacity is evaluated regarding both overfitting and underfitting, which is a known approach from machine learning (Géron 2017; Goodfellow et al. 2016), and the results from the above Table 3 showing predictive capabilities and statistical significance levels. As Table 3 shows, all equations have very similar predictive capabilities and statistical significance levels, and therefore, the capacity will be the deciding factor. Evaluating the capacity is done in two steps. The first was by analyzing the below figure 1, showing NA CLP and DW together with linear regression lines for each dataset. In the second step, the plots from each curve estimation were reviewed (not displayed in the paper). The conclusion is clear; using equations more advanced than the linear equation will lead to overfitting, and thus the linear equation is found to be the best suited to describe the relationship between DW measured on the activity and project level and CLP measured on the national level.

Figure 1. USD per hour for NA, DW for NA and a linear regression line for both.

To investigate further the linear equation as the best choice for describing the relationship between NA CLP and DW, four linear regression analyses were done. In the regression analyses, NA DW is the independent (predictor) variable, and NA CLP is the dependent (response) variable. Table 4 presents the four linear regression analyses. A t-Test to establish the 95% confidence intervals (CI) for the predictor coefficient (a) and constant coefficient (b), the R²-values from the regression analysis and finally the ANOVA result showing the statistical significance level for each linear regression model.
Table 4. Regression analysis, t-Test (95% CI), and ANOVA for the relationship between DW and CLP.

<table>
<thead>
<tr>
<th>Regression models</th>
<th>Regression model</th>
<th>N</th>
<th>a (95% CI)</th>
<th>b (95% CI)</th>
<th>R²</th>
<th>ANOVA p-Value³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw mean values</td>
<td>y=0.399x+66.679</td>
<td>32</td>
<td>(0.1,0.699)</td>
<td>(54.613,82.745)</td>
<td>.198</td>
<td>.011</td>
</tr>
<tr>
<td>Filled gaps 3-year mean</td>
<td>y=0.348x+70.209</td>
<td>39</td>
<td>(0.054,0.643)</td>
<td>(56.450,83.968)</td>
<td>.135</td>
<td>.022</td>
</tr>
<tr>
<td>2-year average</td>
<td>y=0.527x+61.937</td>
<td>20</td>
<td>(0.060,0.995)</td>
<td>(40.464,83.411)</td>
<td>.238</td>
<td>.029</td>
</tr>
<tr>
<td>3-year moving average</td>
<td>y=0.516x+62.605</td>
<td>39</td>
<td>(0.172,0.860)</td>
<td>(46.680,78.531)</td>
<td>.199</td>
<td>.004</td>
</tr>
</tbody>
</table>

³Significance level at 95%, p≤0.05

Further, the following Figure 2 shows a visual representation of the linear regression models and the original non-manipulated yearly data (raw mean values). The four linear regression analyses present a coherent picture of a solid linear relationship between NA CLP and DW. The regression analysis shows R²-values ranging from .135 to .238, which is far above the defined lower limit (R²=0.09) and close to what Cohen (1988); Cohen (1992) have defined as a large effect size (R²=0.25), with the ANOVA process revealing p-values ranging from .029 to .004, which shows that all regression models are statistically significant above the 95% level. This analysis shows that the linear regression model is suitable to describe the relationship between NA CLP and DW. In the following the economic implications of this result will be explored.

The predictive capabilities of the four regression models are now in focus to understand the economic implications of the established relationship between NA CLP and DW. The R² values from the above Table 4 tell us what percent of the change in the dependent (response) variable (NA CLP) the independent (predictor) variable (NA DW) can describe. As shown, the four linear regression models can describe between 13.5% and 23.8% of the changes in the NA CLP, with more than 95% statistical security for all linear regression models. The 95% confidence interval (CI) for the 5 linear regression models’ coefficients reveals a predictive range for all models which match well with the plot in the below Figure 2, showing all 4 models fit well with the unmanipulated data (raw mean values).
To understand the predictive capabilities fully and thereby the economic implications, the regression model based on unmanipulated data is put into a real-world context of NA. This model is chosen because it provides the most realistic output since it is based on unmanipulated data. The model is:

\[
\text{USD per hour} = 0.399 \times \text{NA DW} + 66.679
\]  

(1)

Using model (1), the following example will show what impact small changes in work time efficiency would have on the NA CLP and, thereby, NA construction GDP. Taking the predictor coefficient (0.399) from the chosen model (1) shows that a 1 percentage point increase in NA DW values can increase the output of every worked hour by 0.399 USD (2012 value). To gain a 1 percentage point DW increase, a worker would have to spend 36 seconds more every hour or 22.2 minutes a week (37-hour work week) on DW. To understand the value these 36 seconds or 22.2 min. per week can generate, the 0.399 USD (2012 value) is multiplied by total hours for NA in 2010 (2010 is the limit of the model). This would, in 2010, have given an extra NA construction GDP of 5.4 billion 2012 USD (± 4.1 billion 2012 USD for 95% CI).

**Discussion and limitations**

**Importance of focusing on process optimization**

The results show that a linear regression model is the best to describe the relationship between NA CLP and DW in the period of 1972 to 2010. The linear regression model (1) reveals an economic potential of 5.4 billion 2012 USD (± 4.1 billion 2012 USD for 95% CI) in yearly added GDP for the NA construction industry if the NA DW level is increased by just one percentage point. Despite uncertainty in the model (1), further calculations show that even the lowest value in the 95% CI for equation (1) still contains close to 1.4 billion 2012 USD a year. Even more interesting is the other end of the 95% CI showing a potential of 9.5

billion 2012 USD a year. These outcomes may be conservative since previous research has found that DW on projects can be increased by several percentage points by using the work sampling method as part of an optimization process known as activity analysis (CII 2010; Gouett et al. 2011; Hwang et al. 2018).

The established relationship and identified potential can be used as an argument for both policy makers and industry leaders to invest in methods that can create more efficient, effective and in the end productive work processes for the construction industry. Methods already available is work sampling as part of the activity analysis process, lean construction methods as the Last Planner System (Ballard 2000), Location based Scheduling (Kenley and Seppänen 2010) and Integrated Project Delivery (Fischer et al. 2017), and other proven productivity practices (Caldas et al. 2015; Nasir et al. 2016).

The results are based on data from NA and are thus not directly applicable to the rest of the world. It would, however, be surprising if similar relationships do not exist between DW measured on an activity and project level and CLP measured on a national level, especially if looking towards other very similar economies as Europe and Scandinavia.

Further, a limitation that perhaps limits using the developed model (1) in the near future is its reliance on the latest or new data from NA. Since 2010 not enough work sampling studies have been published to allow the continuation of the model up until today. At least the aggregated NA CLP data for the recent years exist, and clearly show the trend of decline is continuing.

**Indirect measurement of CLP**

Of the investigated studies, 8 found a total of 14 relevant relationships between DW and CLP on the activity and project levels (Al-Ghamdi 1995; Handa and Abdalla 1989; Kaming et al. 1997; Liou and Borcherding 1986; Olomolaiye et al. 1987; Siriwardana et al. 2017; Thomas 1991; Thomas et al. 1984) with $R^2$-values ranging from 0.013 to 0.82, with most of them close to or above $R^2=0.09$, which in the method section was established as the lower limit for a usable result in this research. This reveals that DW is a good indirect measure of CLP on the activity, project, and national levels and that the predictive ability are similar on all levels.

Thomas (1991) article must in this context be addressed because he alone, based on statistical analysis concludes the opposite. Thomas (1991) argue that DW cannot be used to predict CLP when 50% to 75% or more of the variability in models describing DW and CLP’s relationship is unexplained. Unexplained variability in the range of 50% to 75% corresponds to R-values ranging from 0.7 ($R^2=0.5$) to 0.5 ($R^2=0.25$). This perspective stands in stark contrast to Cohen (1988); Cohen (1992) work stating that R-values above 0.5 shows a large effect size and 0.3 a medium effect size. In addition, previous research (Gonzalez et al. 2008; Liu et al. 2011) has used R-values values on 0.5 and 0.318 as acceptable limits for regression analysis made with similar purposes. Based on Cohen (1988); Cohen (1992); Gonzalez et al. (2008); Liu et al. (2011) and the results of this research the authors disagree with Thomas’s (1991) argument that DW cannot predict CLP as long as R-values are significant.

One must though be careful when using indirect measurement such as DW for CLP. Because even though $R^2=0.09$ in this research is chosen as the lower limit for any acceptable relationship and that the chosen linear function, to show the economic implications, has an $R^2=0.198$, an $R^2$ value of 0.198 explain only 19.8% of the changes in the dependent variable (CLP). This means that 80.2% of the changes in CLP is possibly caused by other factors than DW. Thus, DW can be used as a trustworthy indicator and predictor of CLP but not as a replacement for unit rate productivity. Based on the latter and the worrying introduction in Thomas

(1991) the authors of this research must emphasize that WS data cannot be used as an argument in hearings and arbitrations aiming at pinpointing responsibility for productivity losses, loss of labor time etc.

This research does, however, find that the identified relationship, and potentials are so significant that both policy makers and industry leaders cannot ignore it.

The unexplained three quarters

The four linear regression models’ predictive capabilities range from 13.5% to 23.8 %, meaning over 75% of the change in NA CLP is explained by factors other than DW. This makes sense because this research regression model looks only at one factor, namely DW. The only reason for not exploring how other categories (and thereby factors) as, for example, the category of preparation was because the data was not available. One thing to remember when exploring advanced models that potentially could explain more of the change in CLP is the actual value of this. Because having a complicated model made up of multiple categories of, for example, talking, preparation etc. might explain larger parts of the change in the dependent variable (CLP), but the actual applicability can be challenging. This is because a complicated model shows complicated interrelations between factors that make it hard to know where to focus if pursuing the optimization of CLP. Finally, when using correlation analysis to establish statistically significant relationships the matter of causality is an important topic. The correlation between DW and CLP is in this research established on three different levels: 1) the activity, 2) project and 3) national level. Having established statistically significant correlations on three different levels points towards a causal relationship and not random correlation.

Construction price index

Finally, can we even trust historical construction GDP data. Inaccurate price indexes for the construction industry have previously been discussed as being one of the main reasons for the construction industries reported decline, stagnation, and lack of development in CLP (Allen 1985; Allmon et al. 2000; Sveikauskas et al. 2016; Teicholz et al. 2001). Despite the reported challenges of measuring construction GDP at national level and, thereby CLP, the OECD, APO (Asian Productivity Organization), McKinsey and Company, World Economic Forum etc. still report CLP on the national level for policy makers to make decisions. This research recognizes the challenge, but since no better source existed, the national databases were used.

Conclusion

Labor productivity in construction has fallen behind other industries. As reported widely, it has been declining continuously for decades, at least in most parts of the western world. To change this negative trend, the construction industry needs to know where to focus.

It was found that craftsmen efficiency is a crucial factor in changing the trend of stagnation and decline in construction labor productivity. The importance of craftsmen efficiency was found by comparing four decades of published direct work rates measured on activity and project levels, with construction labor productivity data measured on a national level.

The comparison showed that if all construction crafts in North America added just 36 seconds of additional direct work time to each working hour in 2010, 5.4 billion USD (2012 value) would be added to the yearly construction GDP.

The findings have implications for leaders and policy makers in the construction industry because the industry potential of focusing on efficiency has been quantified. The result is based on data from North America but looking towards other economies as Europe and Scandinavia, similar potentials probably exist.

Data Availability Statement

Data generated or analyzed during the study are available from the corresponding author by request. Information about the Journal’s data-sharing policy can be found here: https://ascelibrary.org/doi/10.1061/(ASCE)CO.1943-7862.0001263

Acknowledgment

This research is part of the REVALUE research project, which focuses on the refurbishment of social housing buildings. The study was sponsored by the Innovation Fund Denmark. The authors are responsible for the content of this publication.

References


Alinaitwe, H., Mwakali, J., and Hansson, B. "Efficiency of craftsmen on building sites—studies in Uganda." In Proc., The Int. Conf. on Advances in Engineering and Technology. Elsevier.


CII (2010). "Guide to activity analysis." Construction Industry Institute (CII), University of Texas, TX.


Siriwardana, C. S., Titov, R., and Ruwanpura, J. "A study to investigate the relationship between work sampling categories and productivity for concrete slab formwork." In Proc., 8th Int. Conf. on Structural Engineering and Construction Management.


RIL – Association of Finnish Civil Engineers (www.ril.fi).


