# RISK-3751

# The Case for Parametric Quantification of Systemic Risks for Transportation Projects

John K. Hollmann, PE CCP CEP DRMP FAACE Hon. Life

**Abstract–**AACE® International recommended practices (RPs) address *empirically-based*, risk-driven parametric modeling to quantify systemic risks. These RPs reflect the process industry's embrace of phase-gate project systems and extensive research supporting the methods. While the cost overrun-prone transportation sector is catching up in applying phase-gate systems, it lags in benchmarking and research and has been detoured away from the RP methods. Research showing that fundamental practice failures (i.e., *systemic risks*) cause cost growth has been largely ignored. Instead, an unsupported hypothesis as to the cause of cost overruns called the *planning fallacy* (i.e., optimism bias or lying), and a pessimistic "de-biasing" practice called reference class forecasting (RCF) are being embraced by some in transportation. The danger is that RCF will institutionalize mediocrity of cost outcomes; a detour to an economic dead end.

A principle called the *fifth hand,* reflecting situational-specific, mixed causes of cost overrun, that aligns with research and AACE RPs, is reviewed. However, what is missing from the debate is the fact that too-narrow estimate accuracy range expectations are wired into owner phase-gate procedures. Estimators use subjective risk analysis that underestimates contingency because it meets these expectations; i.e., the *planning fallacy* is institutionalized. Those who benefit from low estimates need not lie; just announce a project early in perfect confidence that estimators have failed to put a price on poor scope definition. The paper subtitle should be "we have met the enemy and he is us".

The paper reviews the cost overrun situation, the theories, the research and the various proposed and recommended methods. Despite all evidence to the contrary, the author is hopeful that empirically-valid risk quantification and contingency setting practices will become more widely used, putting an end to endemic cost overruns in the transportation industry.

**Table of Contents**

## Introduction

### *Background*

Cost Overruns and the News

It will come as no surprise to the cost engineering and risk analysis communities that large transportation industry projects appear to be prone to cost overruns. Some recent examples with >US$1 billion overruns:

- **California High Speed Rail Project**: "California's bullet train to nowhere continues to misfire" (Las Vegas Review, Oct 20, 2019),
- **London Cross Rail Project**: "Crossrail cannot say when the much-delayed line will open, or what the final cost will be" (Reuters, May 3, 2019),
- **Sydney Light Rail Project**: "a 'horror story' of missed deadlines and billions in cost overruns" (news.com.au, April 10, 2018),
- **Seattle ST3 Transit Project**: "transportation plan busting $54B budget" (Construction Dive, March 9, 2019),
- **Melbourne Metro Tunnel**: "$2.7 billion budget blowout revealed" (The Age, December 11, 2020).

Of course, not all large transportation projects overrun as per the following:

- **Woodrow Wilson Bridge**: "Final tally: Wilson Bridge project $86 million under budget" (Washington Post, March 31, 2015)

It is imperative that cost engineers and risk analysts understand the causes of such cost overruns in order to do more effective project risk quantification (PRQ) and improve on these outcomes. The first thing to understand is the definition of the term *cost overrun*, and to distinguish that from *cost growth* which is what estimators attempt to predict (i.e., contingency).

Cost Overrun versus Cost Growth

- *Cost overrun* = amount by which the final actual cost varies from the *announced or approved cost*.
    - Reflects the investor or taxpayer point of view.
    - Estimates perceived as a whole; not concerned with details; only that what was spent differed from what was promised.
    - Negative connotation/failure to meet cost objectives.
    - The actual and announced costs are all that is available to most broad-based studies using public information.
- *Cost growth* = amount by which final actual cost varies from than the *base cost estimate excluding contingency*.
    - Reflects the base estimator and contingency estimator/risk analyst points of view.

- Recognizes that the base and contingency estimates are unique; the former deterministic focused on what is known and the latter probabilistic focused on risks; estimated with separate methods and usually by separate parties.
- Neither good nor bad; where there is risk, there is cost growth; i.e., there will be cost growth but no overrun if appropriate contingency is provided for the risks.
- Knowledge of the base estimate is usually confidential requiring a legally vetted benchmarking regime prior to any broad-based study.

For both of these measures, escalation is usually normalized out for research. The impact of major scope change (i.e., a change in a basic premise of the business case) may also be normalized out; however, publicly available data often lacks scope change information. These metrics may be expressed as an amount of money, a percentage or a ratio of actual cost over the reference cost base. Actual costs may of course be less than the reference cost base (i.e., underrun or negative cost growth).

Cost Overrun-Based Research and the Planning Fallacy

If the only estimate information available is the total announced or approved value, one can only study *cost overrun*. It is not possible to research where cost estimating and budgeting went wrong; in the base estimate? in the contingency setting? in final pronouncement? This can result in cognitive bias in research. If there is only an announced or approved "number", it is tempting to view practices leading to that number as an undifferentiated whole and propose sweeping hypotheses as to why and where the low number originated. This bias was expressed by Maslow (Maslow's hammer) as "*I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail*." [1]. In the case of transportation cost overruns, the hypothesis that wraps all up in a single cause (i.e., nail) is called the "planning fallacy."

The planning fallacy theory holds that cost overruns are caused by the promoter or politician lying about or strategically misrepresenting the total cost to the investor or taxpayer. If the misrepresentation is subconscious, it is called "optimism bias". The base estimate, the contingency estimate, the estimators, the risk analysts, their practices – all are irrelevant in this theory. It simply assumes estimates all are optimistically biased (intentionally or otherwise) and therefore must be *de-biased* using another method. The most common debiasing method is called *reference class forecasting* (RCF) which assumes (too often correctly) that base estimating does not employ validation and risk quantification does not employ empiricism. The planning fallacy and RCF presupposes flawed, biased practices and institutionalizes them. RCF is a kind of corrective lens for genetic estimating myopia.

The transportation sector has been left open to this hammer and nail situation, and layering on of RCF practice, because, unlike in the process industries, industry-wide benchmarking of practices behind the decision has not been done [2]. What *cost growt*h research in the process industries, supported by benchmarking, has shown is that overrun results from various loose screws, rivets and, yes, nails (bias) in the joinery that makes up the investment decision framework.

Cost Growth-Based Research and Risk Drivers

If one has access to data on the practices of how scope was defined, how that scope was translated into a base cost, and then how risks were identified and quantified, the cause of a low number in the end would no longer be left to speculation. In fact, there is a long history of empirical research into these pieces of the number puzzle: the scope, the base estimate, and the risk quantification. The theory supported by this research is that the cause of overruns is failure of fundamental practices including the failure to realistically quantify risks and bias. This has been called the "fifth hand" principle in recent academic literature[1] [3].

The good news is that the primary causes of *cost growth* have been identified by over 60 years of empirical cost growth research. Collectively speaking, that primary cause is *systemic risk.* Per AACE® International (AACE) Recommended Practice (RP) 10S-90 "Cost Engineering Terminology", systemic risks are "*uncertainties that are an artifact of an industry, company or project system, culture, strategy, complexity, technology, or similar over-arching characteristics*." [4]. A key systemic risk driver is the level of definition upon which an estimate or announcement is based (i.e., an artifact of the discipline and governance of the project system). This paper discusses the research in some depth. The best method for quantifying systemic risk is empirically-based parametric modeling. Parametric models are usually developed using multiple linear regression (MLR) or other similar statistical methods. The parameters in the model are ratings of the "risk drivers"; e.g., measures of the level of scope definition, project complexity and so on. It should be noted that one systemic risk, but usually not primary, is bias, optimistic or pessimistic. A benefit of the method is that it uses measures of the qualities of the project system; i.e., it is a governance and quality assurance method supporting continuous process improvement as well as a risk quantification method. Lying and misrepresentation do occur, but as the exception, not the rule [5].

The bad news is that the primary cause of *cost overrun* is also known. The primary cause is the ubiquitous practice of subjective (not empirically-grounded) risk quantification that ignores systemic risks and the research and thereby usually understates contingency [6]. To quote Walt Kelly (in his Pogo comic strip); "*We have met the enemy and he is us*" [7]. The good news is that empirically-based methods for quantifying systemic risks are available in AACE Recommended Practices (RPs) that will be discussed in this paper.

*Purpose*

As was stated, it is imperative that cost engineers and risk analysts understand the causes of *cost overruns*, but more importantly, *cost growth* which is what risk quantification attempts to predict. This requires an understanding of *systemic* risk and *parametric modeling* of systemic risk. The purpose of the paper is to improve these understandings in the transportation industry where benchmarking and research are limited and risk quantification methods are not empirically-valid, resulting in misguided theory as to the cause of cost overruns.

[1] The "fifth hand" refers to a third person cutting into a dance. In this case, the dancers are the causes of cost overrun when described as a false dichotomy of either optimism bias or fundamental practice failure. The fifth hand postulates that both causes may be at work.

*Approach*

The paper reviews the two leading theories or schools of thought on the cause of cost overruns along with their chief academic proponents. It also reviews research supporting each theory. Finally, it reviews the respective quantification methods. These are summarized as follows (full references later in the text):

1. **Planning Fallacy/Optimism Bias**
   - Academic proponent: Dr. Bent Flyvbjerg (University of Oxford; UK)
   - Research: Dr. Bent Flyvbjerg, et. al. (2002)
   - Quantification method: Reference class forecasting (RCF)
2. **Fundamental Practice Failure (Fifth Hand)/Systemic Risks**
   - Academic proponents: (Fifth Hand) Drs. Lavagnon A. Ika (University of Ottawa), Peter E. D. Love (Curtin University; Au), and Jeffrey K. Pinto (Penn State University). Also, Dr. Dominic Ahiaga-Dagbui (Deakin University; Au)
   - Research: John Hackney (1958), Rand Corporation (1981), Construction Industry Institute (1995 & 2011), Independent Project Analysis, Inc. (2012), Grattan Institute (2016). Hollmann (2020)
   - Quantification method: Risk-driven, empirically-based parametric models

The paper concludes by explaining why *fundamental practice failure* (in which bias, optimistic or pessimistic, is one element of practice) is the *only* thesis supported by scientific, empirical research of *cost growth*. It also reviews the research-based AACE RPs for estimate classification systems and for the parametric modeling of systemic risks. The paper explains why the transportation industry's embrace of the planning fallacy and RCF will result in *institutionalized mediocrity* in practice and outcomes in that industry.

**Theory One: Planning Fallacy/Optimism Bias**

*Deception and Lying*

In 2002 a paper was published in the Journal of the American Planning Association that argued that project overruns in the transportation sector "*cannot be explained by error and seems to be best explained by strategic misrepresentation, i.e., lying*." [8] The lead author was Dr. Bent Flyvbjerg who threw down a proverbial gauntlet to those who believe that the failure of fundamental project practices explains cost overruns. Dr. Flyvbjerg (and others of what might be called the "Oxford school" of overrun theory; he is now a professor at the University of Oxford) have continued to publish papers along these lines. His provocative "lying" thesis attracted wide public interest with his research being covered by publications ranging from The Economist to The Wall Street Journal.

Dr. Flyvbjerg's theory depends on accepting the logic that the failure of fundamental project practices could not be the cause of overrun because "*we would expect underestimation to* *decrease over time as better methods were developed and more experience gained through the planning and implementation of more infrastructure projects*." [8] In other words, the practice of estimating and risk quantification by professionals may or may not be appropriate, but it has been rendered pointless by "*tactics in power struggles*" by promoters and politicians. This theory is a more malevolent version of the "planning fallacy" [9] or optimism bias theory of Kahneman and Tversky [10].

*Planning Fallacy Research*

The research shared in the Dr. Flyvbjerg's 2002 article consisted solely of descriptive statistics; i.e., the mean and distributions of percentage cost overrun for 258 published transportation projects. The overrun is the percent that the "*real, accounted construction costs determined at the time of project completion*" varied from the published "*budgeted, or forecasted, construction costs at the time of decision to build*". The costs were normalized to exclude escalation. The term construction here refers to the project capital cost (e.g., it would include engineering costs). Figure 1 shows the distribution of cost overruns and mean value (the median was not reported which is arguably a research bias in its own right).

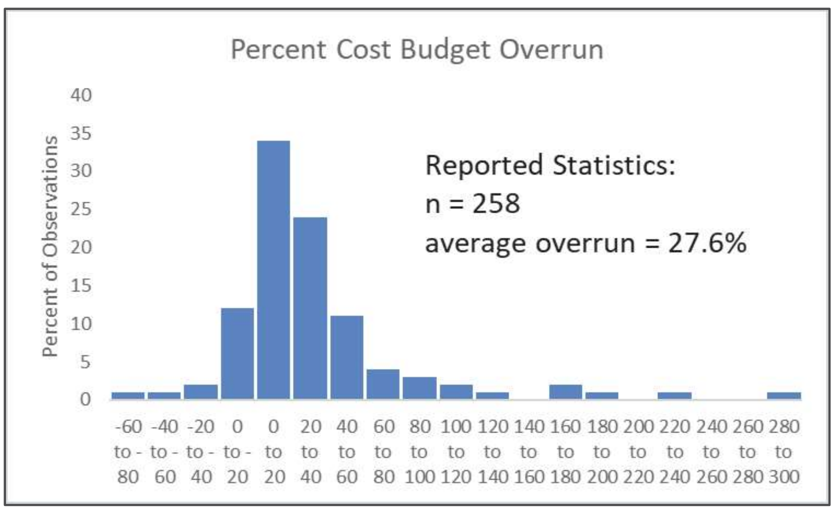


**Figure 1–Percent Cost Overrun Distribution for Transportation Projects (Flyvbjerg [8])**

The study looked at the statistical differences in cost overruns for different types of projects (rail, road, etc.), regions, and year of completion. However, no inferential statistical study of fundamental practices or any other potential cause of overrun was done. Any technical cause of overrun other than "lying" (and perhaps economic) was simply rejected as illogical.

Flyvbjerg was aware of other research (e.g., Rand Corporation to be discussed later) that suggested that cost overruns may result from announcing a project too early when the estimate is based on "incomplete information" (i.e., the primary systemic risk). However, he assumes that estimators do not have effective means to estimate this risk leaving the decision makers free to "*ignore, hide, or otherwise leave out important project costs and risks in order to make total costs appear low*" (examples given are poorly defined environmental and geotechnical conditions). In other words, promoters and politicians do not need to lie about cost or strong-arm the estimator. They only need to announce the cost early and trust that the risks are not yet identified and/or the estimated contingency for poor definition and other systemic risks will be too low. Unfortunately, this is usually a good assumption given that for much of the transportation industry risk quantification practices have been subjective and not empirically-based.

*Planning Fallacy Corrective Method Proposed*

Given the planning fallacy premise that optimism bias is ubiquitous and/or promoters and politicians will actively misrepresent the estimate (i.e., deception and lying), an additional post-estimate practice called reference class forecasting (RCF) is recommended to "de-bias" the estimate by taking an "outside view" of the cost [10].

In 2008, Dr. Flyvbjerg summarized the RCF steps as follows [11]:

1. "Identification of a relevant reference class of past, similar projects. The class must be broad enough to be statistically meaningful but narrow enough to be truly comparable with the specific project.
2. Establishing a probability distribution for the selected reference class. This requires access to credible, empirical data for a sufficient number of projects within the reference class to make statistically meaningful conclusions.
3. Comparing the specific project with the reference class distribution, in order to establish the most likely outcome for the specific project."

The difference between this "most likely outcome" based on the RCF distribution and the cost estimate (including contingency) proposed would then be funded as a management reserve. By doing this, it would presumably add costs for the risks *ignored, hidden or otherwise left out*. Unfortunately, it also adds costs for ineffective practices and performance of all kinds, as well as for risks that may have already been included in the cost. Recall that a problem for the transportation industry, lacking benchmarking, is that the only information with which to select a reference class is that which is publicly available; i.e., scope of the asset, size, location, etc. The class will be mediocre by design and the outcome of the RCF method likely be institutionalized mediocrity.

**Theory Two: Fundamental Practice Failure/Systemic Risks**

*Fundamentals and the Fifth Hand Principle*

Before discussing industry research on the causes of *cost growth* (i.e., research that includes insight to what makes up the published number), the following describes an academic counter-argument to the Flyvbjerg theory that the planning fallacy is the dominant cause of overruns.

In 2018, Dr. Peter E.D. Love and Dr. Dominic D. Ahiaga-Dagbui, published a paper entitled "*Debunking fake news in a post-truth era: The plausible untruths of cost underestimation in transport infrastructure projects*" [12]. The title is perhaps as provocative as Flyvbjerg's use of the term "lying" in the title of his 2002 paper.

The Love/Ahiaga-Dagbui paper explains why the theory of deception and lying as the primary cause of cost overruns is "fake news". It states, and explains in depth, that "*No evidence at all supports the causal claims of delusion and deception as the main explanations for cost underestimation in transport infrastructure projects*." The article points out that the Flyvbjerg paper discarded technical causes of overrun without measuring them and it counters that industry needs to rely on solid research linking causes to cost overrun outcomes.

The 2018 paper also describes a false dichotomy in academic research as to fundamental practice failure versus psycho/political causes of overrun. It states: "*While there is widespread consensus that cost overruns are a pervasive problem, their causes remain matters of contention. This has been, in part, due to the limited access to cost information that is used to produce estimates and the availability of reliable data that can be used to prove causes.*"

Note, Flyvbjerg and co-authors critiqued the Love paper stating in part that if one did not make "*acknowledgment that the root cause of cost overrun is behavioral bias*" or if one did not recommend "*de-biasing cost estimates with reference class forecasting or similar methods based in behavioral science*", that in itself is prime facie evidence that one is guilty of "*bad practice*" [13]; hardly a convincing come-back.

In 2019, Dr. Love and others provided a quantitative case study that supported the "fake news" arguments above [14]. Using a dataset of 85 transportation projects in Hong Kong, they analyzed cost overruns, but with the distinction that they had access to details such as the amount of contingency included in the approved budget. They found a mix of optimism and *pessimism* with approved budgets, and the contingencies therein, often being highly conservative. 47% of the projects were delivered for less than their approved budget. They urge "*agencies that have actively embraced this theory [the planning fallacy] to <u>reconsider</u> their approaches to cost estimating and risk analysis used to deliver their transportation infrastructure to ensure taxpayers are provided with better value for money*". This included ensuring that "*estimators are trained to understand the nature of the uncertainties they are predicting*".

In 2019, Dr. Ahiaga-Dagbui submitted invited written evidence to the May 2019 Government's Management of Major Projects Inquiry, Public Administration and Constitutional Affairs Committee (UK House of Commons). The report title was "Reference Class Forecasting: A clear and present danger to cost-effective capital investment on major infrastructure projects." [15] His evidence "*calls on governments and asset owners to base their investment decision making on empirically-based cost estimation approaches, backed by a disciplined stage-gate practice to alleviate the problem.*" Also, "*The pursuit of predictability over competitiveness and cost-effectiveness in RCF does not solve the problem of cost overruns but it introduces a new one – institutionalized mediocrity.*"

In 2020, Drs. Lavagnon Ika, P. Love, and J. Pinto gave the name "f*ifth hand* principle" to a theory that cost overruns were caused by both fundamental practice failure and bias, but also that the bias may be optimistic or pessimistic. Each project situation has to be assessed for what regime is operative. [3] It is feared that RCF, which presumes that the planning fallacy is controlling, replaces assumed optimism bias with proscribed pessimism bias that funds counter-productive practices and risks that can go unchecked; all in the name of not overrunning.

The above academic papers effectively counter the theory that the planning fallacy is the predominant cause of cost overrun. The next section of this paper reviews the 60 plus years of *cost growth* research in both the private and public sector that reinforces that fundamental practice failure (including the failure to realistically quantify risks and bias) causes both cost growth and cost overruns.

*The Level of Scope Definition/Front-End Loading (FEL)*

As discussed, the publicly funded transportation sector has suffered from a lack of benchmarking of practices and cost outcomes. Lacking data on practices, the Flyvbjerg 2002 research simply reported descriptive *cost overrun* statistics and speculated that the misrepresentation was the primary driver of overruns (later he presented the concept as the "malevolent hiding hand; or planning fallacy writ large" [9]).

Private industry on the other hand has been doing quantitative, practice-focused research of *cost growth* for over 60 years. This research has proven that the *level of scope and planning definition* used as the basis of estimate (but also the level of complexity and technology) is the primary driver of cost growth. Recall that cost growth measures how actual cost differs from the base estimate *without contingency*. Contingency is the part of the estimate that covers the uncertainties and risks, and how that is estimated is the central issue in respect to cover overruns. That practice-focused research resulted in practical, empirically-based contingency estimating tools; i.e., parametric models. Unfortunately, in lieu of validated tools, both private and public industry have continued to rely on either arbitrary (pre-defined) or subjective contingency estimating practice that on average understate the risk. As stated earlier, promoters and politicians do not need to lie about cost; they simply need to report the low estimates they are given.

The large-scale adoption in use of Monte-Carlo simulation (MCS) in the 1980s (spreadsheet add-ons became available) coinciding with owners downsizing their cost engineering staffs and abandoning historical project databases, likely accounts for the loss of focus on empirical-based risk modeling. Few companies or agencies now have data to build or calibrate models, and MCS creates the comforting illusion of realism (MCS can incorporate empirically-based inputs, but this is not commonly done). Flyvbjerg rejected the idea that industry and the profession could fail to learn and improve its practice, yet that is what empirical research has shown has occurred in the estimating and risk analysis fields these last 35 years.

Fortunately, also coinciding with downsizing, owners in the process industry began benchmarking intensely (i.e., outsourced their database activity), so empirical research and modeling continued albeit mostly behind the confidential doors of the benchmarking firms. The central practice that evolved from the benchmarking and empirical research was the *phase-gate project system* which is often called front-end loading (FEL). AACE estimate classification system RPs were developed specifically to support phase-gate/FEL [16]. The main purpose of these systems is to manage systemic risks (e.g., the degree and quality of project scope definition) in a way that results in improved practices and outcomes. Since 1990, phase-gate systems have become ubiquitous in the process industries.

Unfortunately, phase-gate systems came late to the public transportation sector (AACE RP 98R-18 "Cost Estimate Classification – As Applied in Engineering, Procurement, and Construction for the Road and Rail Transportation Infrastructure Industries" reviews some national transport phase gate systems [17]). Hence, there has been little practice data sharing, research or benchmarking between regional or national government agencies and entities that fund and/or manage transportation projects. However, in the research review that follows, there is one enlightening example from the transportation industry.

*Research of Fundamental Practices and Cost Growth/Overruns*

At the 1956 founding of AACE International (then known as the American Association of Cost Engineers), the founders' collective experience had already demonstrated that the accuracy range of a cost estimate, and the forecasted cost growth (i.e., contingency required) were correlated with the level of scope definition: i.e., better definition = better accuracy and less contingency. It was also common practice at the time to phase project scope definition and investment decision making in a stepped manner [18]. The AACE founders captured this experience in their first standard entitled "Estimate Types" in 1958 [19]. The three estimate types/phases, were Order of Magnitude, Preliminary and Definitive.

1958 also marked the publication of the first empirically-based parametric cost risk quantification model by Mr. John Hackney, an AACE founding member [20]. Mr. Hackney went on to consult on or inform other empirical research of cost growth by The Rand Corporation (Rand) [21] and the Construction Industry Institute (CII) [22]. This evolving chain of research extends beyond the process industries into infrastructure, including transportation [23]. There is now over 60 years of evolving empirical research demonstrating that the quality of fundamental project practice, led by the level of scope definition (FEL) drives cost growth, and depending on how well risks are quantified at the time of project announcements, cost overrun.

The following sections review this chain of research. For each study a chart from the owner's perspective is provided showing the relationship of the percentage cost growth (before contingency) or cost overrun (after contingency) versus the level of definition. A purpose of this paper is to drive home the fact that the statistical relationship of <u>cost overrun to the level of definition is essentially the same in every study</u>. This is not theory – it is *settled science*. The fact that this research is not being widely used in day-to-day industry risk analysis is a travesty.

The following sections review high points of FEL research history

A. John Hackney: 1958 (Process)
B. Rand Corporation: 1981 (Process)
C. Construction Industry Institute (CII): 1995 (Industrial) versus 2011 (Infrastructure)
D. Independent Project Analysis, Inc (IPA): 2012 (Process-Longitudinal)
E. Grattan Institute: 2016 (Transportation)
F. Hollmann, et.al.: 2020 (Regulated Utilities)

A. John Hackney: 1958 (Process)

John W. Hackney CCE Hon. Life, a founder of AACE, is perhaps best known for his 1965 landmark text "Control and Management of Capital Projects" [24] (the 2nd edition is available from AACE.). The book defines a detailed project scope definition rating method as well as a parametric risk analysis model using this rating. However, he originally published the model in 1958 [20]. The second edition of his Capital Projects book describes a 1982 data update, noting that estimate accuracy had improved since the original model. Mr. Hackney's model reflected over 30 years of cost experience at process industry companies such as Alcoa, Diamond Alkali, and Mobil, but also active involvement with his peers in AACE and elsewhere. The Hackney model is available as a working Excel® model in AACE RP 43R-08 "Risk Analysis and Contingency Determination Using Parametric Estimating – Example Models as Applied for the Process Industries." [25]

The Hackney model ratings included not just the design status, but the level of technology and complexity. The rating was from 0 (best) to 1600 (worst). Of most interest in respect to transportation is that the rating includes an "Ownership Status" factor that multiplies the overall rating by 1.12 if the facility is in the "public sector" and 1.24 if "mixed private/public" (Private is 1.00). Given that "mixed" ownership is the most severe case, this likely reflects execution complexity (more stakeholder interaction/conflict) as opposed to bias. Figure 2 shows the overall finding of cost growth versus the 0-1600 definition rating. This was based on his study of 30 process plant projects for which he had full data access.

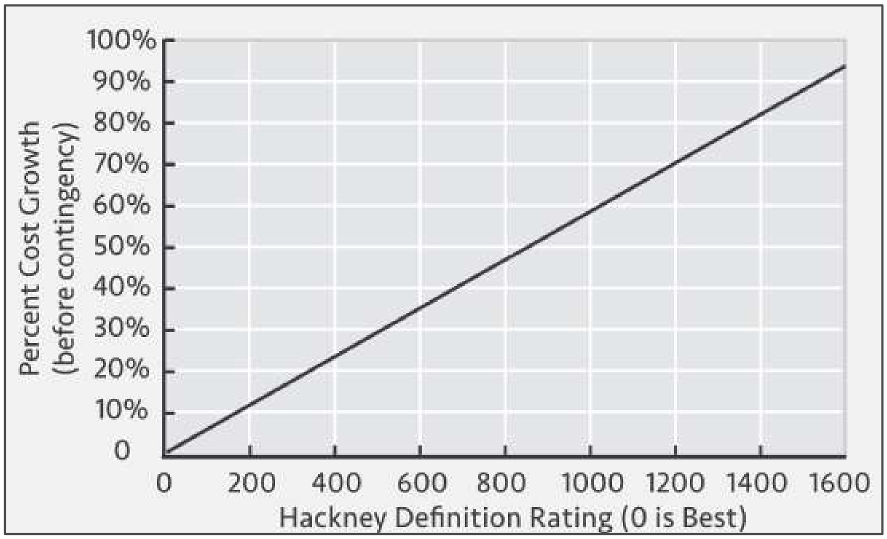


**Figure 2–Hackney Model (1958): Percent Cost Growth versus Level of Definition [20]**

B. Rand Corporation: 1981 (Process)

As part of an analysis program for the U.S. Department of Energy (DOE), The Rand Corporation studied the cost growth of chemical, oil and minerals projects from 34 companies and 106 estimates in North America [21]. The main objective was to understand the cost growth of pioneer process plants; however, to understand the impact of technology and complexity, the dataset covered a range of project types including those with more conventional technology and moderate complexity. Unlike Mr. Hackney's long list of causes based on his experience, Rand only retained *statistically significant* risk drivers in its model. In respect to the planning fallacy hypothesis (not known as such then), the study did drop observations from the analysis where estimators were directed by management to reduce their estimates; however, those were exceptions, not the rule.

The Rand study applied multiple linear regression (MLR) focused on the physical characteristics (e.g., complexity or complicating factors) of the plant, the level of technology and the level of scope definition and their relationship to cost overrun (after contingency). The Rand model rating of the level of project definition ranged from 2 (best) to 8 (worst). The model had a 0.82 correlation between this rating and cost overrun. The cost overrun metric was a ratio of the total estimated cost at that phase (including contingency) and the final total cost. The impact of escalation, major scope change, and major risk events was removed from the actual cost prior to calculating the metric. A working Excel® version of this model is also available in RP 43R-08 [25]. Figure 3 shows the overall finding of cost overrun versus the 0 to 8 definition rating (with all other risk drivers set to mean values; i.e., the cost overrun can be much more or less depending on those drivers).

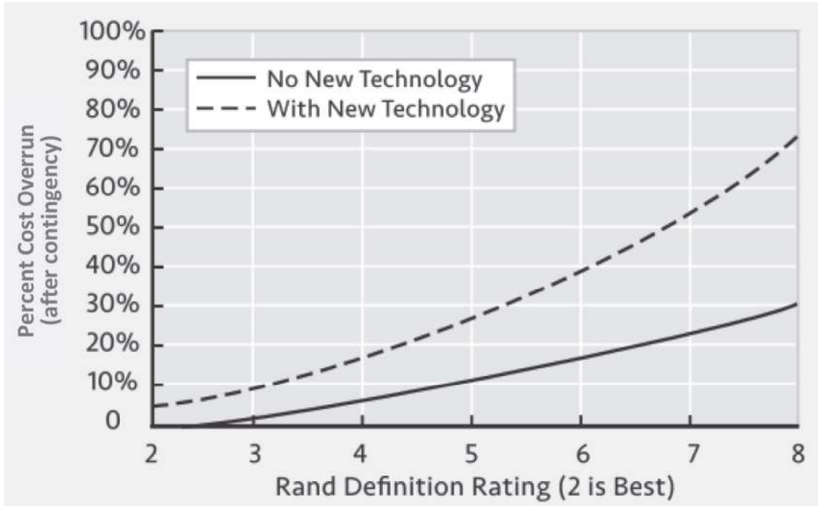


**Figure 3–Rand Model (1981): Percent Cost Overrun versus Level of Definition [21]**

Note that the Hackney model (Figure 2) is of cost growth from the base estimate excluding contingency while the Rand model (Figure 3) is cost overrun of the total cost including contingency. In large part, this explains why the Hackney values are higher. If companies had allowed for contingencies consistent with Hackney's model, the Rand model should not have shown average overrun; i.e., Figure 3 is showing the percentage that contingencies were underestimated.

In 1987, the lead Rand study researcher, Mr. Edward Merrow, launched a private benchmarking and research firm, Independent Project Analysis, Inc. (IPA). The firm started with and built on the data acquired from Rand; originally calling its products "project risk analyses". The phase-gate approach IPA developed working with its industrial clients was called "front-end loading" (FEL); a term that quickly caught on. IPA's measure of scope development was called the FEL Index. IPA is largely responsible for keeping empirically-based parametric contingency modeling alive when companies, lacking their own data, became enamored with MCS. The results of research using the FEL index will be discussed later.

C. Construction Industry Institute: 1995 (Process) versus 2011 (Infrastructure)

Following on the success of FEL, the Construction Industry Institute (CII) implemented a study to develop a ''project definition rating index for industrial projects" (PDRI) [26]. Unlike Rand, which only included statistically significant measures in its FEL Index, the CII team returned to a long list (70) of subjectively weighted elements similar to Mr. Hackney's model. CII weighted its PDRI elements based on a survey of the research team. The PDRI ranges from 0 (best) to 1000 (worst). Using linear regression of actual data, the team found an R-squared value of 0.40 between the industrial PDRI and cost overrun. Later PDRI research confirmed the findings [22]. Figure 4 shows the CII finding of cost overrun versus the 0 to 1000 definition rating (none of the 23 industrial projects tested were rated worse than 500).

Like Rand, the overrun metric is relative to the estimate including contingency. If companies had allowed for contingencies consistent with Hackney's model, the CII model should not have shown average overrun; i.e., Figure 4 shows the percentage that contingencies were underestimated.

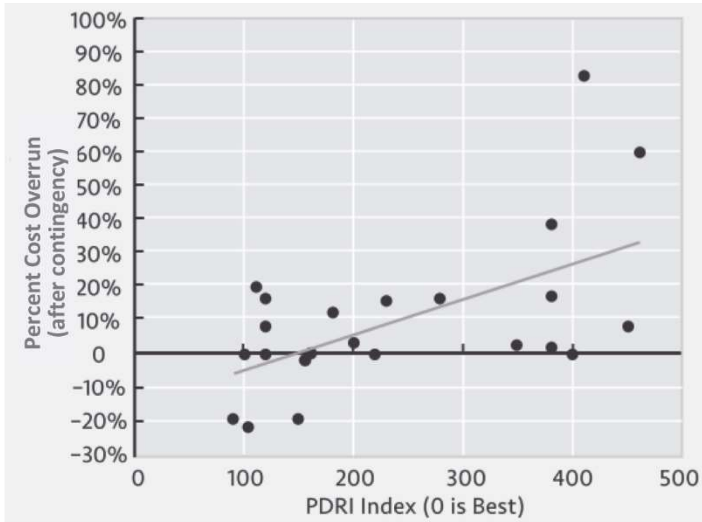


**Figure 4–CII Industrial PDRI (1995): Percent Cost Overrun versus Level of Definition [26]**

There are now PDRIs for multiple industries including one for infrastructure projects [23]. In the CII approach, infrastructure includes linear mode projects such as pipelines and *transportation*. Using linear regression of actual data on 22 projects (including 11 for highway, rail, tunnels and airport runway), the team found an R-squared value of 0.47 between the infrastructure PDRI and cost overrun. Figure 5 shows the CII finding of cost overrun versus the 0 to 1000 definition rating (none of the 22 infrastructure projects tested were rated worse than 500). A comparison of Figures 4 for industrial and 5 for infrastructure indicates the strong similarity in cost overrun versus level of definition for these physically different industry segments. Figure 5 shows somewhat greater overrun, which would be consistent with Hackney's "ownership" factor findings.

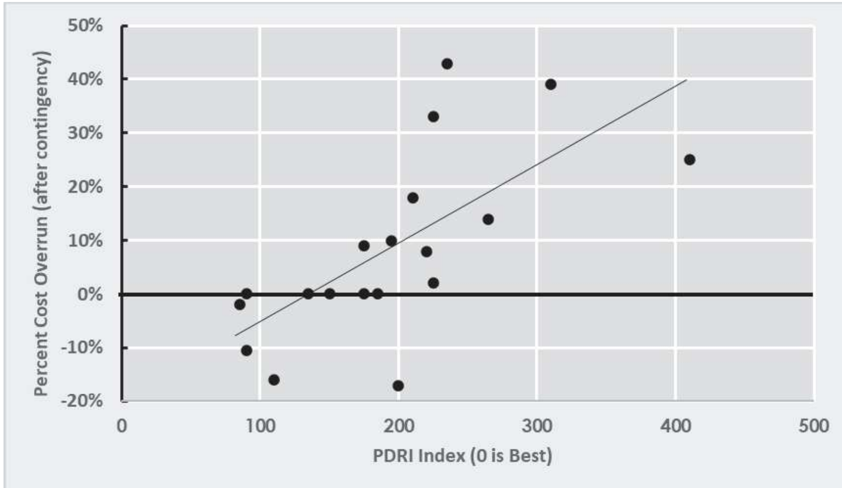


**Figure 5–CII Infrastructure PDRI (2011): Percent Cost Overrun versus Level of Definition [23]**

D. Independent Project Analysis, Inc (IPA): 2012 (Process-Longitudinal)

The Hackney, Rand and CII research examined data retrospectively; i.e., after the project was complete (sometimes ten or more years later). If cost growth or overrun were low or high, this may influence one's judgment of what the level of definition had been when documentation of practices is less than ideal (i.e., if cost growth was high, one might suspect it was poorly defined). IPA's benchmarking measurement of the level of definition is contemporaneous with estimate preparation, thereby avoiding retrospective bias. In addition, IPA often measures the estimate status at each phase of scope development; i.e., longitudinally. Finally, IPA's data is from their benchmarking clients which arguably have better practices than firms that do not benchmark. Also, IPA removes the impact of catastrophic risk events. IPA published a paper in 2012 that showed how the distribution of cost overrun phase-by-phase varied. Figure 6 shows the median (p50) and mean of those distributions. The higher mean values indicate high-side skewing [27]. The overrun indicates that as of 2012, companies were still underestimating needed contingency, particular at FEL 1.

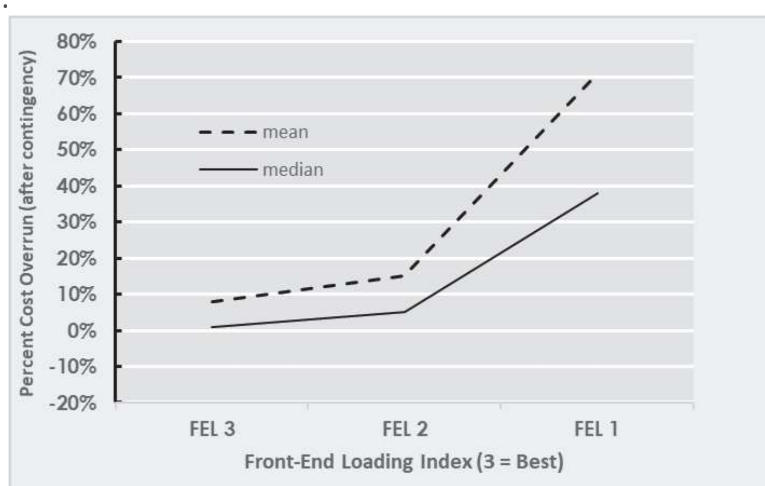


**Figure 6–IPA FEL Index (2012): Percent Cost Overrun versus Level of Definition [27]**

E. Grattan Institute: 2016 (Transportation)

In 2016, the Grattan Institute studied 836 transportation projects valued at A$20 million or more and planned or built since 2001 in Australia [28]. From this dataset, Grattan researched 51 projects valued at greater than A$100 million in more detail. This is one of the few transportation project cost overrun studies that examine cost growth by scope development phase. The phases are 1) from the first project announcement to formal funding commitment; 2) from the formal funding commitment to the start of construction; and 3) from the start to the end of construction. These roughly correspond to AACE Class 4 (FEL 2), 3 (FEL 3) and 2 (Execution) estimates respectively.

The study found that "*premature announcement"* and presumed underestimation of contingency at early phases is the primary cause of overruns. Note that premature announcement is not the same as misrepresentation (i.e., planning fallacy); the overrun is the result of industry's failure to factor poor scope definition and other risks into their contingency. In other words, industry's failure to use the research surveyed in this paper. Figure 7 shows the mean cost growth from the phases studied. The lower curve represents the entire dataset (836 projects) dominated by smaller projects. The upper curve is for the 51 larger projects studied in more detail as to the cause of overruns.

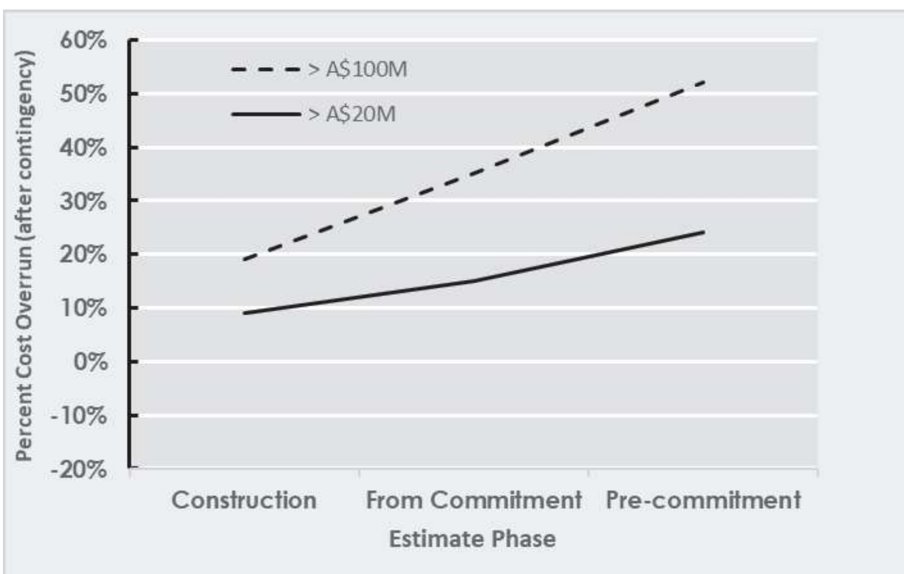


**Figure 7–Grattan Institute (2016): Percent Cost Overrun versus Estimate Phase [28]**

F. Hollmann, et.al.: 2020 (Regulated Utilities)

From 2014 to 2020, a series of cost growth studies were conducted by a group of North American power utility companies. In 2020, the combined results were shared [29]. The combined datasets included 89 hydropower, power generation and power transmission projects and 214 phased estimates. While not transportation projects, these represent regulated infrastructure. The study looked at cost growth from the base estimate excluding contingency (i.e., the outcome reflects the contingency that was needed). Cost growth was measured from each phase using the AACE classification system with Class 5 being the earliest conceptual estimate to Class 3 being the funded amount (some were based on tender pricing; i.e., Class 2). The study also looked at the 80 percent confidence interval (i.e., the accuracy range). Figure 8 shows the median (p50) cost growth and the range at each AACE Class.

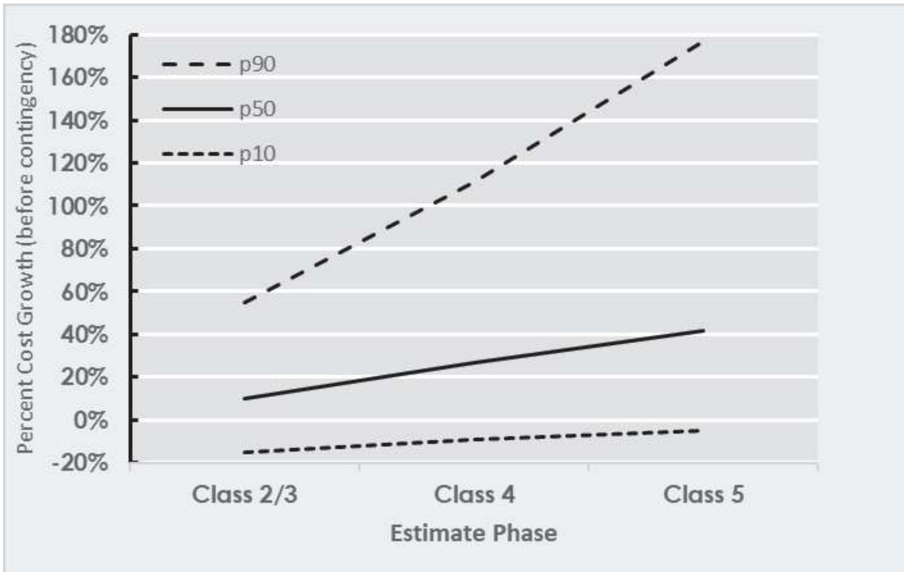


**Figure 8–Hollmann, Power Utilities (2020): Percent Cost Overrun versus Estimate Phase [29]**

The study found the usual underestimation of contingency at each phase is the primary cause of overruns. One thing that stood out in this study was that the actual average contingencies allowed which were 10%, 11% and 16% respectively for Class 3, 4 and 5 estimates respectively. The practices were similar in all the participating companies in many regions working under a variety of regulators. Compare that to the contingency required in Figure 8 of 10%, 27% and 42%. The early estimates allowed about 1/3 the contingency needed. No deception was at play here; these major utility companies were simply using contingency rules-of-thumb or subjective risk quantification methods without empirical basis. The other thing that stands out is the p90 range. As is typical for every study the author has been associated with in the last 25 years, the p90 of reality is 2 to 3 times what the company forecasts the p90 will be [6].

## Planning Fallacy Versus Fundamental Practice Failure

Having presented the two cost overrun cause theories, the next step is to compare the findings of their research. The 2002 Flyvbjerg study, purporting to be one of the most comprehensive studies of transportation data with 258 projects, presented only the mean and distribution of cost overruns. As a fitting first comparison, the 2012 IPA study of mostly process projects is used because it is also based on a very large dataset of 462 projects and it also presents means and distributions, including by estimate phase (i.e., FEL 1, 2 and 3).

Figure 9 overlays the Flyvbjerg and the IPA distributions using the IPA data for FEL 2 (AACE Class 4) estimates. In defined phase-gate systems, FEL 2 is the phase when a single option is selected, and the most likely phase that a cost value would be shared. The means are 28% versus 15% for Flyvbjerg and IPA respectively and the standard deviations are 39% versus 41%. The 13% greater mean of the Flyvbjerg dataset is statistically significant (unpaired t-test, t=4.15).

In terms of what this says about the theories, we know from IPA that the process industry overrun is not the result of deception. As Merrow has stated "*no Machiavellian explanation is required*." [5] IPA's 40 years of quantitative benchmarking and practice research starting from the Rand study in 1981 clearly shows that fundamental project practices (including underestimated contingency) explain process industry cost overruns. These leave the possibility that the 13% difference in the mean values could in part reflect optimism bias in public sector transportation projects. However, IPA (and Rand before that), having access to detailed information, removed the cost of major scope changes and catastrophic risk events (as is standard analysis practice when the cause of cost changes is known) while the Flyvbjerg data is only corrected for escalation.

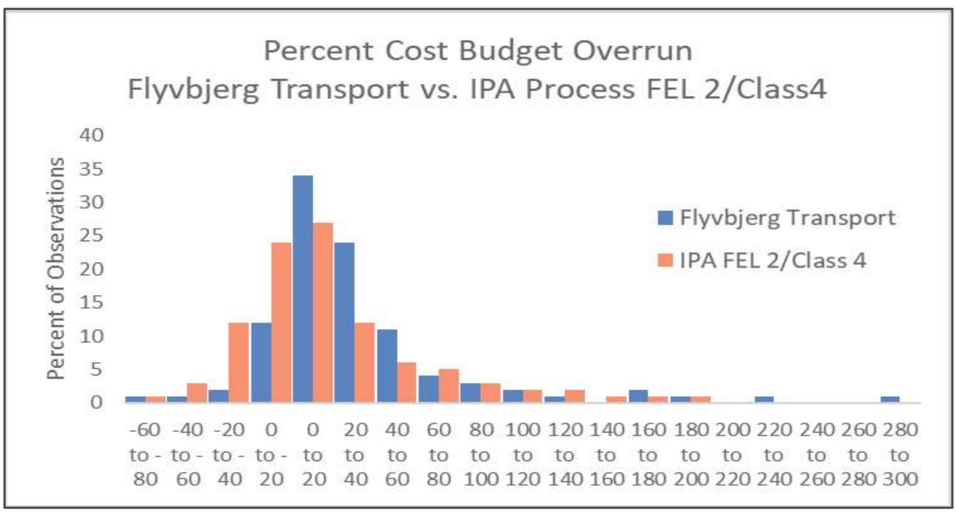


**Figure 9–Percent Cost Overrun Distribution: Flyvbjerg (Transportation) vs. IPA (Process)**

The 2016 Grattan study of 834 transportation in projects in Australia had an average overrun of 24% overrun from the pre-commitment announced value which is similar to the Flyvbjerg dataset's 28% average. However, the Grattan study also measured cost overrun from the *commitment* estimate and that was only 13%. Arguably, this 13% represents public sector optimism bias, at least in part. However, the Grattan study concluded that most of the overrun was explained by "premature announcement". The study also recognized the need to improve risk quantification practices informed by real data (i.e., the overrun represented under-estimated contingency).

Finally, the paper shared the 2019 study of 85 transportation projects in Hong Kong by Dr. Love, et. al. which found a mix of optimism and pessimism with 47% of the projects were delivered for less than their approved budget. There was no sign of the planning fallacy at all.

In summary, ascribing all cost overrun to the misrepresentation ("the planning fallacy writ large [9]") is not justified. However, there may be optimism bias in the publicly funded transportation sector relative to the process industry (perhaps in the neighborhood of 5 to 10%). However, bias can be readily studied and calibrated in a parametric risk model. In fact, the author regularly conducts model calibration (i.e., prediction "bias"; not elicitation bias) studies for industrial clients that implement the parametric method. As found in the Hong Kong study by Love at. al, the bias, if present, can be either optimistic or pessimistic (bias studies will be addressed in a future paper).

However, to informed cost engineers and risk analysts what should leap out from all these charts is that our prevailing contingency estimating practices are, without question, utterly failing to quantify the risk of lagging scope definition and the bias (optimistic or pessimistic). Further, the efficacy of risk quantification practices has been studied, and subjective contingency estimating methods were found to be a "disaster" when the level of scope definition was lagging [30]. What is worse (or better if viewed as an opportunity) is that the underestimation of contingency for Class 4 and 5 (FEL 2 and 1) estimates is *extraordinarily consistent* whether it is in the process or the transportation industry (all industries are in the same boat). This disconnect is not a known-unknown, it is a *known-known*. Industry professionals must use empirically-valid, risk-driven methods; based on the facts presented here; to not do so, at least as a check (the outside view), would be unethical.

**Empirically-Based Risk Quantification Practices**

The chain of Hackney-Rand-CII research has resulted in parametric risk quantification models in which the parameters are risk drivers, and the output is cost growth (and schedule slip which is not covered in this paper) [25]. The empirical research has identified the primary project system attributes and practices that drive cost growth; i.e., systemic risks. These systemic risks include:

- Level of Scope Definition
- Level of Technology
- Level of Complexity
- Team Development
- Project Management/Control Maturity/Capability
- Process or Service Severity
- and Bias

The level of scope definition is the primary risk driver; the following section defines how this can be measured in the transportation industry. For the other risk drivers, an example measurement scheme (usually a Likert scale or similar) and a parametric model have been proposed by the author [31]. Complexity is a perhaps the systemic risk driver of most active industry interest (including non-linear impacts in modeling), but is beyond the scope of this paper [32].

*Phase-Gate/Estimate Classification in Transportation*

For transportation, the measurement of the level of scope definition can be done using AACE RP 98R-18 that defines a classification system for road and rail projects [17]. Class 5 is the earliest conceptual estimate when little is defined beyond overall system capacity and there are multiple options and estimates considered within that broad definition. Class 4 is where enough design is done such that a single project option is selected for more engineering. Class 3, in all industries, is where scope and planning uncertainty are reduced to a level where a prudent decision maker can make a full funds commitment to a project [16]. Class 2 represents the definition when the main construction tender is received. Class 1 usually is only for change order estimates during execution.

A consistent measure of the level of scope definition can be a challenge for transportation projects in that each country or region has its own phase-gate system, often without well evolved deliverable specifications. Figure 10 shows some examples of national systems (subject to change) [17]. Note that Norway has the closest match to AACE Class (perhaps due to the importance of the oil industry where the use of Class is well established). Also note that Class 4 or "pre-study" is when project costs are first publicly known; this is likely the average for most places. However, other regions are less clear as to what the phases represent in respect to AACE Class. Some seem to assess (a defacto announcement if not kept confidential) costs as early as Class 5 which would explain some of the worst cost overruns.

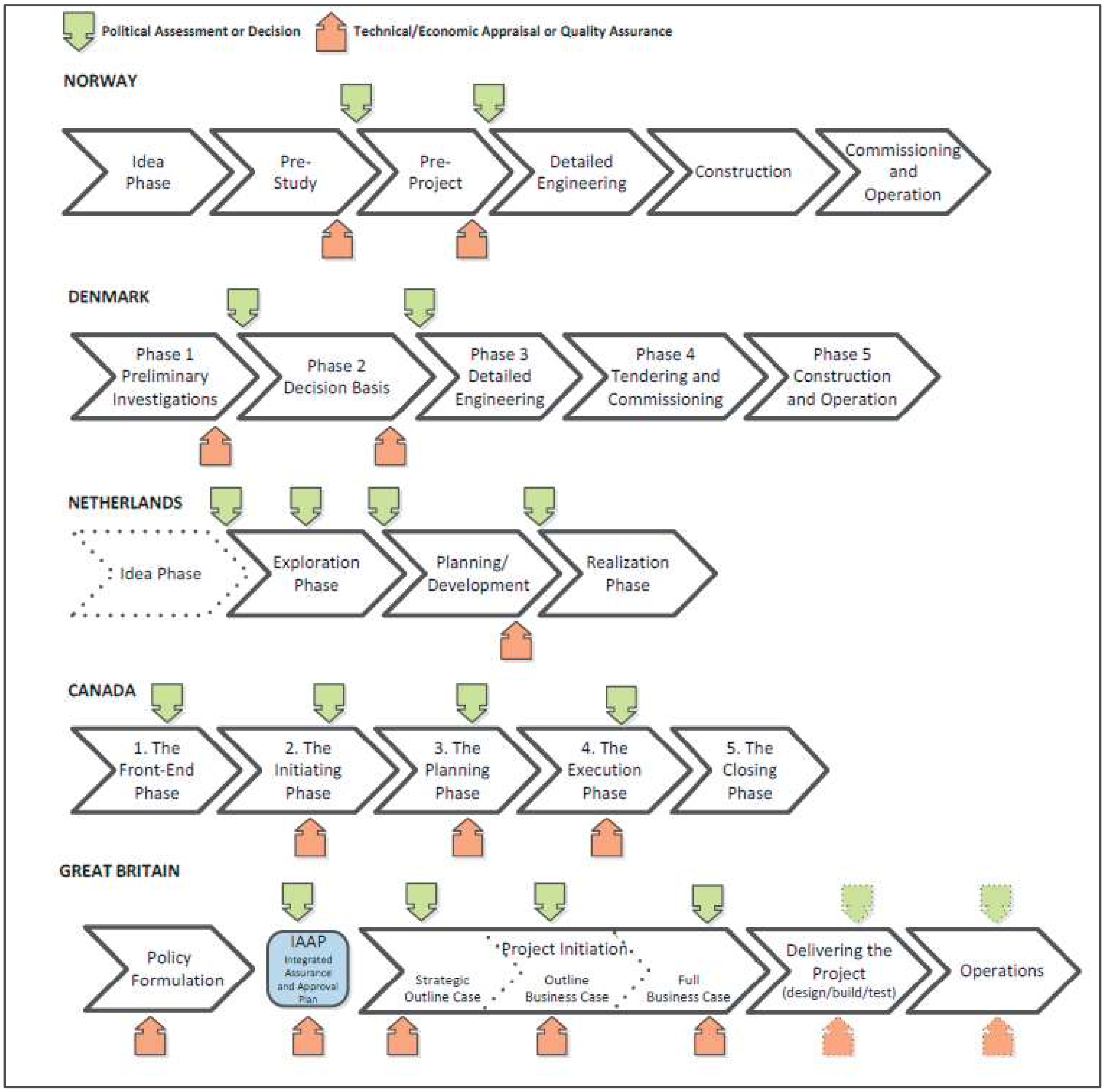


**Figure 10–Example National Transportation Phase-Gate Schemes [17]**

AACE classification is not the only measure of the level of scope definition upon which an estimate or budget can be based. The FEL index developed by IPA is widely used in the process industries. The CII PDRI index, like AACE Classification, has versions for multiple industries including infrastructure of which transportation is a part [23]. Class, FEL and PDRI are interchangeable; it requires only a simple conversion factor.

*Parametric Modeling of Systemic Risks and Hybrid Approaches*

AACE has developed a series of recommended practices (RPs) for risk quantification; they are summarized in Professional Guidance Document PGD-02 [33]. The practices recommended are predicated on a number of principles documented in RP 40R-08 "Contingency Estimating – General Principles" [34]. One principle is that the method be risk-driven; i.e., the risks and impacts are clearly linked. Another is that they employ empiricism (perhaps the most overlooked requirement). In addition, the risk quantification methods should align with the type of risks; i.e., systemic, project-specific (i.e., risk events), escalation or currency. If the principle of empiricism is honored, then the only empirical method for quantifying systemic risk is MLR-based parametric modeling which is the focus of this paper.

For project-specific risks (i.e., risk events), the models are usually MCS-based and the risk probability and impact inputs are mostly subjective (though some risks such as weather-driven uncertainty can be historically-based). MCS can be applied to an expected value model or a critical path method (CPM) schedule model (as with any MCS model, fixed but uncertain values such as probability and impact are replaced with distributions). Both of these models can be integrated with parametric modeling in a hybrid approach; i.e., the systemic and project-specific risks are all quantified in an integrated way resulting in a single probabilistic output distribution. The parametric model can be used alone for Class 5 estimates when no details are available and the level of scope definition is by far the greatest source of uncertainty. Otherwise, hybrid approaches should be used. As was mentioned, risk experts need not fear for their livelihoods because of the ease of parametric models; team and expert input is vital to hybrid, integrated risk quantification methods (and some day for AI approaches). Figure 11 illustrates the hybrid concept.

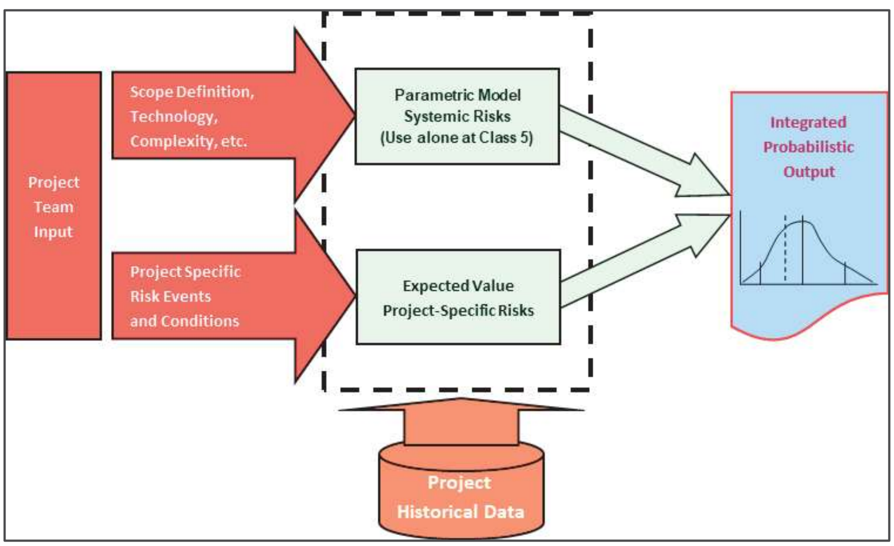


**Figure 11–Use of Parametric Modeling of Systemic Risk in a Hybrid Approach**

The AACE RPs for parametric modeling (Class 5 estimates) and hybrid methods (for Class 4 and better estimates) include the following:

- RP 42R-08: Parametric modeling [35]
- RP 43R-08: Example parametric models (includes Excel® versions of the Hackney and Rand parametric models) [25]
- RP RM-30: Cost Estimate Accuracy Range and Contingency Determination using Tables Derived from Parametric Risk Models [36]
- RP 113R-20: Hybrid Parametric + Expected Value method with MCS [37]

- RP RM-31: Hybrid Parametric + CPM-based with MCS [38]

*A Parametric Model Applied to Transportation (and Comparison to Actual Data)*

Figure 12 is a snapshot of an example demonstration parametric risk model developed by the author (contact the author to obtain a copy). The model is documented in the author's book "Project Risk Quantification" [31]. The model is based on the Rand model [25] supplemented with the latest published research on systemic risks (as addressed in the book) such as team development and project control.

| RISK DRIVER | ENTER PARAMETER (a) | COEFFICIENT (b) | a*b |
|---|---|---|---|
| CONSTANT | | | -30.5 |
| SCOPE | 4 | | |
| PLANNING | 4 | | |
| ENGINEERING | 4 | | |
| SCOPE DEFINITION | 4.0 | 9.8 | 39.2 |
| NEW TECHNOLOGY | 0% | 0.12 | 0.0 |
| SERVICE SEVERITY | 4 | 1 | 4.0 |
| COMPLEXITY | 4 | 1.2 | 4.8 |
| SUBTOTAL BASE (prior to adjustments) | | | 17.5 |
| ADJUSTMENTS | Complex Exec Strategy? > | No | |
| Team Development | Fair | | 0 |
| Project Control | Fair | | 0 |
| Estimate Basis | Poor | | 3 |
| Equipment (rolling stock) | 10% | mix highway/rail | 4 |
| Fixed Price | 0% | not fixed at time of est | 0 |
| TOTAL BASE (prior to bias adjustment; rounded to whole number) | | | 25 |
| Bias | Low | somewhat aggressive | 5 |
| SYSTEMIC COST CONTINGENCY (at shown chance of underrun) | | | |
| MEAN | | | 30% |
| 10% | | | -7% |
| 50% | | | 26% |
| 70% | < enter funding p-level (increments of 10%) | | 43% |
| 90% | | | 71% |

**Figure 12–Example Parametric Risk Model with Example Transportation Inputs**

The way the model is used is that systemic risk ratings are entered in the yellow cells, and the probabilistic output is reported at the bottom (there is no MCS; the range is determined in the tool algorithm). The main analysis challenge is to obtain objective rating information from project management. In the example, the author entered assumed ratings for a typical transportation project as reflected in the Flyvbjerg study [8]. The primary entry is the level of scope definition at the time of the estimate. In the example, this was assumed to be AACE Class 4 which is the most likely stage that estimated costs become known or announced per Figure 10. No new technology is assumed and complexity and service severity (e.g., will trains need to handle 100% capacity 24/7?) are moderate on scale of 0 to 10. Team development and project control capability are rated as "fair", but the estimate basis (e.g., data quality, knowledge of market, etc.) is rated as "poor". The project includes some rolling stock acquisition (lower risk for equipment purchases than for installation). The estimate bias is rated low; i.e., it is somewhat aggressive (i.e., some optimism bias). The bottom of Figure 12 shows the output percentage cost growth, i.e., these are percentages that history says we need to add to the base estimate to address systemic risks (major risk events are extra).

In the example shown in Figure 12, assuming a $100 base estimate, the model suggests a p10/p90 cost range of $93 (-7%) to $171 (+71%). It also suggests a contingency (not counting event risks) of $26 (+26%) at p50. Assume now that the model cost range reflects actual outcomes (as is expected). However, to compare this model outcome to the Flyvbjerg actual outcomes, the project team must be assumed to have set continency based on typical industry practice which is not empirically valid. Assume that instead of 26% contingency, the team assigned 11% for the Class 4 estimate as was the case in the regulated utility study [29] so $111 was announced (notice decision maker has no role in this example; they simply report what they were given by estimating).

If the parametric model reflects the actual outcome, the result of this underestimated contingency would be a cost overrun of 14% ($126 vs. $111) to p50 and a p10/p90 range of -16% ($93 vs. $111) and +54% ($171 vs. $111). The Flyvbjerg actual cost overrun distribution is shown in Figure 13. Based on this histogram, the p50 cost overrun was approximately 20% and the p10/p90 was approximately -20% to +60%.

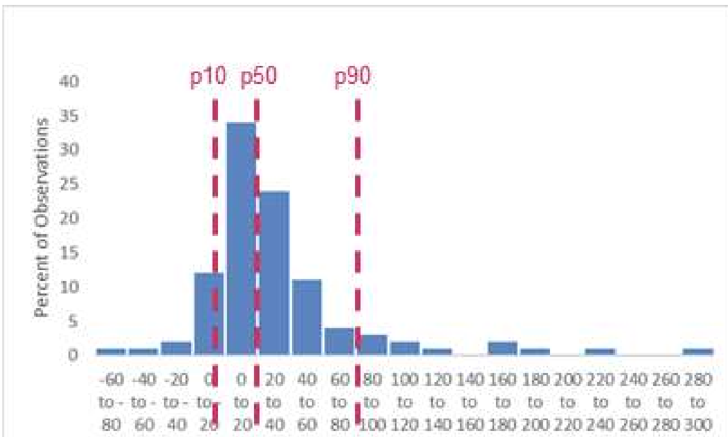


**Figure 13–Flyvbjerg 2002 Cost Overrun Distribution [8] with p10/p50/p90 Added**

Table 1 compares the model prediction (assuming contingency was set at 11%) and the Flyvbjerg actual cost overrun distribution. At first glance, the Flyvbjerg data shows a 6% greater cost overrun than the parametric model at p50; however, the parametric model only accounts for systemic risk! Allowing for approximately 5% cost increase caused by risk events (not untypical), the model predicts the actual outcomes very well. No deception is required: only the underestimation of contingency! Why are industry cost engineers and risk analysts not using simple, empirically valid risk models that are freely available? It is a tragedy of colossal economic proportion given the trillions of dollars spent annually on transportation projects. As professionals, we should not sit back while others are falsely accused of being “liars”. Per Pogo, we have met the enemy and he is us.

| Cost Overrun % | P50 | P10 | P90 |
|---|---|---|---|
| **Parametric Model** | +14% | -16% | +54% |
| **Flyvbjerg 2002** | +20% | -20% | +60% |

**Table 1–Comparison of Parametric Risk Model Prediction vs. Flyvbjerg Cost Overrun Data**

*Institutionalized Bias*

As a final point, it is necessary to discuss a ubiquitous, but also pernicious industry practice of including expected estimate accuracy ranges in company and agency process documentation and standards. In 23 years of benchmarking and cost estimate review consulting, the author has reviewed or studied the phase-gate processes and estimates of scores of companies, agencies and their contractors, including in the public sector. Virtually every company with a phase-gate project system includes “expected” or typical accuracy ranges for estimates at each phase [6]. The classic example is the expectation that Class 3 (funding) estimate accuracy should be +15/-10%, +/-10% or similar. The problem is that this and similar ranges have no empirical basis and the ranges (which reflect an assumed contingency[2]) are almost always too narrow. However, anchoring bias (having been exposed to these quasi-standards one’s whole career) makes these ranges impossible to dislodge from both decision maker and analyst’s minds. Confirmation bias then takes over and infects prevailing subjective risk analyses (expert or otherwise) such that the predicted range and contingency always looks like the expectation published in the procedure; it is uncanny.

AACE recognizes this problem and has fought valiantly to discourage this pre-stated range practice. Its first estimate classification system recommended practice in 1998 shifted to a “range-of-ranges” approach and further stated that ranges should only be based on analysis of a given project’s risks [16]. The communication struggle continues to the present with recommended practice RP 104R-19 “Communicating Expected Estimate Accuracy” [39].

Where this leaves us is back to the fifth hand principle. Fundamental practice failure drives cost growth; this is settled science as shown by the many referenced studies in this paper. But the planning fallacy is also real. However, contrary to Flyvbjerg’s presumption of willful lying or deception on the part of promoters and politicians, the “lie” of (usually) narrow accuracy ranges and low contingency is wired right into company (but not professional) standards. The bias has been institutionalized. Unfortunately, Flyvbjerg’s proposed solution of RCF will, in effect, replace the too optimistic ranges with too pessimistic ones (i.e., every risk that occurred in the reference class is assumed to apply to every project). The only effective solution is empirically-based, *risk driven* (reflecting only those practices and risks that are relevant), risk analysis.

[2] Contingency set at the p50 confidence level for a base estimate with typical bias, the contingency has been observed to be approximately equal to the standard deviation of range. Hence, a stated range expectation is also stating an expected contingency [40].

## Conclusions

This paper makes the case for parametric quantification of systemic risks for transportation projects. The problem of *cost overruns* in this industry is well known. Two main theories as to the cause of cost overruns in the public transportation sector were reviewed: i.e., the *planning fallacy* (lying, misrepresentation, optimism bias) and the *fifth hand* (mix of fundamental practice failure and bias, optimistic or pessimistic). The paper shares decades of empirical research, mainly from the process industry, that unambiguously shows the main cause of *cost growth* is the level of scope definition and other systemic risks such as the level of complexity. However, the research also clearly shows that the cause of *cost overrun* is that our contingency estimates are too low; cost engineers and risk analysts are not applying the research and empirically-based tools at their disposal. Promoters and politicians, who benefit from low-cost estimates, do not need to lie. They can simply announce a project cost early in scope development in perfect confidence that cost and risk professionals have not put a price on the poor definition and other systemic risks in our estimates.

One reason given by the author's peers for not using empirically-based contingency estimating is that it is not based on "their" data so it cannot be trusted. Another is that a method that quantifies overall outcomes is not useful for examining risk at a detail level in a schedule or estimate. Others argue that risks are unique (e.g., black swans or unknown, unknowns) and the past does not foretell the future. None of these arguments stand up to the findings of empirical research. This paper shows how the research-based model is generic to process and transportation projects. Further, details of a schedule or estimate are irrelevant when the main risk is that there is little detail. Finally, research has shown that empirically-based models are the most predictive (i.e., risk events are of relatively minor consequence compared to systemic risks).

However, one final reason for not using empirically-based contingency estimating considering systemic risk is that it often results in forecasts that differ from accuracy range expectations wired into company phase-gate procedures and management's minds. The risk analyst (or consultant) reporting a +54%/-16% range (as in this paper's transportation example) when the company guideline book says a Class 4 estimate range is +30%/-15% is going to have a miserable experience. Estimators, being all too human with careers to consider, will stick to subjective risk analysis methods that almost universally meet management's expectations.

The paper concludes that the failure to use empirically-valid risk quantification methods is a tragedy of colossal economic proportion given the trillions of dollars spent annually on transportation projects. Cost and risk professionals should not sit back while others are falsely accused of being "liars". Per Pogo, "we have met the enemy and he is us".

Despite all evidence to the contrary, the author is hopeful that empirically-valid contingency practices will become more widely recognized and used. AACE recommended practices for empirically-based parametric modeling of systemic risks are shared (for use either alone for Class 5 estimates or in hybrid with methods for event risks for Class 4 and better estimates.) Better yet, parametric models are freely available.

**References**


1. A. Maslow, The Psychology of Science, New York NY: Harper & Row, 1966.
2. B. Ashuri, "Benchmarking in the Infrastructure Sector," Royal Institute of Chartered Surveyors (RICS), London UK, 2020.
3. L. Ika, P. Love, and J. Pinto, "Moving Beyond the Planning Fallacy: The Emergence of a New Principle of Project Behavior," *IEEE Transactions on Engineering Management ,* Accepted for publication November 19, 2020.
4. AACE International, "Recommended Practice No. 10S-90; Cost Engineering Terminology," AACE International, Morgantown WV, Latest revision.
5. E. Merrow, "Why Large Projects Fail More Often; Megaproject Failures: Understanding the Effects of Size," in *AACE National Capital Section Meeting*, Washington DC, April 20, 2011.
6. J. Hollmann, "Estimating Accuracy: Dealing with Reality," *Cost Engineering Journal,* Nov/Dec 2012.
7. W, Kelly, We Have Met the Enemy, and He is Us, Simon & Shuster: New York NY, 1972.
8. B. Flyvbjerg, M. Holm, and S. Buhl, "Underestimating Costs in Public Works Projects Error or Lie," *Journal of the American Planning Association,* vol. 68, no. 3, Summer 2002.
9. B. Flyvbjerg and C. Sunstein, "The Principle of the Malevolent Hiding Hand; or, The Planning Fallacy Writ Large," *Social Research,* vol. 83, no. 4, pp. 979-1004, 2015.
10. D. Lovallo and D. Kahneman, "Delusions of Success: How Optimism Undermines Executives' Decisions," *Harvard Business Review,* July 2003.
11. B. Flyvbjerg, "Curbing Optimism Bias and Strategic Misrepresentation in Planning: Reference Class Forecasting in Practice," *European Planning Studies,* January 2008.
12. P. Love and D. Ahiaga-Dagbui, "Debunking fake news in a post-truth era: The plausible untruths of cost underestimation in transport infrastructure projects," *Transportation Research Part A,* vol. 113, p. 357–368, 2018.
13. B. Flyvbjerg, et.al., "Five things you should know about cost overrun," *Transportation Research Part A: Policy and Practice,* vol. 118, no. December, 2018.
14. P, Love, M. Sing, L. Ika, and S. Newton, "The cost performance of transportation projects: The fallacy of the Planning Fallacy account," *Transportation Research Part A,* vol. 122, pp. 1-20, 2019.
15. D. Ahiaga-Dagbui, "Reference Class Forecasting: A clear and present danger to cost-effective capital investment on major infrastructure projects," The Government's Management of Major Projects Inquiry, Public Administration and Constitutional Affairs Committee (UK House of Commons), London UK, May 2019.
16. AACE International, "Professional Guidance Document No. PDG-01, Guide to Cost Estimate Classification Systems," AACE International, Morgantown WV, Latest revision.
17. AACE International, "Recommended Practice No. 98R-18; Cost Estimate Classification – As Applied in Engineering, Procurement, and Construction for the Road and Rail Transportation Infrastructure Industries," AACE International, Morgantown WV, Latest Revisions.
18. H. C. Bauman, "Ratio Cost Engineering," in *Annual Meeting of the American Association of Cost Engineering* , Cleveland OH, 1958.
19. J. M. Gorey, "Estimate Types: A proposal by the Estimating Committee," *AACE Bulletin,* vol. 1, no. 1, pp. 12-13, 1958.

20. J. W. Hackney, "Special Report on Costs: What Affects Estimate Accuracy?," *Petroleum Refiner,* vol. 37, no. 6, pp. 128-134, 1958.
21. E. Merrow, K. Phillips and C. Myers, "R-2569 DOE: Understanding Cost Growth and Performance Shortfalls In Pioneer Process Plants," The Rand Corporation, Santa Monica CA, 1981.
22. S. Trost and G. Oberlender, "Predicting Accuracy of Early Cost Estimates Using Factor Analysis and Multivariate Regression," *Journal of Construction Engineering and Management,* vol. 129, no. 2, pp. 198-204, 2003.
23. E. Bingham, G. Gibson, R. Stogner, "Research Report 268-11; Development of the Project Rating Definition Index for Infrastructure Projects," Construction Industry Institute, Austin, TX, 2011.
24. J. W. Hackney, Control and Management of Capital Projects, New York, NY: John Wiley & Sons, 1965.
25. AACE International, "Recommended Practice No. 43R-08: Risk Analysis and Contingency Determination Using Parametric Estimating – Example Models as Applied for the Process Industries," AACE International, Morgantown WV, Latest revision.
26. G. Gibson and P. Dumont, "Report 113-11; Project Definition Rating Index for Industrial Projects," Construction Industry Institute, Austin TX, December 1995.
27. A. Ogilive, R. Brown, F. Biery and P. Barshop, "Quantifying Estimate Accuracy and Precision for the Process Industries: A Review of Industry Data," *Cost Engineering Journal,* pp. 28-38, Nov/Dec 2012.
28. M. Terrill and L. Danks, "Cost Overruns in Transport Infrastructure," Grattan Institute, Melbourne, VIC, October 2016.
29. J. Hollmann, D. Clark, et.al., "Variability in Accuracy Ranges: A Case Study in the US and Canadian Power Industry," in *AACE International Transactions*, Morgantown WV, 2020.
30. S. Burroughs and G. Juntima, "Exploring Techniques for Contingency Setting," in *AACE International Transactions*, Morgantown WV, 2004.
31. J. Hollmann, Project Risk Quantification; Chapter 11-Systemic Risks and the Parametric Model, Sugarland TX: Probabilistic Publishing , 2016.
32. J. Hollmann, Project Risk Quantification; Chapter 15-The Tipping Point: Risk Analysis at the Edge of Chaos, Sugarland TX: Probabilistic Publishing, 2016.
33. AACE International, "Professional Guidance Document No. PGD-02: Guide to Quantitative Risk Analysis," AACE International, Morgantown WV, Latest revision.
34. AACE International, "Recommended Practice No. 40R-08: Contingency Estimating - General Principles," AACE International, Morgantown WV, Latest revision.
35. AACE International, Recommended Practice No. 42R-08, Risk Analysis and Contingency Determination Using Parametric Estimating, Morgantown, WV: AACE International, Latest revision.
36. AACE International, "Recommended Practice No. RM-30: Cost Estimate Accuracy Range and Contingency Determination using Tables Derived from Parametric Risk Models," AACE International, Morgantown WV, Latest revision.
37. AACE International, "Recommended Practice No. 113R-20: Integrated Cost and Schedule Risk Analysis and Contingency Determination Using Combined Parametric and Expected Value," AACE International, Morgantown WV, Latest revision.

38. AACE International, "Recommended Practice No. RM-31: Integrated Cost and Schedule Risk Analysis and Contingency Determination Using a Hybrid Parametric and CPM-Based Method," AACE International, Latest revision.
39. AACE International, "Recommended Practice No. 104R-19: Communicating Expected Estimate Accuracy," AACE International, Morgantown WV, Latest revision.
40. G. Rothwell, "Cost Contingency as the Standard Deviation of the Cost Estimate for Cost Engineering," Stanford Institute for Economic Policy Research, Stanford University, Stanford CA, Feb 9, 2004.

John K. Hollmann, PE CCP CEP DRMP FAACE Hon. Life
Validation Estimating LLC
jhollmann@validest.com