International Project-level Comparisons of Construction Industry Performance

Rick Best,
University of Technology Sydney, Australia (email: rick.best@uts.edu.au)

Abstract

This paper reports on further investigations into comparisons of the performance of the construction industries in various countries. It builds on earlier work that analysed construction time, cost and quality data gathered for high rise office building projects in a number of countries and sought to use that data to make valid comparisons of industry performance at the level of individual projects. The previous work produced a simple "performance index" that could be used to compare construction performance across projects. That index is examined in more detail with particular emphasis on the conversion of construction costs from national currencies to a common base.

Keywords: Construction industry performance, international comparisons, purchasing power parity, Big Mac Index.

1. Introduction

The construction industries in various countries contribute greatly to national economies – gross construction output in the period 1945-1990 in North America and Western Europe averaged somewhere between 12% and 15% of GDP [1]. In OECD countries generally the contribution ranges between 5% and 8% [2]. Construction is a key element in investment, and to the future prosperity of nations [3]. It is important that the sector operates with maximum efficiency [4] and that scarce resources are utilised in the best possible way.

Given the importance of construction in measures of national output it is not surprising that many stakeholders are interested in measuring how effective the activities of the construction sector are. This has become increasingly so as national construction industries have sought to expand into international markets [5]. Comparative studies have been attempted since at least the late 1940s [6] and many different methodologies have been used, with varying degrees of success.

This paper reviews some of the more recent attempts at international comparisons and compares them to a Performance Index (PI) proposed by Langston and Best [7]. In particular the paper focuses on the problems of gathering suitable data (time, cost and quality) and the major problem of converting construction costs to a common base.
2. Comparative Studies

Any sort of comparative study has a number of inherent problems, the most common are described as “comparability and representivity” – simply, the comparing of like with like but using data that is representative of typical practice in different places. With construction these problems are particularly severe as output (buildings, bridges, roads and the like) are seldom, if ever, identical. When international comparisons are attempted these problems are exacerbated, as there is no truly “standard” project that can be used as a basis for comparison, however, comparisons are made and do produce some useful insights (e.g. [8, 9, 10, 11, 12]).

The studies have approached international comparisons in many different ways including those based purely on construction cost [13, 14], or labour productivity [15], those using input-output analysis [16], and other macroeconomic data [17] and some that have attempted to analyse and explain differences across a broader range of factors in the whole process of building procurement [7]. These studies have been categorised as pricing studies, macroeconomic studies and case studies [18]. A few have characteristics of more than one type.

2.1 Comparative Methods

Pricing studies compare construction costs between countries, typically on the basis of a standard project (or set of projects). The problems of comparability and representivity of the projects are always evident, and in any case equating industry performance or competitiveness on the basis of cost alone is surely problematic, as apparently lower local construction costs do not mean that the same industry building in another country (and subject to that country’s laws, regulations, labour market and so on) could build for the same cost that it does at home [18]. There is even greater concern when costs are converted from local currency to some common base (usually US$). In one study [19] the authors assessed Australia’s construction costs as second lowest and Germany most expensive (in a field of five) but they had converted costs using money market exchange rates, and commented that “it would appear that Australia with its currency exchange advantage is more competitive” [emphasis added]. Using different methods (discussed below) the same cost data was used by Langston and de Valence [14] - they ranked Germany least expensive, and Australia most expensive. This illustrates just how sensitive these studies are to the methods used for cost conversion.

Macroeconomic studies are generally based on a single input, usually labour. Typical measures include value added per employee or hour worked, or construction volume per employee or hour worked. Many authors are critical of this approach, and question its usefulness for a variety of reasons claiming it to be too simple to be an adequate measure of overall productivity in an economic sector as large and complex as construction [20]; cost effective but with suspect reliability and comparability of data [21]; too lacking in detail to be of use below the whole-of-industry level [9]; an unsatisfactory measure of efficiency of utilisation of resources [22] and noting its to reflect differences in capital intensity in measures such as output per hour worked or per person employed [23]; or to provide data on quality [24].
Croce et al. [3] list a series of reasons why labour productivity is an unreliable measure, including, amongst other things, distortions that emerge when construction times are shortened (and, as a result, costs decrease), thus appearing to decrease output when measured in dollars, as is typically done. They also mention changes in composition of activity (e.g. residential versus non-residential construction) that see increased activity in the residential sector, which typically has lower productivity than the non-residential sector, giving the appearance that overall productivity has levelled out or even declined. Although they repeatedly question the reliability of macroeconomic studies, the authors use labour productivity figures to conclude that labour productivity in Australia “is above that for the US, Japan and all European nations”. They rely on similar analyses of construction costs to conclude that “Australia is more than competitive with other western countries”. On the basis of their analysis, their conclusions about Australia’s performance must be questioned.

Case studies provide useful insights into why there are differences in performance between countries, but they, too, have their drawbacks. They are labour-intensive and are consequently reduced to just a few projects, perhaps as few as two, e.g [25, 8], and while the projects may be quite comparable it is always doubtful that they can be representative of the whole industry in a country.

Some studies are hybrids. Langston and Best [7] gathered project data for 78 high rise office projects across 12 countries. Costs were converted using several methods, including the Big Mac index (see below at 4.4) and relationships between factors such as resource consumption (cost/m²) and productivity (m²/month) were tested. From this a Performance Index (PI) was proposed, which was expressed as a ration of productivity to resource consumption. Simple algebra reduced this expression to a ratio of the square of the building area to cost multiplied by time to construct:

\[
\text{Performance Index (PI)} = \frac{a^2}{ct}
\]  

(Equation 1)

where \(a\) is the area in m², \(c\) is cost (in a common currency) and \(t\) is time to construct (in months). The higher the index, the better the performance.

Xiao and Proverbs [10, 11 12] carried out a more comprehensive study of contractor performance in the UK, the US and Japan. They sought data on various aspects of time, cost and quality performance. They used a standard project but tried to improve representivity (while maintaining comparability) by providing only a verbal description of the project and leaving respondents to provide estimates based on construction methods and materials typical of their region. This study has some similarities with previous pricing studies but it goes far beyond cost alone. The methodology, however, is not necessarily as successful as they claim as they draw very definite conclusions about industry performance in the countries studied based on estimates alone (of construction time and cost particularly) of a hypothetical and loosely defined project, and include some very questionable data (e.g. estimated construction times ranging from an
unlikely 12 weeks up to a glacial 108 weeks for the same “standard” project). They convert estimated construction costs to £UK using OECD purchasing power parities (PPPs – see below at 4), which is reasonable, but when slightly more recent PPPs are used the outcomes change dramatically with the ranking of the UK and Japan (based on average cost/m²) reversed, with the cost in Japan changing from around 5% less than the UK to nearly 12% above that in the UK. The outcomes of such studies must always be treated with care, and this again highlights the importance of valid currency conversions.

3. The Theory of the Performance Index

While previous studies focused on one or two aspects of performance (e.g. cost or time), performance has four aspects: productivity, time, cost and quality, and any measure of performance should embrace all of them [26]. The two studies described above do, to some extent, include all four, although measuring quality remains a difficulty. Productivity can be measured at an industry level (using macroeconomic data) or at a micro level (individual trades or onsite tasks), however, time, cost and quality are best measured at a project level. The Performance Index proposed by Langston and Best [7] includes a measure of productivity and makes this approach very useful as it assesses, to some degree, all aspects of performance at the level of individual projects. If data for a large enough number of projects can be obtained then the sample should be reasonably representative of the industry as a whole.

Productivity is typically expressed as the ratio of inputs to outputs – units of construction output can be expressed as m² of floor space. Inputs to the construction are many and varied, and different inputs may be measured in different units: number of people, items of plant, quantities of materials and so on. The only available common unit, however, is money, so factors of production (inputs) are most conveniently expressed in dollars. The rate of resource use may therefore be expressed as “construction $/month”. Hence productivity can be expressed as the ratio of output (m² of floor area/month) to inputs (resource use expressed as $/m²).

The performance index is defined as

\[ \frac{\text{productivity}}{\text{resource use}} = \frac{m^2}{\text{cost/m}^2} \]

which equals

\[ \frac{m^2}{\text{cost} \cdot \text{month}} \] or \[ \frac{a^2}{ct} \] (see Equation 1)

Cost/m² is obviously influenced by quality standards, i.e. a higher quality output may simply reflect higher quality inputs (such as prestige materials or fittings) so some adjustment to cost is needed to address quality differences. The simplest way is to separate projects into quality groups, e.g. for offices the typical groupings are Prestige, Grade A, B, C and D [27]; alternatively indices can be calculated that deflate costs to a common standard. Other aspects of quality, such as client satisfaction, are less easily accommodated but can be measured if data is available. Current UK KPIs [28] include measures of defects, predictability of cost and time as well as client satisfaction (with both product and process), so this sort of data is now being gathered.
One of the strengths of the PI is that it uses actual data from completed projects rather than estimates based on hypothetical projects. Total cost and time to construct, for example, will include at least some part of any extra money and/or time required to correct defects or carry out rework – contractors would be going out of business more often than they do if these penalties were not included. Costs are best stated exclusive of substructure and siteworks as cost and time associated with these elements can vary markedly between projects purely due to the scope of the work [29]. Projects should be of one type or grouped according to type (as well as quality) and analysed separately – this will improve comparability.

4. Currency Conversions

Converting costs to a common currency using exchange rates only, is a method now generally recognised as flawed [30, 13, 31, 32, 33]. The aim must be to eliminate, as far as possible, differences in prices between countries so that when costs are converted to a common currency they are then valued at the same price levels [34]. This is commonly done using purchasing power parities (PPPs). While PPPs are widely used for a variety of purposes they are not necessarily appropriate in all cases, and the use of general PPPs for industry-level productivity comparisons is not recommended [33]. Stapel [35] also advises against their use in a number of circumstances, including their use “as a measure to generate output and productivity comparisons by industry (unless there are industry specific PPPs)”. To date few, if any, reliable industry-specific PPPs have been published.

4.1 The Law of One Price

Purchasing power parities are based on the so-called Law of One Price [36, 37]. The basis of this is the notion that the cost of a good or service (or a basket of goods and services), once prices are converted to a common currency, should cost the same in different countries [38]. For some commodities that are traded often, such as gold, the law of one price holds well [38]. It does not hold for construction output [13] as built facilities are not tradeable [33], and are produced and consumed locally. A large proportion of the cost of a building is made up of labour costs and basic materials such as bricks, sand and concrete. These are mostly produced locally rather than imported and therefore their costs are little affected by exchange rates [13].

4.2 Calculating Purchasing Power Parities

A range of PPPs are routinely produced. General PPPs are the result of extensive price gathering exercises carried out by Eurostat and the OECD, part of the International Comparison Program (ICP) for the UN and World Bank [39] The OECD’s 1996 comparison used a list of around 4000 items including consumer goods and services, pharmaceutical products, capital goods, motor vehicles, government services, health and education services, and twenty construction projects [39]. Construction prices were gathered by having comprehensive bills of quantities (BQs) priced
by estimators in participating countries. This system has been under review for some years as there are serious doubts about the reliability of results [40, 41] and the level of effort and resources required to gather the data [42, 29, 43]. The ICP method is also under review, with a new framework for collecting construction data about to be tested [39].

### 4.3 Construction Data for PPPs

In the Eurostat program the amount of data collected for construction has been reduced by around 50% in recent rounds [29]. Fewer projects are being priced and the BQs for those projects that are priced have been abbreviated. Meikle [45, 41] advocates a move away from the BQ approach to a weighted basket of goods (BOG) approach. The aim is to make pricing easier and more reliable, and improve the comparability and representivity of the data collected.

The ICP, however, are to trial a method proposed by Walsh and Sawhney [39, 44], using a basket of construction components (BOCC). A pilot study is to be run in 2005 that will also include multiple pricing of some simple inputs (steel and cement) as well as components such as structural concrete columns, simple pad footings and cement render [44]. These input items will be priced at various locations several times each year in a bid to minimise temporal and regional differences. As the ICP includes many more countries than the Eurostat/OECD group (186 countries rather than around 35) the problems of comparability and representivity are multiplied. The key to success in using this approach lies in identifying and specifying a basket of components that are reasonably similar in composition and construction in all countries. It remains to be seen whether a large enough basket can be assembled to allow estimated prices to be meaningful. Meikle [45, 41] suggests that the approach is flawed, as the pricing of the BOCC will lack context (e.g. scale, location, site conditions, access). It can be argued, however, that the perceived lack of context does not necessarily invalidate the concept, as most of these concerns can be shown to be of relatively little consequence. The fact that the best estimates routinely vary by as much as ±10% when compared to outturn costs, supports this view as the sort of cost differences that arise through contextual variations are likely to be insignificant within such an estimate band.

Stapel [29] suggests that “the principal determinants of price level [in construction] are probably scale, complexity and location rather than the type of work”. Meikle [45] is quite definite about it. The two-stage study completed in Australia in 1999 [19, 14] supports the view that the type of project has little impact on cost differentials. Based on that research Langston and Best [7] limited their study to a single building type (high rise commercial offices) as this was considered to be the most “generic” building type common to all the countries in their study.

It could be argued that even with the use of detailed BQs the issue of scale may not be addressed. Meikle [45] shows that the projects currently priced in the Eurostat/OECD program do not properly represent the typical mix of projects (with regard to scale) usually built. He demonstrates that the hypothetical projects priced in the UK are generally unrepresentative of the size range of projects actually constructed. Furthermore it is doubtful that estimators would apply
a level of care that would include pricing differently for projects of varying scale; there is no incentive to price as low as possible in order to win the job, and the estimators can be sure that their companies will not have to build to their estimated cost [46].

Location (geographically speaking) can be addressed as it is now, by asking estimators to give national average prices, or by having estimates produces in several locations within a country. This could be done regardless of whether it is a basket of goods, a basket of components or a set of standard projects that is being priced. Location, (within projects, i.e. ground floor, 10th floor, 50th floor) has some impact on prices but it is common practice for many levels of multi-storey projects to be measured and priced together, particularly in countries such as Australia where there has been a definite move towards “concise” BQs. Hoisting and scaffolding costs will affect unit rates for work at elevation but given that labour and materials generally make up the greater proportion of unit rates, and prices gathered are estimates of national averages, price distortions due to the inclusion or exclusion of within-project location in pricing exercises would be unlikely to vary beyond the normal limits of estimating accuracy. If overheads are spread within prices then fixed crane and hoist charges are likely to be embedded in unit rates anyway.

Stapel’s third determinant, complexity, is more difficult to accommodate but as it is “average” prices that are generally sought, the effects of complexity of construction on some projects can arguably be discounted. Once again it is unlikely that estimates will vary beyond normally accepted limits of estimating accuracy.

If context is indeed a problem with the BOCC method then it is likely to be even more of a concern in a BOG approach. While scale may be addressed by asking for prices for differing quantities of the same item (e.g. 100m³, 10,000m³, 100,000m³ of excavation in other than rock) location and complexity would appear to be impossible to deal without providing very detailed contextual information with every item in the basket. Such information would be very hard to convey accurately and the outcomes would be no more or less reliable than those gained using other methods. Meikle [45] suggests methods for deriving resource weightings in a BOG exercise – similar methods could be applied to a BOCC. More importantly, collecting more than one set of estimates from each country would do much to eliminate, or least ameliorate, the effects of contextual differences.

### 4.4 The Big MacTM Index

The Law of One Price, which provides the foundation for the doctrine of PPPs has been applied in a very simple way since the 1980s in the formulation of the Big Mac Index. In 1986 The Economist magazine first published, somewhat light-heartedly, a comparative index based on a single, tightly-specified manufactured commodity, the Big Mac hamburger [36].

The idea is simple: the Big Mac is offered for sale, with virtually identical specifications, in around 120 countries [49]. By assuming that the value of this commodity must be equal in all countries, as it is an identical product, it was argued that the cost of a Big Mac in any country
could provide a measure of the relative value of various currencies, and so reflected the purchasing power of a unit of currency in each country. For example, if a Big Mac cost, on average, AU$2.50 in Australia, and US$2.00 in the US, then the exchange rate should be AU$1.00 = US$0.80. If the exchange rate was below US$0.80 then the Australian currency was considered to be undervalued, and vice versa. It was further argued that an exchange rate based on the “hamburger standard” would be far less sensitive to the short term fluctuations that characterise open currency markets.

The idea has been expanded in various ways. Gunther [47] suggests a “Braten1 index”, based on a simple meal of meat with vegetables, as a means of assessing the purchasing power of people in Europe in the 18th century. The Economist [49] has compared the costs of a Big Mac in a number of countries with the cost of another standard food item, a “tall latte” as sold in over 30 countries by the Starbucks coffee company.

In essence the Big Mac approach assumes that a single commodity is as representative of domestic consumption as the extensive basket of goods and services used in sophisticated PPP exercises. Pakko and Pollard [48] suggest that the “simple collection of items … does just as well (or just as poorly) at demonstrating the principles and pitfalls of PPP as do more sophisticated measures”.

The index has been tested over the years. Ong [32] concludes that the notion of PPP “as a theory of exchange rate determination is probably the most useful and used of all exchange rate theories, despite its many detractors”, and (at 29) notes that empirical tests, utilising the Big Mac index to track exchange rate movements have been “surprisingly successful”. She concurs with Pakko and Pollard [36], saying that, based on her research, the hamburger index is as reliable as “most other measures of purchasing power parity:”

The index was used the index to compare “real” academic salaries across eight English-speaking countries [32]. The study highlights just how misleading such a comparison can be when based on currency exchange rates. Comparisons of “real” salaries then involved a simple comparison of the number of hamburgers that could be bought in each country. When compared using the exchange rate, a typical Australian salary was only $US38,319; in contrast, when the Big Mac index was applied, the Australian salary had purchasing power equivalent to a salary of $US47,992, making the “real” Australian salary considerably greater than the comparable US salary.

---

1 German for “roast meat”
4.5 Applying “Burgernomics” to Construction

The Big Mac index has been applied in a construction context in only two previous studies [14, 7]. Croce et al. [3] were scathing in their attack on the use of the Big Mac by Langston and de Valence, saying (at 21) that

“it should be obvious that countries where food production is not based on wheat, sesame seeds, beef, dairy products, dill pickles and potatoes and where a Big Mac is a luxury item, available only in major cities to urban elites mimicking Western tastes, rather than a fast food staple, is not any sort of a ‘standard commodity’”.

As discussed above, there is some support for the index in the literature (e.g. [32]). The hamburger actually does represent a variety of inputs provided by a number of industry sectors including agriculture (meat, vegetables, grain), manufacturing (bread, packaging), capital goods (kitchen machinery and buildings), as well as service, finance and direct labour. It also combines a range of tradeable and non-tradeable inputs (rogoff). As it has been estimated that the non-traded component of the cost of a Big Mac is as high as 93% [32], and the proportion of construction inputs traded externally is quite low [13] the index may be particularly well-suited as a deflator of construction costs. While the item may not be fully representative of all countries and regions, it is a commonly available in many places, particularly in large cities, and has been judged (by Ong and others) as a useful tool.

In effect Langston and Best calculated an equivalent costs expressed as “number of hamburgers” by dividing project cost (in local currency) by the local price of a Big Mac. While costs were expressed in US$/m2, costs between countries were actually compared by simply comparing the number of hamburgers that each project “cost”. If the Big Mac is accepted as a representative basket of goods, then such comparisons may be the most reliable available method.

5. Conclusions

International comparisons of performance are not easy but a relatively straightforward method has been described. As performance has a number of dimensions the challenge lies in finding a method that is comprehensive enough to be meaningful, but simple enough that it can be applied readily. The Performance Index described here relies mostly on project data that is routinely available within construction organizations, yet it includes, to some degree, the four components of performance, namely, productivity, time, cost and quality. The conversion of costs to a common base is, however, at least as important as the choice of methodology. Different conversion methods can produce significantly different results, and there are no readily available indices that can be applied to construction without the risk of distorted results. The Big Mac Index has been suggested as a viable tool that minimizes some of the problems inherent in such conversions. The best approach is to use as many methods as possible and test the sensitivity of the outcomes. At best, results should be taken as being indicative and not definitive.
References


186


