Journal of Construction Engineering and Management

On the duration and cost variability of construction activities: an empirical study

--Manuscript Draft--

Manuscript Number:	COENG-8211R1
Full Title:	On the duration and cost variability of construction activities: an empirical study
Manuscript Region of Origin:	SPAIN
Article Type:	Technical Paper
Manuscript Classifications:	3.03000: Pre-project Planning; 5.01000: Estimating; 5.02000: Scheduling; 9.04000: Modeling
Funding Information:	Ministerio de Ciencia, Innovación y Universidades (Ramon y Cajal contract (RYC-2017-22222)) — Dr. Pablo Ballesteros-Pérez Ministerio de Ciencia, Innovación y Universidades (Estancias de movilidad en el extranjero José Castillejo para jóvenes doctores, 2017 (grant ref. CAS17/00488)) — Enoc Sanz-Ablanedo Ministerio de Ciencia, Innovación y Universidades (Estancias de profesores e investigadores senior en centros extranjeros Salvador de Madariaga 2018 (grant ref. PRX18/00381)) — Mª Carmen González-Cruz Junta de Andalucía (PAIDI TEP-955) — Dr. Pablo Ballesteros-Pérez, Alberto Cerezo-Narváez
Abstract:	The unique nature of construction projects can mean that construction activities often suffer from duration and cost variability. As this variability is unplanned it can present a problem when attempting to complete a project on time and on budget. Various factors causing this variability have been identified in the literature, but they predominantly refer to the nature and/or context of the whole project, rather than their specific activities. In this paper, the order of magnitude of and correlation between activity duration and cost variability is analyzed in 101 construction projects with over 5000 activities. To do this, the first four moments (mean, standard deviation, skewness and kurtosis) of actual versus planned duration and cost (log) ratios are analyzed by project, phase of execution and activity type. Results suggest that, contrary to common wisdom, construction activities do not end late on average. Instead, the large variability in the activity duration is the major factor causing significant project delays and cost overruns. The values of average activity duration and cost variability gathered in this study will also serve as a reference for construction managers to improve future construction planning and project simulation studies with more realistic data.
Corresponding Author:	Pablo Ballesteros-Pérez, Ph.D. Universidad de Cadiz Puerto Real, Cádiz SPAIN
Corresponding Author E-Mail:	pablo.ballesteros.perez@gmail.com;pabbalpe@hotmail.com
Order of Authors:	Pablo Ballesteros-Pérez, Ph.D. Enoc Sanz-Ablanedo, PhD Robby Soetanto, PhD Mª Carmen González-Cruz, PhD Graeme Larsen, PhD Alberto Cerezo-Narváez, PhD
Suggested Reviewers:	Mario Vanhoucke, PhD

	Professor, Universiteit Gent Many of his works are cited in our paper. He doesn't have any connection with this piece of research.
	Miklos Hajdu, PhD Senior researcher, Budapest University of Technology and Economics Some of his works are cited in our paper. He doesn't have any connection with this piece of research.
	Keith Molenaar, PhD Professor, University of Colorado, Boulder He is an expert in this area and he was involved in the early stages of this piece of research. He would be capable of assessing its quality while probably remain impartial.
Opposed Reviewers:	
Additional Information:	
Question	**Response**
Authors are required to attain permission to re-use content, figures, tables, charts, maps, and photographs for which the authors do not hold copyright. Figures created by the authors but previously published under copyright elsewhere may require permission. For more information see http://ascelibrary.org/doi/abs/10.1061/9780784479018.ch03. All permissions must be uploaded as a permission file in PDF format. Are there any required permissions that have not yet been secured? If yes, please explain in the comment box.	No
ASCE does not review manuscripts that are being considered elsewhere to include other ASCE Journals and all conference proceedings. Is the article or parts of it being considered for any other publication? If your answer is yes, please explain in the comments box below.	No
Is this article or parts of it already published in print or online in any language? ASCE does not review content already published (see next questions for conference papers and posted theses/dissertations). If your answer is yes, please explain in the comments box below.	No
Has this paper or parts of it been published as a conference proceeding? A conference proceeding may be reviewed for publication only if it has been	No

significantly revised and contains 50% new content. Any content overlap should be reworded and/or properly referenced. If your answer is yes, please explain in the comments box below and be prepared to provide the conference paper.	
ASCE allows submissions of papers that are based on theses and dissertations so long as the paper has been modified to fit the journal page limits, format, and tailored for the audience. ASCE will consider such papers even if the thesis or dissertation has been posted online provided that the degree-granting institution requires that the thesis or dissertation be posted. Is this paper a derivative of a thesis or dissertation posted or about to be posted on the Internet? If yes, please provide the URL or DOI permalink in the comment box below.	No
Each submission to ASCE must stand on its own and represent significant new information, which may include disproving the work of others. While it is acceptable to build upon one's own work or replicate other's work, it is not appropriate to fragment the research to maximize the number of manuscripts or to submit papers that represent very small incremental changes. ASCE may use tools such as CrossCheck, Duplicate Submission Checks, and Google Scholar to verify that submissions are novel. Does the manuscript constitute incremental work (i.e. restating raw data, models, or conclusions from a previously published study)?	No
Authors are expected to present their papers within the page limitations described in <u><i>Publishing in ASCE Journals: A Guide for Authors</u></i>. Technical papers and Case Studies must not exceed 30 double-spaced manuscript pages, including all figures and tables. Technical notes must not exceed 7 double-spaced manuscript pages. Papers that exceed the limits must be justified. Grossly over-length papers may be returned without review. Does this paper exceed the ASCE	Yes

length limitations? If yes, please provide justification in the comments box below.	
If yes, please provide justification in the comments box below. as follow-up to "Authors are expected to present their papers within the page limitations described in <u><i>Publishing in ASCE Journals: A Guide for Authors</u></i>. Technical papers and Case Studies must not exceed 30 double-spaced manuscript pages, including all figures and tables. Technical notes must not exceed 7 double-spaced manuscript pages. Papers that exceed the limits must be justified. Grossly over-length papers may be returned without review. Does this paper exceed the ASCE length limitations? If yes, please provide justification in the comments box below."	The paper is rather short (26 double-spaced pages with separations between all sections) and there is only one figure that can be also presented in a small size (but that we have provided bigger just in case). However, there are four tables out of six that might take a whole page (or maybe half, as we know the final paper format is much smaller). We have really tried very hard to condense the information of this paper to the very maximum (a shallow reading proves that easily). Maybe some of the tables can be presented within the Supplemental online material too, but for a first review I think the reviewers will find it easier if everything is more accessible. Still, we think that, if our paper is longer than the 30-page limit, it must be by a little margin. We are pointing this out just in case.
All authors listed on the manuscript must have contributed to the study and must approve the current version of the manuscript. Are there any authors on the paper that do not meet these criteria? If the answer is yes, please explain in the comments.	No
Was this paper previously declined or withdrawn from this or another ASCE journal? If so, please provide the previous manuscript number and explain what you have changed in this current version in the comments box below. You may upload a separate response to reviewers if your comments are extensive.	No
Companion manuscripts are discouraged as all papers published must be able to stand on their own. Justification must be provided to the editor if an author feels as though the work must be presented in two parts and published simultaneously. There is no guarantee that companions will be reviewed by the same reviewers, which complicates the review process, increases the risk for rejection and potentially lengthens the review time. If this is a companion paper, please indicate the part number and provide the title, authors and manuscript number (if available) for the companion papers along with your detailed justification for the editor in the comments box below. If there is no justification provided, or if there is	This is not a companion paper.

insufficient justification, the papers will be returned without review.	
If this manuscript is intended as part of a Special Issue or Collection, please provide the Special Collection title and name of the guest editor in the comments box below.	Not intended for a SI.
Papers published in ASCE Journals must make a contribution to the core body of knowledge and to the advancement of the field. Authors must consider how their new knowledge and/or innovations add value to the state of the art and/or state of the practice. Please outline the specific contributions of this research in the comments box. To read more about how the JCEM considers each paper's contributions please see the following Editor's Note.	Contributions are stated in the Abstract and Conclusions. A summary of those are: - This is the first extensive analysis measuring what are the average deviations (actual vs. planned) of activity durations and costs in several types of construction projects and phases of execution. -The results tables can be used to improve construction pre-project planning (use more representative data), for enhancing project simulation studies (using more realistic data for modelling activity durations and costs), and many other. - Finally, the study also proves numerically supported with real empirical data that, contrary to common wisdom, construction activities do NOT end late on average, instead it is their large variability what causes construction projects to end late.
JCEM strives to continually provide readers with the most relevant data and research. Authors are asked to consider the impact of their research on the global construction community. Papers with a single-region focus should be submitted as a Case Study. Descriptions of a Case Study can be found in <u><i>Publishing in ASCE Journals: A Guide for Authors</u></i>. Does your paper contain research that is focused on a single region?	No
The flat fee for including color figures in print is $800, regardless of the number of color figures. There is no fee for online only color figures. If you decide to not print figures in color, please ensure that the color figures will also make sense when printed in black-and-white, and remove any reference to color in the text. Only one file is accepted for each figure. Do you intend to pay to include color figures in print? If yes, please indicate which figures in the comments box.	No
The board of JCEM encourages authors to make data used in the current study available to other researchers in order to advance the science and profession. A Data Availability statement is required to appear at the end of the published paper.	Data analyzed during the study were provided by a third party. Requests for data should be directed to the provider indicated in the Acknowledgements.; All data generated or analyzed during the study are included in the submitted article or supplemental materials files.

Please include this information in a section titled “Data Availability Statement” right before the acknowledgements. Please indicate below whether your data is available and how a reader may access the data. Note that selected statements will appear verbatim in your paper.	
If there is anything else you wish to communicate to the editor of the journal, please do so in this box.	Cádiz (Spain), 25th April 2019 Dear Editor Prof. Jesus M. de la Garza, After the first round of reviewers’ comments, we are pleased to send you our revised manuscript within the area of construction productivity and scheduling. This paper we called “On the duration and cost variability of construction activities: an empirical study” analyzes the time and cost performance of over 5000 real construction activities with the intention of measuring to what extent those variables condition the whole project ending late and/or costing more. Nowadays, there is a proliferation of advanced scheduling tools to better plan and monitor project progress. Those scheduling tools assume specific statistical distributions to model the activity durations and costs. However, there is no previous study analyzing what those distributions look like and/or how they vary depending on the project execution stage. This paper offers a reliable first set of values describing these statistical distributions that almost all scheduling tools will be able to use to enhance their future forecasting accuracy. Furthermore, we think that practical implications of this work challenge many assumptions that had been made about construction projects. For example, it is proven that construction activities do not end late on average. This basically rules out the possibility that construction projects end late because most of their activities end late. Instead it is their high variability (both positive and negative) what makes a project end late. Also, it is analyzed how the significantly high correlation between activity duration and cost has a big influence on the final project cost overrun. This means that, very often, most delayed projects will also bring a huge cost increment. Furthermore, the authors would like to state that this manuscript contains an original work and is not under consideration by another journal. We, the authors, approve the current manuscript version and its submission. Finally, we would like to thank you for receiving our manuscript and considering it for review. Please address all correspondence concerning this manuscript to my personal email address should you have any other queries (pablo.ballesteros.perez@gmail.com) Yours sincerely, Pablo Ballesteros-Pérez* Enoc Sanz-Ablanedo Robby Soetanto Mª Carmen González-Cruz Graeme Larsen Alberto Cerezo-Narváez

On the duration and cost variability of construction activities: an empirical study

Ballesteros-Pérez, P. Ph.D.[1]*; Sanz-Ablanedo, E. Ph.D.[2] ; Soetanto, R. Ph.D. [3] ;
González-Cruz, Mª.C. Ph.D.[4] ; Larsen, G.D. Ph.D.[5]; Cerezo-Narváez, A. Ph.D.[6]

Abstract

The unique nature of construction projects can mean that construction activities often suffer from duration and cost variability. As this variability is unplanned it can present a problem when attempting to complete a project on time and on budget. Various factors causing this variability have been identified in the literature, but they predominantly refer to the nature and/or context of the whole project, rather than their specific activities.

In this paper, the order of magnitude of and correlation between activity duration and cost variability is analyzed in 101 construction projects with over 5000 activities. To do this, the first four moments (mean, standard deviation, skewness and kurtosis) of actual versus planned duration and cost (log) ratios are analyzed by project, phase of execution and activity type. Results suggest that, contrary to common wisdom, construction activities do not end late on average. Instead, the large variability in the activity duration is the major factor causing significant project delays and cost overruns. The values of average activity duration and cost variability gathered in this study will also serve as a reference for construction

---

[1]* Senior researcher. Dpto. de Ingeniería Mecánica y Diseño Industrial, Universidad de Cádiz.
Avda. Univ. de Cádiz 10, Puerto Real, 11519 Cádiz, Spain. Phone: +34 956 483 200, pablo.ballesteros@uca.es

[2] Associate Professor. Dpto. Tecnología Minera, Topográfica y Estructuras. Universidad de León.
Avda. Astorga, s/n. 24400 Ponferrada, León. Spain. Phone: +34 987 442 110, esana@unileon.es

[3] Senior Lecturer. School of Architecture, Building and Civil engineering. Loughborough University
Loughborough, LE11 6DF, UK, United Kingdom. Phone: +44 (0) 1509 228 748, r.soetanto@lboro.ac.uk

[4] Associate Professor. GIDDP, Depto. de Proyectos de Ingeniería. Universitat Politècnica de València.
Camino de Vera s/n, 46022 Valencia, Spain. Phone: +34 963 879 866, mcgonzal@dpi.upv.es

[5] Associate Professor. School of the Built Environment. University of Reading
Whiteknights, Reading RG6 6DF, United Kingdom. Phone: +44 (0) 118 378 7185, g.d.larsen@reading.ac.uk

[6] Assistant Professor. Dpto. de Ingeniería Mecánica y Diseño Industrial, Universidad de Cádiz.
Avda. Univ. de Cádiz 10, Puerto Real, 11519 Cádiz, Spain. Phone: +34 956 483 311, alberto.cerezo@uca.es

managers to improve future construction planning and project simulation studies with more realistic data.

Keywords: scheduling; activity variability; merge event bias; network topology; project delays

Introduction

Construction activities usually suffer from variability in terms of both duration and cost. With each construction project being unique, factors of this variability are plentiful (Ballesteros-Pérez et al. 2017). These factors include project location, clients, regulations, labor, equipment, technology, subcontractors, experience, stakeholders, even the project team, are likely to change, at least partially, among projects (Chudley and Greeno 2016). All these factors, plus many other, make of the duration and cost estimation exercise, a challenging task for construction managers.

It may be easy to believe, though, that construction activities are apparently more likely to end later and cost more than the other way around. In fact, this would constitute a compelling reason why so many construction projects end late and exceed their initial budget.

Factors that cause projects to end late or result in cost overruns have been studied in the construction literature for a long time. Some of the most recurrent are poor planning and control practices, deficient construction site management, shortages of labor and/or low productivity, problems with the supply chain and/or procurement practices, contractor’s and/or client’s financial problems, project specifications or design changes, communication and/or co-ordination problems among stakeholders, interferences with onsite services, adverse weather conditions, and legal disputes and contract claims (Ballesteros-pérez et al. 2015, 2018b). Among all these, however, poor planning and control practices are consistently among the most pervasive (AlSehaimi and Koskela 2008).

Ballesteros-Pérez et al. (2018a) recently showed how the most common scheduling techniques (Gantt chart, Critical Path Method and Project Evaluation and Review Technique, PERT) consistently underestimate the actual project duration and cost. One of the major causes of this underestimation came precisely from neglecting activity duration variability.

Apart from the classical scheduling techniques, more advanced techniques for getting improved project duration and/or cost estimates have been proposed over the years (e.g. fuzzy logic, neural network analysis, Monte Carlo simulations, artificial intelligence methods, many variants of PERT, and even more extensions of Earned Value Management (Ballesteros-Pérez, 2017a)). What all these methods have in common, classical and modern alike, is that they all require some prior estimates of the potential activity durations and costs. For example, PERT-related techniques generally resort to three-point estimates (pessimistic, optimistic and most likely durations and costs); Monte Carlo simulations require the statistical distributions of each activity as input; and neural network analysis and artificial intelligence methods require training sets of similar construction projects. Access to this information is often the major limitation of these methods. Similarly, realistic data on the correlation between activity duration and costs is also a rare commodity, which forces these techniques to either assume independence between activities and costs, or resort to subjective correlation factors (Banerjee and Paul 2008; Cho 2009). Consequently, when enough quantity or quality of information is not available, the forecasting accuracy of the actual project duration and/or cost is expected to be unreliable.

Unfortunately, despite its importance, there is a dearth of research into activity duration and cost variability in the construction management literature. Maybe, the only exception would be the work of Trietsch et al. (2012) who attempted to establish a distribution that satisfactorily describes construction activity durations. However, as early suggested by MacCrimmon and Ryavec (1964), trying to find a universal distribution that fits all types of activities is a futile effort because each type of activity is unique. Furthermore, its context might also have a significant influence which is difficult, if not impossible, to parameterize mathematically.

Nonetheless, these difficulties should not be a deterrent to, at least, attempting to measure the average level of variability of construction activity durations and costs. As argued, this would be an extremely valuable input for future project duration and cost forecasting techniques, as well as providing powerful baseline information for enhancing project control and monitoring.

Hence, the present paper precisely attempts to fill this research gap in the construction management literature: measure the average level of activity duration and cost variability. It will also justify how and why, given this level of activity variability in common project networks, it is expected that most construction projects end late and go over budget. To achieve this, the actual/planned (log) ratios of many project and activity durations and costs will be analyzed. The correlation between activity durations and costs will also be studied. Finally, the most common network topologies (descriptors of what the project networks are like, that is, how activities are arranged and connected with each other) will be summarized and the potential impact of activity variability on these networks described in detail.

The paper will be structured as follows. The *background* section will provide an overview of the importance of the first four moments of the activity duration and costs impacting the final project duration and cost. This section will introduce the concept of merge event bias and describe how it may cause project delays and cost overruns depending of each project network topology. The *materials and methods* section will describe how a dataset of 101 projects was classified according to different activity categories, and then their log actual vs planned durations and cost deviations analyzed activity by activity. The *discussion* section will provide insights on to what the numerical results mean and how they are connected to the project network topology in common construction projects. Finally, the *conclusions* will summarize the whole analysis, highlight the major contributions to the body of knowledge, state the study limitations, and propose future research continuations.

Background

There have been numerous studies analyzing delays and cost overruns in construction projects at project level (e.g. (Hamzah et al. 2011; Keane and Caletka 2008; Mahamid et al. 2012; Ogunlana et al. 1996; Orangi et al. 2011; Senouci et al. 2016)). Most studies have focused on either establishing the causes of delays and cost overruns, and/or proposing some regression analyses to avoid slippages in the future. Generally, these studies have been aligned with a more reductionist perspective, seeking to emphasize a particular context (same region, client, type of projects, or a combination of these).

Conversely, there have not been hardly studies measuring the ‘activity’ durations and costs, let alone their variability in real construction projects. With the exception of Trietsch et al. (2012) mentioned earlier, perhaps the closest are a handful of studies analyzing the sensitivity of the project duration to different levels of activity mean duration and dispersion (e.g. Elmaghraby & Taner (1999) and Elmaghraby (2000)).

Additionally, but from a purely mathematical and simulation perspective, some studies have tried to gauge to what extent the adopted activity statistical distributions have a significant repercussion on the final project duration. In this regard, a recent study by Hajdu and Bokor (2014) concluded that the maximum project duration deviation when using alternative activity distributions was generally well below 10%. This finding resonated with observations from an earlier study on the limitations of PERT. MacCrimmon and Ryavec

(1964) showed that, if triangular distributions for modelling activity durations had been chosen instead of Beta distributions, the probabilistic project duration would have produced almost identical results.

The reason why the choice of a particular statistical distribution does not seem that relevant is because the third and fourth moments (skewness and kurtosis) are blurred very quickly in Stochastic Network Analysis (SNA) (Hajdu and Bokor 2016). At the time of writing, SNA is considered the most accurate approach to model project schedule networks (Ballesteros-Pérez, 2017b). In SNA, activity durations and costs are modelled by statistical distributions (with or without correlation with each other). More precisely, distributions are summed when computing the total costs of activities, or the total duration of activities arrayed in series. On the other hand, the maximum of distributions (instead of a sum) is calculated whenever we calculate the total duration of a set of activities placed in parallel. In either case, the third and fourth moments (skewness and kurtosis) have a minor influence on the resulting distribution (of a path or project duration).

However, the first two moments (mean and variance, or alternatively, standard deviation) play a major role in the resulting distribution modelling the total project duration. When there is some correlation between durations and costs (virtually always in construction projects), they also have an indirect but still significant, influence on the final project cost.

To sum up, when two or more distributions are convoluted (summed for computing the project cost or the duration of activities in series) the resulting distribution, by the Central Limit Theorem, quickly converges to a Normal distribution. The mean and variance of this Normal distribution correspond to the sum of means and variances, respectively, of the individual activity distributions. Therefore, the first two moments will mostly determine what the resulting distribution looks like.

When some activities are arranged in parallel and they all need to finish before the project can continue, the resulting distribution quickly converges to an extreme value distribution of maxima (normally a Fréchet or a Gumbel distribution) (Dodin and Sirvanci 1990). Again, the first two moments of the involved activity distributions will determine the location and scale of the resulting extreme value distribution. This phenomenon is commonly known as the ‘merge event bias’ (Khamooshi & Cioffi, 2013; Vanhoucke, 2012) and it is indeed the major source of inaccuracy of all deterministic scheduling techniques.

Real construction project schedules (networks) generally involve many subsets of activities both arranged in parallel and in series. Hence, multiple convolutions (sums) and maxima of distributions need to be computed so that the final project duration and cost can be calculated. The influence of each activity’s first two moments (mean and standard deviation) will be key in this final result. This justifies why an order of magnitude of these two moments is worth collecting from a representative dataset of real construction activities.

Finally, another factor that determines how the activity distributions are merged with each other is dependent on the project network topology itself. Network topology refers to the logical layout of a network (a project schedule). It defines the way different activities (often referred to as nodes) are placed and interconnected with each other. Many metrics have been proposed for describing the network configuration. Some well-known examples are the Coefficient of Network Complexity (Davies 1973; Pascoe 1966), the Order Strength (Mastor 1970) and the Complexity Index (Bein et al. 1992). However, these only capture the project complexity and will not be used here.

Instead, this study will make use of four topology measures that describe the structure of an activity-on-the-node network, not just its complexity. These measures were initially proposed by Tavares et al. (1999) and later improved by Vanhoucke (2008). The four measures (also named *indicators*) used are: serial-Parallel (SP) indicator, Activity

Distribution (AD), Length of Arcs (LA) indicator, and Topological Float (TF) which will be explained in the following sections. All these indicators range between 0 and 1 and constitute simple measures describing to what extent the first two moments of the construction activities may condition the final duration and cost of a project.

Materials and methods

In this section, the characteristics of the projects and activity datasets analyzed are described first. The details of how the activity and project data was filtered and categorized, under multiple levels of analysis, is also presented. Next, the first four moments of activity durations and costs are reported and commented separately. Finally, the correlations between activity durations and costs are reported along with their statistical significance.

Projects and activities dataset

This research used two different project datasets. The first (and main) one is analyzed at both activity- and project-level. The second dataset contains project level information (planned and actual project durations and costs) and will be used for illustrative purposes in the *discussions*.

In order to obtain representative values of the first four moments of the activity durations and costs, a significant amount of activities is necessary. In the first dataset, 101 construction projects are analyzed initially encompassing 5,697 activities.

Projects are classified in four types: Building, Civil engineering, Industrial and Services. Building projects are mostly aimed at constructing a building or parts of a building. Civil engineering refers to infrastructure construction in general. Industrial projects refer to installations and/or electromechanical equipment. Services refer to projects with a significant operational and/or production component.

The 101-project dataset was retrieved from a real projects dataset originally developed by Batselier and Vanhoucke (2015) and Vanhoucke et al. (2016). Although the exact location of those projects is not disclosed in most cases (due to a confidentiality clause with the information donors), it is known that most of them belong to Belgium, the Netherlands, Italy, USA and Azerbaijan.

At the time of writing, the complete project dataset is curated by the Operations Research & Scheduling Research Group at Ghent University and comprises 125 projects. 24 projects out of the 125 were not used as they did not include tracking information (actual activity durations and costs). All 125 projects, however, can be accessed at the website of OR-AS.be (2018). The major features of the 101 construction projects selected for this study are summarized in Table 1. The last four columns of Table 1 include some project network topological information (indicators SP, AD, LA, and TF) that will be used later.

<Insert Table 1 here>

We deem the variety and number of project types, costs, durations, topologies and number of activities as sufficiently representative for a first representative analysis. Yet, further details and specific project information can also be found as individual project cards at OR-AS.be (2018).

Analysis outline

This analysis focuses first on the activity-level deviations of durations and costs. Project-level data will also be analyzed later, but from a complementary point of view to activities analyses. The activity duration and cost deviations are calculated for each activity *i* in the first dataset according to these two expressions, respectively:

$$Activity\ duration\ deviation\ of\ \ activity\ i = LOG_{10}\left(\frac{Actual\ duration\ of\ \ activity\ i}{Planned\ \ duration\ of\ \ activity\ i}\right) \qquad (1)$$

$$Activity\ cost\ deviation\ of\ \ activity\ i = LOG_{10}\left(\frac{Actual\ cost\ of\ \ activity\ i}{Planned\ \ cost\ of\ \ activity\ i}\right) \qquad (2)$$

It is worth emphasizing that both ratios above are expressed in logarithmic scale. This is important, as ratios of variables which are always positive (e.g. durations and costs) are not symmetrical respect to the value 1. The scale distortion of these ratios (they range between 0 and 1 when the denominator is bigger than the numerator, but between 1 and + infinity when the numerator is bigger than the denominator) creates an artificial positive skewness in the data distribution that can only be removed by taking the log ratios beforehand. Additionally, in log scale, the variable variances are additive, rather than multiplicative.

Therefore, we will take the logarithm of every ratio before analyzing their activity duration and cost moments. We resorted to logarithms with base 10 because their orders of magnitude are a little more familiar, but any other base would have been possible.

Lastly, it is important to note that ratios in natural scale from 0 to 1 correspond to values from -infinity to 0 in any log scale. Whereas ratios in natural scale from 1 to +infinity correspond to the (0, +∞) range. Both ranges also have a symmetrical correspondence with each other in log scale (e.g. ratios ½ and 2 in natural scale have the same values with opposite signs in log scale, that is -0.301 and 0.301, respectively) which makes the interpretations of variability results easier. Bearing this in mind, the next step consists of describing how the activities were grouped to analyze their ratios and produce robust results. The progressive classification levels can be found in Table 2.

<Insert Table 2 here>

From top to bottom, three levels of activity classifications are presented. Each level consists of three types of activities:

- Planned and Performed (P&P). These activities correspond to activities that were initially planned and were also finally executed in the projects analyzed. These are the most frequent and the only ones that are considered in the analysis.
- Unplanned but Performed (UbP). These activities correspond to activities that were not initially planned but that were deemed necessary and had to be eventually carried out. These activities were removed from the analysis because their ratios converged to + infinity (as the planned values in the denominators equal 0), and because most of the time they come from planning mistakes or omissions.
- Planned but not Performed (PbnP). These activities correspond to activities that were initially planned, but that were not executed in the end. These activity ratios would equal zero in natural scale but their logarithmic values would converge to - infinity. They also represent bad estimates of the planned schedule like UbP activities, hence, they were also removed from the analysis.

Concerning activity grouping, four levels of analysis (0 to 3) were considered:

- Level 0 comprises all activities analyzed from all projects. This allows drawing general average conclusions without paying attention to proportions nor types of those activities.
- Level 1. Activities are classified under the same four types of projects stated in Table 1 (building, civil engineering, industrial and services). As expected, this level allows analyzing how the activity durations and costs deviations differ by (generic) types of

projects. Some group average and dispersion results of activity durations and costs are also included for reference on the right columns of Level 1 sub-table.

- Level 2. Within the previous four project type categories we further classify activities into three standard phases of the every project lifecycle according to the PMBoK: Planning, Execution and Closure (Project Management Institute 2017). Classifying activities into these three categories is straightforward with the activity descriptions available in almost all projects. The fourth phase considered by the PMBoK (Monitoring and control) is not relevant for this analysis, therefore not considered.
- Level 3. For the *execution* phase of *Building* and *Civil engineering* projects only activities are further classified into five generic groups, called here *activity types* (auxiliary works, substructure, superstructure, specialized works, and facilities). These are also common and relatively straightforward groups of activities in most construction projects. For a more detailed description of the scope of each group the reader is referred to Chudley and Greeno (2016).

Level 3 allowed classifying activities into one last level right above the nature of the activity itself. Activities in this level were classified mostly thanks to the descriptions of the project *summary activities* (that were indeed not used for anything else in the analysis). Finally, as highlighted at the beginning of level 3, only activities from the *execution* phase of *building* and *civil engineering* projects were used. This is due to the number of *execution* activities in *Industrial* projects being considered too low. Also, because *Execution* activities belonging to *Services* projects, despite higher in number, were found too heterogeneous. The latter made hard to classify these activities within similar self-contained categories (*Services* projects are indeed much more varied regarding the nature of its activities).

Activity duration results

The first four moments (average, standard deviation, skewness and kurtosis) of the activities log ratios were analyzed according to the four levels described in Table 2. Table 3 shows now the results for the activity *duration* log ratios (LOG_{10} (*actual* / *planned*)).

<Insert Table 3 here>

For each case and level analyzed, four numerical values are displayed: *n* (the sample size, that is, the number of activities used to calculate the four moments), and the four moment values (in logarithmic scale). However, due to the major relevance of the first two moments (average and standard deviation) these two have also been included in natural scale within parentheses right below their respective logarithmic values. Values in natural scale are expected to help the reader to better grasp the order of magnitude of these moments. With this information, Table 3 is self-explanatory. The number of readings and details in this table are numerous, so attention is given to the most relevant findings.

Concerning *Averages*, it is striking to observe how most values remain very close to 0 (in log values) or 1 (in natural values). Some exceptions may be *Services* projects and the *Planning* phase activities (Level 2) from *Building* and *Civil engineering* projects. Yet, in the latter, average ratios values remain close to 5% (in log values) or 11% (in natural values). Overall, as these log ratios are so close to zero, this suggests that construction activities do *not* end late (on average). This may be an unexpected finding, as the easier explanation for projects ending late was that its activities ended late on average. This result seems to suggest the problem lies somewhere else.

Concerning the *Standard Deviation* (SD) values, results are very different. SD, by definition, can only be positive but it is quite clear that, unlike the averages, SDs are not close to zero. Instead, with a few exceptions, SD values are almost always above 0.15 (in log scale) or 43% (in natural scale) between the actual and planned durations. This is an extremely high level of variability and, despite construction activities do not end late *on average*, they do suffer from wide dispersions which condition to a big extent the project-level delays, as will be justified later. On a secondary note, *Industrial* and *Services* projects also have a bigger variability than the other types of projects. Interpretations by project phase (level 2) and activity type (level 3) are more varied.

The results on *Skewness* are relatively uniform. A common rule of thumb assumes that skewness values ranging from -2 to +2 are indicative of a low distribution asymmetry (George and Mallery 2010). This is the case in Table 3 with very few exceptions. Therefore, the log ratios distribution must be approximately symmetrical and, combined with averages also close to zero, we can conclude that there is approximately the same probability of finding early activities than tardy activities.

Concerning *kurtosis*, the picture is very different. Values are generally well above 3, which would describe the kurtosis corresponding to the Normal distribution. This result means that log ratio duration values resemble a peaked distribution with heavy tails. In other words, the majority of the actual durations are not close to their planned values. As stated earlier, many other readings may be extracted from Table 3. However, for the sake of clarity, only the most relevant high-level interpretations are presented.

Activity cost results

Table 4 represents the first four moments of the activity actual versus planned *cost* log ratios. In parentheses, we can find the antilogarithmic (natural scale) values of the first two moments as well. Table 4 values differ substantially from those found in Table 3.

<Insert Table 4 here>

Concerning *Average* values, most of them are clearly positive and generally above 1.01 (in log values) or alternatively above 3% (in natural scale). A clear exception may be the *Industrial* projects whose average is negative. This may be because *Industrial* projects are frequently composed of electromechanical equipment whose procurement prices are relatively easier to estimate more accurately ex-ante than other types of projects. Additionally, *Civil engineering* and *Services* projects are among the ones whose activities tend to suffer from more cost overruns. This may be due to *civil engineering* projects being (generally) less standard than Buildings whose average log ratios remain closer to 0. On the other hand, *services* projects as indicated in Table 3, suffered from more delays on average than other types of projects. Being these types of projects frequently more labor intensive, it seems logical that those extra durations are correlated with these extra costs.

Concerning *Standard Deviation* (SD), variability is even more evident than in the case of duration log ratios. On level 0 we can appreciate how the average activity SD reaches 0.25 (78% of variability in natural scale). On level 1, no project type has a variability below 0.16 (46% of variability, in the case of *Building* projects) and two of them (*Civil engineering* and *Services*) remain above 0.30 (>100% of variability). SDs on levels 2 and 3 offer similar readings but with wider values.

Concerning *skewness*, cost log ratios are more varied than their duration counterparts. In general, when *average* values are negative, the skewness values are also predominantly negative. Similarly, when the *average* costs are positive, the cost distribution is also positively skewed.

Concerning *kurtosis*, values are much higher than its duration ratios counterpart too. This would be indicative again that most activity actual costs substantially differ from their planned values (a high proportion of the actual costs tend to be substantially different from their planned costs).

Activity duration and cost correlation

Numerical results of the log ratios of the first four moments offered very interesting information about the nature of duration extensions and cost overruns at activity level. It is not the intention of this study to find a distribution that fits these four moments, though. As suggested by other researchers and also discussed earlier, each activity is different in nature and it is quite likely that a fit-for-all distribution does not exist. Indeed, on observing the wide range of skewness and kurtosis values in Tables 3 and 4, that seems to be exactly the case.

However, a pending but also equally relevant issue is to analyze the potential correlation between activity duration variation and cost variation. For this aim, all activities were grouped under the very same levels previously described and linear correlations were calculated among the duration log ratios and the cost log ratios. A summary of this analysis is presented in Table 5. Spearman's rho and Kendall's tau non-linear (rank) correlations were also tested. However, they only very marginally improved the linear correlation results and were considered not worth including as they did not seem to barely depart from the linear case shown in Table 5.

<Insert Table 5 here>

Table 5 is divided in two major blocks. The upper block is devoted to activity-level correlations. The lower block is reserved for project-level correlations. For each correlation it has been specified how many datapoints were used (column labelled as n), Pearson's correlation coefficient (R), the coefficient of determination (R^2), along with the gradient (*slope* column) and *intercept* of the linear regression lines.Statistically highly significant correlations have been marked with two asterisks (**) separately for R^2 tests (with the

Snedecor's *F* distribution) and *slope* tests (with the Student's *T* distribution). Significant statistical correlations have been marked with a single asterisk (*).

In the case of activity-level correlations, almost all correlation values are significant. This mean that values of R^2 are very unlikely to have happened by chance. This is not the case at Project-level correlations where, apart from the level 0 of analysis (all 101 projects grouped together), R^2 values have not been found to be statistically significant. This means we cannot count on the reliability of project-level duration-cost correlations, hence they will be ignored moving forward.

Correlations at activity-level do offer very interesting results. R and R^2 evidence weak to moderate correlations (R^2 ranging between 0.10 and 0.62), but the slopes of such correlations are rather close to 0.50 in some levels and almost all of them are significant (marked with ** or *). More precisely, when there is no differentiation among activities (level 0), the slope is as high as 0.704. This means that a 100% activity duration extension (in log scale) would cause a 70.4% cost increment on that activity. This is quite a high gradient.

Differentiating by project type (level 1), the slopes become more informative. *Building* and *civil engineering* projects boast a gradient close to 0.5, that is, every 100% of duration increment is likely to cause a 50% of cost increment for that activity. For the other two types of projects we have no statistically significant slopes, despite it seems clear that *industrial* projects (probably due to the higher component of electromechanical equipment in the project budget) have lower slopes. On the contrary, *Services* projects, being more labor intensive, have higher slopes.

Results by project phase (level 2) seem more homogeneous. However, only the *execution* activities' slope is statistically significant. This level of correlation seems to replicate the results previously provided for level 0.

Results at level 3 are again not that heterogeneous and they all are statistically significant. However, there is nothing remarkable that has not been highlighted before.

A last note concerns the regression line intercepts (last column in Table 5). As can be seen, these values remain above 0.02 (in log scale) most of the time. That is approximately equivalent to an intercept of 5% in natural scale, which means that, no matter whether activity duration extensions are materialized or not, costs are likely to increase around 5% by default. These values are in line with the log ratio cost *averages* found in Table 4.

Discussion

So far, almost all analyses have focused on individual activities. Yet, it is acknowledged that the construction process is not an exact science and construction managers are often ‘judged’ upon their capability to manage activity variability. Hence, the key concern is the whole project suffering from delays and cost overruns, not just some of its activities. It was proposed earlier that this is because activities suffer from variability (both positive and negative), not because they are delayed on average. This section is devoted to analyze whether this speculation seems acceptable.

Let us start by approaching the problem from a graphical perspective first. For that purpose, a second dataset of 746 road construction projects from the Florida Department of Transportation (USA) is used. Given the number of contracts, no descriptive table is included in the paper, but the complete dataset can be found as *supplemental online material*. This additional project dataset has been used here because they represent relatively similar (homogeneous) contracts, from the same client, and during a short period of time. Arguably, this is the closest to assuming that these projects are 746 different realizations (possible outcomes) of the same generic type of project (in this case a road construction, that is, a *civil engineering* project).

Figure 1 represents the distributions of the log deviation ratios for durations and costs for the 746 contracts (using expressions (1) and (2) at project-level, not activity-level).

<Insert Figure 1 here>

Concerning project duration deviations (curve with black circles), it closely resembles an extreme value distribution of maxima (both Fréchet and Gumbel fits have been provided for comparison in black colors). This means that the merge event bias takes an important role when determining the actual project duration. Results in natural scale are, in this occasion, almost identical but they have not been provided to avoid curve cluttering.

Furthermore, it is worth noting that the average project duration extension is around 0.21 (in log scale). For *Civil Engineering* projects in Table 3, the average of the duration log ratio was negative (-0.008). This means the activities from civil engineering projects ended sooner than planned (on average). It is unlikely then, that the projects represented in Figure 1 could have ended later because a significant proportion of their activities ended late. However, the activity duration variability (the standard deviation) was 0.20. In extreme value theory, the mean of the highest order statistic distribution of a Normal distribution with three or four draws is approximately one standard deviation. The Normal distribution represents very well the distribution of the durations of each path (before they merge) (Ballesteros-Pérez, 2017a). Therefore, the average of the duration distribution coincides very closely with what is to be expected from the data from Table 3 for civil engineering projects ($0.21 \approx 0.20$). Later it will be shown how more than three paths are quite common in civil engineering construction schedules.

Concerning project costs (curve with grey crosses), the situation is very different. The distribution of costs (log scale) resembles a Gamma distribution. It is worth noting that when a random variable *X* follows a Pareto distribution with parameter λ, the logarithm of *X* follows an Exponential distribution with the same parameter λ. This is relevant because the costs of individual activities are well known to resemble a Pareto distribution in almost all construction projects (Love et al. 2014; Love and Sing 2013). Hence, as cost ratios are being processed here in log values, our distribution should also resemble an Exponential distribution (continuous grey line in Figure 1). Additionally, as the sum of exponential distributions is a Gamma distribution, that would offer some explanation, to why we are observing a Gamma distribution (dashed grey line) fitting almost perfectly the log cost deviations in Figure 1. In this case the exponential distribution also provides a good fit, but that is not always the case in other construction project datasets.

Having approached the problem from a graphical and statistical perspective, it will be addressed now from a topological perspective. Network topology describes the layout of project schedules. The values of four representative topological indicators are displayed on the last four columns of Table 1 for the 101 projects analyzed. Table 6 now shows the average values of each topological indicator listed in Table 1, but categorized by Project type (*building, civil engineering, industrial* and *services*), as well as for all projects together (last row).

<Insert Table 6 here>

The Serial-Parallel (SP) indicator is probably the most relevant of the four indicators for the purpose of this study. This indicator measures the closeness of a network to a serial or parallel network. Namely, SP = (*m*-1)/(*n*-1); where *n* is the total number of project activities in a project schedule, and *m* is the number of activities in the path with a higher number of activities (which may not necessarily be the longest in duration, as topological measures ignore the activity durations). Hence, SP=0 means all activities are in parallel, whereas SP=100% means all activities are in series. This indicator can also be considered as an estimate of the amount of critical and non-critical activities in a network (Vanhoucke and Vandevoorde 2009). Therefore, rounded up values of the inverse of the SP (that is $\lceil 1/SP \rceil$) provide us with an estimate of the minimum number of paths of a project schedule. Values of SP below 50% would mean that construction schedules have (approximately) at least three paths. This agrees with what we appreciated in the black curve of Figure 1. *Industrial* projects, despite having on average at least two paths, generally have a dominant one (which condenses, on average, 55% of the activities). In *service* projects schedules there are at least five paths (on average), as only 20% (a fifth) of the activities are critical.

Activity Distribution (AD) measures the distribution of project activities along the levels of the project. In network topology, the number of project levels can be loosely defined as the number of activities that are arrayed in parallel in a project schedule. Hence, AD measures the width of the network. However, it is worth noting that activities arrayed in parallel do not necessarily have be executed simultaneously (because they may have different time lags and/or activity durations). When AD=0 all levels contain a similar number of activities and the number of activities is uniformly distributed over all levels. When AD=100% there is one level with a maximal number of activities, and all other levels contain a single activity. All four types of projects average AD values are close to 58% indicating that the longest path has more activities than other paths, but still those other paths contain a significant number of activities, that is, they can potentially cause project delays.

The Length of Arcs (LA) indicator measures the tightness of each precedence relationship between two activities as the distance between two activities in the project network. When LA=0 the network has many precedence relationships between two activities on levels far from each other such that the activity can be shifted further in the network.

When LA=100%, many precedence relationships have a length of one, resulting in activities with immediate successors on the next level of the network and with little freedom to shift. Average LA values are much closer to 0 than to 100% (overall average of 14.1%). This means that activities tend to have many predecessors (on average) from different levels (paths), which would reinforce the merge event bias effect.

Finally, the Topological Float (TF) measures the degrees of freedom per activity as the amount of slack or float an activity has. When TF=0 the network structure is 100% dense and no activities can be shifted within its structure. When TF=100% the schedule consists of a single chain of activities without topological float. The average TF indicator value of 40.5% means that the average activity structure of construction projects is rather dense.

Therefore, the highlights of this brief topological analysis above for construction projects are that: construction schedules are relatively dense (activity-wise), usually composed of at least three major paths, and with activities whose predecessors usually come, not just from activities located on the same path, but also from other paths. This means that the merge event bias plays a very important role in construction schedules. And, precisely thanks to the high level of duration variability existing at activity level, many delays are expected to cumulate every time two or more paths merge into a single successor.

However, mergers are much more frequent towards the end of the project compared to the earlier stages of execution. This is as, for any paths to close, they have to open first. Therefore, it is not a surprise that many construction projects get off to a good start (on time and on budget), but half way across their duration, (local) delays start being detected (whenever two or more paths are merged into one). As delays emerge, the cost of activities will also increase proportionally as the correlation between duration deviations and cost deviations was quite substantial on average. As a result, it is not that surprising that projects end later and cost more than initially anticipated.

Conclusions

The activity duration and cost variability of construction projects has been analyzed in this research by different types of projects (*building*, *civil engineering*, *industrial* and *services*), project phase (*planning*, *execution* and *closure*), and activity type (*auxiliary works, substructure*, *superstructure, specialized works,* and *facilities*). Correlation factors between activity duration deviations and activity cost deviations have also been studied under the same activity categories. The research is novel because it describes the first four moments (average, standard deviation, skewness and kurtosis) of how actual versus planned durations and costs differ at activity level in construction projects. A set of 101 projects and 5289 activities, plus another set with 746 projects have been used.

The first contribution of this study is providing construction managers with a first, yet rather complete, set of actual-vs-planned average activity durations and costs deviations with application in multiple contexts (project types, execution phases and types of activity). From now on, a construction manager will be able to more realistically (thus accurately) anticipate how likely and how much the activities in the project schedule will vary, that is, last or cost something different. This might potentially improve the quality and robustness of all construction schedules, for example allowing them to feed more advanced (non-deterministic) scheduling and simulation tools with more representative data. These techniques generally need a substantial amount of information from previous similar projects which is rarely available. With the set of moments provided here, these techniques will be able to resort to average values for their activity durations and cost distribution parameters depending on the type and/or execution phase of the project. These distributions will also be able to assume non-independence between the stochastically-generated activity durations and costs values (thanks to the set of duration-cost correlation values also published in this study).

This is expected to enhance future construction project monitoring and control, but also actual project duration and cost forecasting accuracy.

However, the analysis developed has also provided some interesting insights from its numerical perspective. One of the most relevant is that it has been shown that construction activities do not end late on average. Instead, it is their high level of variability (around 60% of its average duration) the key factor eventually causing project-level delays. Such high levels of activity variability exacerbate the merge event bias, a phenomenon by which whenever two or more schedule paths converge into a single one, the average completion times exceed the maximum average path durations.

Actual activity costs, on the other hand, do tend to be higher than what was planned (around 7%). This cannot be the result of price adjustments or inflation, as hardly any project lasted longer than a year. Instead, the major project-level cost overruns are expected to occur as a consequence of delayed start of activities located nearer the end of the project. This, as it has been demonstrated how most duration-cost correlation factors range within 0.40 and 0.70. The latter would cause that those activities that cannot start until their predecessors have finished, start incurring in costs before their actual execution.

Many other interpretations can arise from the numerical results of the four moments describing activity duration and cost variability that refer to specific types of projects, phases of execution or activity types that have not been recounted here. The reader is invited to refer to Tables 3 and 4 for such a purpose.

A limitation of this study is mostly connected to the composition and sample size of the construction projects analyzed. 101 projects have been used here with a varied composition. However, this sample size could have been bigger. It must be clarified, though, that accessing actual duration and cost information is ontologically questionable and certainly methodologically challenging. Companies are not open to share this information because it would clearly indicate how competent and efficient their operations are. Under that perspective, the current sample size probably seems satisfactory, at least for a first representative analysis.

A second limitation arises from having removed at the outset the Unplanned but Performed (UbP) and Planned but Unperformed (PbU) activities. This was necessary as the ratios (either in natural or log scale) converged to infinity causing a distortion in the moments calculation. However, we acknowledge that these activities can be found in almost all real projects. Frequently, they are the consequence of scope changes, works reorganization or changes in the available resources. Obviously, UbP and PbU activities add to the total project variability (beyond the activity duration and cost variability analysed here). In our analysis, though, there were only 279 UbP + 129 PbU = 408 activities out of the initial 5,697 (7% in total). Hence, while we believe the influence of UbP and PbU activities needs to be duly investigated, our analysis (with 93% of the activities) can still be considered representative enough to draw valid conclusions. Additionally, it is also expected that some degree of cancellation will occur among those 7% of activities (as frequently new activities replace others which are not eventually performed).

In the same vein, there are many potential future research continuations after this piece of research. Again, this study might be extended to analyze other types of projects and/or other more specific types of activities (maybe at trade-level: concrete, steel, asphalt, earthworks, etc.). The network topologies for other types of projects may also be studied to anticipate to what extent current levels of activity variability might impact their final schedules. The statistical distribution of activity (duration and cost) variability may also be analyzed. This was not possible at the general activity-level as discussed in this paper, but it should be possible for activities at their trade level.

A last conclusion derived from this research is that activity duration variability is the actual foe in project monitoring and control. This may not sound new to Lean Construction researchers and practitioners. However, this research has provided compelling empirical evidence suggesting that we do really need to start taking activity variability more seriously. There is a need to develop more techniques that can effectively handle/restrain this variability. Value stream mapping and Last planner have been some attempts to address this problem, but more are needed. This will open the door to new and more effective approaches for tackling the widespread phenomenon of construction projects ending late.

Acknowledgements

The first author acknowledges the Spanish Ministry of Science, Innovation and Universities for his Ramon y Cajal contract (RYC-2017-22222) co-funded by the European Social Fund. This work was also supported by the second author's "Estancias de movilidad en el extranjero José Castillejo para jóvenes doctores, 2017 (grant ref. CAS17/00488)" and the fourth author's "Estancias de profesores e investigadores senior en centros extranjeros Salvador de Madariaga 2018 (grant ref. PRX18/00381)", both from the Spanish Ministry of Science, Innovation and Universities. The first and last authors also acknowledge the help received by the research group TEP-955 from the PAIDI (Junta de Andalucía, Spain).

Data availability

All data generated or analyzed during the study are included in the submitted article or supplemental materials files.

References

AlSehaimi, A., and Koskela, L. (2008). “What Can be Learned from Studies on Delay in Construction?” *Proceedings for the 16th Annual Conference of the International Group for Lean Construction*, 95–106.

Ballesteros-Pérez, P. (2017a). “M-PERT: Manual Project-Duration Estimation Technique for Teaching Scheduling Basics.” *Journal of Construction Engineering and Management*, 143(9).

Ballesteros-Pérez, P. (2017b). “Modelling the boundaries of project fast-tracking.” *Automation in Construction*, 84(December), 231–241.

Ballesteros-Pérez, P., Campo-Hitschfeld, M. L., González-Naranjo, M. A., and González-Cruz, M. C. (2015). “Climate and construction delays: case study in Chile.” *Engineering, Construction and Architectural Management*, 22(6), 596–621.

Ballesteros-Pérez, P., Rojas-Céspedes, Y. A., Hughes, W., Kabiri, S., Pellicer, E., Mora-Melià, D., and del Campo-Hitschfeld, M. L. (2017). “Weather-wise. A weather-aware planning tool for improving construction productivity and dealing with claims.” *Automation in Construction*, 84(December), 81–95.

Ballesteros-Pérez, P., Larsen, G. D., and González-Cruz, M. C. (2018a). “Do projects really end late? On the shortcomings of the classical scheduling techniques.” *Journal of Technology and Science Education*, 8(1), 86–102.

Ballesteros-pérez, P., Smith, S. T., Lloyd-papworth, J. G., and Cooke, P. (2018b). “Incorporating the effect of weather in construction scheduling and management with sine wave curves : application in the United Kingdom.” *Construction Management and Economic*s, 36(12), 666–682.

Banerjee, A., and Paul, A. (2008). “On path correlation and PERT bias.” *European Journal of Operational Research*, 189(3), 1208–1216.

Batselier, J., and Vanhoucke, M. (2015). “Construction and evaluation framework for a real-life project database.” *International Journal of Project Management*, 33(3), 697–710.

Bein, W. W., Kamburowski, J., and Stallmann, M. F. M. (1992). “Optimal Reduction of Two-Terminal Directed Acyclic Graphs.” *SIAM Journal on Computing*, 21(6), 1112–1129.

Cho, S. (2009). “A linear Bayesian stochastic approximation to update project duration estimates.” *European Journal of Operational Research*, 196(2), 585–593.

Chudley, R., and Greeno, R. (2016). Building Construction Handbook. *Routledge*; 11th edition. ISBN-10: 113890709X

Davies, E. M. (1973). “An Experimental Investigation of Resource Allocation in Multiactivity Projects.” *Operational Research Quarterly (1970-1977), Journal of the Operational Research Society*, 24(4), 587.

Dodin, B., and Sirvanci, M. (1990). “Stochastic networks and the extreme value distribution.” *Computers & Operations Research*, 17(4), 397–409.

Elmaghraby, S. E. (2000). “On criticality and sensitivity in activity networks.” *European Journal of Operational Research*, 127(2), 220–238.

Elmaghraby, S. E., Fathi, Y., and Taner, M. R. (1999). “On the sensitivity of project variability to activity mean duration.” *International Journal of Production Economics*, 62(3), 219–232.

George, D., and Mallery, P. (2010). SPSS for Windows Step-by-Step: A Simple Guide and Reference, 17.0 Update. (Pearson, ed.), *Pearson*, Boston.

Hajdu, M., and Bokor, O. (2014). "The Effects of Different Activity Distributions on Project Duration in PERT Networks." *Procedia - Social and Behavioral Sciences*, 119, 766–775.

Hajdu, M., and Bokor, O. (2016). "Sensitivity analysis in PERT networks : Does activity duration distribution matter ?" *Automation in Construction*, Elsevier B.V., 65, 1–8.

Hamzah, N., Khoiry, M. A., Arshad, I., Tawil, N. M., and Che Ani, A. I. (2011). "Cause of Construction Delay - Theoretical Framework." *Procedia Engineering*, 20, 490–495.

Keane, P. J., and Caletka, A. F. (2008). Delay analysis in construction contracts. *Wiley-Blackwell*, Oxford, U.K.

Khamooshi, H., and Cioffi, D. F. (2013). "Uncertainty in Task Duration and Cost Estimates: Fusion of Probabilistic Forecasts and Deterministic Scheduling." *Journal of Construction Engineering and Management*, 139(5), 488–497.

Love, P. E. D., Lopez, R., Kim, J. T., and Kim, M. J. (2014). "Probabilistic Assessment of Design Error Costs." *Journal of Performance of Constructed Facilities*, 28(3), 518–527.

Love, P. E. D., and Sing, C.-P. (2013). "Determining the probability distribution of rework costs in construction and engineering projects." *Structure and Infrastructure Engineering*, 9(11), 1136–1148.

MacCrimmon, K. R., and Ryavec, C. A. (1964). "An Analytical Study of the PERT Assumptions." *Operations Research*, 12(1), 16–37.

Mahamid, I., Bruland, A., and Dmaidi, N. (2012). "Causes of Delay in Road Construction Projects." *Journal of Management in Engineering*, 28(3), 300–310.

Mastor, A. A. (1970). "An Experimental Investigation and Comparative Evaluation of Production Line Balancing Techniques." *Management Science*, 16(11), 728–746.

Ogunlana, S. O., Promkuntong, K., and Jearkjirm, V. (1996). "Construction delays in a fast-growing economy: Comparing Thailand with other economies." *International Journal of Project Management*, 14(1), 37–45.

OR-AS.be. (2018). "Real Project Data." Last accessed on 23th April 2019: http://www.projectmanagement.ugent.be/research/data/realdata

Orangi, A., Palaneeswaran, E., and Wilson, J. (2011). "Exploring Delays in Victoria-Based Astralian Pipeline Projects." *Procedia Engineering*, 14, 874–881.

Pascoe, T. (1966). "Allocation of resources: CPM." *Revue Française de Recherche Opérationnelle*, 38, 31–38.

Project Management Institute. (2017). PMBOK® Guide – Sixth Edition. Newtown Square, Pa. ISBN-10: 1628251840

Senouci, A., Ismail, A., and Eldin, N. (2016). "Time Delay and Cost Overrun in Qatari Public Construction Projects." *Procedia Engineering*, 164, 368–375.

Trietsch, D., Mazmanyan, L., Gevorgyan, L., and Baker, K. R. (2012). "Modeling activity times by the Parkinson distribution with a lognormal core: Theory and validation." *European Journal of Operational Research*, 216(2), 386–396.

Valadares Tavares, L., Antunes Ferreira, J., and Silva Coelho, J. (1999). "The risk of delay of a project in terms of the morphology of its network." *European Journal of Operational Research*, 119(2), 510–537.

Vanhoucke, M. (2008). “Measuring time using novel earned value management metrics.” *Proceedings of the 22nd IPMA World Congress* (Rome), vol 1, 99–103.

Vanhoucke, M. (2012). Project Management with Dynamic Scheduling. *Springer* Berlin Heidelberg, Berlin.

Vanhoucke, M., Coelho, J., and Batselier, J. (2016). “An Overview of Project Data for Integrated Project Management and Control.” *The Journal of Modern Project Management*, 3(3), 6–21.

Vanhoucke, M., and Vandevoorde, S. (2009). “Forecasting a project’s duration under various topological structures.” *The Measurable News*, (Spring), 26–30.

Project ID	Project Name	Project Type	Planned Cost (€)	Actual Cost (€)	Planned Dur. (d)	Actual Dur. (d)	Nº activ.	SP (%)	AD (%)	LA (%)	TF (%)
C2011-05	Telecom System Agnes	Service	180,485.27	180,485.27	43	53	20	60	58	38	9
C2011-07	Patient Transport System	Service	180,759.44	191,065.06	389	444	49	70	70	7	8
C2011-10	Building a House	Building	484,398.41	494,947.71	195	203	32	51	47	27	10
C2011-12	Claeys-Verhelst Premises	Building	3,027,133.19	3,102,395.91	443	453	49	41	50	5	43
C2011-13	Wind Farm	Civil Eng.	21,369,835.51	26,077,764.74	525	600	107	27	36	0	48
C2012-13	Pumping Station Jabbeke	Industrial	336,410.15	350,511.31	125	140	74	64	59	3	27
C2012-15	The Master Project	Service	185,472.45	185,113.10	32	32	121	17	66	0	84
C2012-17	Building a Dream	Building	241,015.00	314,856.14	145	204	33	65	61	35	19
C2013-01	Wiedauwkaai Fenders	Civil Eng.	1,069,532.42	1,314,584.58	152	152	39	48	45	0	68
C2013-02	Sewage Plant Hove	Civil Eng.	1,236,603.66	1,146,444.38	403	408	175	12	38	0	62
C2013-03	Brussels Finance Tower	Building	15,440,865.89	16,338,027.20	425	426	55	3	82	0	87
C2013-04	Kitchen Tower Anderlecht	Building	2,113,684.00	2,512,524.00	333	453	244	47	59	0	63
C2013-05	PET Packaging	Service	874,554.28	874,554.28	521	632	28	14	69	0	80
C2013-06	Govmnt. Office Building	Building	19,429,810.51	21,546,846.18	352	344	275	10	36	0	34
C2013-07	Family Residence	Building	180,476.47	175,030.65	170	174	46	40	44	3	25
C2013-08	Timber House	Building	501,029.51	576,624.05	216	235	41	29	42	0	47
C2013-09	Urban Develop.Project	Civil Eng.	1,537,398.51	1,696,971.79	291	360	71	34	51	6	16
C2013-10	Town Square	Civil Eng.	11,421,890.36	15,218,926.38	786	785	186	18	36	0	62
C2013-11	Recreation Complex	Building	5,480,518.91	5,451,028.00	359	277	159	27	44	0	32
C2013-12	Young Cattle Barn	Building	818,439.99	879,853.17	115	188	27	64	77	6	54
C2013-13	Office Finish. Works (1)	Building	1,118,496.59	955,929.22	236	217	11	20	49	33	6
C2013-14	Office Finish. Works (2)	Building	85,847.89	75,468.30	80	88	9	62	80	66	47
C2013-15	Office Finish. Works (3)	Building	341,468.11	308,343.78	171	115	17	25	43	21	35
C2013-16	Office Finish. Works (4)	Building	248,203.92	198,567.00	196	108	7	33	62	0	75
C2013-17	Office Finish. Works (5)	Building	244,205.40	203,605.97	161	107	23	36	38	20	32
C2014-01	Mixed-use Building	Building	38,697,822.73	39,777,643.30	474	448	41	50	38	3	49
C2014-02	Playing Cards	Industrial	191,492.70	190,266.50	124	146	21	81	94	0	14
C2014-03	Organizational Develop.	Service	43,170.15	83,712.15	229	260	112	9	31	0	36
C2014-04	Compres. Station Zelzate	Industrial	62,385,597.58	65,526,930.04	522	844	24	95	100	0	100
C2014-05	Apartment Building (1)	Building	532,410.29	591,410.53	228	274	25	58	71	35	18
C2014-06	Apartment Building (2)	Building	3,486,375.47	3,599,114.11	547	611	29	57	75	46	15
C2014-07	Apartment Building (3)	Building	1,102,536.78	1,289,696.78	353	404	25	58	71	35	18
C2014-08	Apartment Building (4)	Building	1,992,222.09	2,380,299.86	233	275	39	44	29	11	14
C2015-01	Young Cattle Barn (2)	Building	612,769.44	646,473.65	131	210	27	57	73	0	46
C2015-02	Railway Station (1)	Civil Eng.	1,121,316.94	967,988.79	417	501	216	8	66	1	80
C2015-03	Industrial Complex (1)	Building	2,244,090.74	1,868,796.28	257	278	135	16	43	0	58
C2015-04	Apartment Building (5)	Building	2,750,938.00	2,590,796.73	160	205	56	27	37	0	57
C2015-06	Family Residence (2)	Building	143,673.20	186,107.00	260	290	184	18	0	30	38
C2015-07	Industrial Complex (2)	Building	5,999,600.00	5,414,544.00	297	313	138	27	38	0	49
C2015-08	Garden Center	Building	467,297.21	461,900.17	191	186	186	14	52	0	79
C2015-09	Railway Station (2)	Civil Eng.	1,457,424.00	2,145,682.26	354	569	340	4	48	0	75
C2015-10	Tax Return System (1)	Service	18,990.00	8,010.00	85	85	15	10	82	23	21
C2015-11	Staff Authoriz. System	Service	14,400.00	9,105.00	55	55	7	25	66	0	52
C2015-12	Premium Payment System	Service	132,570.00	58,410.00	184	184	35	19	63	9	61
C2015-13	Broker Acc.Conv. System	Service	12,735.00	9,990.00	117	117	16	19	60	7	51
C2015-14	Sup. Pensions Database	Service	34,260.00	18,285.00	124	124	17	17	55	3	50
C2015-15	FACTA System	Service	11,700.00	7,035.00	57	57	13	22	57	8	18
C2015-16	Generic Doc. Output Syst.	Service	64,620.00	64,125.00	270	270	22	10	61	12	26
C2015-17	Insurance Bundling Syst.	Service	281,430.00	281,070.00	208	236	86	6	77	8	41
C2015-18	Tax Return System (2)	Service	39,450.00	25,380.00	128	128	15	10	66	16	11
C2015-19	Receipt Numb. System	Service	43,800.00	37,530.00	182	182	20	21	46	8	31
C2015-20	Policy Numbering System	Service	12,645.00	11,100.00	171	161	6	20	62	20	13

C2015-21	Investment Product (1)	Service	4,020.00	3,240.00	37	37	12	18	35	2	36
C2015-22	Risk Profile Questionnaire	Service	29,880.00	17,400.00	151	151	22	16	70	9	40
C2015-23	Investment Product (2)	Industrial	46,920.00	32,805.00	122	120	33	17	53	5	39
C2015-24	CRM System	Service	44,130.00	36,870.00	233	233	21	7	59	7	29
C2015-25	Beer Tasting	Service	1,210.00	1,780.00	14	14	18	16	40	21	19
C2015-26	Debt Collection System	Service	458,112.37	512,546.15	148	154	214	9	43	0	61
C2015-27	Railway Station Antwerp	Building	22,703.52	25,313.12	68	81	18	23	40	-2	54
C2015-28	Web. Tennis Vlaanderen	Service	219,275.00	382,475.00	201	212	20	15	54	0	67
C2015-29	Fire Station	Building	1,874,496.82	1,887,087.25	284	298	204	48	34	0	41
C2015-30	Social Apts. Ypres (1)	Building	440,940.89	440,940.89	244	254	40	25	51	-1	76
C2015-31	Social Apts Ypres (2)	Building	1,310,723.46	1,282,185.98	271	364	29	32	49	23	43
C2015-32	Social Apts Ypres (3)	Building	2,509,031.42	2,509,031.42	358	265	48	38	63	3	59
C2015-33	IJzertoren Memor. Square	Civil Eng.	214,417.71	224,789.67	50	94	12	63	57	0	14
C2015-34	Roadworks Poperinge	Civil Eng.	511,325.86	440,394.16	120	193	13	91	99	0	18
C2015-35	Retirement Apartments	Building	14,956,314.25	16,068,878.30	850	951	11	48	57	21	35
C2016-01	Railway Bridge (1)	Civil Eng.	671,383.50	703,703.50	225	274	26	51	71	0	86
C2016-02	Railway Bridge (2)	Civil Eng.	962,181.56	972,341.56	229	239	23	63	71	0	82
C2016-03	Railway Bridge (3)	Civil Eng.	926,888.01	910,728.01	203	220	25	16	37	0	56
C2016-04	Railway Bridge (4)	Civil Eng.	906,253.87	906,253.87	248	242	26	64	62	0	71
C2016-05	Railway Bridge (5)	Civil Eng.	832,497.46	832,497.46	195	197	32	77	74	0	51
C2016-06	Defense Building	Service	4,331,260.49	4,331,260.49	252	232	96	14	55	0	76
C2016-07	Shop. Village Walkways	Civil Eng.	930,179.09	932,757.25	224	316	110	95	98	0	99
C2016-08	SCM System	Service	375,253.34	438,741.66	725	725	99	49	59	8	52
C2016-09	Data Loss Prevent. System	Service	584,951.77	1,425,155.96	195	189	113	10	36	1	51
C2016-10	Biofuel Refinery	Industrial	14,362,625.00	14,466,100.00	360	375	23	18	22	6	21
C2016-11	Residential House (1)	Building	162,472.00	163,189.00	241	254	55	57	77	52	16
C2016-12	Residential House (2)	Building	222,858.00	226,285.00	291	291	59	56	72	50	19
C2016-13	Residential House (3)	Building	367,952.00	379,300.00	306	330	51	64	81	54	14
C2016-14	Residential House (4)	Building	218,366.00	222,021.78	321	320	48	68	78	42	10
C2016-15	Resid. House Struct. Work	Building	95,694.00	100,763.00	126	130	13	66	75	100	0
C2016-16	Resid. Finish. Works (1)	Building	54,577.76	64,526.76	90	90	24	69	68	50	28
C2016-17	Resid. Finish. Works (2)	Building	54,703.17	64,580.17	86	86	24	69	68	50	28
C2016-18	Resid. Finish. Works (3)	Building	51,115.52	60,829.52	91	91	25	66	62	27	31
C2016-19	Resid. Finish. Works (4)	Building	51,303.38	53,351.38	91	91	25	66	62	27	31
C2016-20	Resid. Finish. Works (5)	Building	52,021.28	53,783.28	91	91	25	66	62	27	31
C2016-21	Resid. Finish. Works (6)	Building	54,324.22	54,996.22	101	101	24	69	68	50	28
C2016-22	Resid. Finish. Works (7)	Building	56,969.40	57,822.40	101	101	24	69	68	50	28
C2016-23	Resid. Finish. Works (8)	Building	56,182.71	56,645.71	101	101	24	69	68	50	28
C2016-24	Resid. Finish. Works (9)	Building	52,262.83	53,176.83	101	101	24	69	68	50	28
C2016-25	Resid. Finish. Works (10)	Building	54,580.33	56,748.33	91	91	24	69	68	50	28
C2016-26	Resid. Finish. Works (11)	Building	51,286.24	53,319.24	91	91	24	69	68	50	28
C2016-27	Apt. Build. Foundat. (1)	Building	813,663.06	879,701.06	78	88	16	66	59	0	48
C2016-28	Apt. Struct. Work (1)	Building	569,177.85	586,086.85	71	79	19	55	29	0	30
C2016-29	Apt. Struct. Work (2)	Building	1,797,873.62	1,860,330.62	129	148	19	72	69	0	35
C2016-30	Apt. Struct. Work (3)	Building	1,319,736.29	1,353,361.29	85	96	23	81	83	0	31
C2016-31	Apt. Struct. Work (1)	Building	488,936.00	498,473.00	105	117	23	31	40	0	11
C2016-32	Apt. Struct. Work (2)	Building	477,381.00	496,991.00	89	97	22	52	72	0	27
C2016-33	Apt. Struct. Work (3)	Building	377,282.00	394,829.00	116	129	23	50	72	0	30
C2016-34	Apt. Struct. Work (4)	Building	362,476.00	383,871.00	83	92	23	40	43	0	26
		Avg.	2,647,861.81	2,837,446.83	221	240	Σ=5,697	41.0	58.2	14.1	40.5

Table 1. First projects dataset summary

Level 0 (All activities*)

Nº activities		
Planned & Performed	Unplanned but Performed	Planned but not performed
5289	279	129

Level 1 (by Project type*)

Project Type	n	**Nº activities**			Actual Cost (10^3€)		Actual Dur. (days)	
		Planned & Performed	Unplanned but Performed	Planned but not performed	Avg.	SD	Avg.	SD
Building	56	2894	18	12	48.88	267.20	11.35	29.78
Civil Eng.	15	1092	250	59	40.43	161.26	12.92	15.92
Industrial	5	170	0	5	473.92	1225.85	21.60	48.60
Services	25	1133	11	53	8.03	31.15	11.13	31.48
Sum	101	5289	279	129				

Level 2 (by Project phase*)

	Nº activities							
Proj. type >	**Building**		**Civil Engineering**		**Industrial**		**Service**	
Project phase v	Plan. & Perform.	Unplan. but Perform.	Plan. & Perform.	Unplan. but Perform.	Plan. & Perform.	Unplan. but Perform.	Plan. & Perform.	Unplan. but Perform.
Planning	49	0	38	0	10	0	81	0
Execution	2810	18	1034	250	154	0	990	11
Closure	35	0	20	0	6	0	62	0

Level 3 (by Activity type *&**)

	Nº activities					
Project type >	**Building**			**Civil Engineering**		
Activity type v	Planned & Performed	Unplan. but Performed	Planned but not perform.	Planned & Performed	Unplan. but Performed	Planned but not perform.
Auxiliary works	139	1	0	207	27	9
Substructure	171	2	0	229	11	4
Superstructure	654	1	0	257	104	20
Specialized works	1272	11	10	264	88	25
Facilities	574	3	2	77	20	1

*Only Planned & Performed activities are used for later analyses

** Only for 'Execution' activities from Building and Civil Engineering projects

Table 2. Summary of activities analyzed

Level 0 (All Activities)					**Level 1** (by Project Type)						**Level 2** (by Project phase)						**Level 3** (by Activity type)					
n	Avg	SD	Skew	Kurt	Type	n	Avg	SD	Skew	Kurt	Phase	n	Avg	SD	Skew	Kurt	Type	n	Avg	SD	Skew	Kurt
5289	0.010 (1.023)	0.19 (1.56)	0.91	9.90	**Building**	2894	0.004 (1.009)	0.15 (1.43)	-0.36	9.88	**Planning**	49	0.035 (1.083)	0.21 (1.62)	1.51	8.13	*(insufficient data sample)*					
											Execution	2810	0.003 (1.007)	0.15 (1.42)	-0.46	9.92	**Auxiliary Works**	139	0.017 (1.040)	0.16 (1.46)	0.36	5.54
																	Substructure	171	0.035 (1.083)	0.14 (1.39)	1.89	8.38
																	Superstructure	654	-0.018 (0.960)	0.16 (1.45)	-0.98	9.60
																	Specialized Works	1272	0.004 (1.010)	0.15 (1.42)	-0.87	10.02
																	Facilities	574	0.011 (1.026)	0.14 (1.39)	0.53	11.71
											Closure	35	0.022 (1.052)	0.16 (1.45)	1.33	3.91	*(insufficient data sample)*					
					Civil Eng.	1092	-0.008 (0.982)	0.20 (1.58)	0.53	9.61	**Planning**	38	0.052 (1.126)	0.18 (1.53)	2.38	12.88	*(insufficient data sample)*					
											Execution	1034	-0.010 (0.977)	0.20 (1.58)	0.49	9.36	**Auxiliary Works**	207	0.013 (1.030)	0.13 (1.36)	1.75	9.43
																	Substructure	229	-0.005 (0.990)	0.18 (1.53)	-0.39	8.57
																	Superstructure	257	-0.030 (0.934)	0.20 (1.57)	0.90	10.73
																	Specialized Works	264	-0.012 (0.972)	0.24 (1.72)	0.28	5.81
																	Facilities	77	-0.018 (0.959)	0.26 (1.82)	1.09	10.65
											Closure	20	-0.011 (0.975)	0.05 (1.12)	-4.47	20.00	*(insufficient data sample)*					
					Industrial	170	-0.010 (0.977)	0.22 (1.65)	-0.76	3.37	**Planning**	10	0.001 (1.003)	0.05 (1.12)	0.43	4.59	*(insufficient data sample)*					
											Execution	154	-0.009 (0.981)	0.23 (1.68)	-0.76	3.17	*(insufficient data sample)*					
											Closure	6	-0.090 (0.813)	0.25 (1.77)	0.04	0.81	*(insufficient data sample)*					
					Services	1133	0.045 (1.110)	0.26 (1.83)	1.53	5.87	**Planning**	81	0.055 (1.134)	0.22 (1.67)	1.05	3.55	*(insufficient data sample)*					
											Execution	990	0.048 (1.118)	0.27 (1.86)	1.58	5.70	*(insufficient data sample)*					
											Closure	62	-0.014 (0.969)	0.19 (1.54)	-1.36	7.74	*(insufficient data sample)*					

Table 3. Activity actual/planned duration log ratios (natural values stated between parentheses)

Level 0 (All Activities)					**Level 1** (by Project Type)						**Level 2** (by Project phase)						**Level 3** (by Activity type)					
n	Avg	SD	Skew	Kurt	Type	n	Avg	SD	Skew	Kurt	Phase	n	Avg	SD	Skew	Kurt	Type	n	Avg	SD	Skew	Kurt
5289	0.031 (1.074)	0.25 (1.78)	2.49	15.56	**Building**	2894	0.015 (1.035)	0.16 (1.46)	2.02	25.27	**Planning**	49	-0.002 (0.996)	0.19 (1.56)	-1.66	10.45	*(insufficient data sample)*					
											Execution	2810	0.015 (1.035)	0.16 (1.46)	2.12	25.91	**Aux. Works**	139	0.027 (1.065)	0.11 (1.29)	2.73	14.82
																	Substruct.	171	0.014 (1.034)	0.10 (1.26)	-0.21	8.37
																	Superstruct.	654	0.010 (1.023)	0.10 (1.26)	0.01	10.14
																	Spec. Works	1272	0.014 (1.034)	0.21 (1.62)	2.12	18.98
																	Facilities	574	0.020 (1.046)	0.12 (1.33)	0.53	13.85
											Closure	35	0.041 (1.098)	0.16 (1.43)	2.01	6.60	*(insufficient data sample)*					
					Civil Eng.	1092	0.057 (1.139)	0.30 (2.01)	1.78	6.28	**Planning**	38	0.322 (2.099)	0.43 (2.71)	0.99	-0.75	*(insufficient data sample)*					
											Execution	1034	0.048 (1.116)	0.30 (1.98)	1.77	6.87	**Aux. Works**	207	0.059 (1.147)	0.32 (2.08)	2.89	11.31
																	Substruct.	229	0.057 (1.140)	0.30 (1.98)	0.63	2.26
																	Superstruct.	257	0.057 (1.141)	0.32 (2.11)	1.40	3.11
																	Spec. Works	264	0.016 (1.038)	0.24 (1.74)	1.93	13.70
																	Facilities	77	0.067 (1.166)	0.31 (2.04)	2.03	7.66
											Closure	20	0.011 (1.026)	0.01 (1.02)	-0.95	-1.24	*(insufficient data sample)*					
					Industrial	170	-0.011 (0.975)	0.20 (1.59)	-2.05	11.12	**Planning**	10	0.02 (1.046)	0.05 (1.03)	0.74	0.71	*(insufficient data sample)*					
											Execution	154	-0.004 (0.99)	0.18 (1.50)	-1.65	12.65	*(insufficient data sample)*					
											Closure	6	-0.27 (0.536)	0.67 (4.63)	-0.12	-2.71	*(insufficient data sample)*					
					Services	1133	0.052 (1.128)	0.36 (2.29)	2.25	9.01	**Planning**	81	0.021 (1.05)	0.18 (1.53)	0.40	3.89	*(insufficient data sample)*					
											Execution	990	0.059 (1.145)	0.37 (2.37)	2.23	8.27	*(insufficient data sample)*					
											Closure	62	-0.007 (0.985)	0.29 (1.93)	1.28	13.16	*(insufficient data sample)*					

Table 4. Activity actual/planned costs log ratios (natural values stated between parentheses)

Group of analysis		**n**	**R**	**R^2**	**Slope**	**Intercept**
Activity-level *(duration-cost correlations)*						
Level 0	All activities	5289	0.55	0.30**	0.704**	0.024
Level 1	Building	2894	0.46	0.21**	0.488**	0.013
	Civil Engineering	1092	0.33	0.11**	0.502**	0.061
	Industrial	170	0.11	0.01	0.106	-0.010
	Services	1133	0.79	0.62**	1.074	0.004
Level 2	Planning	178	0.38	0.15**	0.534	0.055
	Execution	4988	0.55	0.30**	0.706**	0.024
	Closure	123	0.60	0.36**	0.534	0.055
Level3	Auxiliary Works	349	0.34	0.12**	0.601*	0.037
	Substructure	400	0.25	0.06**	0.343*	0.035
	Superstructure	912	0.32	0.10**	0.289**	-0.028
	Specialized Works	1609	0.44	0.20**	0.566*	0.013
	Facilities	654	0.53	0.28**	0.522**	0.021
Project-level *(duration-cost correlations)*						
	All Projects	101	0.22	0.05*	0.156*	0.029
	Building	56	0.52	0.27	0.957	0.006
	Civil Engineering	15	0.01	0.00	0.017	0.080
	Industrial	5	0.56	0.31	0.629	0.083
	Services	25	0.23	0.05	0.039	0.014
	Road projects (Figure 1)	746	0.34	0.11	0.108	0.016

******Snedecor's F test (for R^2) or student's T test (for slopes) significant at $\alpha<0.001$

***** Snedecor's F test (for R^2) or student's T test (for slopes) significant at $\alpha<0.05$

Table 5. Duration vs Cost (log ratios) linear correlations

Project type	**n**	**SP (%)**	**AD (%)**	**LA (%)**	**TF (%)**
Building	56	48.2	57.4	21.4	35.2
Civil Eng.	15	44.7	59.3	0.5	44.7
Industrial	5	55.0	65.6	2.8	55.0
Service	25	20.1	57.6	8.3	20.1
All	101	41.0	58.2	14.1	40.5

Table 6. Average network topological values by project type

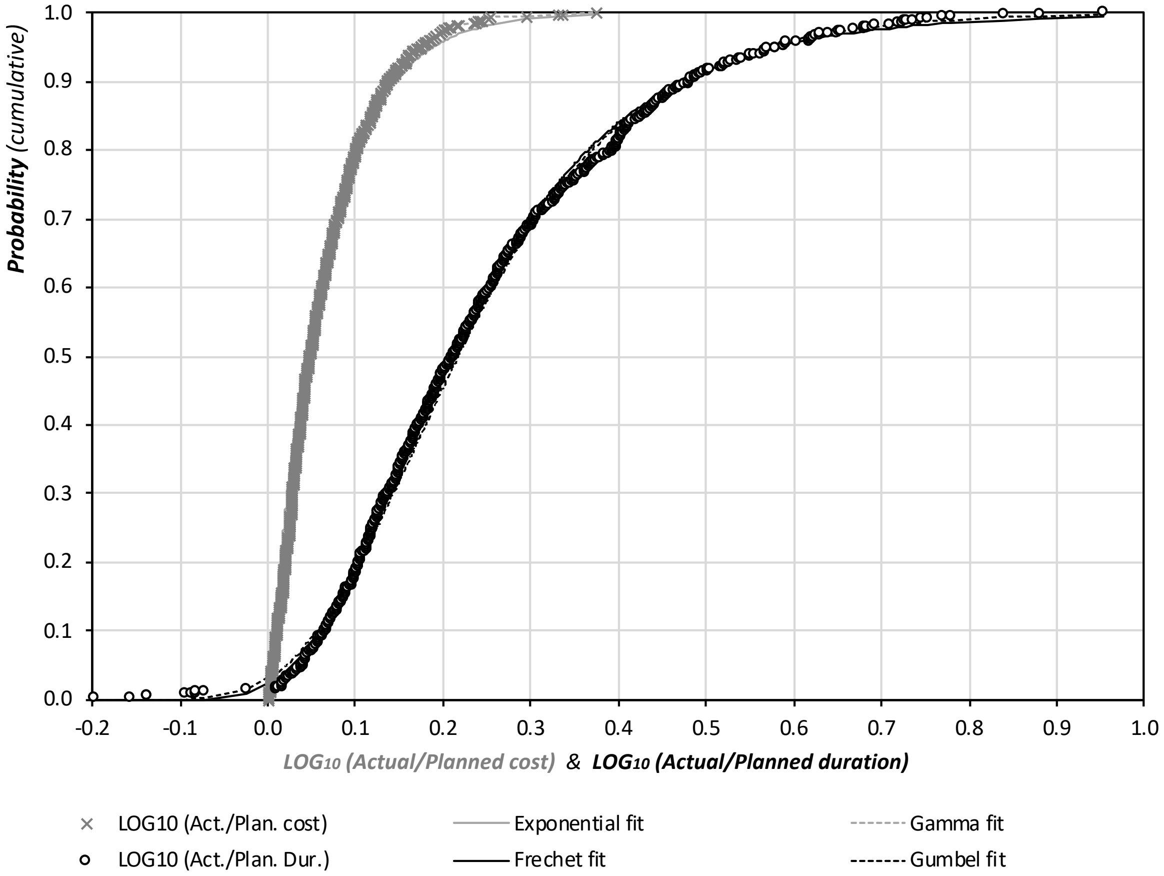

Fig 1. Duration and Cost overrun probability distribution of 746 road construction projects from the Florida Department of Transportation

LIST OF FIGURE CAPTIONS

Figure 1. Duration and Cost overrun probability distribution of 746 road construction projects from the Florida Department of Transportation

Permissions

One new co-author was added in this revised manuscript (Dr Alberto Cerezo-Narváez).

I, **Dr Pablo Ballesteros-Perez**, as co-author and corresponding authors of the manuscript "On the duration and cost variability of construction activities: an empirical study" (COENG-8211R1) submitted to Journal of Construction Engineering and Management (ASCE), acknowledge that Dr Alberto Cerezo-Narváez has contributed to the revised version of this manuscript and I agree to include him as co-author.

The **rest of the co-authors' statements (including the new co-authors') are appended over the next pages** (copy from emails) as instructed in your email.

Thank you and sorry for the inconvenience,

Dr Pablo Ballesteros-Pérez

Subject: New co-author Permission in COENG-8211R1 - from enocsanz_unileon.es
Date: Saturday, 27 April 2019 at 20:12:25 Central European Summer Time
From: Enoc Sanz Ablanedo
To: Pablo Ballesteros Perez

Hi Pablo, I send you the statement about including Alberto as a new author. Kind regards. Enoc.

I, Dr Enoc Sanz-Ablanedo, as co-author of the manuscript “On the duration and cost variability of construction activities: an empirical study” (COENG-8211R1) submitted to Journal of Construction Engineering and Management (ASCE), acknowledge that Dr Alberto Cerezo-Narváez has contributed to the revised version of this manuscript and I agree to include him as co-author.

--

Enoc Sanz Ablanedo
Dpto. Tecnología Minera Topográfica y Estructuras
Universidad de León
Avda. Astorga s/n
24400 Ponferrada
LEON

T:987442110, F:987442070, M:605874159
enocsanz@gmail.com, esana@unileon.es
http://enocsanzablanedo.unileon.es

Subject: New co-author Permission in COENG-8211R1
Date: Saturday, 27 April 2019 at 12:16:06 Central European Summer Time
From: Robby Soetanto
To: Pablo Ballesteros Perez

Dear Corresponding Author,

I, Dr Robby Soetanto, as co-author of the manuscript “On the duration and cost variability of construction activities: an empirical study” (COENG-8211R1) submitted to Journal of Construction Engineering and Management (ASCE), acknowledge that Dr Alberto Cerezo-Narváez has contributed to the revised version of this manuscript and I agree to include him as co-author.

Yours faithfully,
Robby

Dr Robby Soetanto
Senior Lecturer
School of Architecture, Building and Civil Engineering
Loughborough University
Leicestershire, LE 11 3TU
United Kingdom

Goggle Scholar; @soetanto_robby
BIM-Hub: BuGIS, @resilientBuGIS, UKIREN
Recent publications:
Lean approach in precast concrete component production
The perceptions of social responsibility for community resilience to flooding
Online Learning for STEM Subjects

Subject: New co-author Permission in COENG-8211R1
Date: Friday, 26 April 2019 at 22:36:11 Central European Summer Time
From: María Del Carmen Gonzalez Cruz
To: pablo.ballesteros@uca.es

I, Dra. M. Carmen González-Cruz, as co-author of the manuscript "On the duration and cost variability of construction activities: an empirical study" (COENG-8211R1) submitted to Journal of Construction Engineering and Management (ASCE), acknowledge that Dr Alberto Cerezo-Narváez has contributed to the revised version of this manuscript and I agree to include him as co-author.

Best Regards,
M.Carmen

Subject: “New co-author Permission in COENG-8211R1”
Date: Monday, 29 April 2019 at 12:04:30 Central European Summer Time
From: Graeme Larsen
To: Pablo Ballesteros Perez
CC: Graeme Larsen
Priority: High
Attachments: image005.jpg, image006.gif, image007.gif, image008.gif, image009.gif

To whom is may concern,

I, Dr Graeme D. Larsen, as co-author of the manuscript “On the duration and cost variability of construction activities: an empirical study” (COENG-8211R1) submitted to Journal of Construction Engineering and Management (ASCE), acknowledge that Dr Alberto Cerezo-Narváez has contributed to the revised version of this manuscript and I agree to include him as co-author.

Kind regards
Graeme

Dr Graeme D. Larsen FCIOB FHEA
Associate Professor
Research Group Lead: Organization, People and Technology
0118 378 7185

School of the Built Environment
Architecture | Construction | Energy | Environmental Engineering
University of Reading, PO Box 219, Reading, RG6 6AW
http://www.reading.ac.uk/CME/about/staff/g-d-larsen.aspx

For freedom of information requests please contact imps@reading.ac.uk

---

Coordinator of W065 Organization and Management of Construction

International Council
for Research and Innovation
in Building and Construction

From: Pablo Ballesteros Perez [mailto:pablo.ballesteros@uca.es]
Sent: 29 April 2019 09:19
To: Graeme Larsen
Subject: Permission (urgent)
Importance: High

Hi Graeme,

Subject: Contribution to COENG-8211R1
Date: Saturday, 27 April 2019 at 00:51:45 Central European Summer Time
From: Alberto Cerezo Narvaez
To: Pablo Ballesteros-Pérez

I, Dr Alberto Cerezo-Narváez, have been working on the revised version of the manuscript "On the duration and cost variability of construction activities: an empirical study" (COENG-8211R1) submitted to Journal of Construction Engineering and Management (ASCE). I agree to be listed as co-author when this paper is resubmitted for re-revision.--

UCA

Dr. Alberto Cerezo Narváez
Project Manager Professional (PMI-PMP)
Profesional en Dirección de Proyectos (IPMA Nivel C)
Área de Proyectos de Ingeniería
Departamento de Ingeniería Mecánica y Diseño Industrial
Universidad de Cádiz
Avenida de la Universidad de Cádiz, Nº 10 - 11519
Puerto Real (Cádiz)
alberto.cerezo@uca.es
Página Web http://orcid.org/0000-0003-1404-5771

ASCE Authorship, Originality, and Copyright Transfer Agreement

Publication Title: Journal of Construction Engineering and Management

Manuscript Title: On the duration and cost variability of construction activities: an empirical study

Author(s) – **Names, postal addresses, and e-mail addresses of all authors**

Pablo Ballesteros-Pérez, & Alberto Carezo-Narváez; Univ. de Cádiz (Spain) ; Avda. Univ. de Cádiz 1
Enoc Sanz-Ablanedo; Univ. de León (Spain); Avda. Astorga, s/n. 24400 Ponferrada, León; esana@ur
Robby Soetanto; Loughborough University (UK); Loughborough, LE11 6DF; r.soetanto@lboro.ac.uk
Mª Carmen González-Cruz; Univ. Politècnica de València (Spain); Camino de Vera s/n, 46022 Valenc
Graeme Larsen; Univ. of Reading (UK); Whiteknights, Reading RG6 6DF; g.d.larsen@reading.ac.uk

I. Authorship Responsibility

To protect the integrity of authorship, only people who have significantly contributed to the research or project and manuscript preparation shall be listed as coauthors. The corresponding author attests to the fact that anyone named as a coauthor has seen the final version of the manuscript and has agreed to its submission for publication. Deceased persons who meet the criteria for coauthorship shall be included, with a footnote reporting date of death. No fictitious name shall be given as an author or coauthor. An author who submits a manuscript for publication accepts responsibility for having properly included all, and only, qualified coauthors.

I, the corresponding author, confirm that the authors listed on the manuscript are aware of their authorship status and qualify to be authors on the manuscript according to the guidelines above.

Dr. P. Ballesteros-Pérez (26/April/2019) 26/4/19
Print Name Signature Date

II. Originality of Content

ASCE respects the copyright ownership of other publishers. ASCE requires authors to obtain permission from the copyright holder to reproduce any material that (1) they did not create themselves and/or (2) has been previously published, to include the authors' own work for which copyright was transferred to an entity other than ASCE. Each author has a responsibility to identify materials that require permission by including a citation in the figure or table caption or in extracted text. Materials re-used from an open access repository or in the public domain must still include a citation and URL, if applicable. At the time of submission, authors must provide verification that the copyright owner will permit re-use by a commercial publisher in print and electronic forms with worldwide distribution. For Conference Proceeding manuscripts submitted through the ASCE online submission system, authors are asked to verify that they have permission to re-use content where applicable. Written permissions are not required at submission but must be provided to ASCE if requested. Regardless of acceptance, no manuscript or part of a manuscript will be published by ASCE without proper verification of all necessary permissions to re-use. ASCE accepts no responsibility for verifying permissions provided by the author. Any breach of copyright will result in retraction of the published manuscript.

I, the corresponding author, confirm that all of the content, figures (drawings, charts, photographs, etc.), and tables in the submitted work are either original work created by the authors listed on the manuscript or work for which permission to re-use has been obtained from the creator. For any figures, tables, or text blocks exceeding 100 words from a journal article or 500 words from a book, written permission from the copyright holder has been obtained and supplied with the submission.

Dr. P. Ballesteros-Pérez (26/April/2019) 26/4/19
Print name Signature Date

III. Copyright Transfer

ASCE requires that authors or their agents assign copyright to ASCE for all original content published by ASCE. The author(s) warrant(s) that the above-cited manuscript is the original work of the author(s) and has never been published in its present form.

The undersigned, with the consent of all authors, hereby transfers, to the extent that there is copyright to be transferred, the exclusive copyright interest in the above-cited manuscript (subsequently called the "work") in this and all subsequent editions of the work (to include closures and errata), and in derivatives, translations, or ancillaries, in English and in foreign translations, in all formats and media of expression now known or later developed, including electronic, to the American Society of Civil Engineers subject to the following:

- The undersigned author and all coauthors retain the right to revise, adapt, prepare derivative works, present orally, or distribute the work, provided that all such use is for the personal noncommercial benefit of the author(s) and is consistent with any prior contractual agreement between the undersigned and/or coauthors and their employer(s).
- No proprietary right other than copyright is claimed by ASCE.
- If the manuscript is not accepted for publication by ASCE or is withdrawn by the author prior to publication (online or in print), this transfer will be null and void.
- Authors may post a PDF of the ASCE-published version of their work on their employers' ***Intranet*** with password protection. The following statement must appear with the work: "This material may be downloaded for personal use only. Any other use requires prior permission of the American Society of Civil Engineers."
- Authors may post the ***final draft*** of their work on open, unrestricted Internet sites or deposit it in an institutional repository when the draft contains a link to the published version at www.ascelibrary.org. "Final draft" means the version submitted to ASCE after peer review and prior to copyediting or other ASCE production activities; it does not include the copyedited version, the page proof, a PDF, or full-text HTML of the published version.

Exceptions to the Copyright Transfer policy exist in the following circumstances. Check the appropriate box below to indicate whether you are claiming an exception:

☐ **U.S. GOVERNMENT EMPLOYEES:** Work prepared by U.S. Government employees in their official capacities is not subject to copyright in the United States. Such authors must place their work in the public domain, meaning that it can be freely copied, republished, or redistributed. In order for the work to be placed in the public domain, ALL AUTHORS must be official U.S. Government employees. If at least one author is not a U.S. Government employee, copyright must be transferred to ASCE by that author.

☐ **CROWN GOVERNMENT COPYRIGHT:** Whereby a work is prepared by officers of the Crown Government in their official capacities, the Crown Government reserves its own copyright under national law. If ALL AUTHORS on the manuscript are Crown Government employees, copyright cannot be transferred to ASCE; however, ASCE is given the following nonexclusive rights: (1) to use, print, and/or publish in any language and any format, print and electronic, the above-mentioned work or any part thereof, provided that the name of the author and the Crown Government affiliation is clearly indicated; (2) to grant the same rights to others to print or publish the work; and (3) to collect royalty fees. ALL AUTHORS must be official Crown Government employees in order to claim this exemption in its entirety. If at least one author is not a Crown Government employee, copyright must be transferred to ASCE by that author.

☐ **WORK-FOR-HIRE:** Privately employed authors who have prepared works in their official capacity as employees must also transfer copyright to ASCE; however, their employer retains the rights to revise, adapt, prepare derivative works, publish, reprint, reproduce, and distribute the work provided that such use is for the promotion of its business enterprise and does not imply the endorsement of ASCE. In this instance, an authorized agent from the authors' employer must sign the form below.

☐ **U.S. GOVERNMENT CONTRACTORS:** Work prepared by authors under a contract for the U.S. Government (e.g., U.S. Government labs) may or may not be subject to copyright transfer. Authors must refer to their contractor agreement. For works that qualify as U.S. Government works by a contractor, ASCE acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce this work for U.S. Government purposes only. This policy DOES NOT apply to work created with U.S. Government grants.

I, the corresponding author, acting with consent of all authors listed on the manuscript, hereby transfer copyright or claim exemption to transfer copyright of the work as indicated above to the American Society of Civil Engineers.

Dr. Pablo Ballesteros-Pérez

Print Name of Author or Agent

26th April 2019

Signature of Author of Agent

Date

More information regarding the policies of ASCE can be found at http://www.asce.org/authorsandeditors

RESPONSE TO THE EDITOR'S AND REVIEWERS' COMMENTS

Journal: Journal of Construction Engineering and Management (ASCE)
Manuscript Ref.: COENG-8211
Title: On the duration and cost variability of construction activities: an empirical study

The authors wish to thank the Editor and Reviewers for their time and effort in reviewing our manuscript, as well as for their suggestions and observations. We have addressed all their comments in the following responses (item-by-item) providing detailed explanations, as well as the consequent modifications of the paper. We are confident that the paper has improved after your suggestions.

EDITOR

Point 0.1. Please spell out all acronyms at their first mention in the text. (PERT)

We thank the Editor for providing us with an opportunity to improve our manuscript. All acronyms have been spelled out the first time they have been used. Particularly, the word PERT has been spelled out on page 2 (Program Evaluation and Review Technique).

Point 0.2. Footnotes/Endnotes. Please remove all footnotes and endnotes (except those used in identifying author affiliation) and incorporate them into the text of your manuscript.

All footnotes, but the one with the authors' affiliations, have been incorporated into the text as requested by the Editor.

Point 0.3. Please sign and upload a copy of our Copyright Transfer Agreement in ink or with a verified Adobe signature. The CTA can be found at: https://ascelibrary.org/pb-assets/images/CUSTOM%20PAGES/FILES/Revision%20ASCE%20Authorship%20originality%20and%20CTA%20form_09282017-1551711012447.pdf

The CTA has been filled out, signed, scanned and uploaded along with the other manuscript files.

Point 0.4. Also, please note in order to clarify math for copyeditors, please ensure that you use boldface for matrixes, vectors, tensors; italics for all variables, including variables that are subscript and superscript; roman for all numerals and Greek characters, and mathematical operators; and Helvetica for all dimensionless numbers (Froude, Weber, Prandtl, etc.).

There are currently two equations and they only contain variables. They have all been represented in italics as requested.

Point 0.5. The journal is requiring a clear explanation of the primary contributions this research makes to the Body of Knowledge. These claims need to be present in the Abstract and Conclusion of the manuscript and function as a summary of the unique value the work contributes to the construction engineering and management global community. If not already present, please include these statements in your revised manuscript.

The contributions to the body of knowledge were already present in the Abstract and Conclusions of the original manuscript. However, after some reviewers' comments, they have now been emphasized and extended even more, particularly in the *Conclusions* section.

REVIEWER 1

Point 1.0. The paper examines the difference between estimated and actual data for activity duration and cost.

The authors appreciate the comments from Reviewer 1. We have revised the paper and/or provided appropriate rebuttal reasons as described in the points below.

Point 1.1. In overall, the paper does not provide a solid novel research question and its importance. The whole paper is all about the statistics of the difference between estimate and actual data.

We disagree with both statements. Regarding, the second, we concede that there is a significant proportion of the paper devoted to measuring the differences between the actual and estimated activity duration and cost estimates. Namely, there are two subsections (*Activity duration results* and *Activity cost results*). Overall, these two subsections encompass 100 lines plus two tables out of the approx. 600 lines of the whole revised manuscript (once references, acknowledgements and data availability sections have been excluded). Arguably, in this computation, the *Analysis outline* subsection (with another 75 lines) should not be considered. This, as this subsection sets a common framework for other subsections, particularly, the *Discussions*.

However, even if we accepted that the three subsections above are just devoted to measuring actual vs planned estimates, there is much more content in the paper. For example, the subsection *Activity duration and cost correlation* (approx. 60 lines plus one table) is devoted to analyze (also for the first time) the empirical correlations between activity duration and cost extensions. These had been measured at project level many times, but never at activity level, neither with a thorough classification of activities (as described in the *Analysis outline* subsection), nor with such a representative project dataset.

Furthermore, the major aim of the paper (clearly stated at the outset in the second half of the *Abstract* and at the end of the *Introduction*) is using all these measurements and correlations to make a more relevant point: that construction projects end late and cost more, not because their activities end late (on average), but because they mostly suffer from an extraordinarily high level of duration variability. The only way to make such a claim credible is: first, to measure this variability soundly, and, second, to explain how from this activity variability construction projects can, indeed, end late.

Derived from the first point (to measure the activity duration variability soundly) is that we need to build a solid measurement framework beforehand, which is what the *Analysis outline* is about. Without it, nothing can be built upon later. Regarding the second point (how we connect activity variability with projects ending late) is that we wrote some extensive *Discussions*. Namely, in this section, it is shown (and in two different ways: graphically and topologically at a project network level) how activity duration variability is causing many construction projects to end late and cost more. This section on its own is approx. 100 lines plus one table and one figure. Hence, this section alone takes the same extension as the two subsections devoted to measuring the actual vs planned activity duration and cost estimates.

In conclusion, we can’t agree with reviewer 1’s second statement. We do not feel the contents of the paper are dominated by the comparison of estimated versus actual activity durations and costs. Yet, we acknowledge things can always be written in a clearer manner to ensure the argument can make sense to all readers. Multiple parts of the papers have been slightly rewritten with the intention of highlighting the aims of the paper. These are found in green text in the revised manuscript. However, many other have also been shortened respect to the original version. This reviewer will appreciate that the revised manuscript has hardly grown respect to the original submission. With this we have tried to make the arguments clearer.

On the other hand, regarding reviewer 1's first statement (that the paper does not provide a solid novel research question and its importance), we also disagree. Indeed, the research question is built upon the statement of a problem existence and its importance. Both aspects are fully addressed in the *Introduction*. Let us present an abridged version of this section which intentionally excludes some secondary statements and references. In this summary, the problem statements are colored in blue and the problem importance is colored in red:

"*Construction activities usually suffer from variability in both duration and cost dimensions. Sources of this variability are plentiful […] All these factors, plus many other, make of the duration and cost estimation exercise, a challenging task for construction managers. […]*

Factors that cause projects to end late or result in cost overruns have been studied in the construction literature for a long time. Some of the most recurrent are […]. Among all these, however, poor planning and control practices are consistently among the most pervasive.

Ballesteros-Pérez et al. (2018a) recently showed how the most common scheduling techniques […] consistently underestimate the actual project duration and cost. One of the major causes of this underestimation came precisely from neglecting activity duration variability.. […]

Apart from the classical scheduling techniques, more advanced techniques for getting improved project duration and/or cost estimates have been proposed over the years […]. What all these methods have in common […] is that they all require some prior estimates of the potential activity durations and costs. […]. Access to this information is, often, the major limitation of these methods. Similarly, realistic data on the correlation between activity duration and costs is also a rare commodity, which forces these techniques to either assume independence between activities and costs, or resort to subjective correlation factors […]. Consequently, when enough quantity or quality of information is not available, the forecasting accuracy of the actual project duration and/or cost is expected to be unreliable.

Unfortunately, despite its importance, there is a dearth of research into activity duration and cost variability studies can be found in the construction management literature. […] Nonetheless, these difficulties should not be a deterrent to, at least, attempting to measure the average level of variability of construction activity durations and costs. […] this would be an extremely valuable input for future project duration and cost forecasting techniques, as well as providing a powerful baseline information for enhancing project control and monitoring."

(And with the problem existence and importance now appropriately addressed we are in a position to adequately formulate our research question, in this case near the end of the *Introduction* section)

"Hence, the present paper precisely attempts to fill this research gap in the construction management literature: measure the average level of activity duration and cost variability. It will also justify how and why, given this level of activity variability in common project networks, it is expected that most construction projects end late and go over budget.

(And, still in the *Introduction*, how we plan to achieve it)

To achieve this, the actual/planned (log) ratios of many project and activity durations and costs will be analyzed. The correlation between activity durations and costs will also be studied. Finally, the most common network topologies (descriptors of what the project networks are like, that is, how activities are arranged and connected with each other) will be summarized and the potential impact of activity variability on these networks described in detail. […]"

The research question is repeated in other parts of the manuscript as well (second part of the *Abstract* and *Conclusions*, for example), but we will not present those statements here as they just reiterate the same messages. We believe our point is clear. The research question has been formulated and its need justified from different perspectives. Yet, we are open to suggestions in case reviewer 1 considered something else should be added to complement what has already been said.

Point 1.2. There is no novel model or framework proposed in the paper.

We agree and are glad this came across clearly in the paper. There is no model, nor a novel framework proposed in the paper, this was not the aim of the research. This is an eminently empirical paper as observed from our reply to point 1.1, which is also mentioned in the paper itself. The paper is built upon two major blocks. The first block is devoted to gathering, classifying, measuring and presenting the results of a wide set of construction activity durations and costs. Then, in light of new (empirical) evidence, the second block elaborates on how those measurements translate at the whole project level. We will go over them separately.

The novelty of the first block (gathering and processing the empirical data) is not a new model or framework. The novelty is that this is the first serious and representative attempt to measure activity variability (versus the well-known project variability). This is achieved by resorting to an extensive project and activity dataset like no analysis had counted on before. This is also evidenced by a representative depth of analysis (up to four levels of activity classifications: all-in, by project type, project phase, and activity type). No previous study had been able to generalize its results (not even down to different types of activities), much less to calculate the second, third and fourth moments (for which a substantially big sample size is needed).

All these ideas are mentioned in several parts of the manuscript, but in the *Conclusions* we can see them all together:

"*…The research is novel because it describes the first four moments (average, standard deviation, skewness and kurtosis) of how actual versus planned durations and costs differ at activity level in construction projects. A set of 101 projects and 5289 activities, plus another set with 746 projects have been used.*

The first contribution of this study is providing construction managers with a first, yet rather complete, set of actual-vs-planned average activity durations and costs deviations with application in multiple contexts (project types, execution phases and types of activity). […] This might potentially improve the quality and robustness of all construction schedules, for example allowing them to feed more advanced (non-deterministic) scheduling and simulation tools with more representative data […] With the set of moments provided here, these techniques will be able to resort to average values for their activity durations and cost distribution parameters depending on the type and/or execution phase of the project. These distributions will also be able to assume non-independence between the stochastically-generated activity durations and costs values (thanks to the set of duration-cost correlation values also published in this study)..."

The novelty of the second block (analyzing how the activity variability causes the projects to end late and cost more) is not resorting to a new model or framework either. It is resorting to a project network perspective to explain how projects whose activities are so variable, can end late. Project network topology is not new, but certainly it is not common in construction management research either. Maybe the only exception is in project scheduling within Operations Research. Resorting to the tools and techniques that project topology offers was not possible in the absence of a very rich and varied activity dataset. This, as, in order to apply these tools, we need a lot of activity-level data, but also all the activities precedence network information (the predecessors and successors of all activities, generally contained in the project schedules).

Again, these points are emphasized in different parts of the manuscript, but summarized in the *Conclusions*:

"*... One of the most relevant [insights] is that it has been shown that construction activities do not end late on average. Instead, it is their high level of variability (around 60% of its average duration) the key factor eventually causing project-level delays. Such high levels of activity variability exacerbate the merge event bias, a phenomenon by which whenever two or more schedule paths converge into a single one, the average completion times exceed the maximum average path durations…"*

As a matter of conclusion, this is an empirical paper. It does not contain a novel model or method proposal because it is not necessarily in the nature of an empirical paper to contain these. Its distinctive point is to have addressed an as-yet unsolved problem with an unprecedented amount of information in an unconventional way. This has allowed understanding the problem of construction projects ending late from a different perspective. With all certainty, construction projects will keep ending late, but now we have provided strong evidence suggesting that activity variability is one of the biggest causes of this problem. We need shift our attention to develop tools that can handle/restrain activity duration variability much more effectively. If we manage to do achieve this, construction projects will surely end a little sooner.

Again, all these points are mentioned in several parts of the paper, but have been summarized them at the very end of the *Conclusions* after this and other reviewer's suggestions. A new paragraph says:

"*A last conclusion derived from this research is that activity duration variability is the actual foe in project monitoring and control. This may not sound new to Lean Construction researchers and practitioners. However, this research has provided compelling empirical evidence suggesting that we do really need to start taking activity variability more seriously. There is a need to develop more techniques that can effectively handle/restrain this variability. Value stream mapping and Last planner have been some attempts to address this problem, but more are needed. This will open the door to new and more effective approaches for tackling the widespread phenomenon of construction projects ending late.*"

This paragraph highlights again the empirical nature of this paper, plus some of its contributions and suggested lines of research. New models and frameworks can be expected in the future. This paper has just provided a strong case suggesting that this is the right path to follow.

<u>Point 1.3.</u> One major finding of this study is probably that the current practices of duration and cost estimation have a low accuracy. We all know about this. What the body of knowledge needs is a new advanced method with higher accuracy which the paper does not contribute.

Thank you for this comment, however we have not said this anywhere in our paper, nor have we tried to make it implicit. As such, we accept that we need to make of argument much clearer. For this reason we have partially rewritten this statement at the beginning of the *Conclusions*:

"...The research is novel because it describes the first four moments (average, standard deviation, skewness and kurtosis) of how actual versus planned durations and costs differ at activity level in construction projects."

However, also in the *Conclusions*, for example, we emphasize more that the activity duration estimates (the planned activity durations) are generally *quite* precise (*on average*). Construction managers are indeed doing a good job when estimating the average costs and durations. However, what we have also said is that, when projects cost more and end late is partially because some of their activities have had their start date delayed (not because they last longer or cost more per se). As a consequence, those activities with a delayed start will also cost more. All these ideas are more clearly stated now in several parts of the paper, but also summarized in the *Conclusions*:

"*However, the analysis developed has also provided some interesting insights from its numerical perspective. One of the most relevant is that it has been shown that construction activities do not end late on average. Instead, it is their high level of variability (around 60% of its average duration) the key factor eventually causing project-level delays […]*

Actual activity costs, on the other hand, do tend to be higher than what was planned (around 7%). This cannot be the result of price adjustments or inflation […] Instead, the major project-level cost overruns are expected to occur as a consequence of delayed start of activities located nearer the end of the project. This, as it has been demonstrated how most duration-cost correlation factors range within 0.40 and 0.70. The latter would cause that those activities that cannot start until their predecessors have finished, start incurring in costs before their actual execution."

However, construction managers clearly have a problem assessing the activity durations and costs variability, but that is shared with many project managers from other industries. In point 1.2 we agreed on the fact that new techniques for more effectively monitoring and restraining activity variability were needed. All this was made clearer in the new last paragraph of the *Conclusions*:

"*A last conclusion derived from this research is that activity duration variability is the actual foe in project monitoring and control. This may not sound new to Lean Construction researchers and practitioners. However, this research has provided compelling empirical evidence suggesting that we do really need to start taking activity variability more seriously. There is a need to develop more techniques that can effectively handle/restrain this variability. Value stream mapping and Last planner have been some attempts to address this problem, but more are needed. This will open the door to new and more effective approaches for tackling the widespread phenomenon of construction projects ending late.*"

Point 1.4. **In addition, the paper is written in a way that the reviewer find hard to follow and stay focused.**

The subject is complex and that is reflected in the paper. We have tried very hard to offer a logical structure. The level of appeal for the paper will clearly be influenced by the philosophical and theoretical orientation of the reader. We agree that the paper is indeed quite numerical and it can be very dense in some sections where the empirical results are presented (subsections *Activity duration results* and *Activity cost results* mostly). Conversely, some readers may have the opposite problem and find the *Discussions* very dense. In the *Discussions* we take the numerical results and build a strong case defending that, because of the merge event bias, projects whose activities suffer from such a significant variability will also be very prone to end late. Either way, we tried to shorten the paper slightly after this revision to provide the reader exclusively with the most relevant results and interpretations. A proof of this is that the paper is no longer than average papers in this journal (around 10.000 words) even after the revision. Additionally, most sections and subsections of this paper can be read almost independently from each other. This, as each section at the outset concisely summarizes the previous take-away data and results.

Therefore, apart from the multiple trimmings performed across the paper, we can't perform any major changes in this regard unless reviewer 1 provides us with more specific guidance on what aspects/excerpts he/she finds hard to follow. The other two reviewers did not raise concern about the paper being poorly written, either. Still, we are open to further exchanges if this reviewer suggests a specific way of improving the paper.

Point 1.5. **The paper uses strange terms such as 'moments' rather than 'statistical measures'.**

We have to disagree regarding this comment. *Statistical measures*, or rather *statistical measurements*, is a loose term that can refer to many different things (generally to a wide range of *summary statistics*, among which we can also find the mean, standard deviation, skewness and kurtosis).

Moments, on the other hand, is more appropriate and quite common for researchers with a primer in statistics. The term *moment* refers to specific quantitative measures of the shape of a function. If the function is a probability distribution, the first moment is the mean, the second (central) moment is the variance, the third (standardized) moment is the skewness, and the fourth (standardized) moment is the kurtosis. The word *moment* offers no ambiguity regarding what we want to say.

Nevertheless, for those readers who may be confused, we have been careful enough to list what we mean next to the word *moment*, and to do it more than once. For example, in the *Introduction*:

"*To do this, the first four moments (mean, standard deviation, skewness and kurtosis) of actual versus planned duration and cost (log) ratios are analyzed…*"

Or in the *Background* section:

"*The reason why the choice of a particular statistical distribution does not seem that relevant is because the third and fourth moments (skewness and kurtosis) are blurred […] However, this is not the case for the first two moments (mean and variance, or alternatively, standard deviation). …*"

Even in the *Conclusions* now:

"*The research is novel because it describes the first four moments (average, standard deviation, skewness and kurtosis) of…*"

There are more examples in the manuscript. We believe this is a compromised agreement between using the right word and allowing those uninitiated readers to understand what we want to refer to.

Point 1.6. **Still, there are lots of grammatical errors in the paper.**

We appreciate this observation and the paper has been proofread again after this revision. Among the authors we count on some native English speakers who have paid special attention not to pass any grammar errors this time. If reviewer 1 still finds some errors, we will be very grateful if they let us know where they are so that we can correct them. Furthermore, if eventually accepted for publication, then the journal will issue a 'final proof' for approval and thus yet another round of proof reading will take place.

Point 1.7. **This paper should be submitted as a forum paper rather than a technical paper.**

We believe this paper should be submitted as a technical paper (as it is currently submitted). The other two reviewers also believe the paper has been submitted under the right article type. Consequently, we do not think we must change it.

The authors appreciate the comments made by Reviewer 1. Despite we have not agreed with some of them, we hope we have justified why it was more appropriate not to act upon them. We believe your suggestions and observations have greatly helped us to improve the manuscript.

REVIEWER 2

Point 2.0. **Overall, it is a very well written paper. The topic is pretty interesting. The authors used an empirical study to examine the variability of the construction schedule and cost at both project level and activity level. The results are quite interesting and the authors used network theory to explain the findings from the statistical analysis of 101 real construction projects.**
The authors have some minor mistakes for the authors to address:

The authors appreciate the positive comments from Reviewer 2. We have revised the paper as suggested in the next lines.

Point 2.1. **Lines 380-382 : The line reads "This is not the case at Project-level correlations, where apart from the level 0 of analysis (all 101 projects 382 grouped together), R2 values have not been found to be (statistically) significant." According to Table 5, the R2 is 0.05 and it is statistically significant. Please double check.**

Thank you for this observation. We believe the sentence had a slight punctuation problem (the comma had been put one word earlier). Now it reads like this:

"*This is not the case at Project-level correlations where, apart from the level 0 of analysis (all 101 projects grouped together), R^2 values have not been found to be (statistically) significant.*"

This revised version duly acknowledges that the R^2 was indeed statistically significant for the level 0 when all projects were grouped together.

However, on double-checking Table 5 results we also found one mistake (one * was not supposed to be there at Project-level results). This error has now been corrected.

Point 2.2. **Lines 435- 436: "For Civil Engineering projects in Table 3, the average of the duration ratio was -0.008, hence, very unlikely to have caused this." This sentence is not clear. What does "this" refer to? The reviewer is not clear about "what causes what."**

Reviewer 2 is right. This sentence was difficult to understand and probably open to different interpretations. We have now rewritten the whole paragraph to make the point clearer (new text in green):

"Furthermore, it is worth noting that the average project duration extension is around 0.21 (in log scale). For Civil Engineering projects in Table 3, the average of the duration log ratio was negative (-0.008). This means the activities from civil engineering projects ended sooner than planned (on average). It is unlikely then, that the projects represented in Figure 1 could have ended later because a significant proportion of their activities ended late. However, the activity duration variability (the standard deviation) was 0.20. In extreme value theory, the mean of the highest order statistic distribution of a Normal distribution with three or four draws is approximately one standard deviation. The Normal distribution represents very well the distribution of the durations of each path (before they merge) (Ballesteros-Pérez, 2017a). Therefore, the average of the duration distribution coincides very closely with what is to be expected from the data from Table 3 for civil engineering projects (0.21≈0.20). Later it will be shown how more than three paths are quite common in civil engineering construction schedules."

Point 2.3. **Lines 470-471: "In service projects schedules there must be (on average) at least five paths in the schedule, as only 20% (a fifth) of the activities are critical." Does SP = 20% really**

mean 20% of the activities are critical? The reviewer believes that theoretically, it is possible that the network can have five paths and all the activities can be on the critical paths. The authors probably need to present this statement more rigorously.

Technically speaking, network topology neglects activity durations. The SP is just an estimate of the proportion of critical activities, but it is not exact. Namely, the SP counts the number of activities in the longest chain (in number of activities, not necessarily in duration) and divides it by the total number of activities. This might mean, theoretically, that the longest chain might contain activities with very short duration and another path might contain, for example, just one activity with an extraordinarily long duration. But, in real contexts (and that is why we include so many times the warnings: “approximately” or “on average”) there is a high correlation between the relative number of activities in a path and the probability that this path becomes critical. Hence, despite all these are just approximations, when analyzed together, they all seem to match and be telling the same story.

As a consequence, we have offered a little more of details in the paragraph where the sentenced was extracted from. We believe this version makes things a little clearer, without being too overwhelming for those not initiated in network topology:

“The Serial-Parallel (SP) indicator is probably the most relevant of the four indicators for the purpose of this study. This indicator measures the closeness of a network to a serial or parallel network. Namely, SP = (m-1)/(n-1); where n is the total number of project activities in a project schedule, and m is the number of activities in the path with a higher number of activities (which may not necessarily be the longest in duration, as topological measures ignore the activity durations). Hence, SP=0 means all activities are in parallel, whereas SP=100% means all activities are in series. This indicator can also be considered as an estimate of the amount of critical and non-critical activities in a network (Vanhoucke and Vandevoorde 2009). Therefore, rounded up values of the inverse of the SP (that is $\lceil 1/SP \rceil$*) provide us with an estimate of the minimum number of paths of a project schedule. Values of SP below 50% would mean that construction schedules have (approximately) at least three paths. This agrees with what we appreciated in the black curve of Figure 1. Industrial projects, despite having on average at least two paths, generally have a dominant one (which condenses, on average, 55% of the activities). In service projects schedules there are at least five paths (on average), as only 20% (a fifth) of the activities are critical.”*

<u>Point 2.4.</u> Lines 472-473: "Activity Distribution (AD) measures the distribution of project activities along the levels of the project" Can the authors define the level of the project here? It can create confusion if the reader is not familiar with the terminology of network models.

Yes, we gave little details as the more concepts we defined, the more complicated the argument becomes. We have tried to find a compromised agreement here. We have provided a loose definition of level so that the average reader can understand the message. However, a much more rigorous definition would involve bringing up other concepts we don’t deem necessary. As a result, this is the rewritten paragraph:

“Activity Distribution (AD) measures the distribution of project activities along the levels of the project. In network topology, the number of project levels can be loosely defined as the number of activities that are arrayed in parallel in a project schedule. Hence, AD measures the width of the network. However, it is worth noting that activities arrayed in parallel do not necessarily have be executed simultaneously (because they may have different time lags and/or activity durations). When AD=0 all levels contain a similar number of activities and the number of activities is uniformly distributed over all levels. When AD=100% there is one level with a maximal number of activities, and all other levels contain a single activity. All four types of projects average AD values are close to 58% indicating that the longest path

has more activities than other paths, but still those other paths contain a significant number of activities, that is, they can potentially cause project delays. ”

We would have liked to have put the “However” sentence above as a footnote, but JCEM does not allow footnotes (as can be seen from the Editor’s point 0.2). Therefore, we have integrated it with the rest of the text as best as we can.

Point 2.5. **Lines 518- 523: The claim of the contribution of this paper sound a little bit trivial here. For a construction manager, knowing the variability does not necessarily help him/her have a more accurate project schedule or cost estimate. Instead, the authors should focus on more the contribution of the paper to the theory instead of practice.**

Yes, probably the initial paragraph seemed weak regarding its contribution. We have merged it with the next paragraph and partially rewritten both. This reinforces the contribution statement and reads like this:

“*The first contribution of this study is providing construction managers with a first, yet rather complete, set of actual-vs-planned average activity durations and costs deviations with application in multiple contexts (project types, execution phases and types of activity). From now on, a construction manager will be able to more realistically (thus accurately) anticipate how likely and how much the activities in the project schedule will vary, that is, last or cost something different. This might potentially improve the quality and robustness of all construction schedules, for example allowing them to feed more advanced (non-deterministic) scheduling and simulation tools with more representative data. These techniques generally need a substantial amount of information from previous similar projects which is rarely available. With the set of moments provided here, these techniques will be able to resort to average values for their activity durations and cost distribution parameters depending on the type and/or execution phase of the project. These distributions will also be able to assume non-independence between the stochastically-generated activity durations and costs values (thanks to the set of duration-cost correlation values also published in this study). This is expected to enhance future construction project monitoring and control, but also actual project duration and cost forecasting accuracy.*”

The authors appreciate the comments made by Reviewer 2 and believe his/her suggestions and observations have greatly improved the manuscript.

REVIEWER 3

Point 3.0. The manuscript explores the variability of construction activity costs and especially durations from a quite large dataset. Various statistical measures were used to derive findings. This reviewer has the following comments:

Thank you for reviewing our original manuscript. We have revised the paper according to your suggestions. The changes implemented and some additional comments are explained below.

Point 3.1. More caution should be considered in discussing the results. "Prove" has used a lot throughout the manuscript when discussing the results. For example, "Overall, these log ratios so close to 0, prove that construction activities do not end late on average." (line 301-302). The reason is that "[…] statistics neither proves anything nor disproves anything" (Saksena, A Hand Book of Statistics, 1981, p.6, https://eur01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fbooks.google.com%2Fbooks%3Fid%3DfHJ0ZWtp1kEC&data=02%7C01%7C%7Ca22ca68399044a72f46208d6aed5215b%7C84df9e7fe9f640afb435aaaaaaaaaaaa%7C1%7C0%7C636888627737467179&sdata=DsSPT%2BkwuKgScxElHBs15xNvVkFgu%2Fz8Ugn%2FbzCjHZc%3D&reserved=0)
or "statistics can never 'prove' anything." https://eur01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwebstat.une.edu.au%2Funit_materials%2Fc5_inferential_statistics%2Fprobability_and_proof.html&data=02%7C01%7C%7Ca22ca68399044a72f46208d6aed5215b%7C84df9e7fe9f640afb435aaaaaaaaaaaa%7C1%7C0%7C636888627737467179&sdata=RkSebFhdrDDK%2BrkrWWv39dwRyEQI%2Bq93ZWrIUYBgs9M%3D&reserved=0

We agree. The word "prove" conveys absolute certainty. We have replaced that verb with "suggest" or "show". For example, in the excerpt used by the reviewer we say now:

"…Overall, as these log ratios are so close to zero, this suggests that construction activities do not end late (on average). This may be an unexpected finding, as the easier explanation for projects ending late was that its activities ended late on average. This result seems to suggest the problem lies somewhere else."

Other later sentences have also changed the word "prove". These have not been shown here but can be easily found in "green" in the tracked changes version.

Point 3.2. The study found that construction activities did not end late on average. This may not be surprising if one looks at the activity data (from the Supplemental Data provided) closer. More than 70% of activities (4917 out of 6841 activities) in the dataset took 10 days or less. The short-duration activities could be extracted from short-term schedules (e.g., three-week look ahead schedules). The short-term plans tend to be more reliable, including activity duration. In addition, while an activity is not extended in duration but its start date may be delayed (the manuscript also mentions this in lines 546-547) due to other reasons such as delay in mobilization, approval of shop drawings, etc. A few if not at all in the list of 6,841 activities in the dataset reflect those types of activities. The bottom line is that it is hard to make a solid conclusion based only on the statistics without knowing the context of data.

In the case of this project dataset, we know for certain that a significant proportion activities did not come from short-term look ahead schedules. We know this for several reasons. First, detailed explanations on how these projects information were gathered are explained in these two papers referenced in our paper:

- Vanhoucke, M., Coelho, J. and Batselier, J., 2016, "An overview of project data for integrated project management and control", Journal of Modern Project Management, 3(2), 6–21.

- Batselier, J. and Vanhoucke, M., 2015, "Construction and evaluation framework for a real-life project database", International Journal of Project Management, 33(3), 697–710.

Our paper refers to these papers in the *Projects and activities dataset* subsection when saying:

"…The 101-project dataset was retrieved from a real projects dataset originally developed by Batselier and Vanhoucke (2015) and Vanhoucke et al. (2016)."

In the papers above, the specific data gathering framework can be better understood by those interested readers. Additionally, most projects of this dataset have been used repeatedly in other recent studies on Earned Value Management (EVM). For example in:

- J. Batselier, M. Vanhoucke, Evaluation of deterministic state-of-the-art forecasting approaches for project duration based on earned value management, International Journal of Project Management. 33 (2015) 1588–1596.
- J. Batselier, M. Vanhoucke, Improving project forecast accuracy by integrating earned value management with exponential smoothing and reference class forecasting, International Journal of Project Management. 35 (2017) 28–43.

In EVM research, the baseline activity costs and durations are not generally altered during the project execution (e.g. they are not updated so that the initial estimates change). In other words, in these projects, it prevails whatever has been stated as "planned" at the outset.

The second reason is that many of these projects count on a project card where it can be found out in which context each project data was collected and (partially) processed. This is also mentioned in this excerpt of the paper within the *Projects and activities dataset* subsection:

"…Yet, further details and specific project information can also be found as individual project cards at OR-AS.be (2018).."

These cards are quite succinct and they are not available in all projects, but they are available in many of them. These cards can be found in this link which is also provided in the references of our paper:

http://www.projectmanagement.ugent.be/research/data/realdata

Overall, the way the project dataset has been collected may have some deficiencies, but we are confident that this dataset is quite an excellent and abundant source of reliable and representative information. At least, unlike many other construction projects currently used by many researchers, the way their information was retrieved and processed has gone through peer-review process in two good project management journals. These criteria can also be found summarized at the top of the link provided above.

Finally, it is true that many activities lasted less than 10 days. Among those 6841, though, many activities were excluded (as stated in column E). Also, it is well known that activity durations and costs resemble quite closely a (Generalized) Pareto distribution. This means there are always a few activities (around 20 to 30%) that cost a lot and/or last quite long (in duration), whereas the rest of activities (from 70% to 80%) last and cost significantly less. Therefore, the alert raised on many activities lasting less than 10 days is not worrying, unless the reviewer is suggesting we should have ruled out some of the activities and/or weighted them somehow (depending on their activity cost and/or duration, for example). Since we don't think neither option should have been the right approach, we still believe we did our analysis with representative data.

Point 3.3. **The fact that variability (in activity duration) and merged bias play significant roles in causing project delays and cost overruns is not new. The Lean Construction community has studied and implemented various initiatives/tools (e.g., Last Planner) to reduce variability in construction planning and production for 30 years.**

Agreed. One of the contributions of this paper is that we have measured how big this variability can be (on average) and for different types of projects, phases of execution and activity types. These measurements had not been performed before, and certainly they had not been put under the perspective of network topology either. Only with both actions we have been able to provide mathematical evidence on the important effect that such a big activity duration variability may have in the merge event bias of real construction projects.

Still, we believe we must acknowledge that our concern has been shared with the Lean Construction community for a long time. For this reason, we have added a new paragraph just at the end of the paper (in the *Conclusions*) where all these ideas are highlighted. We believe this paragraph will also duly accommodate reviewer 3's suggestion:

"*A last conclusion derived from this research is that activity duration variability is the actual foe in project monitoring and control. This may not sound new to Lean Construction researchers and practitioners. However, this research has provided compelling empirical evidence suggesting that we do really need to start taking activity variability more seriously. There is a need to develop more techniques that can effectively handle/restrain this variability. Value stream mapping and Last planner have been some attempts to address this problem, but more are needed. This will open the door to new and more effective approaches for tackling the widespread phenomenon of construction projects ending late.*"

Point 3.4. **As the manuscript correctly states in lines 40-41, "poor planning and control practices are consistently among the most pervasive [in project delays and cost overruns]." The project delays and cost overruns can be due to scope change (or scope creep), work out of sequence, etc. These are other causes in addition to variability in activity duration and merge bias. To some small extent the scope change may reflect in "unplanned but performed" activities in the dataset, which were however removed form analysis in this study. The analysis of the activities' durations alone cannot adequately explain the cause of the project delays and cost overruns.**

We agree on this point too. Unplanned but Performed (UbP) and Planned but Unperformed (PbU) activities being excluded was a limitation of our study. This needed to be highlighted in the limitations, emphasizing exactly the same ideas that reviewer 3 has suggested. We have included a new paragraph in the *Conclusions* which solves this problem:

"A second limitation arises from having removed at the outset the Unplanned but Performed (UbP) and Planned but Unperformed (PbU) activities. This was necessary as the ratios (either in natural or log scale) converged to infinity causing a distortion in the moments calculation. However, we acknowledge that these activities can be found in almost all real projects. Frequently, they are the consequence of scope changes, works reorganization or changes in the available resources. Obviously, UbP and PbU activities add to the total project variability (beyond the activity duration and cost variability analysed here). In our analysis, though, there were only 279 UbP + 129 PbU = 408 activities out of the initial 5,697 (7% in total). Hence, while we believe the influence of UbP and PbU activities needs to be duly investigated, our analysis (with 93% of the activities) can still be considered representative enough to draw valid conclusions. Additionally, it is also expected that some degree of cancellation will occur among those 7% of activities (as frequently new activities replace others which are not eventually performed)."

<u>Point 3.5.</u> **Table 2 looks like a figure to this reviewer.**

Table 2 is indeed composed of several tables. We may have to leave it as it is, divide it in sub-tables or, alternatively, present it as a single figure as suggested by reviewer 3. The editor has not brought our attention to this issue, though. We think it is more sensible to wait for the copy-editors to provide us with further instructions (should the paper is accepted). Hence, we don't think it is advisable to act upon this comment for the time being.

The authors appreciate the comments made by Reviewer 3 and believe his/her suggestions and observations have greatly improved the manuscript.