



Systematiska fel i kostnads kalkyler

En tillämpning av statistiska metoder för att studera kostnadskontroll vid investeringsprojekt

Herman Andersson och Peter Wikman

Herman Andersson och Peter Wikman

Systematiska fel i kostnads kalkyler

En tillämpning av statistiska metoder för att studera
kostnadskontroll vid investeringsprojekt

Titel	Systematiska fel i kostnadskalkyler – En tillämpning av statistiska metoder för att studera kostnadskontroll vid investeringsprojekt
Title	Biases in cost estimates – An application of statistical methods to study cost control in investment projects
Rapportnr	FOI-R--5269--SE
Månad	April
Utgivningsår	2022
Antal sidor/Pages	39
ISSN	1650-1942
Uppdragsgivare	FOI
Forskningsområde	Försvarsekonomi
FoT-område	Inget FoT-område
Projektnr	A12251
Godkänd av	Malek Finn Khan
Ansvarig avdelning	Försvarsanalys

Bild: Shutterstock

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk, vilket bl.a. innebär att citering är tillåten i enlighet med vad som anges i 22 § i nämnd lag. För att använda verket på ett sätt som inte medges direkt av svensk lag krävs särskild överenskommelse.

This work is protected by the Swedish Act on Copyright in Literary and Artistic Works (1960:729). Citation is permitted in accordance with article 22 in said act. Any form of use that goes beyond what is permitted by Swedish copyright law, requires the written permission of FOI.

Sammanfattning

Att genomföra större samhällsprojekt är en viktig del i den offentliga sektorn och utgör en signifikant del av statsbudgeten. En bedömning av projektens relevans är därmed ett avgörande moment för verksamheten. För att beslutsfattare ska kunna prioritera mellan de projektförslag som bereds är det viktigt att de har tillgång till ett rättvisande beslutsunderlag. Ett centralt sådant är de kostnadskalkyler som bland annat ska svara på frågan hur mycket projekten förväntas kosta. För att öka kunskapen inom detta område avser denna rapport att beskriva och problematisera det systematiska framtagandet av kostnadskalkyler för offentliga investeringsprojekt. Rapportens första del presenterar en statistisk analys som metod för bedömning om systematiska fel i kostnadskalkyler. Därefter beskrivs och diskuteras kunskapsläget, vilket visar hur vanligt det är med systematiska kostnadsunderskattningar för en mängd olika investeringsprojekt. Avslutningsvis presenteras en metod som, med hjälp av historisk data, minimerar potentiella kostnadsökningar och kalibrerar kostnadskalkyler. För användning av metoden inom offentliga myndigheter så föreslås att man inför erfarenhetsdatabaser som grund för att utvärdera metodiken i sina beslutsunderlag.

Nyckelord: systematiska fel, kostnadskalkyler, strategisk snedvridning, optimism bias, prediktioner, statistisk analys.

Summary

A significant part of the public spending is designated to investment projects within the public sector. Hence, it is crucial to properly prioritize among the different projects. Cost calculations is an essential tool in the process of evaluating and prioritizing projects. This report describes and discusses systematic errors in cost estimations of public investment projects.

First, the report starts with a description of the statistical method used to identify systematic errors in cost calculations. Second, a literature review shows that systematic cost underestimations that are both statistically and economically significant are common in a broad spectrum of investment projects. The last part of the report presents a method for adjusting cost calculation with data from previous projects in order to minimize the discrepancies. Finally, the report suggests that government authorities develop a database with previous projects in order to support their methods for cost calculations.

Key words: systematic errors, cost calculations, strategic misrepresentation, optimism bias, predictions, statistic analyse

Förord

Kostnadsökningar i projekt är ett vanligt problem inom den offentliga sektorn. Dessa ökningar leder till en rad olika svårigheter, till exempel att planerad verksamhet måste ställas in eller skjutas på framtiden. Trots att problemen är välkända saknas det kunskap och forskning om hur problemen ska lösas. Genom att komplettera inifrånperspektivet, där enstaka projekt analyseras var för sig, med ett utifrånperspektiv, där systematiska fel i ett stort antal projekt står i fokus, är förhoppningen att FOI ska kunna bidra med kunskap som minskar kostnadsökningarna. Även om de metoder och verktyg som diskuteras i rapporten inte löser alla problem är det vår övertygelse att forskning om systematiska fel i kostnadskalkyler är en del av lösningen.

Målgruppen för rapporten är beslutfattare, yrkesverksamma och andra intresserade av projekt- och budgetfrågor inom totalförsvaret. Rapporten vänder sig till en bred målgrupp och har ett övergripande perspektiv. Den internationella litteraturen inom systematiska fel i kostnadskalkyler har sin tyngdpunkt inom transportinfrastruktur, vilket också speglas i den här rapporten. Även om typen av projekt skiljer sig åt inom totalförsvaret är den grundläggande problematiken med långa och komplexa projekt den samma: kostnadskalkylerna som används som beslutsunderlag är behäftade med stor osäkerhet. Det gör att övergripande insikter och slutsatser kan överföras mellan områdena.

Den breda ansatsen gör att kapitel 2 kan upplevas som obefogat beskrivande för läsare med goda kunskaper om grundläggande statistiska begrepp. Dessa läsare kan med fördel gå direkt till kapitel 3 efter det inledande kapitlet.

Rapporten är en del av projektet *Ekonomiska risker i offentlig verksamhet* och är finansierad genom FOI:s anslagspost för egeninitierad forskning. Syftet med den egeninitierad forskning är att ge möjlighet att bedriva forskning med ett strategiskt nydanande perspektiv, som inte kan härledas till identifierade förmågebehov, och som inte inriktas i de processer som styr forskning och utveckling mot sådana kända behov. Projektet är tvåårigt och detta är den första rapporten i projektet efter ett års arbete.

Författarna vill rikta ett särskilt tack till Mattias Johansson för värdefulla kommentarer och synpunkter under granskningsprocessen. Vidare så bidrog Erik Lundberg, Evelina Bonnier och Josefin Haaga genom att ingå i projektets referensgrupp. Slutligen skulle författarna vilja tack Sanna Dalberg, Ulf Jonsson, Tobias Junerfält, Monia Lougui, Anders Strömgren, Jan Henningsson, Kent Ehliasson och Susanne Oxenstierna för att ni tog er tid att lämnat skriftliga synpunkter i samband med granskningen.

Stockholm, mars 2022
Herman Andersson, Projektledare

Innehållsförteckning

1	Inledning	7
1.1	Syfte och frågeställning	9
1.2	Metod.....	9
1.3	Avgränsningar	9
1.4	Disposition	10
2	Kostnadskalkyler — osäkra prediktioner	11
2.1	Osäkra kostnader och sannolikhetsfördelningar	12
2.2	Aggregerade risker.....	15
2.3	Systematiska fel i kostnadskalkyler	16
3	Hur ser kostnadskontrollen ut i investeringsprojekt?	21
3.1	När beslutas det att ett projekt ska genomföras?	21
3.2	Hur vanligt är det att underskatta projektkostnader?	23
3.3	Faktorer som påverkar storleken på underskattningen	26
3.4	Förbättringar över tid	27
4	Orsaker till och metoder mot systematiska fel	29
4.1	Orsaker.....	29
	Tekniska förklaringsmodeller — Inte enligt empirin.....	30
	Psykologiska förklaringsmodeller — Optimism bias.....	30
	Ekonomiska/politiska förklaringsmodeller — Strategic misrepresentation.....	31
4.2	Motverkan — prediktioner i referensklasser	31
5	Slutsatser och fortsatta studier	34
6	Referenser	36

1 Inledning

Kostnaden för offentliga investeringsprojekt utgör en signifikant del av statsbudgeten. Exempelvis uppgår utvecklandet av transportinfrastrukturprojekt till 30 miljarder kronor per år och förväntas öka de kommande åren (Riksrevisionen, 2021a). För att beslutsfattare ska kunna prioritera mellan de projektförslag som bereds och besluta att genomföra det som ger mest nytta för skattepengarna är det viktigt att de har tillgång till rättvisande beslutsunderlag. Ett sådant underlag är projektens kostnadskalkyler som bland annat ska svara på frågor som gäller de planerade kostnaderna för projekten. Riksrevisionen (2021a) menar att rättvisande kostnadskalkyler är en *nödvändig förutsättning* för att regeringen och Trafikverket ska kunna prioritera effektivt mellan potentiella projekt.

Ett problem som inte är helt ovanligt inom den offentliga sektorn är att kostnaderna i kostnadskalkylen är underskattade. Under senare år har exempelvis konsekvenserna av kostnadsökningarna för inköpet av helikopter 14 till Försvarsmakten och byggandet av Nya karolinska sjukhuset debatterats flitigt i media (se exempelvis SVT, 2018 och Expressen, 2019). Det kan dock vara svårt att utifrån enskilda exempel få en rättvisande bild av hur vanligt det är att kostnaderna ökar jämfört med vad som ursprungligen budgeterats. Är exemplen som tas upp i media enstaka undantag eller ökar kostnaderna systematiskt?

Med kostnadsökningar avses i den här rapporten skillnaden mellan de predikterade och det faktiska kostnadsutfallet. Med andra ord så avses oförutsedda ökningar av kostnader utöver de som planerades för när beslutet att genomföra projektet fattades.¹

Kostnadsökningar riskerar att leda till flera negativa effekter. Om inte den offentliga verksamheten håller sig inom givna budgetramar kan det leda till att annan verksamhet helt eller delvis måste ställas in eller skjutas på framtiden. Statskontoret (2018) konstaterar exempelvis i en rapport att förstärkningen av Försvarsmaktens operativa förmåga har påverkats negativt på grund av kostnadsökningar. När projekten blivit dyrare har viktig verksamhet tvingats ställas in eller nedprioriterats. På liknande sätt konstaterar Riksrevisionen (2021b) i en granskning av arméns tillväxt att oregelbundna beslut och prioriteringar, bland annat framtvingat av kostnadsökningar, har lett till negativa konsekvenser för övningsverksamheten, renovering av materiel, inköp av materiel med mera. Detta har i sin tur försvårat uppbyggandet av förmågan i krigsförbanden.

Flyvbjerg m.fl. (2002) visar i en betydelsefull studie, som utvärderat kostnadskalkyler för ca 250 investeringsprojekt, att nästan nio av tio av dem blev dyrare än

¹ Se Olsson (2019) för en bredare definition och diskussion om kostnadsökningar. Olsson (2019) delar upp kostnadsökningar i tre delar: kunddrivna kostnadsökningar, inflation och fördyrningar. De två förstnämnda kostnadsökningarna går att planera och förutsäga. Denna rapport fokuserar på den typen av kostnadsökningar som inte går att planera eller förutsäga. Enligt Olssons (2019) definition så behandlar rapporten därmed endast fördyrningar och inte kunddrivna kostnadsökningar eller inflation.

kalkylerat. Med grund i det stora dataunderlaget har författarna använt sig av statistiska analysmetoder som visar att kostnadskalkylerna för projekten systematiskt underskattar den slutgiltiga kostnaden. Sedan studien publicerades har det börjat växa fram en empirisk litteratur som utvärderar kostnadskalkyler med liknande metoder. Resultaten av dessa studier indikerar att det finns ett tydligt problem med systematiska och ekonomiskt signifikanta kostnadsunderskattningar inom olika offentliga investeringsprojekt, från transport- och energinfrastrukturprojekt till vapen- och rymdprojekt.

Inom statistiken refererar termen *systematiska fel* till fel i en mätning, beräkning eller tolkning som beror på att information den baseras på är snedvriden. Tillämpar man den på kostnadskalkyler för projekt är det när de faktiska kostnadsutfallen i likartade projekt avviker på ett systematiskt sätt från vad som predicerats, antingen i form av kostnadsunder- eller överskattningar. Ett problem med systematiska fel kan vara att kostnadskalkylerna är snedvridna och inte tar höjd för de kostnadsutvecklingar som ofta sker. Då kostnadskalkyler är behäftade med osäkerhet är det svårt att dra slutsatser om felaktigheter i enskilda fall utan ett större antal projekt behöver undersökas. Med andra ord: ska man undersöka systematiska fel är det den genomsnittliga avvikelser som är intressant snarare än utfallet av enskilda projekt.

Då kostnadsökningar både är vanliga och leder till negativa effekter är det viktigt att utveckla kunskap och metoder för att minska dessa problem. I denna rapport är vi intresserade av kostnadskalkyler som beslutsunderlag och att undersöka riktigheten i kostnadskalkyler som gjorts i samband med beslutet om att projektet ska genomföras. Trots en omfattande diskussion om problematiken i den internationella litteraturen finns det relativt lite vetenskapligt dokumenterade analyser och studier som undersöker svenska förhållanden. Anmärkningsvärt är att vi endast hittat ett fåtal formella granskningsrapporter i den svenska kontexten.

För att minska problemen med systematiska fel i kostnadskalkyler krävs sannolikt ett antal olika insatser av betydande omfattning beroende på vilken typ av organisation och vilken kontext den verkar i. Att till exempel hitta en lösning där alla projekt håller sin budget för respektive budgetår kan vara utmanande då slutkostnaden alltid är behäftad med någon form av genuin osäkerhet.

Målet bör istället vara att genomföra prediktioner där kostnader ibland överstiger och ibland understiger de faktiska kostnaderna; men även att de negativa konsekvenserna är begränsade och hanterbara för organisationen. Den här rapporten diskuterar således hur systematiska fel kan minskas med statistiska metoder och utgår ifrån erfarenheter och slutsatser från svensk och internationell litteratur inom området. Vi menar att uppbyggandet av erfarenhetsdatabaser är ett första steg för att på lång sikt kunna bidra till att riktigheten i kostnadskalkyler förbättras och därmed minska problemen med kostnadsökningar.

1.1 Syfte och frågeställning

Syftet med denna rapport är att beskriva och problematisera konceptet systematiska fel tillämpat på framtagandet av kostnadskalkyler för offentliga investeringsprojekt. Tanken är att rapporten ska bidra till bättre förståelse och tillvägagångssätt för att motverka de systematiska kostnadsunderskattningar som har kommit att känneteckna kostnadskalkyler i sådana projekt.

För att uppnå detta syfte ämnar denna rapport svara på följande frågeställningar:

1. Hur kan riktigheten i kostnadskalkyler analyseras med hjälp av statistiska metoder?
2. Hur stor är och vad påverkar den dokumenterade skillnaden mellan predicerade och faktiskt utfall i investeringsprojekt?
3. Hur kan uppskattningen av systematiska fel användas för att justera kostnadskalkyler för att bli ett mer rättvisande beslutsunderlag?

1.2 Metod

Studien utgår från empirisk forskning som visar att kostnadskalkyler är behäftade med systematiska underskattningar av slutkostnader för offentliga investeringsprojekt (se exempelvis Miranda Sarmiento och Renneboog 2017 för en översikt av litteraturen). För att förstå den statistiska metodiken som dessa studier utgår ifrån har en översikt av metodiken genomförts i Kapitel 2. De statistiska metoder som presenteras sträcker sig från grundläggande metoder för statistisk analys som återfinns i de flesta introducerande läroböcker inom ämnet till mera avancerade områden rörande prediktionsmetodik.

1.3 Avgränsningar

Rapporten är avgränsad till att undersöka systematiska fel i kostnadskalkyler för investeringsprojekt. Inom ramen för denna avgränsning definieras investeringsprojekt, hädanefter projekt, som en temporär satsning för att framställa en unik vara eller tjänst med en början och ett slut för att åstadkomma ett förutbestämt resultat. Rapporten kommer att fokusera på offentliga investeringar och exkluderar privata investeringsprojekt. Det görs dels för att huvuddelen av litteraturen inom området gör motsvarande avgränsning. Vidare så bedöms offentliga investeringar vara mest relevanta för Totalförsvaret och FOI:s verksamhet. Vi har även gjort en avgränsning avseende frågor som rör livscykelkostnader och fokuserar endast på kostnader associerade med framställandet av projektet. Vi har även valt att fokusera uteslutande på kostnadssidan av projekten och exkluderar frågor som rör nyttan. En anledning till detta är att kostnader är observerbara kvantiteter som är jämförbara mellan projekt. Till skillnad från kostnader är nyttan med projekt ofta icke-observerbara kvantiteter som måste skattas var för sig och är svåra att jämföra

mellan olika typer av projekt. Vi är medvetna om att detta innebär att vi inte kan säga något om huruvida en kostnadsökning är befogad som resultat av att exempelvis nyttan ökat i motsvarande grad.

1.4 Disposition

Rapporten inleds med en grundläggande redogörelse för vad en kostnadskalkyl fyller för funktion och varför osäkerheten i projekten gör att kalkylerna är svåra att utvärdera. Därefter diskuteras hur statistiska metoder kan användas för att skatta storleken på systematiska kostnadsavvikelser. Läsare med goda kunskaper om grundläggande statistiska begrepp kan med fördel lämna denna del och gå direkt till kapitel 3. I Kapitel 3 görs ett antal nedslag i den empiriska litteraturen angående systematiska kostnadsavvikelser i olika projekt. Kapitel 4 diskuterar olika förklaringsmodeller till de systematiska kostnadsökningarna och presenterar en metod för att justera dem. Det avslutande kapitlet sammanfattar rapporten och ger förslag på framtida studier.

2 Kostnads-kalkyler — osäkra prediktioner

En *kostnads-kalkyl* för ett investeringsprojekt är ett viktigt underlag för beslutsfattare som svar på frågor som: ”Vad förväntas projektet kosta?”, ”Vad är sannolikheten att kostnaderna överstiger ett visst tak?” och ”Vad är osäkerheterna och hur påverkar dessa kostnaden för projektet?”. Ett sådant underlag är en förutsättning för att beslutsfattarna för projektet effektivt ska kunna prioritera mellan potentiella investeringsprojekt. I den offentliga sektorn måste alltid avvägningar mellan olika projekt och andra satsningar göras eftersom myndigheters verksamhet ska bedrivas inom fasta budgetramar. För att försöka maximera samhällsnyttan behöver beslutsfattarna därför jämföra nyttan och kostnaden för projekten de väljer mellan. Men eftersom budgetramarna och inte samhällsnyttan sätter begränsningarna för vilka projekt som kan genomföras, innebär det, *även om den totala samhällsnyttan är korrekt beräknad* för varje enskilt projekt, att prioriteringarna riskerar att bli fel om kostnads-kalkylen för ett eller flera projekt är fel.²

Kostnads-kalkyler tas främst fram genom så kallade kostnadsosäkerhetsanalyser. Dessa är processer som bedömer kostnadspåverkan av osäkerheter associerade med investeringens tekniska definition och de kostnadsskattningsmetoder som använts (Garvey m.fl., 2016). I och med de avgränsningar som gjorts i rapporten är syftet med en kostnadsosäkerhetsanalys här att ta fram och beskriva den osäkerhet som är kopplad till slutkostnaden för projektet.

Eftersom fokus i denna rapport ligger på en övergripande nivå kommer vi inte gå in i detalj på vilka typer av osäkerheter som kan leda till att slutkostnaden varierar i olika typer av projekt. Istället kommer vi fokusera på att presentera den konceptuella apparaten som behövs för att förstå grundproblematiken. Denna problematik utgår ifrån synen på kostnads-kalkyler som prediktioner för kostnaderna för ett projekt över en viss tidshorisont. Riktigheten i dessa prediktioner kan utvärderas efter projektens slut och dessa utvärderingar kan sedan användas för att öka träffsäkerheten för framtida prediktioner. En viktig poäng är att prediktioner i stort, och kostnads-kalkyler mer specifikt, är behäftade med genuin osäkerhet i den meningen att det inte går att helt gardera sig emot den. Det är därför svårt att utvärdera enskilda prediktioner mot enskilda utfall. Istället kan och bör metodiken som används för att generera prediktioner utvärderas med hjälp av en representativ mängd historiska data från tidigare utfall och prediktioner utförda med metodiken i fråga.

Vi börjar detta kapitel med att i Avsnitt 2.1 beskriva hur kostnadsosäkerheter, och därmed kostnads-kalkyler, kan beskrivas med hjälp av sannolikhetsfördelningar. Därefter beskriver vi i Avsnitt 2.2 hur denna osäkerhet kan kopplas till facktermen

² Se Simon (1976) för en diskussion om hur handlingsalternativ bör jämföras. Samhällsnyttan kan definieras som nyttan med projekten beskrivet i monetära termer minus kostnaden.

risk. I Avsnitt 2.3 går vi sedan igenom hur kostnadskalkyler kan utvärderas med hjälp av statistiska metoder, vad kostnadskalkyler ska predicera, och hur systematiska fel definieras.

2.1 Osäkra kostnader och sannolikhetsfördelningar

Den slutliga kostnaden för att framställa ett projekt är förknippad med osäkerhet, speciellt för stora projekt. Dessa kostnader är känsliga för de många förutsättningar och antaganden som varierar frekvent över projektets livslängd. Om kostnadskalkyler ska kunna jämföras med varandra på ett rättvist sätt är det därför viktigt att endast jämföra kalkyler som gjorts vid samma tidpunkt i projektet. Kalkyler som görs i ett senare skede i ett projekt tenderar att vara mer träffsäkra än kalkyler som görs tidigare eftersom de förstnämnda görs när mer osäkerheter kring projektet har realiserats.

Valet av tidpunkt att utgå ifrån för kostnadskalkylen bör reflektera vad prediktionen är tänkt att förutsäga och vad som ska mätas. Om man vill förstå hur behjälpta beslutsfattarna är av en kostnadskalkyl bör man utgå från kostnadskalkylen som presenterades när det skulle avgöras om projektet skulle genomföras. Detta eftersom beslutsfattarna använder kostnadskalkylen som ett beslutsunderlag. Om man istället exempelvis vill avgöra hur väl entreprenörer levererar givet ett visst kontrakt, bör man istället utgå ifrån kostnadskalkylen som användes när projektet kontrakterades. Detta eftersom kontraktet är skrivet med den kostnadskalkylen i åtanke. Som nämnts tidigare är fokus för denna rapport kostnadskalkyler som beslutsunderlag vilket motiverar varför vi har valt att fokusera på den förstnämnda tidpunkten.

Slutligen är det också viktigt att specificera när projektet anses vara färdigt. Detta för att kunna särskilja vilka kostnader som relaterar till färdigställandet av projektet och vilka som relaterar till vidmakthållandet av den färdiga produkten (Flyvbjerg m.fl., 2002; Flyvbjerg m.fl., 2018).

De flesta projekt kan delas upp i olika delmoment som var för sig är associerade med olika osäkra kostnader. Därför kommer det faktiska kostnadsutfallet av ett projekt endast vara ett utav ett spektrum av potentiella utfall, vilka beror på en mängd större och mindre kostnadsutfall som summerar till slutkostnaden. Ett vanligt sätt att beskriva detta spektrum av potentiella slutkostnader är med hjälp av så kallade *sannolikhetsfördelningar*. En sannolikhetsfördelning, eller fördelning mer kortfattat, tilldelar varje möjligt kostnadsutfall en sannolikhet mellan noll och ett. Även om begreppet är kopplat till en teknisk definition med ursprung i matematisk statistik kan en fördelning enkelt beskrivas som en organiserad lista över sannolikheter kopplat till de olika utfall som anses möjliga.

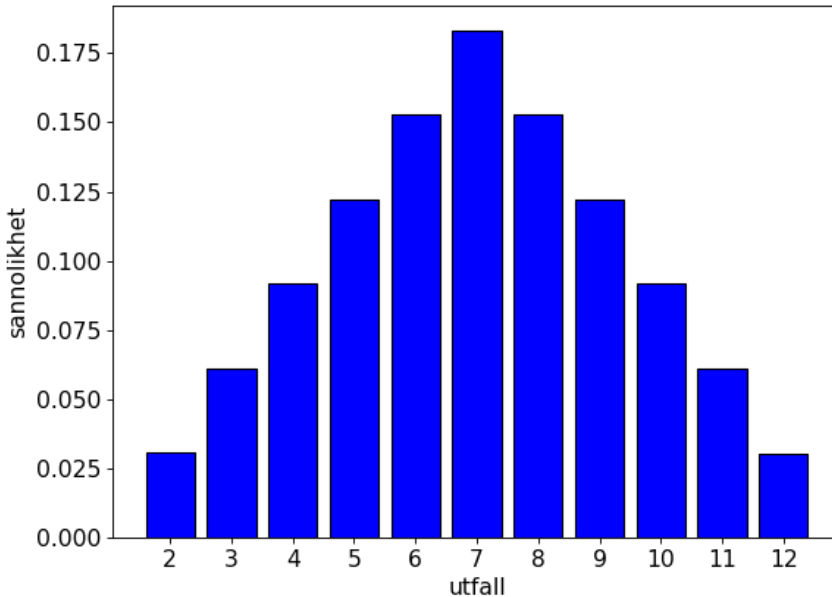
Sannolikheterna för olika kostnadsutfall i ett projekt är okända och behöver därför uppskattas (ofta förkortat som *skattning*). För att göra sådana skattningar används ofta metoder från sannolikhetsteorin. En sådan metod är att bryta ned projekt i delkomponenter där historiska värden och tekniska analyser kan användas för att skatta kostnadsfördelningen av de olika delkomponenterna. Efter att kostnadsfördelningen av de olika delmomenten har skattats kan den övergripande kostnadsfördelningen för projektet beräknas med hjälp av en så kallade Monte Carlo-metod, vilken enkelt beskrivet är ett sätt att beräkna den aggregerade fördelningen med hjälp av slumpade, eller simulerade, kostnader utifrån de beräknade delkomponenternas fördelningar.³ Den aggregerade fördelningen kan därefter beskrivas genom ett diagram.

För att få en övergripande förståelse för hur en sådan skattning kan gå till med hjälp av en Monte Carlo-metod presenterar vi ett stiliserat exempel. Anta att kostnaden för ett projekt ges av två oberoende delkomponenter där kostnadsutfallet av varje delkomponent kan beskrivas av utfallet av en sexsidig tärning som kastas. Totalkostnaden av projektet ges därför av summan av de två tärningarna. För att skatta fördelningen av totalkostnaden kastas de båda tärningarna, t.ex. en miljon gånger som här, och förekomsten av varje givet utfall summeras, för att sedan divideras med det totala antalet kastade tärningspar. Det här ger en skattning av sannolikheten för utfallen. Resultatet presenteras i Figur 1 på nästa sida.⁴

När kostnadskalkyler väl genomförts är det första som kommuniceras inte en sannolikhetsfördelning utan snarare ett *centralmått*, vilket är tänkt att förenklat beskriva fördelningen med endast ett nummer. Centralmått beräknas utifrån den skattade fördelningen av kostnaden. Ett vanligt sådant centralmått är *medelvärdet* för kostnaden som är summan av alla utfall delat med antalet utfall, ett annat är *typvärdet* vilket är utfallet som har störst sannolikhet att inträffa, och ett tredje är *medianen* som är det mittersta kostnadsutfallet när utfallen sorterats i storleksordning (Garvey m.fl., 2016).

³ Se Garvey m.fl. (2016) för en genomgång av avancerade metoder för skattning av kostnadskalkyler, vilka inkluderar Monte Carlo-metoder.

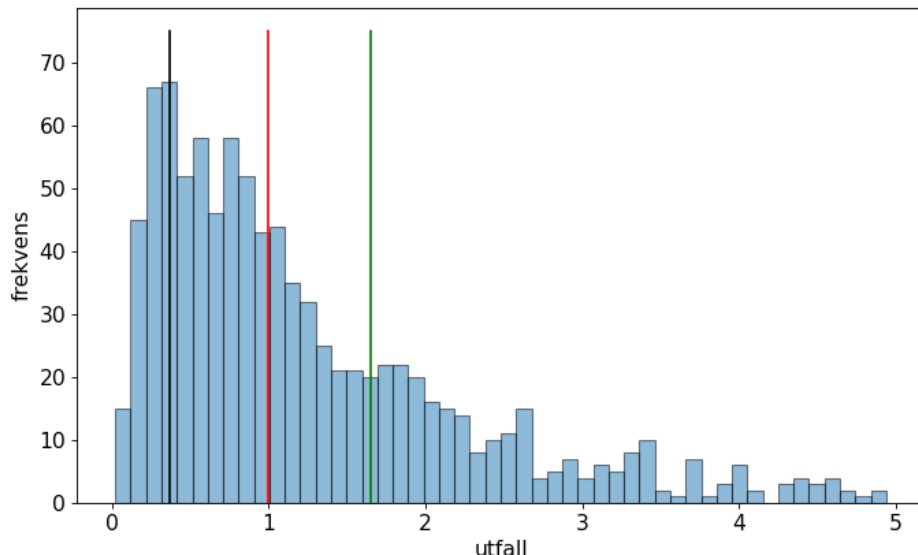
⁴ Notera att den teoretiska fördelningen av summan av två tärningar kan enkelt fås genom andra metoder. Exempelvis kan man räkna samman alla kombinationer som leder till samma utfall och dela dem med 36 (de möjliga utfallen av de kastade tärningarna). Till exempel kan utfallet tre fås på två olika sätt: den ena tärningen visar ett och den andra tärningen visar två och vice versa, vilket ger en sannolikhet på 2/36.



Figur 1. Sannolikhetsfördelningen av summan av två tärningar kastade en miljon gånger.

Ett viktigt val som görs i samband med framställandet av en kostnadskalkyl är valet av centralmått för att kommunicera den skattade kostnaden. Detta val är desto viktigare ju mindre komplementär information om den skattade fördelningen som kommuniceras tillsammans med kostnadsuppskattningen. En ofta förbisedd poäng är att valet av centralmått för en kostnadsuppskattning bör bero på hur man utvärderar *situationer där prediktionen är fel*. Som en tumregel kan följande rekommendationer ges. Om större prediktionsfel är mer allvarliga än små bör medelvärdet användas. Om det är lika allvarligt att exempelvis predicera kostnaden 1 MSEK fel jämfört med den faktiska kostnaden i hälften av fallen som att predicera kostnaden 2 MSEK fel i en fjärdedel av fallen bör medianen istället användas. Om man endast bryr sig om att prediktionen är fel eller inte, men inte felets storlek, bör typvärdet användas (SCB, u. å.; Flyvbjerg m.fl., 2018).

I det stiliserade tärningsexemplet ovan sammanfaller väntevärdet med typvärdet och medianen. Empiriskt sett brukar sannolikhetsfördelningar av kostnader vara skeva åt höger, vilket innebär att medelkostnaden är både större än medianen och typvärdet (De Reyck m.fl. 2017). Figur 2 nedan beskriver slumpmässiga dragningen från en (lognormal) fördelning som är skev åt höger och tenderar att väl beskriva fördelningen av kostnadsutfall. För denna typ av fördelningar är typvärdet lägre än medianen som i sin tur är lägre än medelvärdet.



Figur 2. Slumpvis genererade utfall från en lognormal fördelning. Det svarta sträcket beskriver typvärdet, det röda medianen, och det gröna medelvärde (utfallen sträcker sig längre än i grafen).

2.2 Aggregerade risker

Som har poängterats tidigare kommer det alltid finnas en variation i slutkostnaden för varje enskilt projekt. Det kan handla om alltifrån förändrade prisnivåer på varor som används i projektet till tidsåtgång för utvecklingsarbete. Denna osäkerhet leder inte bara till att vissa projekt blir dyrare än beräknat men också till att andra blir billigare. Osäkerheten kan beskrivas som genuin i den meningen att det inte går att helt gardera sig emot den.

När osäkerhet kan kvantifieras i termer av sannolikheter brukar det beskrivas med facktermen *risk* (Garvey m.fl., 2016). Termen konceptualiseras oftast som ”kombinationen av sannolikheten för en viss händelse och dess konsekvenser” (ISO/IEC Guide 73). Ofta refererar en risk till en konsekvens som är negativ för aktören av intresse. En av fördelarna med att riskerna som analyseras i denna rapport beskrivs i termer av monetära konsekvenser är att enskilda kostnadsrisker kan *aggregeras* till ett mer omfattande mått och beskrivas av en kostnadsfördelning.

Den aggregerade risken beskrivs oftast i den tillämpade litteraturen av ett spridningsmått. Spridningsmått används ofta för att på ett förenklat sätt representera storleken på osäkerheten associerat med en viss sannolikhetsfördelning. En aggregerad risk med en fördelning med ett större spridningsmått kan sägas vara mer riskabel än om spridningsmättet hade varit lägre. Det absolut vanligaste spridningsmättet är

kostnadskalkylens standardavvikelse vilket beror på spridningen av utfall relativt fördelningens medelvärde.⁵

2.3 Systematiska fel i kostnadskalkyler

Om prediktionerna som görs för olika investeringsprojekt följer en likvärdig metodologi och projekten är jämförbara kan riktigheten i metodologin utvärderas. Detta kan göras genom att fördelningen av de historiska utfallen jämförs med deras prediktioner.

För projekt av olika storlekar, associerade med kostnadsfördelningar med olika storlekar på utfallen, behöver kostnaderna normaliseras för att utvärderingen av prediktionsmetodologin ska vara rättvisande. Ett vanligt sätt är att titta på *kostnadsavvikelsen* av projekten, vilket definieras som

$$ka_i = \frac{K_i - E[K_i]}{E[K_i]}$$

där K_i är det faktiska kostnadsutfallet av projekt i och $E[K_i]$ är medelvärdet för den skattade kostnadsfördelningen av projektet, eller i denna rapports kontext den kommunicerade kostnadsskattningen.⁶ En positiv kostnadsavvikelse, $ka_i > 0$, implicerar därmed att det faktiska kostnadsutfallet i projekt i översteg prediktionen $E[K_i]$ med $ka_i * 100$ procent och vice versa för $ka_i < 0$ (se exempelvis Arena m.fl. 2006; De Reyck m.fl., 2017).

Det så kallade mätfelet av ett mått (så som medelvärdet) definieras inom statistiken som skillnaden mellan ett skattat värde och det sanna värdet för hela populationen av potentiella projekt. Mätfelet brukar delas upp i två komponenter: 1) slumpvisa mätfel, som vi tidigare har refererat till i kostnadskalkylssammanhang som den aggregerade risken, och 2) systematiska mätfel som är huvudfokus för denna rapport. Om det genomsnittliga kostnadsutfallet avviker från noll på ett signifikant sätt indikerar detta att prediktionsmetodologin skattar kostnaden på ett *systematiskt felaktigt* sätt. Med andra ord, om kostnadsavvikelserna, uppåt och nedåt, i genomsnitt inte tar ut varandra för ett antal likartade kostnadskalkyler kan de sägas vara systematiskt felaktiga (se exempelvis Wooldridge, 2013).

En viktig poäng är att det i praktiken är väldigt osannolikt att snittkostnadsavvikelsen är noll även om kostnadskalkylerna är korrekt skattade. Det är därför statistiska

⁵ Standardavvikelsen i kvadrat (den så kallade variansen) definieras som summan av den kvadrerade avvikelsen från varje observerat utfall från medelvärdet dividerat med antalet observationer, där avvikelsen är skillnaden mellan kostnadsutfallet och det skattade medelvärdet för utfallet. En fördel med standardavvikelsen är att den tillsammans med medelvärdet helt beskriver en skattning av kostnadsfördelningen i de fallen den teoretiska fördelningen kan beskrivas av en normalfördelning (Garvey m.fl., 2016).

⁶ Det går även bra att använda andra centralmått istället för medelvärdet för att beskriva den förväntade kostnaden.

metoder måste användas för att avgöra om prediktionsmetodiken visar tecken på *statistiskt signifikanta* systematiska fel.

För att göra statistiska analyser måste den faktiska fördelningen av kostnadsavvikelser skattas för typen av projekt som prediktionerna är gjorda för. Denna skattning kan underlättas genom att göra antaganden om hur fördelningen kan tänkas se ut. I det ursprungliga exemplet ovan är antagandet att fördelningen kan beskrivas i enlighet med Figur 1 som summan av två tärningar (där utfallen är utbytta mot kostnadsavvikelser). Ett vanligt antagande är annars att kostnadsavvikelsena är normalfördelade.⁷ Givet ett sådant antagande kan hypotesen att snittkostnadsavvikelsen är noll, med andra ord att prediktionerna inte har systematiskt fel, testas.

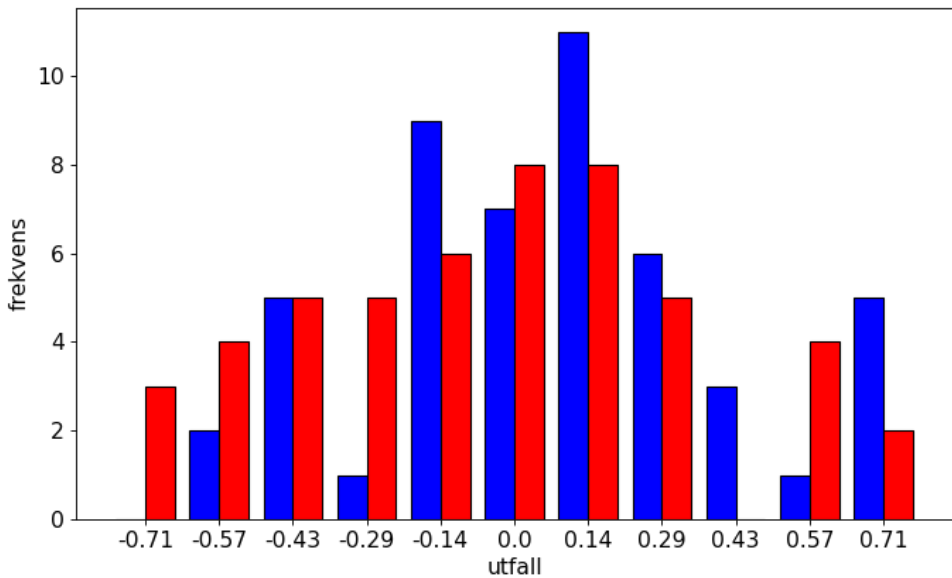
Det vanligaste tillvägagångssättet i litteraturen är att testa hypotesen att snittkostnadsavvikelsen är överskattad (Flyvbjerg m.fl., 2018). Om denna hypotes kan förkastas finns det statistisk evidens för att kostnadsutfallen är systematiskt underskattade. Detta görs genom att utvärdera hur väl insamlade kostnadsdata står sig mot förklaringsmodellen som valts (med andra ord, antaganden som har gjorts angående kostnadsfördelningen). Det vanligaste är att förkasta hypotesen om sannolikheten för att den stämmer, givet förklaringsmodellen, är mindre än fem procent. Denna sannolikhet kallas inom statistiken för hypotesens *p-värde* (Wasserstein och Lazar, 2016). I linje med litteraturen benämner vi i denna rapport en siffra, som motsäger en hypotes angående storleken på siffran med *p-värde* mindre än fem procent som statistiskt signifikant skild från hypotesavgränsningen.

För ett konkret exempel på denna problematik och hur en sådan utvärdering av kalkyler kan komma att se ut återgår vi till det stiliserade tärningsexemplet. I detta exempel antar vi nu att världen kan se ut på två olika sätt vilket gör att kostnadsfördelningen för projekten ter sig olika i de två världarna. I den första världen är kostnadsfördelningen i varje projekt exakt som i exemplet ovan, där kostnaden antas kunna beskrivas som ett kast med två sexsidiga tärningar. I den andra världen är kostnadsfördelningen i varje projekt också beskriven av ett kast med två sexsidiga tärningar. Skillnaden är nu att en av två sexsidiga tärningar är en "fusktärning" där sannolikheten för de tre högsta kostnaderna nu är dubbelt så stor jämfört med sannolikheten för de tre lägsta kostnaderna.

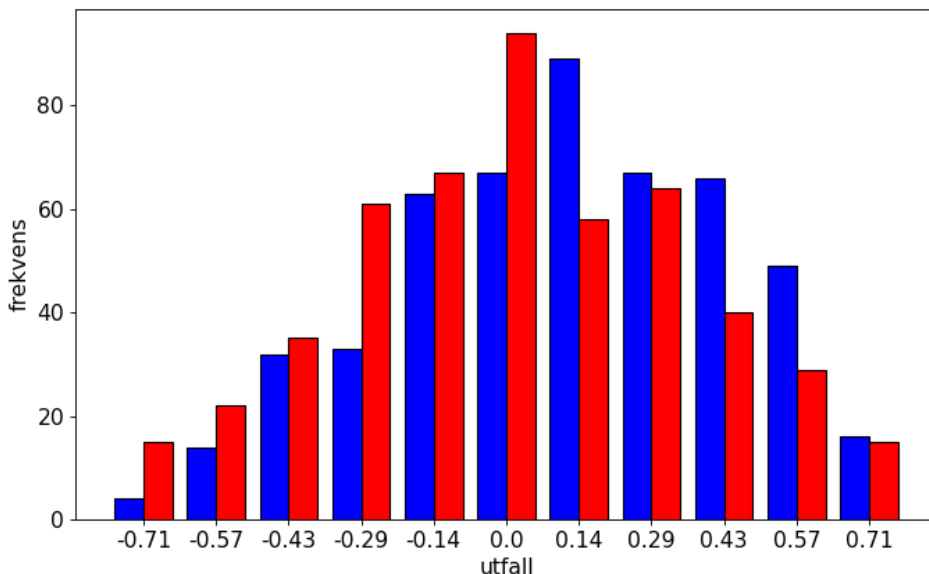
För vårt exempel antar vi först att 50 projekt i båda världarna har implementerats och att kostnadskalkylerna har gjorts i enlighet med fördelningen för Värld 1. Detta innebär att prediktionerna för Värld 2 innefattar systematiska fel men kostnadskalkylerna för Värld 1 är så bra de kan bli givet den faktiska osäkerheten. Det simulerade resultatet ges av Figur 3 nedan där kostnadsavvikelsena för utfallet har beräknats givet den predicerade kostnaden $E[K_i] = 7$. I genomsnitt är kostnadsavvikelsen för de simulerade fallen -5,2% för Värld 1 och 7,7% för Värld 2. Poängen

⁷ Ett annat vanligt alternativt antagande är att den naturliga logaritmen av $ka + 1$ är normalfördelad, vilket innebär att kostnadsavvikelsernas fördelning ser ut som fördelningen i Figur 2 ovan. Men se också Flyvbjerg m.fl. (2018) för en kritik av beskrivningen av kostnadsavvikelser som normalfördelade.

är här att även om prediktionsmetodologin är korrekt för Värld 1, givet ett litet urval av färdigställda projekt, slumpar det sig att slutkostnaden till synes underskattas med 5,2% i genomsnitt. Om antalet färdigställda projekt utökas till 500 är bilden mer rättvisande. Figur 4 nedan visar de simulerade resultaten för de två världarna i detta fall. Givet de 500 simulerade fallen är nu kostnadsökningen 1,15% i genomsnitt för värld 1 och 11,3% i genomsnitt för värld 2.



Figur 3. Sannolikhetsfördelningar av kostnadsavvikelsen från medelvärdet 7 av summan två tärningar för de olika världarna kastade 50 gånger. Utfallen i Värld 1 representeras av röda staplar och utfallen i Värld 2 representeras av blå staplar. Antalet utfall visas på y-axeln.



Figur 4. Sannolikhetsfördelningar av kostnadsavvikelsen från medelvärdet 7 av summan av två tärningar för de olika världarna kastade 500 gånger. Utfallen i Värld 1 representeras av röda staplar och utfallen i Värld 2 representeras av blå staplar. Antalet utfall visas på y-axeln.

Risken för att inte kunna förkasta en felaktig hypotes är större ju färre observationer som finns och ju mindre avvikelserna är relativt hypotesen. I exemplen ovan med de två världarna är det möjligt att exakt beräkna sannolikheten för snittkostnadsökningen för de olika simulerade projektkostnaderna, givet antagandet att kostnadsfördelningen ges av summan av två vanliga sexsidiga tärningar. I det första fallet (Värld 1) där 50 projektkostnader dras från den korrekta fördelningen är sannolikheten för att snittavvikelsen överstiger -5,2% större än 75%. När istället 500 projektkostnader dras är sannolikheten att snittavvikelsen överstiger 1,15% större än 20%. I båda fallen är varken kostnadsöverskattningen eller -underskattningen statistiskt signifikanta.

För Värld 2, där prediktionerna de facto är behäftade med systematiska fel (eftersom den faktiska fördelningen inte stämmer överens med den antagna fördelningen), är sannolikheten för att snittavvikelsen överstiger 7,7% ungefär 5,9% när de 50 första projektkostnaderna har dragits, vilket bara är statistiskt signifikant på 10 procents nivå ($p < 0.1$). När istället 500 projektkostnader har dragits är sannolikheten att snittavvikelsen överstiger det skattade värdet 11,3% obetydligt skild från noll och hypotesen att snittavvikelsen inte är större än noll kan förkastas på alla konventionella nivåer ($p < 10^{-6}$). Med andra ord är det i detta fall möjligt att säkerställa att prediktionen är behäftad med systematiska fel när 500 projekt är utförda och snittöverskattningen, 11,3%, är statistiskt signifikant större än noll.

Som vi kommer diskutera mer senare i rapporten finnas det många anledningar till varför systematiska fel uppstår i kostnadskalkyler. Det kan exempelvis handla om förändringar i kravspecifikationen, missbedömningar eller att kostnadsdelar fallit ur projektkalkylen. Problemet med systematiska fel är inte att det görs misstag i enskilda kostnadskalkyler, problemet är att det finns en tendens att missbedömningarna går i en viss riktning oftare än i den andra.

3 Hur ser kostnadskontrollen ut i investeringsprojekt?

Detta kapitel fokuserar på att besvara frågor som rör storleken på kostnadsavvikelserna i offentliga projekt. En vanlig uppfattning är att projekt i offentlig sektor inte håller budget och att kostnadskontrollen är bristfällig. Det kan samtidigt vara förknippat med svårigheter att få en helhetsbild i enskilda projekt, vilket diskuterats tidigare. I detta kapitel avser vi därför att göra en genomgång av såväl svensk som internationell litteratur som undersökt den genomsnittliga kostnadsavvikelsen i ett stort antal projekt.

Vi börjar kapitlet med att gå igenom Riksrevisionens analys av kostnadskontrollen i Trafikverkets infrastrukturprojekt. Fokus på genomgången är de svårigheter som är förknippade med att fastställa den faktiska tidpunkten när beslutet tas för att genomföra projektet. Vi går sedan igenom litteratur som undersökt storleken på kostnadsavvikelser i civila infrastrukturprojekt och militära materielprojekt. Därefter beskrivs faktorer som påverkar storleken på kostnadsavvikelser. Slutligen diskuterar vi huruvida det finns lärandeeffekter som gör att problemen med systematiska kostnadsökningar minskat över tid.

3.1 När beslutas det att ett projekt ska genomföras?

I Kapitel 2 konstaterades att om beslutsfattande står i fokus för utvärdering av kostnadskalkyler bör tidpunkten för den predicerade kostnaden som utvärderas vara när det faktiska beslutet att genomföra projektet fattas. I praktiken kan det dock vara svårt att avgöra vid vilken tidpunkt som det verkliga beslutet fattades. För att illustrera denna svårighet använder vi den svenska processen för infrastrukturinvesteringar som exempel.

Den svenska infrastrukturplaneringen startar med infrastrukturpropositionen där regeringen bestämmer de ekonomiska ramarna för de kommande 12 åren. Infrastrukturpropositionen läggs vanligtvis vart fjärde år och utgör grunden för bland annat Trafikverkets arbete med att ta fram den nationella planen för den statliga infrastrukturen. Den nationella planen innehåller de större infrastrukturprojekten och innehållet i planen beslutas av regeringen (Trafikverket, 2021). Planen delar in projekten i tre kategorier beroende på hur långt i planeringsprocessen projekten har kommit. De tre kategorierna är 1) planering, 2) förberedelser för byggstartsbeslut, och 3) byggstartsbeslut. Den första delen i processen brukar pågå mellan 7-12 år innan byggstart, den andra 4-6 år och den tredje 1-3 år. Dessa kategorier beskriver ungefär var i planeringsprocessen ett objekt befinner sig, men ska inte tas som en faktisk tidsangivelse över hur lång tid som återstår till byggstart. Mellan de olika

stegen i planen förbättras och förfinas beslutsunderlaget för att projektet vid beslutet om byggstart ska vara tillräckligt berett för att genomföras inom den planerade budget- och tidsramen (Riksrevisionen, 2021; Prop. 2011/12:118).

Riksrevisionen (2021a) påpekar att det kan krävas upp till tre regeringsbeslut (ett för varje kategori i den nationella planen) innan projekten kan genomföras. Det är dock först med beslutet om byggstart som det formella beslutet att genomföra projektet fattas. Det kan därför argumenteras för att det är först vid denna tidpunkt som beslutet att genomföra projektet fattas. Riksrevisionen (2021a) konstaterar dock att när ett projekt har kommit in i den nationella planen är det ovanligt att det inte genomförs. När den senaste planen reviderades togs endast nio av 194 namngivna projekt bort. Av de nio borttagna projekten var det tre som Riksrevisionen bedömer togs bort på grund av kostnadsökningar. Besluten motiveras dock inte vilket gör det svårt att fastställa de exakta orsakerna.

Som nämnt ovan kan det dröja mer än tio år mellan det att ett projekt kommer in i den nationella planen och det formella beslutet om byggstart genomförs. När Riksrevisionen (2021a) jämför kostnadskalkylerna mellan de nationella infrastrukturplanerna kan det noteras att det är relativt vanligt med stora kostnadsökningar. Bland de projekt som fanns med i både 2014 och 2018 års plan ökade kostnaderna med i genomsnitt 39%, med ett medianvärde på 15%. I 92 av de 142 projekten som fanns med i båda planerna ökade slutkostnaden för projektet. I resterande projekt minskade kostnaderna. För de 85 projekten som ingick i både 2010 och 2018 års plan ökade kostnaden i genomsnitt med 68%. Slutligen undersöktes kostnadsavvikelsen för de projekt som fanns med i både 2004 och 2018 års plan. I de 34 projekten ökade kostnaderna i genomsnitt med 165%. Notera att samtliga kostnadsökningar är i fasta priser omräknade till 2017 års prisnivå med hjälp av väg- respektive järnvägsprisindex.

När Riksrevisionen (2021a) undersökte skillnaden mellan den uppskattade kostnaden vid beslut om byggstart och kostnaden vid färdigställt projekt konstaterades det att den genomsnittliga kostnadsökningen var 2%, med en medianökning på minus 3%. Det betyder inte att alla projekten höll budget. Medianprojektet avvek 14% från den budgeterade kostnaden. Kostnadsavvikelseerna uppåt och nedåt tog dock i stort sett ut varandra, vilket gjorde att den genomsnittliga kostnadsavvikelsen var nära noll. Utifrån detta drogs slutsatsen att det inte förekom några systematiska kostnadsavvikelser i byggfasen.

Det kan därmed konstateras att det är under tiden som projekten detaljplanerades som kostnadsökningarna skedde, snarare än under byggfasen. Trots detta är det få objekt som plockas bort. När ett projekt väl har kommit in i den nationella planen verkar regeringen vara relativt okänslig för kostnadsökningar. Det gör att det går att argumentera för att det faktiska beslutet att genomföra projektet fattades när projektet inkluderades i den nationella planen. Givet detta perspektiv så finns det tydliga tecken att kostnadskalkylerna för infrastrukturprojekt är behäftade med systematiska

kostnadsökningar. För att kunna dra denna slutsats behövs det dock göras en statistisk analys av de data som samlats in, vilket Riksrevisionen (2021a) inte gör.

Slutsatsen som argumenterades för i Kapitel 2 att det är tidigt i projekten snarare än under bygg- och produktionsfasen som kostnaderna ökar är inte unikt för Sverige utan kan bekräftas av flera studier. Slutsatsen delas bland annat av Welde och Odeck (2017) och Cantarelli m.fl. (2012a), vilka gör motsvarande observation för norsk respektive nederländsk infrastruktur. Samma mönster hittas också i andra branscher. Arena m.fl. (2006) konstaterar att för amerikanska krigsmaterielprojekt är den genomsnittliga kostnadsökningen lägre för de faser som följer efter konceptframtagningsfasen och utvecklingsfasen. Med andra ord minskar kostnadsökningarna ju närmare den färdiga produkten projektet kommer.

Sammanfattningsvis kan vi konstatera att tidpunkten för utförandet av kostnadskalkylen har stor betydelse för den slutgiltiga kostnadsavvikelsen. En tidigare tidpunkt ger generellt större kostnadsökningar jämfört med en senare. Även om Flyvbjerg m.fl. (2002) argumenterar för principen att det är det första beslutet att genomföra projektet som bör användas som referenstidspunkt, kan det i praktiken vara svårt att avgöra när det faktiska beslutet fattades. Exemplet från den svenska processen för infrastrukturinvesteringar visar att resultatet kan variera från kostnadsavvikelser nära noll till relativt stora kostnadsunderskattningar, beroende på när beslutet att genomföra projekten anses vara taget.

3.2 Hur vanligt är det att underskatta projektkostnader?

Kunskaperna om kostnadsavvikelser i offentliga projekt har ökat under 2000-talet. Flyvbjerg m.fl. (2002) är en av de första studierna som samlat ihop ett tillräckligt stort dataunderlag för att möjliggöra statistiskt signifikanta slutsatser angående systematiska kostnadsavvikelser. Tidigare studier har endast undersökt ett mindre antal projekt, vilket gör det svårt att skilja slumpvisa variationer i urvalet från systematiska fel. Flyvbjerg m.fl. (2002) dataunderlag består av 258 infrastrukturprojekt från 20 länder som genomfördes under perioden mellan 1927 och 1998.

Flyvbjerg m.fl. (2002) konstaterar att det är betydligt vanligare att ett infrastrukturprojekt går över än under budget. I nästan 9 av 10 undersökta projekt var kostnaden högre än prediktionen. Författarna visar även att hypotesen att storleken på kostnadsunderskattningarna är lika stora som kostnadsöverskattningarna kan förkastas med god evidens (p -värde $< 0,001$). Utifrån dessa två resultat drar Flyvbjerg m.fl. (2002) slutsatsen att kostnadsunderskattningar inte bara är vanligare utan också större jämfört med överskattningar. Den genomsnittliga kostnadsökningen för de 258 infrastrukturprojekten var 28%.

Efter Flyvbjerg m.fl. studie har ett flertal studier med liknande ansats publicerats. För en litteraturgenomgång se exempelvis Miranda Sarmiento och Renneboog

(2017). Bland annat utökar Cantarelli m.fl. (2012b) dataunderlaget i Flyvbjerg m.fl. (2002) till att omfatta totalt 806 globala infrastrukturprojekt. Skillnaderna mellan resultaten i studierna är små med möjligt undantag för järnvägsprojekt. Flyvbjerg m.fl. (2002) rapporterar kostnadsökningar för väg-, järnvägs- samt bro- och tunnelprojekt på 20%, 45% respektive 34%. Motsvarande siffror i det utökade underlaget var 20%, 34% och 33%. Jämförelsen av järnvägsprojekten kan dock vara missvisande av fler skäl. Ett exempel är att järnvägsprojekt från Nederländerna är överrepresenterade i Cantarelli m.fl. (2012b). Rensat för dessa är kostnadsökningen 38%. Vidare är observationer från övriga världen, det vill säga inte Europa eller Nord Amerika, överrepresenterade i Flyvbjerg m.fl. (2002) jämfört med Cantarelli m.fl. (2012b). Projekten i övriga världen hade klart störst kostnadsökningar vilket kan ha påverkat det globala genomsnittet för järnvägsprojekten i Flyvbjerg m.fl. (2002).

I en liknande studie undersöker Cantarelli m.fl. (2012a) kostnadsavvikelser för 78 väg-, järnvägs-, bro- och tunnelprojekt i Nederländerna. För de 37 undersökta vägprojekten visar resultatet att kostnaderna ökat i 62% av de undersökta fallen och minskat i 38% av fallen. Andelsskillnaden mellan under- och överskattningar var dock inte statistiskt signifikant (p-värde 0,188). Däremot var kostnadsökningarna statistiskt signifikant större än minskningarna (p-värde 0,009), vilket resulterade i en genomsnittlig kostnadsökning på 19% för de undersökta vägprojekten. För järnvägsprojekt samt bro- och tunnelprojekt (eng. *fixed link projects*) var det ungefär lika många projekt som gick över som gick under budget. Kostnadsökningarna var dock även här statistisk signifikant större än kostnadsminskningarna i de båda kategorierna (p-värde järnvägsprojekt 0,006, p-värde broar och tunnlar 0,011). Sammantaget var den genomsnittliga kostnadsökningen i nederländska infrastrukturprojekt 17%.

En annan typ av projekt som har undersökts med motsvarande metod är infrastrukturprojekt för produktion av elektricitet. Ansar m.fl. (2014) undersöker 245 dammprojekt i 65 olika länder och finner att den genomsnittliga kostnadsökningen var 96%. Författarna drar slutsatsen att kostnadsökningarna var signifikant större än kostnadsminskningarna (p-värde < 0,01). Sovacool m.fl. (2014) tar i sin studie ett bredare grepp på kraftproduktion och inkluderar utöver dammar även kärnkraftverk, vindkraftverk, solkraftverk och annan tillhörande infrastruktur lokaliserade över hela världen. Den genomsnittliga kostnadsavvikelsen bland de 401 undersökta projekten var 66%.

Projekt för att ta fram och driftsätta nya vapensystem skiljer sig delvis från ovan nämnda infrastrukturprojekt då det är en industriell utveckling ”på teknikens framkant” och produktion snarare än byggnation som står i centrum. Projekten liknar dock varandra i att det ofta är långa och komplexa processer. Arena, m.fl. (2006) undersöker riktigheten hos kostnadskalkyler för vapenutvecklingsprogram som antingen tillhör eller liknar de inom det amerikanska flygvapnet. Det innebär

att andra typer av materielprojekt som rör exempelvis stridsfartyg och ubåtar inte ingår i det undersökta urvalet.

Det totala urvalet består av 46 vapenprogram. Beslut om amerikanska vapenprogram fattas i tre steg: 1) konceptutvecklingsfasen, 2) utvecklingsfasen och 3) produktionsfasen. Arena, m.fl. (2006) undersöker i sin studie riktigheten i kostnads kalkyler från det andra beslutet, dvs. när utvecklingsfasen påbörjas. Författarna fann att majoritet av de undersökta programmen var behäftade med kostnadsökningar. Den genomsnittliga kostnadsökningen var 46% och nästan 7 av 10 program blev 25% dyrare eller mer. Frågan om man genom att kategorisera program kan dra slutsatser om kostnadsökningar kommer att diskuteras i kapitel 3.3 och 4.

Noterbart är att vi har funnit relativt få publicerade studier som undersökt kostnadsökningar i svenska projekt. Studier som har undersökt den genomsnittliga kostnadsökningen likt ovan nämna studier är ovanliga i den svenska kontexten. Utöver Riksrevisionen (2021a) som diskuterats ovan har endas tre ytterligare studier identifierats.⁸ Vidare kan det noteras att samtliga av dessa är skrivna inom transportinfrastruktur. Inom exempelvis försvarsmaterielprojekt har vi inte kunnat identifiera någon tidigare studie med likande ansats som exempelvis Arena m.fl. (2006)⁹. Tabell 1 sammanfattar resultaten från ovannämnda studier. Tabellen visar att det är vanligt med relativt stora snittkostnadsökningar i de olika typerna av projekt som undersöks. Detta verkar även vara ett globalt problem då systematiska kostnadsökningar har dokumenterats världen över. Problemet med systematiska kostnadsökningar verkar alltså vara ett frekvent och robust fenomen.

⁸ De tre övriga studierna är Riksrevisionen (2010), Riksrevisionen (2011) och Lundberg m.fl. (2011). Resultatet från de tre studierna skiljer sig inte på något avgörande sätt från Riksrevisionen (2021a).

⁹ Tidigare studier har undersökt liknanden frågeställningar utifrån andra ansatser, se exempelvis Statskontoret (2018) och ESV (2019) för analys av ekonomistyrningen, Nordlund m.fl. (2015) för en diskussion om incitamentsstrukturer i materieförsörjningsprocessen samt Olsson (2019) och Olsson m.fl. (2018) för en diskussion om ökade kostnader för försvarsmateriel.

Tabell 1: Snittkostnadsavvikelser i olika typer av projekt

Studie	Typ av projekt	Geografiskt område	Genomsnittlig kostnadsökning (p-värde)
Flyvbjerg m.fl. (2002)	Infrastruktur: vägar, järnvägar, broar och tunnlar	Global	28% (<0,001)
Cantarelli m.fl. (2012a)	Infrastruktur: vägar, järnvägar, broar och tunnlar	Nederländerna	17% (<0,001)
Riksrevisionen (2021a)	Infrastruktur: vägar och järnvägar,	Sverige	9%*
Ansar m.fl. (2014)	Infrastruktur: dammar	Global	96% (<0,001)
Sovacool m.fl. (2014)	Infrastruktur: elektricitet	Global	66%
Arena, m.fl. (2006)	Krigsmateriel	USA	46% (<0,001)

*Riksrevisionen rapporterar inte som de andra studierna en siffra från första beslut till slutkostnad. Från sista nationella planen till slutkostnad var kostnadsavvikelsen 9%. Som diskuterats i 3.1 kan de diskuterats när det faktiska beslutet att genomföra projektet togs och om den verkliga kostnadsavvikelsen därför inte bör vara större.

3.3 Faktorer som påverkar storleken på underskattningen

Som framgår av tabell 1 varierar storleken på kostnadsunderskattningarna beroende på typen av projekt. För att skapa bättre förståelse för denna variation görs i detta avsnitt några nedslag i litteraturen kring faktorer som kan påverka storleken på de systematiska felen i kostnadskalkylerna.

Flyvbjerg och Cowi (2004) konstaterar att projektlängd och komplexitet är två viktiga faktorer som påverkar storleken på kostnadsunderskattningar. De fann i linje med det vi diskuterat tidigare att projekt som pågår under längre tid är mer känsliga för förändringar i koncept, design, och ambition, som i sin tur driver kostnaderna för projektet uppåt. På liknande sätt konstaterar Lee (2008) i sin studie av sydkoreanska infrastrukturprojekt att projektförändringar är en del av de drivande faktorerna för kostnadsökningar inom transportsektorn. Exempel på förändringar som Lee noterar är fler filer vid vägbyggen, förlängningar av järnvägslinjer samt nya stationer vid järnvägsprojekt, vilka alla är relativt stora och kostnadsdrivande förändringar.

Riksrevisionen (2021a) och Welde och Odeck (2017) gör motsvarande observation för svenska respektive norska förhållanden, dvs. att projektförändringar är en del-förklaring till de systematiska kostnadsökningarna. Welde och Odeck (2017) framhäver att det kan vara stora skillnader mellan de projektförslag som politiken först tar ställning till och den slutliga utformningen av projektet som sedan utförs. Under planerings- och projekteringsfasen lobbbar intresseorganisationer och politiker för förändringar för att gynna olika särintressen. Resultatet blir ofta relativt stora

förändringar i projekten. Detta är särskilt problematiskt när det faktiska beslutet att genomföra projektet tas tidigare än det formella beslutet. Denna förklaringsmodell diskuteras närmare i Kapitel 4.

Det bör dock noteras att det inte bara är lobbygrupper som kan driva igenom förändringar i projekten. Andra faktorer som Flyvbjerg och Cowi (2004) lyfter fram är förändrade normer och regleringar. Nya lagar och regleringar gör att projektet måste anpassas efter de nya förutsättningarna vilket ofta leder till fördyringar i projekten. Som nämnt ovan befinner sig ett typiskt svensk infrastrukturprojekt i den nationella infrastrukturplanen i tolv år. Under den tiden kan exempelvis byggnormer ändras och miljöregleringar tillkomma, vilka tvingar fram designändringar i projekten. Ju längre projekt är desto större blir risken för att nya regleringar förändrar förutsättningarna.

Andra faktorn som Flyvbjerg och Cowi (2004) tar upp är komplexiteten i projekten. Nya och oprövade tekniker ökar komplexiteten i projekten då det finns liten eller ingen tidigare erfarenhet av att använda och implementera tekniken. Younossi m.fl. (2007) jämför kostnadsökningar i projekt med olika komplexitet. Författarna konstaterar att kostnadsökningarna är lägre för transportinfrastruktur jämfört med krigsmateriel och civila rymdprojekt, där ökningen för de två sistnämnda projektyperna är av jämförbara magnituder. Författarna förklarar skillnaden med att den teknik som används i transportinfrastruktur, till skillnad från de andra typerna av projekt, är välkänd och ny teknik införs långsamt i steg. I kontrast till detta sker utveckling av nya vapensystem och rymdprojekt med införandet av fler nya tekniker, vilket tydliggör varför kostnaderna ökar mer i dessa projekt.

Ett illustrativt exempel på hur projektlängd och komplexitet kan påverka kostnadsavvikelserna hos projekt är Sovacool m.fl. (2014) undersökning av kraftproduktionsprojekt. Den genomsnittliga kostnadsökningen för samtliga 401 projekt i studien var 66%. Vattenkraftverks- och kärnkraftsprojekt som karakteriseras av komplexitet och längd har en genomsnittlig kostnadsökning på 71 respektive 117%. Detta kan jämföras med vindkraftverk och solcellsparkar som hade kostnadsökningar på 8 respektive 1%. Författarna förklarar skillnaden med att vind- och solkraft bygger på mer standardiserade och massproducerade komponenter. Vidare gör en stor andel av förtillverkade delar att byggtiden blir betydligt kortare. Den relativt höga andel standardiserade delar och byggmoment gör med andra ord att projektlängden blir kortare och komplexiteten lägre.

3.4 Förbättringar över tid

Kostnadsökningar i offentliga projekt har varit ett känt problem under lång tid. Flyvbjerg m.fl. (2003) påpekar i sin studie om civil infrastruktur att om kostnadsavvikelserna beror på bristande kunskaper eller erfarenheter borde de minska över tid. Med andra ord borde det finnas en lärandeeffekt, där organisationerna och projektledarna lär sig av tidigare projekts kostnadsavvikelser. Flyvbjerg m.fl. (2003)

konstaterar dock att denna lärandeeffekt inte kan observeras i data och att kostnadsavvikelserna inte verkar minska över tid.

Ett annat exempel är amerikanska vapenprogram där problemen med kostnadsökningar varit kända sedan åtminstone 1960-talet (se exempelvis Marshall and Meckling (1962); Scherer (1964); Williamson (1967)). Den amerikanska kongressen har tagit flera initiativ för att stävja problemen. Sedan 1970 har det nästan årligen genomförts en reform för inköpsprocessen för militär materiel.¹⁰ Huvuddelen av dessa har fokuserat på att minska kostnadsökningarna i projekten (Ritschel, 2012).

Younossi m.fl. (2007) bygger vidare på Arena m.fl. (2006) ovan nämnda studie om vapenprogram i det amerikanska flygvapnet för att undersöka om trenden i kostnadsökningarna har förändrats över tid. Resultatet i studien visar att den genomsnittliga kostnadsökningen har minskat kraftigt under både 1980-talet och 90-talet, jämfört med 70-talet. Författarna lyfter dock blandat annat två faktorer för varför en rak jämförelse av genomsnittet mellan årtiondena kan vara missvisande. För det första så har inte de längre projekten från 90-talet avslutats vid studiens genomförande. För det andra så har sammansättningen av projekttyper sett olika ut under tidsperioden. Tidigare forskning indikerar att det är faktorer som har en påverkan på kostnadsökningar (se exempelvis Arena m.fl., 2006; McNicol, 2004, Drezner m.fl., 1993). Efter att ha justerat för dessa två faktorer drar Younossi m.fl. (2007) slutsatsen att det inte finns några statistiskt signifikanta skillnader i genomsnittliga kostnadsökningar under den undersökta perioden.

Younossi m.fl. (2007) är inte ensamma om slutsatsen att kostnadsökningarna i amerikanska vapenprogram inte minskat över tid. Ritschel m.fl. (2019) undersöker likt Arena m.fl. (2006) och Younossi m.fl. (2007) program inom det amerikanska flygvapnet. Ritschel m.fl. (2019) kompletterar Arena m.fl. (2006) och Younossi m.fl. (2007) studier genom att inkludera program som ännu inte hade avslutats innan de två tidigare studierna. Det gör att Ritschel m.fl. kan undersöka om den senaste större reformen Weapons Systems Acquisition Reform Act (WSARA), från 2009 har haft någon effekt. Författarna konstaterar att kostnadsökningarna inom forskning och utvecklingsprojekt minskat efter WSARA. För övriga undersökta delområden fann författarna dock oförändrade eller större kostnadsökningar. Utifrån det dras slutsatsen att WSARA-reformen från 2009 inte har haft någon statistisk signifikant effekt på kostnadsavvikelserna. Ritschel m.fl. (2019) konstaterar att WSARA likt fyra andra liknande stora reformerna som bäst har haft ett blandat resultat.

Givet ovanstående studier verkar det inte finnas några lärandeeffekter som minskar de systematiska kostnadsunderskattningarna över tid. Faktumet att den amerikanska kongressen har genomfört ett antal större reformer utan att minska kostnadsökningarna visar på hur svårlöst problemet verkar vara.

¹⁰ Fem av reformerna har varit större och mer omfattande. Övriga har varit mindre justeringar. Se Ritschel (2012) för en genomgång av de större reformerna.

4 Orsaker till och metoder mot systematiska fel

I detta kapitel går vi först igenom olika förklaringsmodeller till varför kostnads-kalkyler tenderar att vara behäftade med systematiska underskattningar. Därefter går vi igenom en metod som kan användas för att justera prediktioner för att minimera dessa systematiska fel.

4.1 Orsaker

För att avgöra vad som driver de systematiska felen som görs i kostnads-kalkyler har bland annat Flyvbjerg m.fl. (2002, 2004) och Flyvbjerg och Cowi (2004) testat såväl tekniska som ekonomiska, politiska och psykologiska förklaringsmodeller. Några av de potentiella orsakerna har vi redan varit inne på i Avsnitt 3.3 medan andra är nya. Tekniska förklaringsmodeller försöker förklara kostnadsavvikelser i tekniska termer, vilka innefattar otillräcklig data, oförfarna prediktionsmakare, ofullständiga tekniker, den genuina osäkerheten kopplad till att förutspå framtiden samt ärliga misstag. Ekonomiska och politiska förklaringsmodeller fokuserar istället på de involverade aktörernas incitament till att göra kalkyler tillförlitliga, medan de psykologiska förklaringsmodellerna istället förklarar prediktionsfel genom att peka på systematiska brister i tänkandet hos beslutsfattare och prediktionsmakare (Flyvbjerg m.fl., 2003).

Det är viktigt att poängtera att vi i detta avsnitt försöker reda ut varför kostnads-kalkyler tenderar att systematiskt underskatta kostnader för projekt. Denna fråga skiljer sig från frågan huruvida kostnadsökningarna är befogade. Exempelvis skulle en systematisk målglidning uppåt kunna göra att både nyttan och kostnaden av projekten systematiskt stiger på ett sätt så att samhällsnyttan av projekten inte minskar kostnadsökningarna till trots. Liknande kan oförutsedda händelser kräva befogade kostnadsökningar för att den ursprungliga nyttan ska kunna realiseras. Eftersom det inte är möjligt att uttala sig om kostnadsökningen är befogad eller inte utan att ha koll på nyttosidan kan vi inte uttala oss om detta.

Med det sagt är det av åtminstone två anledningar viktigt att förstå varför kostnaden systematiskt underskattas, även om man inte kan säga något om den mer komplicerade nyttosidan av projekten. För det första, som diskuterats i Kapitel 2, prioriteras projekt utifrån fasta budgetramar som endast tar hänsyn till kostnadsdelen. Detta gör att otillförlitliga kostnads-kalkyler är problematiska trots att den beräknade samhällsnyttan av projekten skulle kunna vara korrekt.¹¹ För det andra är det fortfarande problematiskt att både nytta och kostnader systematiskt underskattas (som måste

¹¹ Det ska dock nämnas att studier av till exempel Flyvbjerg m.fl. (2005) visat på att nyttan med infrastrukturprojekt systematiskt överskattas.

vara fallet om samhällsnytta ska förbli densamma) utan att observeras och att prediktionerna förbli felaktiga över tid.

Tekniska förklaringsmodeller — Inte enligt empirin

De tekniska förklaringsmodellerna är vanliga i den äldre litteraturen (se exempelvis Vantson och Vantson, 2004). I den senare litteraturen argumenterar Flyvbjerg m.fl. (2003) och Flyvbjerg (2008) istället för att sådana förklaringsmodeller endast implicerar att kostnadsavvikelse är approximativt normalfördelade med en snittavvikelse nära noll. Detta innebär att tekniska förklaringsmodeller endast kan förklara varför den aggregerade kostnadsrisken är stor men inte varför kostnadskalkylerna är behäftade med systematiska fel. Dessutom bör man se en förbättring i träffsäkerheten av prediktionerna eftersom felen och deras källor borde ha uppmärksamats och åtgärdats över tid. Som vi har framfört i Kapitel 3 verkar inte detta ha skett. Vidare har intervjuer gjorts med personer ansvariga för framtagandet av kostnadskalkyler som understryker bristerna i de tekniska förklaringsmodellerna (Flyvbjerg och Cowi, 2004; Wachs, 1990).

Psykologiska förklaringsmodeller — Optimism bias

Termen *optimism bias*, eller på svenska optimistisk snedvridning, syftar till den vanligt förekommande kognitiva snedvridningen som får personer att systematiskt tro att framtida händelser kommer vara mera fördelaktiga för dem än vad de har skäl till att tro givet deras tidigare erfarenheter (Flyvbjerg, 2008). Termen kommer ursprungligen från ”planning fallacy”, eller planeringsmässig villfarelse, som myntades av Kahneman och Tversky (1979b). Planning fallacy relaterar enbart till optimistiska prediktioner angående hur lång tid det kommer ta att slutföra ett projekt. Termen utvecklades till att bli optimism bias först när Lovallo och Kahneman (2003) använde den för att även innefatta kostnadsunder- och fördelsöverskattningar.

Lovallo och Kahneman (2003) argumenterar för att den optimism som kännetecknar fenomenet beror på att aktörer tenderar att ta ett inifrånperspektiv och ser på projekten de jobbar med som unika. Detta lyfts också fram av Flyvbjerg (2008) som anledningen till varför optimism bias förväntas leda till systematiska fel i kostnadskalkyler eftersom det får planerare och prediktionsmakare att fokusera för mycket på detaljer angående det enskilda projektet, snarare än att ta ett utifrånperspektiv och jämföra det med liknande tidigare. Aktörernas optimism kan därmed förklara varför kostnadskalkyler tenderar att systematiskt underskatta projektens slutkostnad. Utöver detta har Kahneman och Tverskys (1979a) arbete visat att även om personer görs medvetna om sina kognitiva snedvridningar tenderar de fortfarande att tycka att deras resonemang eller agerande är korrekt. Detta kan förklara varför prediktionerna inte förbättras över tid.

Ekonomiska/politiska förklaringsmodeller — Strategic misrepresentation

En annan förklaringsmodell som är baserad på ekonomiska och politiska orsaker har framförts av bland annat Wachs (1989, 1990) och Flyvbjerg m.fl. (2002, 2005). Denna modell har kommit att benämnas *strategic misrepresentation*, eller strategisk förvrängning på svenska. Termen används för att beskriva fenomenet när prediktionsmakare medvetet och strategiskt underskattar kostnader för sitt projekt för att öka sannolikheten att projektet faktiskt genomförs. Exempelvis vittnar Wachs (1990) om att flera planerare, ekonomer, och ingenjörer berättat att det tvingats göra om sina kostnadskalkyler för att göra sina överordnade glada. Dessa överordnade har haft specifika kostnader i åtanke som de ansett vara tillräckligt låga för att få igenom projekten i konkurrens med andra.

Som poängterats av Wachs (1990) är en väsentlig egenskap hos prediktioner att de inte kan bevisas vara fel förrän efter de har spelat ut sin rätt. Detta innebär att prediktioner riskerar att snedvridas till att visa det utfall som prediktionsmakaren har incitament till att predicera. En av anledningarna till att detta tillåts fortgå är den genuina osäkerheten som projektet är behäftade med. Som vi har varit inne på tidigare kan en prediktion inte utvärderas av ett enskilt utfall på ett trovärdigt sätt utan många prediktioner måste utvärderas med hjälp av statistiska metoder. En annan anledning till att prediktionerna inte förbättras över tid är att prediktionsmakarna inte har incitament till att förbättra dem. Om prediktioner snedvrids för att få igenom strategiska projekt är det inte säkert att de som är ansvariga för prediktionerna vill att de ska förbättras över tid (Flyvbjerg, 2008).

Flyvbjerg (2008) argumenterar för att förklaringsmodeller baserade på optimism bias och strategic misrepresentation kompletterar varandra och att den ena förklaringsmodellen fungerar bättre i situationer där den andra fungerar sämre. Till exempel har det argumenterats för att optimism bias är en mer trolig orsak när trycket från politiskt och organisatoriskt håll är mindre medan strategic misrepresentation är en bättre förklaring när detta tryck är högt. Dock finns det även många situationer där båda fenomenen tillsammans bidrar till systematiska kostnadsunderskattningar (Flyvbjerg, 2008).

4.2 Motverkan — prediktioner i referensklasser

Sedan 2005 rekommenderar American Planning Association planerare att inte enbart förlita sig på ingenjörsmetoder som traditionella kostnadsosäkerhetsanalyser när de gör kostnadskalkyler. Istället bör sådana metoder kompletteras med en så kallad *referensprediktionsmetod* (eng. *reference class forecast*). Denna metod utvecklades och utvärderades av Kahneman och Tversky (1979b) och senare Lovallo och Kahneman (2003) för att motverka just systematiska fel i prediktioner.

I stort föreslår författarna att i princip ignorera anledningen bakom de felaktiga prediktionerna och istället fokusera direkt på utfallen för att justera prediktioner.

Metoden fungerar övergripande beskrivet som ett sätt att få prediktionsmakaren att se prediktionsproblematiken från ett utomstående perspektiv. Detta för att, som tidigare diskuterats, motverka tendensen hos personer som är involverad i ett projekt att bara se de unika delarna men inte likheterna med andra projekt som redan har realiserats. Som har påpekats av Kahneman och Tversky (1979b) bör en prediktionsmakare använda sig av all möjlig information som kan säga något om sannolikheten för utfallet som de försöker predicera, med andra ord med alla möjliga medel försöka skatta fördelningen av utfallen. Metoden som beskrivs närmare nedan, har testats i flera laboratorieexperiment och visat sig vara mer exakt än traditionella prediktionsmetoder (Kahneman och Tversky, 1979b; Lovallo och Kahneman, 2003).

Referensklassprediktioner utförs i tre steg. I det första steget identifieras en relevant referensklass av tidigare liknande projekt. I det andra steget skattas sannolikhetsfördelningen för kostnadsavvikelse i referensklassen. Till sist jämförs projektet vars prediktion man vill justera, vanligtvis i form av ett centralmått, med referensklassfördelningens prediktionsavvikelser. Givet det empiriska centralmålet från kostnadsavvikelse i referensklassen kan prediktionen för projektet justeras uppåt eller nedåt. Detta kan liknas med metoden för att utvärdera prediktioner som presenterades i Avsnitt 2.3. Skillnaden är nu att detta görs innan projektet har sjösatts istället för i efterhand när slutkostnaden redan är beräknad.

För att konkretisera hur metoden kan användas presenterar vi först hur en referensklassupplyftning kan gå till tillämpat på tärningsexemplet från Kapitel 2. Därefter presenteras två faktiska tillämpningar av metoden.

För att tillämpa metoden på tärningsexemplet antar vi nu att 500 prediktioner har gjorts baserat på (det felaktiga) antagandet att världen är som beskriven av Värld 1, där kostnadsutfallet ges av summan av två vanliga tärningar. Dessa prediktioner utvärderas utifrån de 500 simulerade referensgruppsprojekten som gjordes i Kapitel 2 givet att den faktiska världen är Värld 2. För att justera snittkostnadsavvikelsen görs först en initial prediktion givet antaganden för de föregående prediktionerna (alltså att Värld 1 är korrekt), vilket ger en prediktion baserat på snittkostnaden för projekten som är 7. Sedan justeras detta värde upp med den procentuella kostnadsavvikelsen som har skattats utifrån referensgruppen. Som tidigare noteras i Avsnitt 2.3, uppskattades detta värde till 11,3%. Det justerade prediktionsmedelvärdet för slutkostnaden blir då ca 7,79, vilket inte är långt ifrån den teoretiska snittkostnaden för projekten i Värld 2 på ca 7,74.

Flyvbjerg och Cowi (2004) och De Reyck m.fl. (2017) står för två tillämpningar av referensklassprediktioner. I båda fallen används metoden för att justera optimistiska kostnadskalkyler hos den brittiska transportstyrelsen. Flyvbjerg och Cowi fokuserar på medianen av kostnadsfördelningen som prediktionsvärde medan De Reyck m.fl. fokuserar på medelvärdet av kostnadsfördelningen. Till exempel finner

Flyvbjerg och Cowi att kostnadskalkylerna för järnvägsprojekten underskattar mediankostnaden med 40% i genomsnitt. Detta innebär att om prediktionsmakaren är villig att ta en 50% risk att kostnaden överstiger prediktionen (dvs. att prediktionen fångar medianen av kostnadsutfallen), och anser att de tidigare järnvägsprojekten utgör en bra referensklass för det nya projektet, bör nya prediktioner systematiskt höjas med 40%. De Reyck m.fl. tittar istället på kostnadsjusteringar baserat på kostnadskalkyler gjorda för järnvägsprojekt vid olika tidpunkter i projektet. De finner att medelvärdet för kostnadskalkylen för järnvägsprojekt behöver justeras upp med 64% i det första steget av projektet för att snittkostnadsavvikelsen inte ska vara behäftad med systematiska underskattningar av slutkostnaden.

När en referensklass ska utses till ett projekt är det ofta en avvägning mellan att ta med få projekt som i stor utsträckning liknar projektet för vilket prediktionen görs, och att referensgruppen blir för liten för att kunna göra meningsfull statistisk analys (dvs. den statistiska analysen kräver en tillräckligt stor datamängd för att statistiskt kunna förkasta felaktiga hypoteser). Som vi har sett i det föregående kapitlet verkar systematiska fel i kostnadskalkyler för en typ av projekt inte skilja sig nämnvärt från andra typer av till synes helt orelaterade projekt. Exempelvis om det saknas observationer skulle långa och komplexa infrastrukturprojekt kunna ingå i referensklassen för projekt som avser att utveckla krigsmateriel, vilka också tenderar att vara långa och komplexa.

Eftersom referensklassprediktioner inte fokuserar på att eliminera de underliggande orsakerna till att kostnadsfördyningarna sker finns det inget som säger att metoden skulle vara sämre att använda när den underliggande orsaken är ekonomiska och politiska snarare än psykologiska. Det kan dock vara svårt att införa metoden i sammanhang där beslutsfattarna och prediktionsmakarna agerar under stora politiska påtryckningar. Detta eftersom prediktionsmakarna kan ha svaga incitament att faktiskt förbättra kostnadskalkylerna.

En annan viktig aspekt som behöver hanteras när referensklassprediktioner används för att justera upp kostnadskalkyler är att det budgetutrymme som finns tillgängligt inom ett projekt tenderar att användas. Därför är det viktigt att säkerställa att uppjusteringarna av kostnadskalkyler inte leder till ökade incitament att spendera då detta skulle kunna leda till ännu värre fördyringar. Ett sätt att försöka förhindra detta är att inte kommunicera kostnadsuppjusteringarna till projektansvariga för att undvika att sända signalen att fördyringar är väntade.

5 Slutsatser och fortsatta studier

I den här rapporten har vi diskuterat systematiska fel i prediktioner tillämpat på kostnadskalkyler i investeringsprojekt. Vi började med att slå fast att kostnadskalkyler är viktiga beslutsunderlag behäftade med genuin osäkerhet. Osäkerheten gör det svårt att utvärdera enskilda kalkyler var för sig. Istället bör riktigheten i prediktionsmetodiken utvärderas genom statistiska analyser av ett representativt urval av liknande prediktioner, där projekten redan är slutförda och slutkostnaden fastställd. Vi har visat hur det är möjligt att med ett representativt dataunderlag och statistiska metoder testa hypoteser om huruvida kostnadskalkyler som gjorts är behäftade med systematiska fel.

En genomgång av litteraturen visar att kostnadskalkyler för offentliga projekt tenderar att systematiskt underskatta slutkostnaden för projekten. En konsekvens är att uppskattningar av lönsamheten och samhällsnyttan av projekten sannolikt ofta blir missvisande. Detta kan medföra att lämpligare projekt blir bortprioriterade eller försenade då beslutsunderlagen som används för prioriteringar mellan projekten är missvisande. Litteraturen visar även att underskattningen av slutkostnaderna i kostnadskalkylerna inte verkar bli lägre över tid fast problemet med kostnadsunderskattningar är välkänt sedan länge. En annan slutsats är att längre och komplexare projekt tenderar att vara behäftade med större fördyrningar.

Orsaken till de optimistiska kostnadskalkylerna brukar förklaras av främst psykologiska och ekonomiska/politiska faktorer. De psykologiska förklaringarna pekar på den inneboende optimism hos människan och tendensen att se varje projekt som unikt som en bidragande faktor till varför kostnadskalkyler är behäftade med systematiska fel. De ekonomiska och politiska förklaringarna pekar istället på ett betydande problem med aktörernas incitament där planerare och de som utvecklar prediktionerna tjänar på att göra optimistiska prediktioner. Dessa två förklaringsmodeller tycks komplettera varandra då den ena modellen verkar mer trolig i situationer där den andra modellen verkar mindre trolig.

I rapporten har vi även presenterat ett tillvägagångssätt för att justera kalkyler för att motverka systematiska fel med hjälp av prediktionsdata från referensklasser. Som har poängterats av American Planning Association redan 2005 bör planerare inte enbart förlita sig på ingenjörsmetoder som traditionella osäkerhetsanalyser när de gör kostnadskalkyler. Istället bör referensklasser med data över kostnadskalkylers riktighet i snarlika projekt användas för att kalibrera kalkylerna för projekten.

En viktig slutsats är att oavsett förklaringsmodell innebär problematiken med optimistiska och missvisande kostnadskalkyler att aktörer i den offentliga sektorn bör lägga kraft på att dokumentera kalkyler och faktiska kostnadsutfall i utvärderings- och kalibreringssyfte. Eftersom den empiriska litteraturen pekar på att det finns stora, systematiska, och statistiskt signifikanta kostnadsunderskattningar i investeringsprojekt är det svårt att acceptera att enskilda myndigheter inte kan

redogöra för hur väl de utför sina kalkyler. Lärdomarna från de amerikanska försöken att minska fördyrningarna i militära materielprojekt visar på hur komplicerad denna uppgift är. Detta medför att alla verktyg för att motverka problemen bör tas i beaktning. Anmärkningsvärt är att vi endast hittat ett fåtal formella granskningsrapporter i den svenska kontexten. Till exempel har det rapporterats om kostnadsökningar i ett fåtal stora materielprojekt men ännu inte systematiskt utretts hur stora kostnadsökningarna är i de projekt som Försvarets materielverk (FMV) ansvarar för (DN, 2021). Framtida studier ämnar därför ta fram dataunderlag för referensklasser genom att kartlägga kostnadsavvikelser i investeringsprojekt i den svenska offentliga sektorn.

En fördel med att systematiskt följa upp prediktioner och utfall för att använda i kalibreringssyfte är att det inte behöver påverka incitamentsstrukturen för prediktionsmakarna. Detta medför att riskerna att ingreppet oavsiktligt snedvrider beslutsunderlag ytterligare kan antas vara låga. Ett uppenbart problem med ett sådant arbete är att få fram ett tillräckligt stort dataunderlag för att kunna göra meningsfulla analyser. Därför kommer en viktig del av det framtida arbetet bestå i att utvärdera om det finns lämpliga referensgrupper för projekt där högkvalitativ data redan finns insamlad. Oavsett om det redan finns ett bra dataunderlag att systematisera finns det dock ett stort värde i att börja insamlandet av prediktioner och kostnadsutfall för att förbättra grunden för framtida utvärderingar.

6 Referenser

- Ansar, A., Flyvbjerg, B., Budzier, A., & Lunn, D. (2014). Should we build more large dams? The actual costs of hydropower megaproject development. *Energy policy*, 69, 43-56
- Arena, M. V., Leonard, R. S., Murray, S. E., & Younossi, O. (2006). Historical cost growth of completed weapon system programs. RAND Corporation, RAND.
- Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, 253(1), 1-13.
- Cantarelli, C. C., van Wee, B., Molin, E. J., & Flyvbjerg, B. (2012a). Different cost performance: different determinants?: The case of cost overruns in Dutch transport infrastructure projects. *Transport Policy*, 22, 88-95.
- Cantarelli, C. C., Flyvbjerg, B., & Buhl, S. L. (2012b). Geographical variation in project cost performance: the Netherlands versus worldwide. *Journal of Transport Geography*, 24, 324-331.
- De Reyck, B., Grushka-Cockayne, Y., Fragkos, I., Harrison, J. & Read, D. (2017) Optimism Bias Study - Recommended Adjustments to Optimism Bias Uplifts. UK Department of Transport. gov.uk, 20017-01-24
<https://www.gov.uk/government/publications/rail-infrastructure-optimism-bias-study> [hämtad 2021-09-30]
- DN. (2021). Notor för stridsflyg och ubåtar spräcker plan. dn.se, 2021-06-02
<https://www.dn.se/sverige/notor-for-stridsflyg-och-ubatar-spracker-plan/> [hämtad 2021-12-20]
- Drezner, J. A., Jarvaise, J. M., Hess, R. W., Hough, P. G., & Norton, D. (1993). An analysis of weapon system cost growth. Air force mobility enter Charleston AFB SC
- ESV. (2019). Utveckling av Försvarsmaktens investeringsplan. Ekonomistyrningsverket. ESV 2019:31, Stockholm 2019-04-29
- Expressen. (2019). Skandalbygget Nya Karolinska nomineras till pris – helt absurt. Expressen.se, 2019-10-29
<https://www.expressen.se/halsoliv/kronikorer/anna-basen/skandalbygget-nya-karolinska-nomineras-till-pris-helt-absurt/> [hämtad 2021-12-10]

- Flyvbjerg, B., Holm, M. S., & Buhl, S. (2002). Underestimating costs in public works projects: Error or lie?. *Journal of the American planning association*, 68(3), 279-295.
- Flyvbjerg, B., Skamris Holm, M. K., & Buhl, S. L. (2003). How common and how large are cost overruns in transport infrastructure projects?. *Transport reviews*, 23(1), 71-88.
- Flyvbjerg, B. & Cowi, (2004). *Procedures for Dealing with Optimism Bias in Transport Planning: Guidance Document*. London: UK Department for Transport.
- Flyvbjerg, B., Skamris Holm, M. K., & Buhl, S. L. (2005). How (in) accurate are demand forecasts in public works projects?: The case of transportation. *Journal of the American planning association*, 71(2), 131-146.
- Flyvbjerg, B. (2008). Curbing optimism bias and strategic misrepresentation in planning: Reference class forecasting in practice. *European planning studies*, 16(1), 3-21.
- Flyvbjerg, B., Ansar, A., Budzier, A., Buhl, S., Cantarelli, C., Garbuio, M., ... & van Wee, B. (2018). Five things you should know about cost overrun. *Transportation Research Part A: Policy and Practice*, 118, 174-190.
- Garvey, P. R., Book, S. A., & Covert, R. P. (2016). *Probability methods for cost uncertainty analysis: A systems engineering perspective*. New York: CRC Press.
- Kahneman, D., & Tversky, A. (1979a). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 363-391.
- Kahneman, D. & Tversky, A. (1979b) Intuitive prediction: Biases and corrective procedures, in: S. Makridakis & S. C. Wheelwright (Eds) *Studies in the Management Sciences: Forecasting*, p. 12 (Amsterdam: North Holland).
- Lee, J. K. (2008). Cost overrun and cause in Korean social overhead capital projects: Roads, rails, airports, and ports. *Journal of Urban Planning and Development*, 134(2), 59-62.
- Lovaglio, D. & Kahneman, D. (2003) Delusions of success: How optimism undermines executives' decisions, *Harvard Business Review*, July, pp. 56-63.
- Lundberg, M., Jenpanitsub, A., & Pyddoke, R. (2011). Cost overruns in Swedish transport projects. *Scandinavian Working Papers in Economics*

- Marshall, A.W., Meckling, W. H. (1962). Predictability of the Costs, Time, and Success of Development. In National Bureau of Economic Research, The Rate and Direction of Incentive Activity. Princeton University Press NJ.
- McNicol, D. L. (2004). Growth in the costs of major weapon procurement programs. IDA Paper P-3832
- Miranda Sarmiento, J., & Renneboog, L. (2017). Cost overruns in public sector investment projects. *Public Works Management & Policy*, 22(2), 140-164
- Nordlund, Peter, Bäckström, Peter & Johansson, Mattias (2015) Effektiv materielförsörjning – Ett incitamentsperspektiv på materielförsörjningen, FOI-R--4063--SE, februari 2015, FOI: Stockholm
- Olsson, Per, Bäckström, Peter, Johansson, Mattias, Lehman, Jens, Lusua, Jens, Ädel, Maria och Öhrn-Lundin, Josefin, Strukturella utmaningar inom det militära försvarets materiel- och personalförsörjning, FOI-R--4593—SE, maj 2018, FOI: Stockholm
- Olsson, Per (2019) Pang för pengarna – En ESO-rapport om Sveriges militära materielförsörjning. Statens offentliga utredningar, ESO-rapport 2019:7, ESO: Stockholm
- Riksrevisionen. (2011) Kostnadskontroll i stora järnvägsinvesteringar? Riksrevisionen 2011:06, Stockholm 2021-01-10
- Riksrevisionen, (2021a). Kostnadskontroll i infrastrukturinvesteringar. Riksrevisionen 2021:22, Stockholm 2021-06-08.
- Riksrevisionen. (2021b). Att bygga nationell försvarsförmåga – statens arbete med att stärka arméstridskrafterna. Riksrevisionen 2021:23, Stockholm 2021-08-19
- Ritschel, J. D. (2012). Efficacy of US Legislation in Military Acquisition Programmes: Nunn-McCurdy Act Unveiled. *Economic Papers: A journal of applied economics and policy*, 31(4), 491-500.
- Ritschel, J. D., Lucas, B., White, E., & Mrla, D. (2019). The impact of WSARA on the cost of air force weapon systems. *Journal of Public Procurement*.
- SCB. (u. å.). Central- och spridningsmått, scb.se
<https://www.scb.se/dokumentation/statistikguiden/grundlaggande-rakneregler-i-statistiken/central--och-spridningsmatt/> (hämtad 21-11-16).
- Scherer, F. 1964. The Weapons Acquisition Process: Economic Incentives. Boston: Harvard

- Simon, H. (1976). *Administrative behavior*. Nova Iorque: The Macmillan
- Sovacool, B. K., Gilbert, A., & Nugent, D. (2014). An international comparative assessment of construction cost overruns for electricity infrastructure. *Energy Research & Social Science*, 3, 152-160.
- Statskontoret. (2018). När planeringen möter verkligheten– Försvarsmaktens interna ledning, styrning och uppföljning. Statskontoret 2018:27, Stockholm 2018-12-03
- SVT. (2018). Analys: Mest misslyckade försvarsaffären i modern tid. svt.se, 2018-02-08 <https://www.svt.se/nyheter/inrikes/mest-misslyckade-forsvarsaffaren-i-modern-tid> [hämtad 2021-12-10]
- Trafikverket. (2021). Långsiktig planering av infrastruktur. trafikverket.se, 2021-02-10 <https://www.trafikverket.se/for-dig-i-branschen/Planera-och-utreda/langsiktig-planering-av-infrastruktur/> [hämtad 2021-07-22]
- Vanston, J. H., & Vanston, L. K. (2004). Testing the tea leaves: Evaluating the validity of forecasts. *Research-Technology Management*, 47(5), 33-39.
- Wachs, M. (1989). When planners lie with numbers. *American Planning Association. Journal of the American Planning Association*, 55(4), 476.
- Wachs, M. (1990). Ethics and advocacy in forecasting for public policy. *Business & professional ethics journal*, 141-157.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: context, process, and purpose. *The American Statistician*, 70(2), 129-133.
- Welde, M., & Odeck, J. (2017). Cost escalations in the front-end of projects– empirical evidence from Norwegian road projects. *Transport Reviews*, 37(5), 612-630.
- Williamson, O. E. (1967). The economics of defense contracting: Incentives and performance. In *Issues in defense economics* (pp. 217-278). NBER.
- Wooldridge.(2013). *Introductory Econometrics : A Modern Approach*. 5. Uppl. Mason, OH: South-Western Cengage Learning.
- Younossi, O., Arena, M. V., Jain, A., Leonard, R. S., & Roll Jr, C. R. (2007). Is weapon system cost growth increasing?: a quantitative assessment of completed and ongoing programs. RAND Corporation, RAND.

