TOWARDS A LEAN UNDERSTANDING OF RESOURCE ALLOCATION IN A MULTI-PROJECT SUB-CONTRACTING ENVIRONMENT

(or 'What makes sub-contractors tick?')

Rafael Sacks

ABSTRACT

For many reasons, subcontracting is prevalent throughout the construction industry. Multiple subcontractors each perform work on multiple projects simultaneously. Each subcontractor strives to win sufficient jobs so as to ensure a ready supply of available work; each project manager strives to control the subcontractors on his/her project to maintain cost- and schedule-compliance, often pushing them to perform even when conditions preclude efficient or quality work. Construction subcontracting suffers the ills of traditional push workflows, but the problem for individual projects is exacerbated by subcontractors optimizing their workload across multiple projects, creating a snowball effect of growing instability for each individual project. Much of the research and implementation of lean construction has been carried out within the conceptual boundary of a single project or a single value stream. This paper proposes a multi-project, multi-subcontractor approach aimed at developing better understanding of workflow from the subcontractors' point of view, and contributing to development of a multi-project and multi-sub-contracting theory of production in construction projects. It attempts to define the questions, scope and methods for the research. A solution approach using an economic model is proposed. The long-term goal is to enable development of sub-contracting relationships and management procedures that harness the potential for both contractors and sub-contractors to benefit.

KEYWORDS

Subcontracting, multiple projects, theory, resource allocation, lean construction.

---

1 Senior Lecturer, Faculty of Civil and Env. Eng., 840 Rabin Building, Technion – Israel Institute of Technology, Haifa 32000, Israel, Phone +972-4-8293190, Fax +972-4-8293190, cvsacks@technion.ac.il
INTRODUCTION

Subcontracting has become prevalent in construction. For example, the proportion of construction employees employed by subcontractors in the UK rose from 25% in 1983 to 45% in 1998 (Edwards 2003). Some of the reasons are:

- increasing sophistication and specialization of trades, which requires long-term investment in personnel and equipment,
- increasing prefabrication off site, which similarly requires large-scale investment in fixed facilities,
- fluctuating demand for the services of general construction contractors, which demands agility in adjusting capacity,

The work environment that results is one in which each subcontractor must perform work on multiple projects simultaneously. From the subcontractor's point of view, the focus is not on any one specific project, but on multiple projects. Each construction project manager, on the other hand, is highly focused on the project for which he/she is responsible. The work environment can be described as a meta-project environment, as described in Figure 1.

![Figure 1. Meta-project subcontracted work environment](image)

Theoretical research in construction project management has focused largely on single project environments. Until the advent of lean thinking applied to construction (Koskela 1992), the primary theoretical bases were schedule optimization (using the critical path method - CPM) and local optimization of productivity (using measures such as Labor Utilization Factors) (Oglesby et al. 1989). Scheduling is mostly performed independently for...
each project, while productivity methods focus more narrowly on individual activities. To date, most lean construction research has retained the single project focus.

This paper proposes that attempting to understand the behavior of independent subcontractors in construction projects requires looking beyond the boundaries of a single project and taking the subcontractor's multi-project viewpoint. Subcontractors perceive optimization of their productivity across numerous projects to be of primary importance, leading them to adjust resource allocations between projects from time to time in accordance with their perception of the work load that will be made available at each project. From the project point of view, subcontractors' attempts to achieve full capacity utilization by shifting resources have a negative impact on workflow reliability.

OBSERVATIONS ON SUBCONTRACTING IN CONSTRUCTION

The importance and extent of subcontracting is familiar to all involved in construction. Its prevalence has been documented in numerous studies (Edwards 2003; Hinze and Tracey 1994; Hsieh 1998). (Walsh et al. 2003) reported an example of large-scale residential housing construction in which the subcontracted project paradigm has been taken to an extreme level of fragmentation and specialization. The subcontractors are termed 'hyper-specialized', and the resulting project management strategy has been to generate fixed construction schedules with large time buffers between subcontractors; the result is that work-in-progress inventory is high and project durations are far longer than the net sum of the actual working times required.

The main characteristic of the meta-project work environment is that multiple subcontractors each perform work on multiple projects simultaneously. Each subcontractor strives to maximize its workload so as to ensure a ready supply of available work (a backlog) at any given time for optimum resource utilization. This is commonly done without any consideration for the interests of any other subcontractor (Mathews et al. 2003). In projects run with 'tight' centralized control (Kim and Paulson 2003), project managers strive to control the subcontractors on their projects to maintain stability and schedule-compliance, most commonly by pushing them to perform even when conditions do not allow the work to be done efficiently or at the required level of quality. Similarly, this is done largely without consideration of the interests of the subcontractors. In many cases, in fact, subcontractors withhold true resource availability and scheduling information from the general contractor (Choo et al. 1999). If loose control is exercised, subcontractors perform activities subject to their individual resource constraints, rather than in compliance with the schedule, with the result that projects have long durations (Kim and Paulson 2003).

Common wisdom regarding the relationship between a general contractor (GC) and its subcontractors (subs) in construction is revealed in the following: "Undoubtedly, however, the provision having paramount importance to the success of the job is the schedule. A schedule that is unrealistic for one or more parties can be a disaster for all who plan on it." ... "the GC must monitor closely each day the activities of each sub and point out any evidence of slippage in schedule...the GC can frequently assist the sub in guidance and advice on the best use of manpower in order to meet the schedule. The more detailed the effort on the part of the GC in planning each phase of the project – and specifically the operations of each sub – the more likely it is the overall schedule will be realized." (Proctor
1996). But how can the GC incorporate considerations such as capacity costs of resource allocations, which are critical for subcontractors but encompass considerations across projects (O'Brien and Fischer 2000)? While calling for consideration of the scheduling interests of all parties at the outset, the approach leaves no room for schedule flexibility or for coping with variability once a project has begun. Only the subcontractor can plan its own operations to create a schedule that is realistic economically.

An examination of productivity in subcontracting in Taiwan revealed that various and extended efforts to improve subcontractor productivity using traditional labor utilization improvement approaches were fruitless (Hsieh 1998). A simple high-level graphic model was proposed for better understanding of barriers to productivity. However, this model does not recognize the importance of workflow and variability. It, too, adopts a project-centric view. Kim and Paulson (2003) proposed a distributed methodology and software intended to enable negotiation between subcontractors in order to adjust project schedules to optimize global utility by allowing subcontractors to trade schedule adjustments between them – including monetary compensation from one to another. This appears to be a sophisticated approach that may be applicable with highly select groups of subcontractors in specific projects, but may be unworkable for the majority of subcontractors and projects.

Finally, it should be noted that subcontracting is sometimes more than one level deep, especially in large scale projects. As the contractual chain grows longer, the relationship between the sub-sub-contractor and the client grows weaker. We might call the lowest level contractors in subcontracting chains – the ones whose workers actually do the work – the 'last contractors'.

HYPOTHESES AND RESEARCH QUESTIONS

The following three hypotheses express possible explanations for conditions that have been observed on construction projects. They are proposed on the basis of the observations of subcontracting practice in the literature and in the author's experience.

A. In sub-contracted projects, the ills common in traditional construction push workflows are exacerbated by the instability in process flow that begins when subcontractors are delayed at any point in the job, and elect to withdraw their crews to other projects where work is available. A common result is that the crews cannot then be withdrawn from the alternative projects just in time to resume work on the original project, which then increases instability even further. This creates a 'snowball' effect, with impact across projects.

B. The standard contractual arrangements between subcontractors and general contractors make it very difficult to implement practical steps intended to improve flow according to lean construction principles. Most contracts have extensive provisions for dealing with non-conformance or non-performance on the part of the subcontractor, but very few provisions – if any – for creating a stable workflow. Specifically, it is difficult to

- reorganize work cells across organizational boundaries (neither workers nor work packages can be moved from one sub-contractor to another),

4
• create a pull system with in which subcontractors remain available for continuous periods when work is required,

• improve flexibility by establishing conditions for shifting workload and/or labor and equipment between subcontractors as conditions demand at any given time,

• improve stability and reduce variability in terms of the number of workers, the arrival times of crews on site, the availability of core equipment, etc., where there is no direct ownership of the equipment and/or employment of the staff of the subcontractor on the part of the general contractor. (In fact, for reasons related to liabilities for compensation, most contracts go to great lengths to avoid establishing any kind of employer-employee relationship between a general contractor on the one hand and the subcontractor and its employees on the other hand).

C. Remuneration of sub-contractors based solely on measurement of work quantities does not achieve the goal of providing value to the end user of a building.

• How can the workers of a "last contractor" be informed of the value proposition for the client, and align their work to providing value for that client?

• Sub-contractor selection on the basis of lowest price often leads to a sub-optimal project result. Some researchers have proposed using key performance indicators, measured for sub-contractors over time in previous projects, as an addition to, or a substitute for, a price-based selection strategy. This approach recognizes the costs caused by unreliability, waste and poor quality, but makes no attempt to understand their root causes.

On the basis of a model, it should be possible to test hypotheses such as these by setting up appropriate experiments, simulations, surveys, or other research methods. The basic question that the hypotheses arouse, and which the model should represent, can be stated as "What are the factors that motivate a sub-contractor in assigning resources to the various projects the company is working on at any given time?" It can be detailed further by asking:

• Under what conditions will a sub-contractor increase, decrease or withdraw all resources from any particular project?

• How do the contractual relationships between a general contractor and a subcontractor influence the behavior of the subcontractor in assigning resources?

• Are sanctions, rewards, profit-sharing or other remuneration arrangements effective? What is their proportion and potential impact when evaluated within the context of the economic and other pressures acting on a subcontracting firm at any given point in time?

• How do market forces affect the willingness and ability of a subcontractor to commit and assign appropriate resources to projects?
The immediate goal of the research needed is not to answer these questions directly, but rather to establish an economic and a behavioral model of subcontractor decision-making that would enable prediction of the range of decisions that may be made, and the likelihood of each. Deterministic models for optimization of resource allocation for multiple tasks are available in traditional production operations research; however, the significant differences to industrial production suggest that a construction-specific theoretical model is required.

AN EXAMPLE OF THE APPROACH

Why does local optimization work against flow? This section presents a simplistic example of how the economic aspect of this tenet of lean thinking can be explained in the context of the subcontracted construction environment. The economic model of the environment begins with statement of the basic goal from the subcontractor’s point of view; to maximize profitability over some period of time. Assuming that the subcontractor is remunerated only according to quantity of work performed in each project, this can be detailed as follows:

\[
\text{max } P_T : P_T = \sum_i I_i - E_F ; \quad I_i = W_i U_i - W_i C_{M_i} - b (W_p - W_i) C_{M_i} - \frac{W_p}{r} C_s - C_0
\]

The parameters for these and the following equations are defined in Table 1. The terms on the right hand side of the second equation represent income for work performed, cost of materials actually consumed, cost of excess material, resource costs and fixed overheads. For sake of simplicity, it is assumed that only one type of work is performed by each subcontractor at each project, that \(U_i\) and \(C_{M_i}\) are constant for any project \(I\), and that fixed expenses (including salaries and overheads) are constant over time period \(T\). Under these assumptions, only the quantities of actual work performed and work planned are variable. The subcontractor’s challenge in any period \(T\) is to set \(W_{pi}\) for each project to maximize \(P_T\), subject to uncertainty about the outcome for \(W_i\) for each project.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P_T)</td>
<td>pre-tax profit over time (T)</td>
</tr>
<tr>
<td>(I_i)</td>
<td>net income from project (i) during any period (T)</td>
</tr>
<tr>
<td>(E_F)</td>
<td>fixed expenses: salaries and overheads.</td>
</tr>
<tr>
<td>(W_i)</td>
<td>actual work performed on project (i) during any period (T)</td>
</tr>
<tr>
<td>(U_i)</td>
<td>unit price for the works at project (i)</td>
</tr>
<tr>
<td>(C_{M_i})</td>
<td>unit cost of the materials for the works at project (i)</td>
</tr>
<tr>
<td>(W_{pi})</td>
<td>work promised/demanded by general contractor of project (i) in period (T)</td>
</tr>
<tr>
<td>(W_{pi})</td>
<td>work planned by the subcontractor in project (i) in period (T)</td>
</tr>
<tr>
<td>(W_{pi})</td>
<td>work actually made available in period (T)</td>
</tr>
<tr>
<td>(T_p)</td>
<td>net working time in period (T).</td>
</tr>
</tbody>
</table>

Table 1. Annotation.
cost/unit of time for one unit of resources allocated by the subcontractor to project \( i \), assumed constant over period \( T \)

| \( R_i \) | number of units of resource allocated by the subcontractor to project \( i \) over period \( T \). In reality, there are practical limits to the range in which resources can be assigned, such as the minimum number required to perform work at all and the ratio of flexible to fixed resources (O'Brien 2000). |
| \( r(t) \) | work rate (units/time) for a single unit of resource \( R \), which varies over time. For this simplified model, which ignores learning curve and other productivity influencing effects such as space available, the average work rate is defined as \( r = \frac{1}{T} \int_0^T r(t) \). |
| \( b \) | Waste factor for materials that remain unused at the end of any period \( T \) |
| \( E_{si} \) | cost for the resources assigned to project \( i \) over period \( T \): \( E_{si} = R_i C_{si} T \) |
| \( k \) | Ratio of resources supplied to resources demanded. |

(Tommelein et al. 1999) showed that variability in the rate at which work is supplied from one trade to another leads to degradation in performance of the downstream trades, and to lengthening of the project duration. An implicit assumption in that work was that average production capacity (the resources) of each trade was maintained at a constant level. However, in a subcontracted project environment, it is reasonable to assume that each subcontractor will set the quantity of resources applied to each project at each distinct time period (at least monthly, if not weekly). Determination of the correct resource level must take into account the expected amount of work that will be made available, because the maximum work that can actually be performed in period \( T \) on any project \( i \) is always the lesser of the work that can be performed given the resources assigned and the work that is actually provided by the general contractor.

The maximum work that can actually be performed is limited to the work actually made available: \( W_i \leq W_{si} \). Work is only available for execution when the work area is free (precedent work teams/subcontractors have completed their work and cleared the area), the materials have been delivered, the information that controls the work is provided, and project management has signaled that work can begin (Koskela (2000) lists seven resource flows that are prerequisite for successful execution of a construction task). The quantity of work actually made available also has a second order impact on profitability, because productivity itself is a function of work quantity and space (O'Brien 2000). Here, the focus is on the subcontractors' strategy in allocating resources; for sake of simplicity, the second order effect, material waste and overheads will be ignored. Thus the simplified expression for the subcontractor's net income for any single project in period \( T \) is stated as follows:

\[
I_i = W_i (U_i - C_{M_{li}}) - \frac{W_{si}}{r} C_{si}
\]
If the sub-contractor assumes a stable, deterministic workflow (i.e. \( T_n = T \), with no downtime, or in other words \( W_{At} = W_{Di} \)) and therefore assigns resources sufficient to meet the general contractor's demand, then the resources supplied are:

\[
R_i = \frac{W_{R_i}}{Tr} = \frac{W_{Di}}{Tr} \quad \text{and the resource cost over period } T \text{ is fixed at } E_{S_i} = R_i TC_{S_i} = W_{Di} \frac{C_{S_i}}{r}.
\]

In this case, the net income for project \( i \) over period \( T \) is:

\[
I_i = W_i (U_i - C_{M_i}) - W_{Di} \frac{C_{S_i}}{r}
\]

However, the actual net income is dependent on the availability of work:

If \( W_i = W_{At} \leq W_{Di} \) then

\[
\frac{I_i}{W_i} = (U_i - C_{M_i}) - \frac{W_{Di} C_{S_i}}{W_{At} r} \quad \text{.........................(1)}
\]

Alternatively, if \( W_{At} > W_{Di} \), then \( W_i = W_{Di} \), since the resources become fully utilized, and so

\[
\frac{I_i}{W_i} = (U_i - C_{M_i}) - \frac{C_{S_i}}{r} \quad \text{.........................................................................................(3)}
\]

Thus a theoretical upper limit is imposed on the unit profitability. This relationship is plotted in Figure 2 (curve \( k=1.0 \)).

![Figure 2](image)

Figure 2. Typical plot of subcontractor's net income as a function of the ratio of work performed to work planned for.

As can be seen, the sub-contractor's profitability is extremely sensitive to the ratio of the quantity of work that can be performed to the quantity of work that was demanded. However, the sub-contractor can reduce this sensitivity by reducing the resources assigned to the project. In this case, the resources allocated to the project are less than what is required to meet the general contractor's demand:

\[
R_i' \leq \frac{W_{Di}}{Tr} \quad \text{or} \quad R_i' = \frac{W_{R_i}}{Tr} = k \frac{W_{Di}}{Tr} = k R_i \text{ where } 0
\]
The total cost of the resources is now $E_i' = R_i'TC_i = kW_{D_i}C_{S_i}/r$ and the net income will depend on whether the actual work made available is smaller or greater than the capacity provided:

If $W_i = W_{A_i} \leq R_i'Tr$ then \( I_i = W_{A_i}(U_i - C_{M_i}) - kW_{D_i}C_{S_i}/r \) or

\[
\frac{I_i}{W_i} = (U_i - C_{M_i}) - k\frac{W_{D_i}C_{S_i}}{W_{A_i}/r}
\] ................................................................. (4)

If $W_i = R_i'Tr \leq W_{A_i}$ then \( I_i = W_i(U_i - C_{M_i}) - R_i'TC_i \) or

\[
\frac{I_i}{W_i} = (U_i - C_{M_i}) - \frac{R_i'C_{S_i}}{R_i'r} = (U_i - C_{M_i}) - \frac{C_{S_i}}{r}
\] ................................................................. (5)

This expression is independent of $W_{A_i}$, which implies that the sub-contractor can achieve full confidence in its level of profitability. It follows that if the sub-contractor can estimate $W_{A_i}$, then income can be optimized by setting $W_{p_i} = W_{A_i}$, i.e. $R_i' = \frac{W_{A_i}}{Tr}$, which can be expressed as $R_i' = kR_i = k\frac{W_{D_i}}{Tr} = \frac{W_{A_i}}{Tr}$, yielding $k = \frac{W_{A_i}}{W_{D_i}}$.

If the subcontractor were able to assess the probability profile for the actual work that it expects will be made available in terms of the demand stated by the general contractor, then the value $k$ can be set according to any given desired level of confidence. Figure 3 shows a theoretical probability profile estimated for $P\left[\frac{W_{A_i}}{W_{D_i}}\right]$, with the maximum at 1 and the mean less than one – representing a relatively high degree of confidence in the general contractor's ability to maintain a schedule. Nevertheless, if the subcontractor wishes to have an 80% probability of achieving the planned unit profitability for the work in time period $T$, it must set $k$ = 0.95. For a 90% probability, $k$ = 0.80. This suggests that:

- At the start of a subcontractor's activity at any construction site, it will assign resources to the project according to its perception of the general contractor's reliability in supplying work at the rate that has been demanded of them. In the majority of cases, the capacity assigned is likely to be lower than that demanded.

- As the project progresses, the subcontractor is likely to revise the resource assignment in response to the actual performance of the general contractor in supplying work. However, its tendency will always be to err on the side of caution – to supply slightly less than the quantity demanded. In this way they attempt to ensure that the resources are always fully utilized.
In practice, the rate of supply of work is dependent on the subcontractor performing work immediately upstream of the subcontractor in question. It follows that if each subcontractor follows the same strategy – starting with fewer resources than required by the work rate demanded by the project manager, then the project will be delayed at the start of activity of each subcontractor, as shown in Figure 4.

An alternative strategy that a subcontractor may adopt is to delay the start of work at a project beyond the start time demanded by the project manager, with the goal of ensuring that a sufficient inventory of work accumulates to buffer it from the upstream trade. In effect, and in the absence of any explicit pull mechanism, the subcontractor increases its confidence in the supply of work by simply waiting for inventory of work in progress to accumulate. Common wisdom in construction project management dictates that subcontractors be required to appear on site on the date stipulated by the project manager in terms of the contract, with the result that subcontractors attempt to create a buffer through what might be called 'unconventional' means. A good theoretical model might enable calculation of the relationship between time buffers and subcontractor confidence levels.

Both of the above cases support the theory presented by (Howell et al. 2001) which states, inter alia, that independent attempts to achieve full resource utilization will result in longer wait times for assignments, ultimately reducing workflow reliability.
Project managers who recognize this behavior often counteract it by demanding more resources than needed in fact; this has the predictable result of lowering the subcontractor’s level of confidence over time, which exacerbates the problem.

An additional aspect that will affect a subcontractor's strategy in allocating work teams to projects is whether the work load demanded from the subcontractor, over all of its current projects, is greater or smaller than its total capacity available. Assuming perfect balance is rare, the two situations are:

\[ \sum_{i=1}^{n} W_{D_i} > r \sum R \quad \text{or} \quad \sum_{i=1}^{n} W_{D_i} \leq r \sum R \rightarrow \sum_{i=1}^{n} R_i \leq \sum R \]

In the first case, the contractor cannot satisfy the demands of all of the projects, but full capacity utilization is achieved. In the second case, the subcontractor's profit is reduced because not all of its resources can be gainfully employed. Therefore, subcontractors strive to accumulate and maintain a backlog of work, even if by definition this implies an inability to provide good service. Workloads are not constant over time: uncertainty in demand for work at existing projects coupled with fluctuations in market demand, encourage subcontractors to minimize fixed resources. The wider the fluctuations are, the smaller a staff the subcontractor will seek to maintain, and thus the lower the level of service that can be provided.

Subcontractors often have fixed production facilities supporting their on-site work. Examples are custom-made windows and doors, carpentry, curtain walls, and precast concrete. In these cases, subcontractors must also balance their allocations of assembly crews with the workflow in their production plants.
CONCLUSIONS

Much of the research and implementation of lean construction has been carried out within the conceptual boundary of a single project or a single value stream. This paper proposes to approach the problem at a multi-project, multi-subcontractor level. The goal is to develop a theoretical model of the subcontracting paradigm, in order to identify ways to align subcontractor's inherent interest with behavior that enhances flow in individual construction projects. The model should consider the social, character, organizational, and other aspects of contractors that dictate their behavior as well as the economic and financial aspects, while recognizing that different projects have different environments (strategic, physical, and contractual).

The starting point should be definition of the correct goal function to be optimized. Lean thinking suggests that it be the value as defined by the building's occupants (the final construction clients). This contrasts with commonly accepted goals in construction practice (for project managers, it is the benefit on a single project for the main contractor; for a manager of a GC company, it is the aggregate benefit across multiple interdependent projects; for a subcontractor, it is the aggregate benefit across multiple activities in separate projects). Adopting a broader view suggests that the correct goal may be the global benefit achievable by multiple clients, contractors and sub-contractors across multiple projects.

The small example economic model presented indicates one possible approach. Others, such as game theory, are likely to be useful, and discrete event simulation may be appropriate for testing predicted outcomes. At the same time, the example economic model has underlined the shortfalls of considering too narrow a scope in a model. It demonstrates how local optimization works against flow, and it ignores waste generated at the project level by hiding it within the price paid per unit of work completed. By extension, limiting the scope of a model to any typical single project will result in ignoring system wide effects. The scope should therefore be set at the meta-project level described in Figure 1.

In reality, the meta-project environment is extremely large but finite, since potentially all projects may be interrelated through subcontractors. In practical terms, it will need to be modeled either as an open-ended system or as a smaller idealized system considering a 'closed shop' model in which a certain subset of subcontractors, and only that subset, performs in a subset of projects built by a subset of general contractors. The applicability of research results established considering a closed-shop model will need to be shown to be extensible to the broader general case.

ACKNOWLEDGMENT

The author gratefully acknowledges the contribution of Micha Harel, whose firsthand experience and keen curiosity were invaluable in developing the ideas presented.

REFERENCES


