Using Public Data for Validation

Alex King, Amplio Services

October 2016

info@amplioservices.com

Problem Statement

Why do we need to expand our verification options?

- The burden of proof is on the analyst
- Overspends remain a major organisational issue
- A lot of models remain rubbish-in, rubbish-out

This Presentation

How this will help you…

- Identify public data to assist validation
- Measure the data properly
- Convert measurements to actionable intelligence

info@amplioservices.com

Validation Definition

Validation is the process, or act, of demonstrating the complex model's ability to function as a credible estimating tool. Validation ensures: [1]

1. The model is a good predictor of costs
2. Estimating system policies and procedures are established and enforced
3. Key personnel have proper experience and are adequately trained.

[1] Parametric Estimating Handbook (2008). International Society of Parametric Analysts

Theory Overview

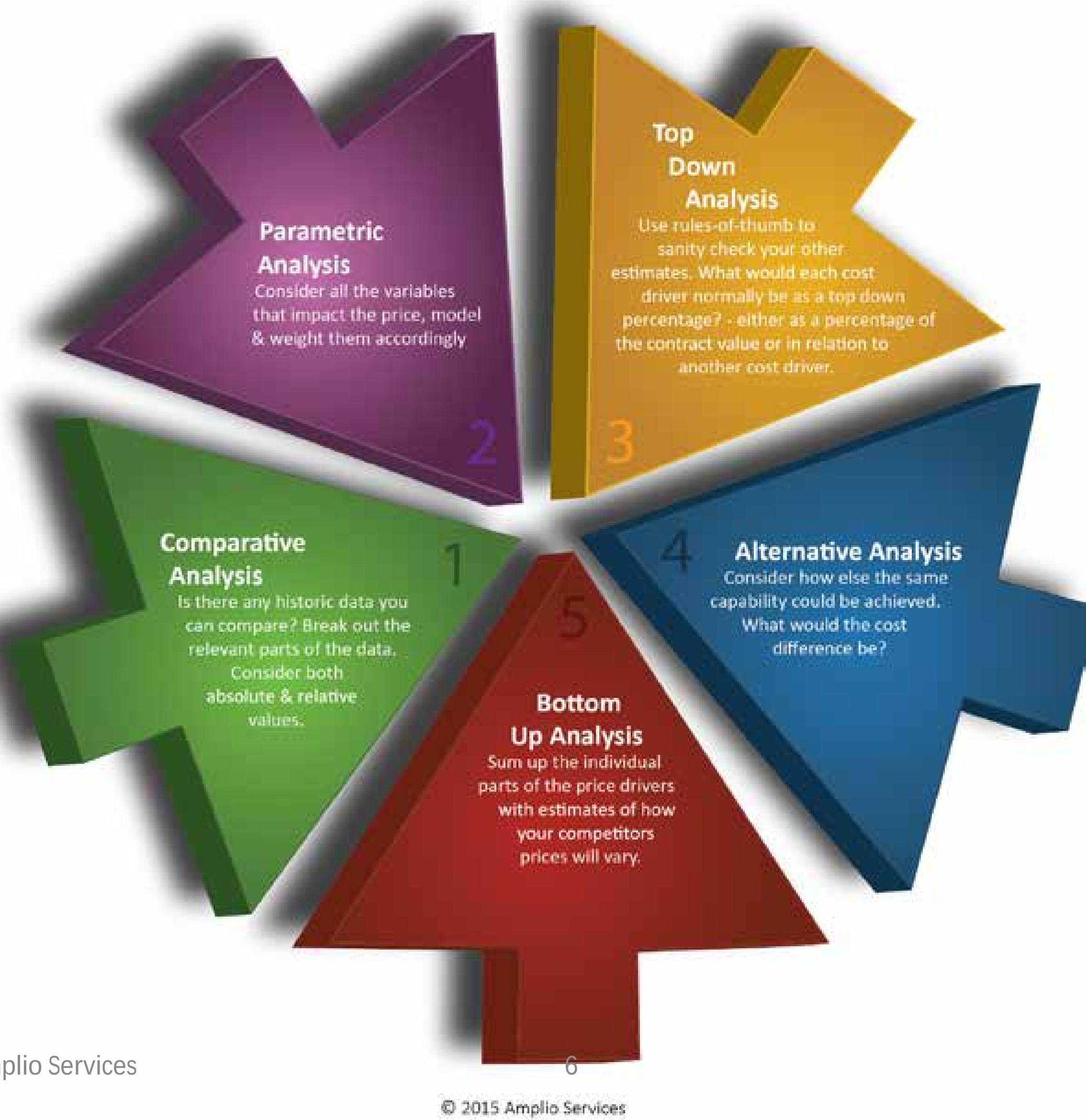
Using Public Data

Multiple Estimating Methods

If your bottom up estimates are 70% accurate. That means there they are 30% inaccurate.

If parametric estimates are 85% accurate and show a similar prediction, the combined accuracy is now 95.5%.

Add a comparative estimate with 60% accuracy that also supports our answer and our combined accuracy is now 98.2%.

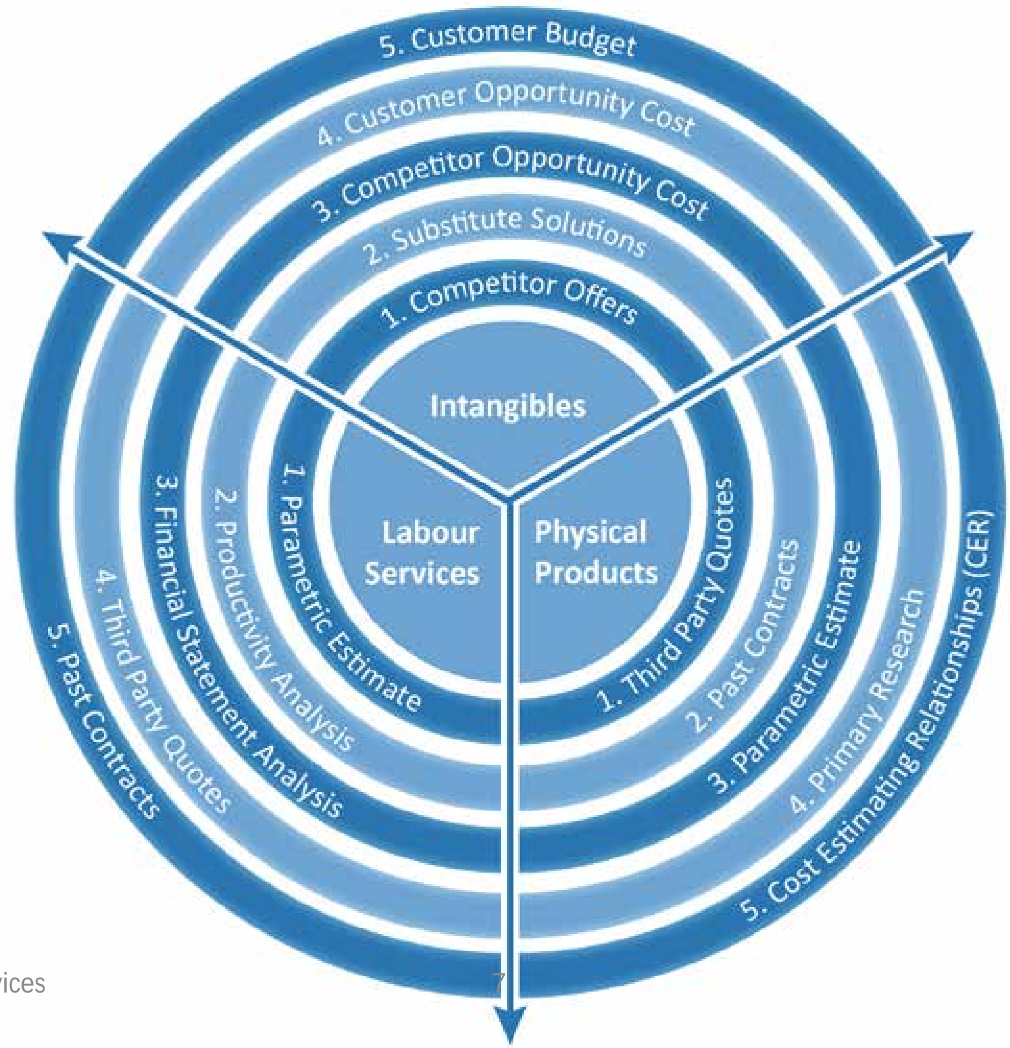


Finding Data

amplio

There is a hierarchy of data that we look for when producing estimates.

The best quality data is shown nearer the middle of the circle, however we usually gather as much as possible.

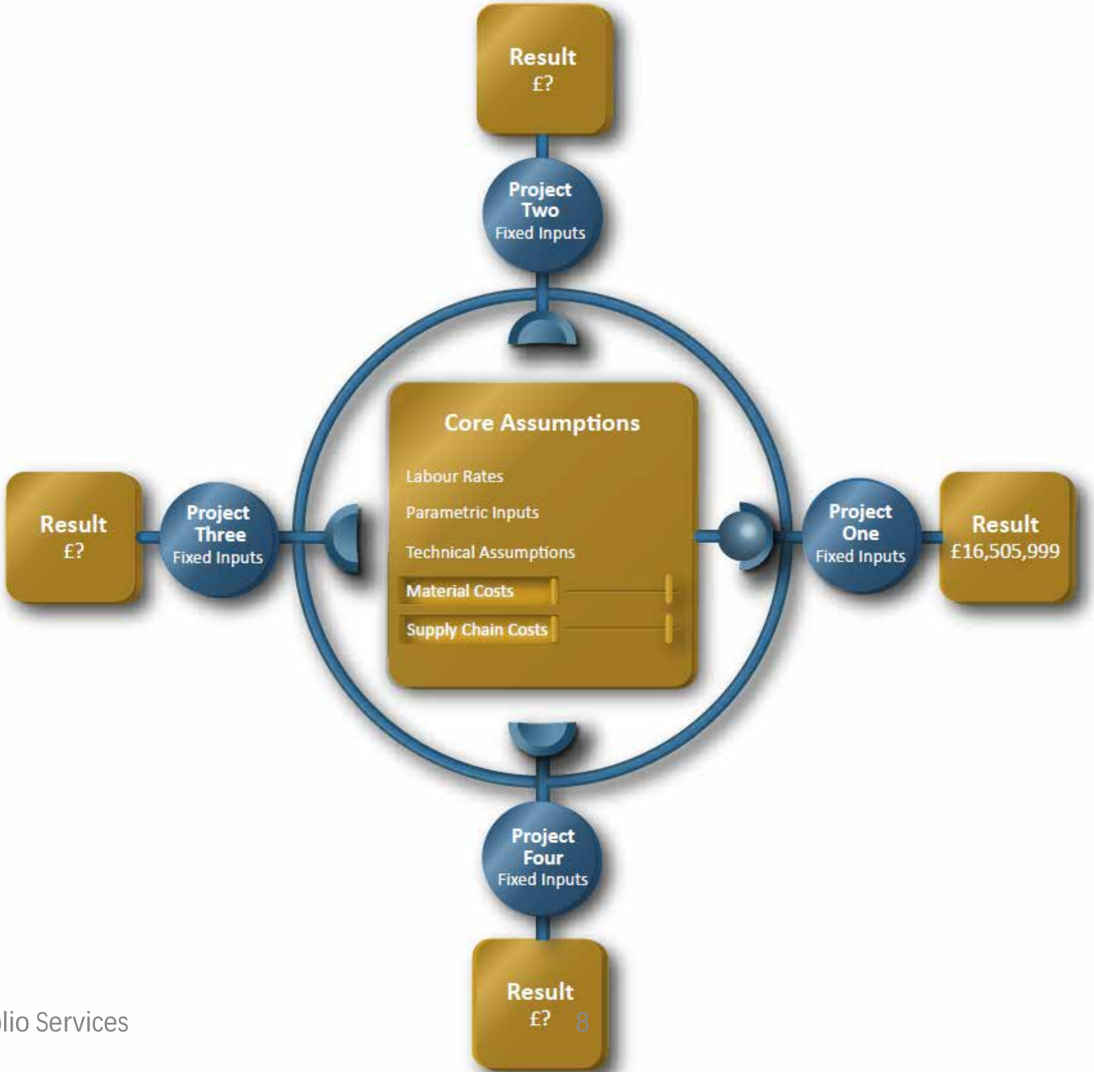


info@amplioservices.com

Data Pivoting

Data pivoting allows us to check for errors by re-using assumptions throughout multiple projects.

By maintaining a set of core assumptions, but varying project based inputs; we can identify potential errors.



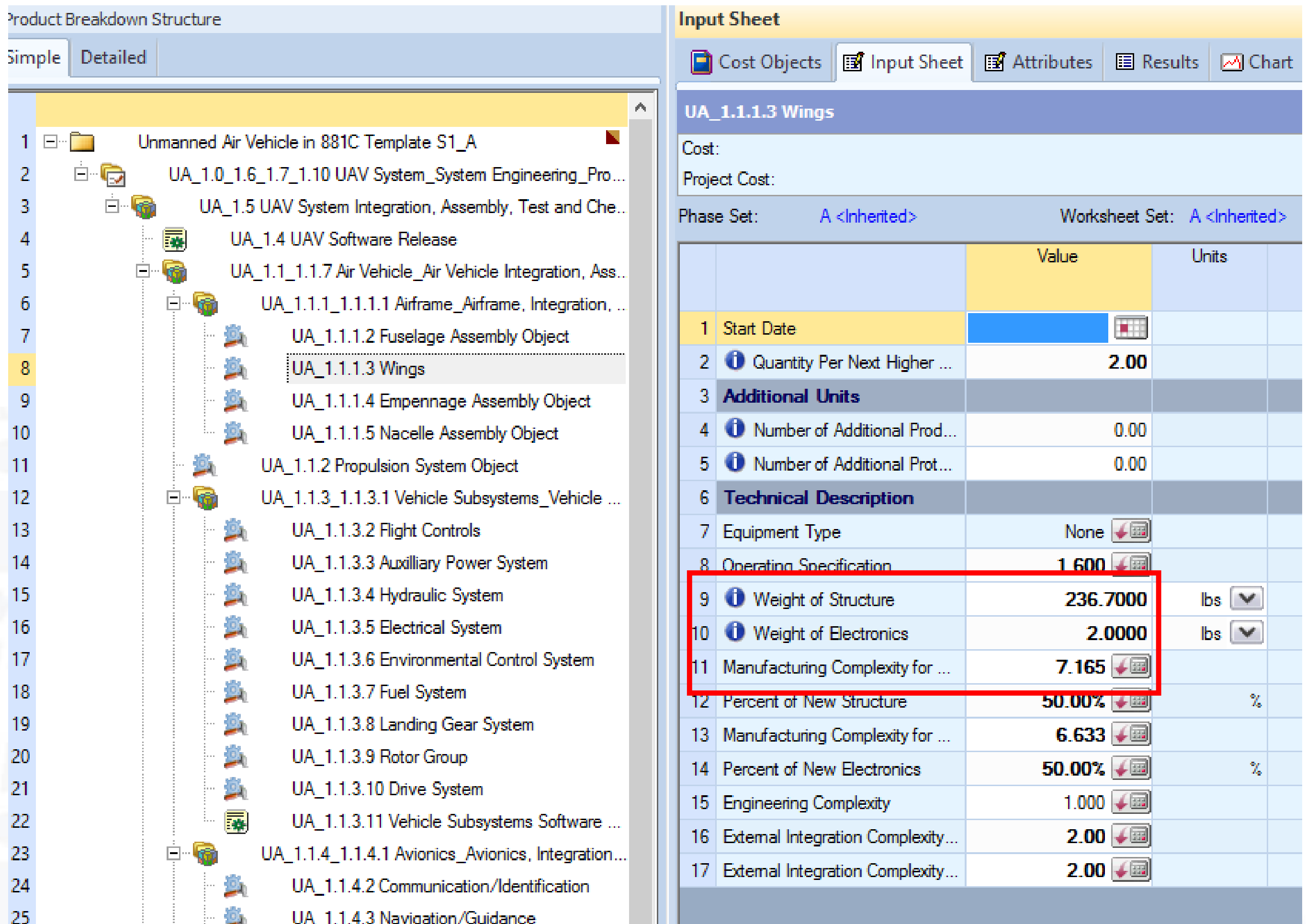
info@amplioservices.com

UAV Example

Validation Example

Parametric Estimate

If we have a **parametric estimate**, we are required to estimate inputs that have high uncertainty



info@amplioservices.com

US Procurement Reports

The US Department of Defence publishes some very detailed cost information that we can use as a **comparative estimate**

info@amplioservices.com

US Procurement Reports

amplio

Whilst we need to adjust for labour rates, currency, inflation & quantities – the upside is the sheer depth of data available in the US

UNCLASSIFIED

Exhibit P-5, Cost Analysis: PB 2014 Army

Date: April 2013

Appropriation / Budget Activity / Budget Sub Activity: 2031A: Aircraft Procurement, Army / BA 01: Aircraft / BSA 10: Fixed Wing

P-1 Line Item Nomenclature: A00005 - MQ-1 UAV

Item Nomenclature (*Item Number - Item Name, DODIC*): A00005 - MQ-1 UAV

Resource Summary	Prior Years	FY 2012	FY 2013#	FY 2014 Base	FY 2014 OCO##	FY 2014 Total	FY 2015	FY 2016	FY 2017	FY 2018	To Complete	Total
Procurement Quantity *(Units in Each)*	55	29	19	15	-	15	15	-	-	-	-	133
Gross/Weapon System Cost *($ in Millions)*	895.501	550.798	518.088	518.460	-	518.460	232.321	1.000	14.000	100.334	-	2,830.502
Less PY Advance Procurement *($ in Millions)*	-	-	-	-	-	-	-	-	-	-	-	-
Net Procurement (P1) *($ in Millions)*	895.501	550.798	518.088	518.460	-	518.460	232.321	1.000	14.000	100.334	-	2,830.502
Plus CY Advance Procurement *($ in Millions)*	-	-	-	-	-	-	-	-	-	-	-	-
Total Obligation Authority *($ in Millions)*	**895.501**	**550.798**	**518.088**	**518.460**	**-**	**518.460**	**232.321**	**1.000**	**14.000**	**100.334**	**-**	**2,830.502**
	(The following Resource Summary rows are for informational purposes only. The corresponding budget requests are documented elsewhere.)											
Initial Spares *($ in Millions)*	-	-	-	-	-	-	-	-	-	-	-	-
Gross/Weapon System Unit Cost *(Units in Thousands)*	16,281.830	18,993.030	27,267.780	34,564.000	-	34,564.000	15,488.067	-	-	-	-	21.282

FY 2013 Program is from the FY 2013 President's Budget, submitted February 2012

The FY 2014 OCO Request will be submitted at a later date

Cost Elements († indicates the presence of a P-5A)	ID CD	All Prior Years			FY 2012			FY 2013			FY 2014 Base			FY 2014 OCO			FY 2014 Total		
		Unit Cost *($ K)*	Quantity *(Each)*	Total Cost *($ M)*	Unit Cost *($ K)*	Quantity *(Each)*	Total Cost *($ M)*	Unit Cost *($ K)*	Quantity *(Each)*	Total Cost *($ M)*	Unit Cost *($ K)*	Quantity *(Each)*	Total Cost *($ M)*	Unit Cost *($ K)*	Quantity *(Each)*	Total Cost *($ M)*	Unit Cost *($ K)*	Quantity *(Each)*	Total Cost *($ M)*
Flyaway Cost																			
Recurring Cost																			
† Aircraft		4,406.000	55	242.347	4,206.000	29	121.964	5,286.000	19	100.429	5,396.000	15	80.944	-	-	-	5,396.000	15	80.944
Ground Control Station (GCS)		3,172.000	10	31.724	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Portable Ground Control Station (PGCS)		1,128.000	4	4.511	631.000	6	3.785	942.000	4	3.768	-	-	-	-	-	-	-	-	-
Universal Ground Control Station (UGCS)		3,202.000	15	48.037	2,757.000	14	38.593	4,060.000	8	32.478	3,537.000	8	28.296	-	-	-	3,537.000	8	28.296
Ground Data Terminal (GDT)		1,229.000	14	17.207	-	-	-	2,871.000	8	22.969	-	-	-	-	-	-	-	-	-
Universal Ground Data Terminal (UGDT)		1,189.000	15	17.842	1,218.000	21	25.585	-	-	-	-	-	-	-	-	-	-	-	-
Portable Ground Data Terminal (PGDT)		379.000	4	1.517	-	-	-	2,034.000	4	8.137	-	-	-	-	-	-	-	-	-
Automatic Take-Off & Landing Sys (ATLS)		713.000	16	11.412	815.000	7	5.704	1,535.000	4	6.138	1,046.000	6	6.274	-	-	-	1,046.000	6	6.274
Satellite Ground Data Terminal (SGDT)		1,632.000	17	27.739	1,719.000	7	12.030	2,947.000	4	11.786	2,205.000	8	17.640	-	-	-	2,205.000	8	17.640
Ground Support Equipment Kits (GSE)		2,554.000	6	15.325	2,127.000	7	14.886	2,974.000	4	11.896	2,768.000	6	16.608	-	-	-	2,768.000	6	16.608

UK NAO Reports

Again, if we adjust our models for the same settings, we can compare to a similar UK system. However, the level of detail here is generally lower

Watchkeeper

The Capability

Watchkeeper will provide the operational commander with a 24-hour, all weather, intelligence, surveillance, target acquisition and reconnaissance capability supplying accurate, timely and high quality imagery to support decision making. The system will consist of unmanned air vehicles, sensors, data links and ground control stations. Watchkeeper is planned to be delivered through an incremental programme to allow the system to benefit from both existing and developing sensors and air vehicle technology.

Overview of Cost, Time and Performance

	Approved	Forecast/Actual	Variation	IY Variation
Cost of Assessment Phase	£52m	£65m	+£13m	–
Cost of Demonstration & Manufacture Phase	£847m	£839m	-£8m	-£4m
Cost of Support Phase	£55m	£53m	-£2m	+£3m
Duration of Assessment Phase		68 months		
In-Service Date	June 2010	February 2012	+20 months	+12 months
Support Contract Go-Live	January 2010	January 2010	0 months	0 months
Support Contract End	May 2013	September 2014	+16 months	+16 months

UK Financial Statements

amplio

Where the UK does offer more information than the US is that private company financial statements are freely available.

In this instance, a joint-venture company carried out the contract. So the project revenue & company revenue are the same.

Profit and loss account

for the year ended 31 December 2014

	Notes	2014 £000	2013 £000
Turnover	2	46,684	61,787
Cost of sales		(31,199)	(46,841)
Gross Profit		15,485	14,946
Administrative expenses		(953)	(999)
Operating Profit	3	14,532	13,947
Interest receivable and similar income	6	23	46
Profit on ordinary activities before taxation		14,555	13,993
Tax	7	(2,653)	(3,392)
Profit for the financial year	12	11,902	10,601

All amounts relate to continuing activities.

UK Financial Statements

Usually, the notes to the accounts show the labour costs which further allows us to cross-reference labour / material split assumptions

5. Staff costs

	2014	*2013*
	£000	*£000*
Wages and salaries	4,632	5,886
Social security costs	475	551
Other pension costs	153	169
	5,260	6,606

The average monthly number of employees (excluding executive directors) for the year was:

	No.	*No.*
Programme	66	91
Administration	6	7
	72	98

Labour Rates Example

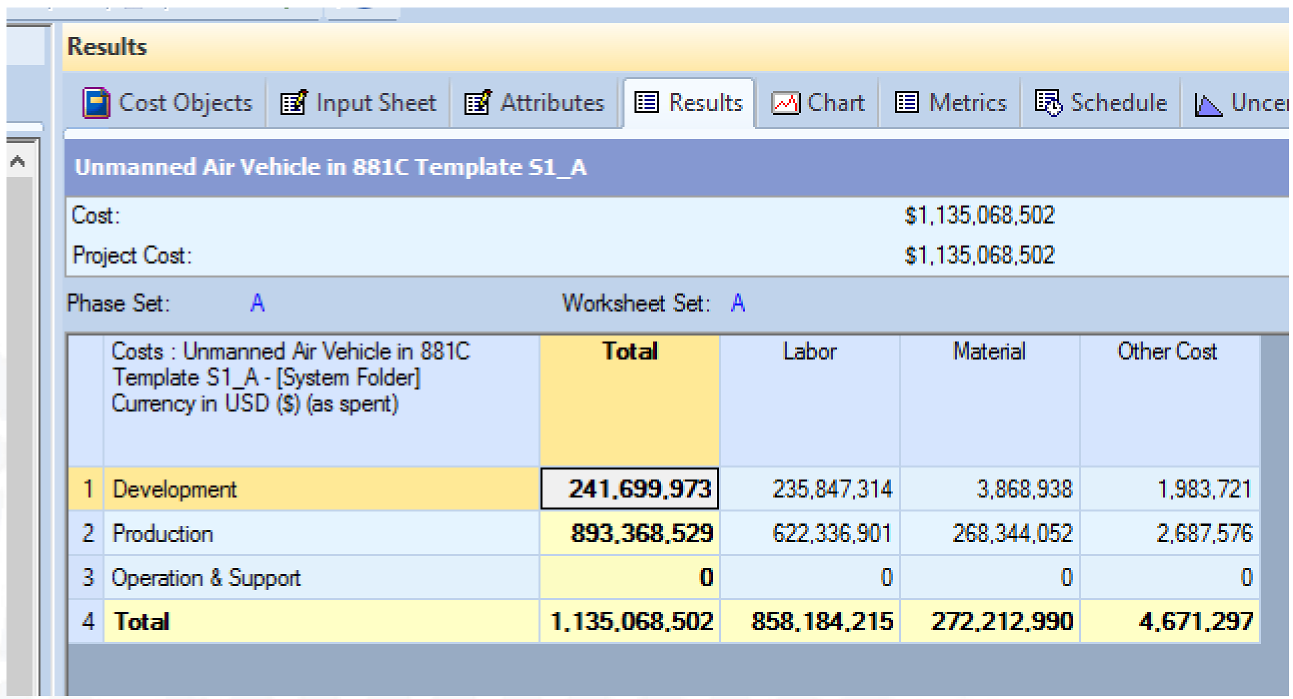
Financial Statements provide many of the inputs we need to calculate labour rates. However, we still need to make assumptions about other inputs. We can measure the standard deviation of the errors to show how good our assumptions really are.

Company A	CY 2012	CY 2013	CY 2014	Average
Implied Wages and Salaries	£173,628,832	£209,371,748	£202,970,111	£195,323,564
Delta from Actual	**-8.3%**	**0.0%**	**-5.4%**	**-4.6%**
Implied Wages of Contractors	£20,954,387	£25,287,496	£24,499,157	£23,580,346
Implied COGS	£919,966,380	£932,833,624	£1,077,984,767	£976,928,257
Delta from Actual	**1.3%**	**1.6%**	**2.0%**	**1.6%**
Implied Revenue	£980,928,169	£1,076,482,521	£1,057,488,135	£1,038,299,608
Actual Revenue	£994,917,691	£970,449,535	£1,164,712,010	£1,043,359,745
Delta from Actual	**-1.4%**	**10.9%**	**-9.2%**	**0.1%**
Variance of Effort				0.2%
Standard Deviation of the Error				**1.6%**

Labour / Material Split

amplio

Even if the top-level results of our data pivoting appear correct, we may dig into deeper analysis to cross-reference assumptions. The split between labour and materials is very useful and informs our '**value added ratio**'


Results

Cost Objects | Input Sheet | Attributes | Results | Chart | Metrics | Schedule | Uncer

Unmanned Air Vehicle in 881C Template S1_A

Cost: $1,135,068,502

Project Cost: $1,135,068,502

Phase Set: A   Worksheet Set: A

	Costs : Unmanned Air Vehicle in 881C Template S1_A - [System Folder] Currency in USD ($) (as spent)	Total	Labor	Material	Other Cost
1	Development	241,699,973	235,847,314	3,868,938	1,983,721
2	Production	893,368,529	622,336,901	268,344,052	2,687,576
3	Operation & Support	0	0	0	0
4	**Total**	**1,135,068,502**	**858,184,215**	**272,212,990**	**4,671,297**

info@amplioservices.com

Metrics

Using data in innovative ways

The Jobs-to-be-done Framework

This framework from Clayton Christensen provides a framework for the parameters we measure in order to quantify value

" Customers aren't interested in buying products or services per se. They have problems they want to solve and goals they want to achieve. These are jobs they want to do " [3]

What job is this customer trying to get done

What is the system the customer uses to get the job done

What is the weakest part of that System?

[3] The Innovator's Solution (2003). Clayton Christensen

The Jobs-to-be-done Framework

amplio

Transport for London (TfL) is investing in several major projects including HS2, Crossrail & Northern Line Extension

What job are they trying to get done?

Currently, the Docklands Light Railway service operates driverless trains but it:

'Improved' capacity

For the first time on deep-level sections of the Tube, the 250 trains, which are mechanised air-cooling system built in.

They will also have improved accessibility, with step-free access from the platf

London Underground said the trains would improve capacity by:

The Central line by 25% (the equivalent of up to 12,000 customers per hour)
The Bakerloo line by 25% (the equivalent of up to 8,000 customers per hour)
The Waterloo & City line by 50% (the equivalent of up to 9,000 customers per hour)
The Piccadilly line by 60% (the equivalent of up to 19,000 customers per hour)

It is hoped that the trains will remain in service for more than 40 years.

London Overground capacity

We are extending most London Overground trains from four to five cars. The five-car fleet will roll out from late 2014.

- Longer trains
- Gospel Oak to Barking
- Next steps
- Contact us
- Managing the impact of work

Since we took over the network in 2007, London Overground passenger numbers have quadrupled. To meet this increasing demand we plan to introduce longer trains, to provide an additional 25% capacity and reduce crowding.

Value Add Ratio

amplio

Added Value

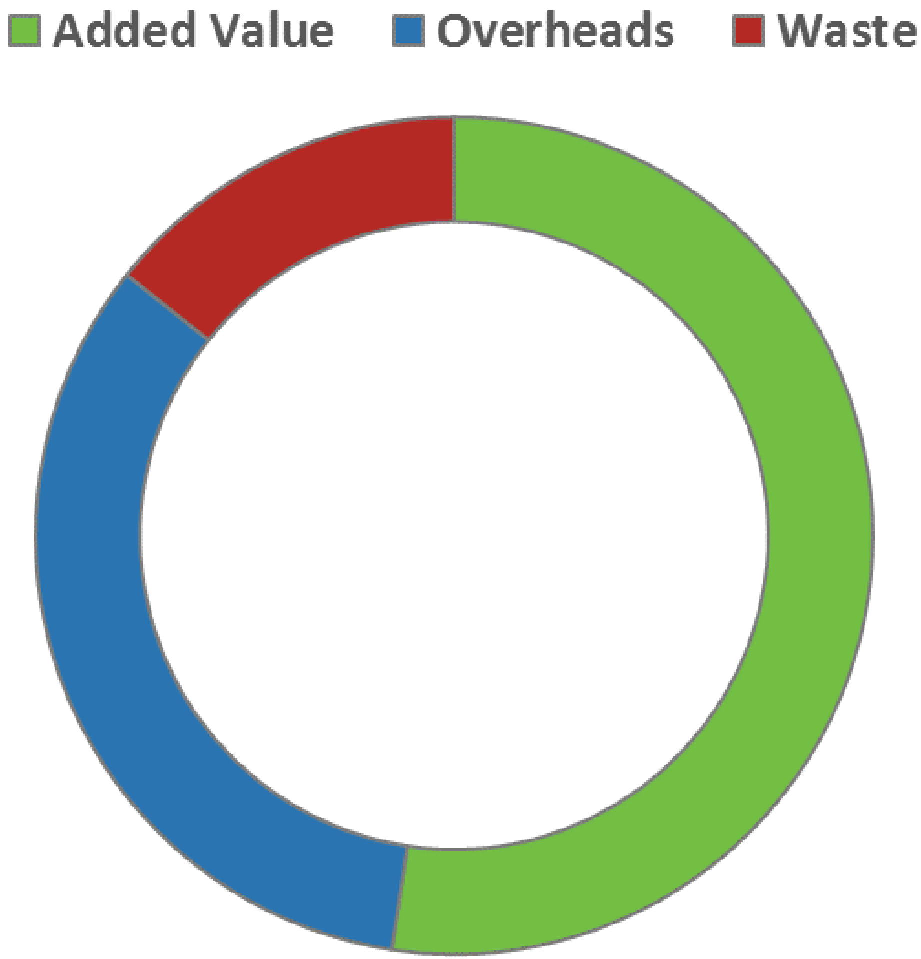
- Finished Goods
- Software as a Service
- Raw Materials
- Capability Enhancement

Overheads

- Project Management
- Training
- Certification
- Design Engineering

Waste

- Rework
- Inefficient Yields
- Emergency Maintaince
- Risk



info@amplioservices.com

COSYSMO

The Constructive Systems Engineering Cost Model (COSYSMO) is a method of quantifying system size and complexity, it was developed by Ricardo Valerdi

A	B	C	D	E	F	G	H	I	J
	Easy	Weight Easy	Nominal	Weight Nominal	Dificult	Weight Dificult	Weight of Category		Size for Category
Number of System Requirements	100	0.1	300	0.8	100	0.1	25.00%		65
Number of System Interfaces			30	1			25.00%		8
Number of System Specific Algorithms	100	0.4	280	0.5	200	0.1	25.00%		50
Number of Operational Scenarios			8	1			25.00%		2
Size of System:	125								
Calibration Constant A =	1.3								
Economy/Desiconomy of Scale E=	0.5								
Person Months:	12								

We have first determined the correlation between the average profit of the company for the last years and the win of a major contract, using Pearson product-moment correlation coefficient:

$$r = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2}}$$

Then we have made a prediction model to determine the probability for a company with certain profit margin to win the major contract using the least squares normal equation:

$$C = (X^T X)^{-1} X^T Y.$$

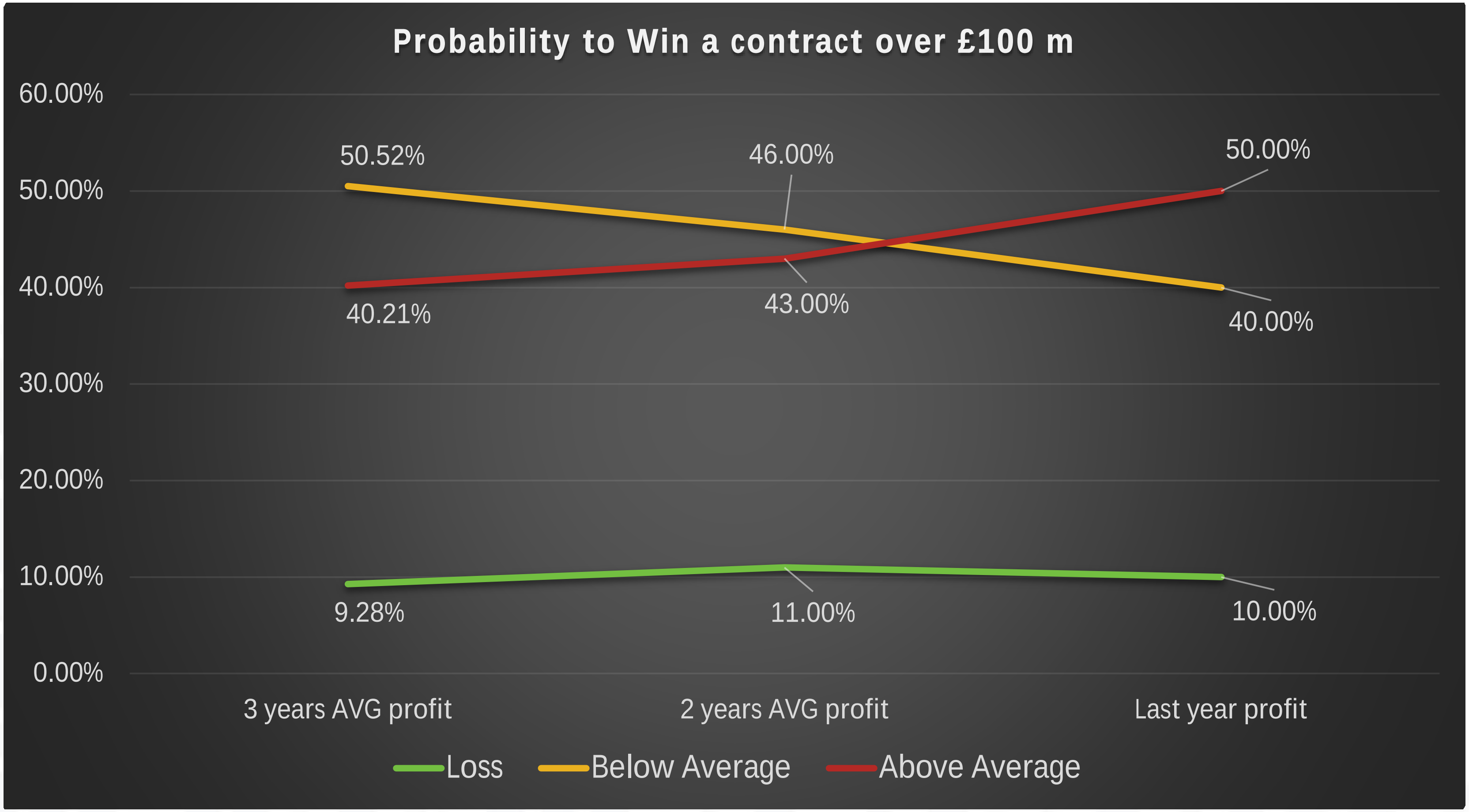
Where:

C is the model coefficients vector,

X is the data matrix and

Y is the result vector.

We have obtained the following probabilities:



Interpretation Issues

The challenge of how to read your data!

info@amplioservices.com

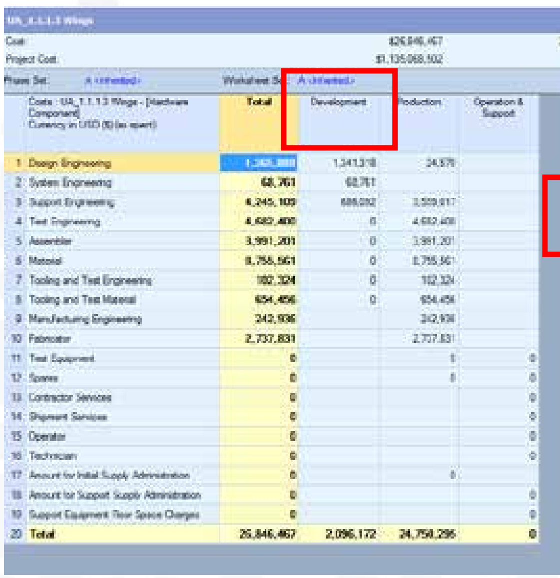
Aligning Terminology

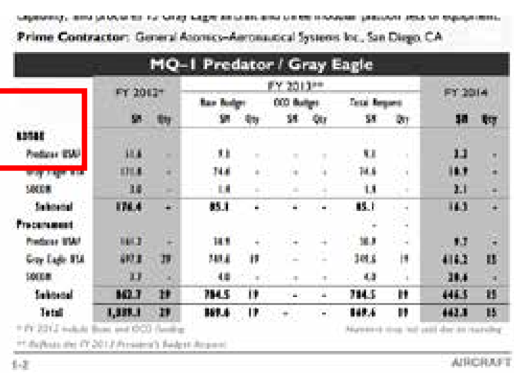
amplio

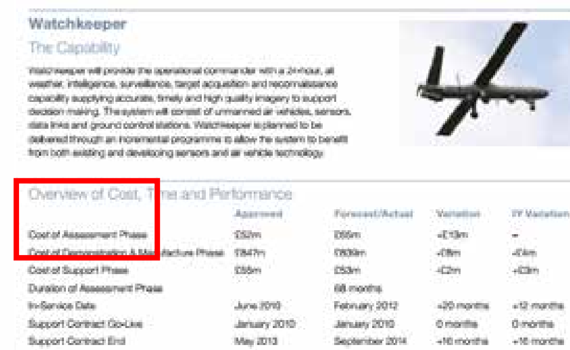
Parametric Estimate: ‘Development’

Department of Defense: “RDTE”

Ministry of Defence: “Concept, Assessment, Demonstration”





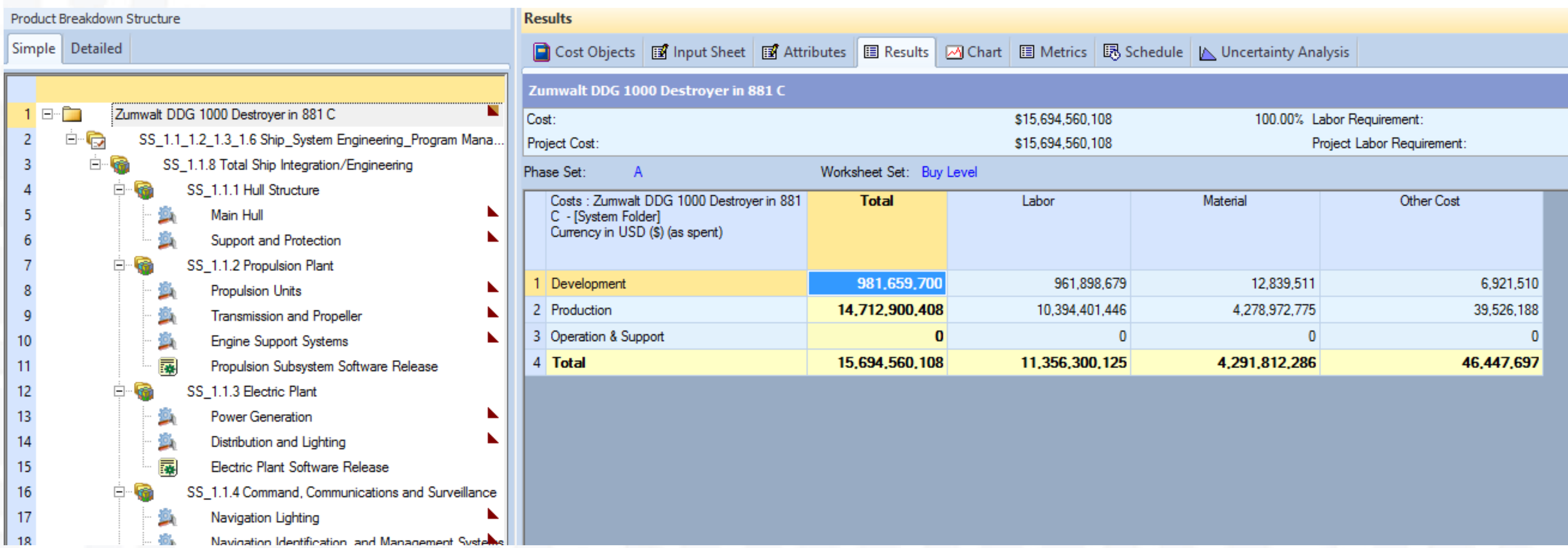


info@amplioservices.com

Interpreting Terminology

amplio

If parametric algorithms are based on a database of actual historical data. What should we assume about the results for a bid?



info@amplioservices.com

Gaps in the data

amplio

We use terms like ‘risk’ but ‘risk’ is not a category you will find on a company balance sheet.

Balance sheet
at 31 December 2014

	Note	2014 £000	2013 £000
Non-current assets			
Property, plant and equipment	8	90,835	93,266
Intangible assets	7	334,573	331,181
Investments in subsidiaries	9	200,459	189,968
Retirement benefit asset	18	208,675	95,506
Deferred tax asset	27	1,594	417
		836,136	710,338
Current assets			
Inventories and contracts in progress	10	101,733	136,248
Construction contracts	11	61,700	76,294
Trade and other receivables	12	190,330	160,676
Derivative financial instruments	20	4,307	8,623
Financial assets	13	186,917	112,941
Cash and cash equivalents		30,043	23,478
		575,030	518,260
Total assets		1,411,166	1,228,598
Current liabilities			
Trade and other payables	19	244,875	224,208
Advances from customers	11	164,050	156,027
Derivative financial instruments	20	9,747	2,167
Financial liabilities	14	177,142	190,710
Current tax liabilities	15	7,554	7,787
Provisions for liabilities and charges	17	24,852	23,319
		628,220	604,218
Non-current liabilities			
Non-current financial debts	14	27	124
Deferred tax liabilities	27	55,195	33,788
		55,222	33,912
Total liabilities		(683,442)	(638,130)
Net assets		727,724	590,468
Equity			
Ordinary shares	16	270,000	270,000
Retained earnings		462,431	316,467
Cash flow hedge reserve	16	(4,707)	4,001
Total equity		727,724	590,468

Explaining the Results

Most people in our organisations do not think probabilistically! Some people even object to it!

At best, most people have three understandings of probabilistic results: “it will happen, it won’t happen, it might”

Summary

Using the lessons in this presentation, you should

- Be able to identify sources public data to assist validation and improve your estimates
- Use measures such as data pivoting, the jobs-to-be-done framework, COSYSMO and 'Value Added Ratio' to convert data to actionable intelligence
- Anticipate some of the interpretation challenges you will encounter when you practice these skills