Value of Information
in
Spreadsheet Monte Carlo Simulation Models

INFORMS 2010 Austin

Michael R. Middleton, Ph.D.
Decision Toolworks
Mike@DecisionToolworks.com
415.310.7190

Background

- Spreadsheet models are often used for what-if sensitivity analysis
- Probability distributions may be assigned to input assumptions, and Monte Carlo simulation can develop the corresponding distribution for the output performance measure
- **Problem**: How can we compute Value of Information using the results of Monte Carlo simulation in a spreadsheet model?

Value of Information

- Expected Value of Perfect Information, EVPI
- Value of Clairvoyance, Value of Information, VOI

- Important for
  - determining an upper bound on the value of actual information-gathering activities
  - comparing the value of information for multiple uncertainties

- Requires a model of a decision problem with alternatives, probabilities for uncertainties, outcome values, and willingness to determine certain equivalents

- Monte Carlo simulation may be used to describe uncertainty associated with a single alternative or strategy

Why is Value of Information important? Answer #1

- Felli and Hazen, 1999
- "In realistically sized problems, simple one- and two-way SAs may not fully capture parameter interactions, raising the disturbing possibility that many published decision analyses might be overconfident in their policy recommendations."
- "… we re-examined 25 decision analyses …"
- "While we expected EVPI values to indicate greater problem sensitivity than conventional SA due to revealed parameter interaction, we in fact found the opposite: compared to EVPI, the one- and two-parameter SAs accompanying these problems dramatically *overestimated* problem sensitivity to input parameters."

Why is Value of Information important? Answer #2

- Hubbard (2009) computed EVPI “on more than 60 large decision models with a total of over 4,000 uncertain input variables”
- “calculations of information values justify some empirical measurement of uncertain values most of the time.”
- “a VBA macro in Excel was used to slice the continuous distribution into thousands of discrete units, computing EVPI for each unit, and totaling the unit EVPIs.”
- Here I show how to compute EVPI for the results of Monte Carlo simulation in Excel without using a VBA macro.

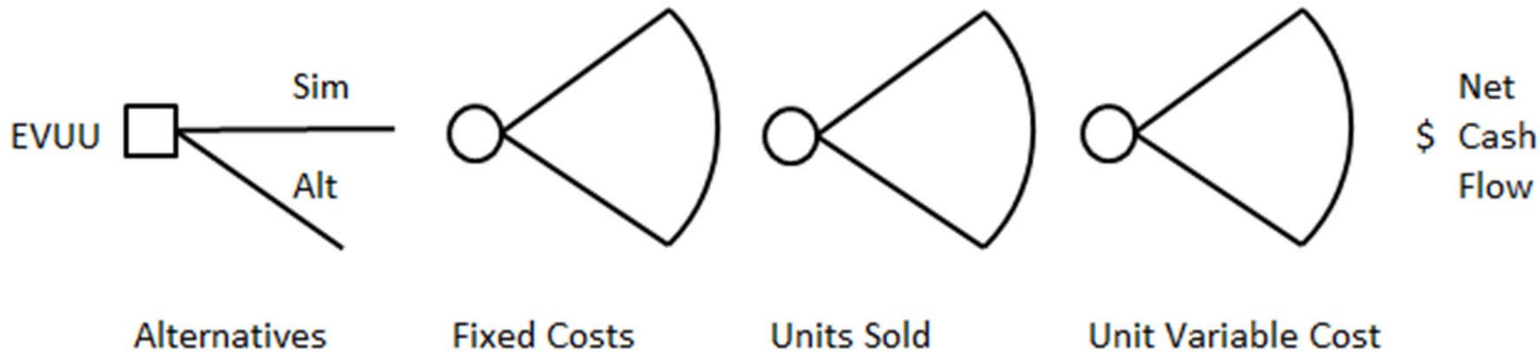
Terminology

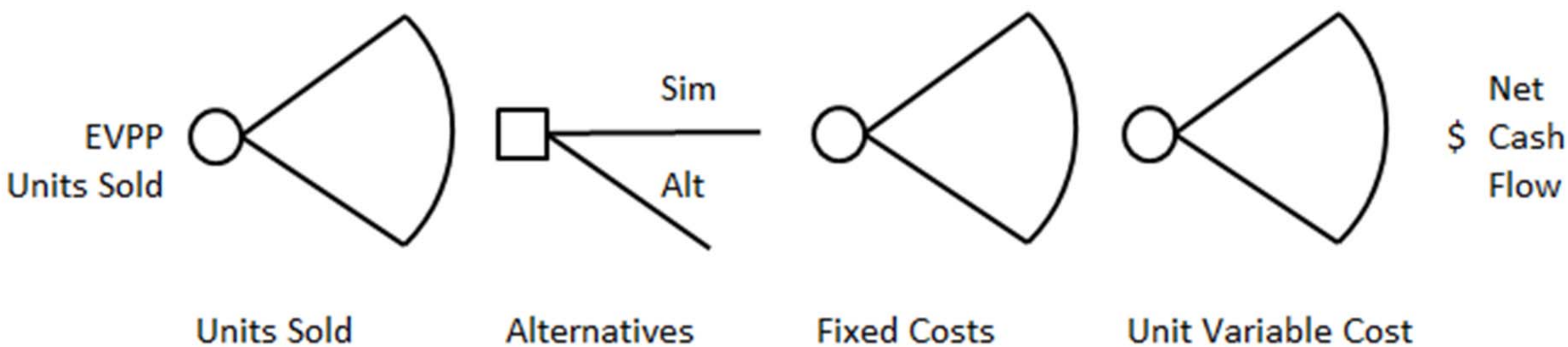
- Risk-neutral certain equivalent
- Probability-weighted average, Expected Value, EV

- EVUU, Expected Value Under Uncertainty
- EVPP, Expected Value with Perfect Prediction
- EVPI, Expected Value of Perfect Information

- EVPI = EVPP - EVUU

Computing EVPI by rearranging a decision tree





EVPI(Units Sold) = EVPP(Units Sold) - EVUU

Simple spreadsheet model with base case and extreme inputs

Software project
Predetermined Unit Price, constant for these analyses
Judgments about Units Sold, Unit Variable Cost, and Fixed Costs
Other alternative (not shown) yields $80,000

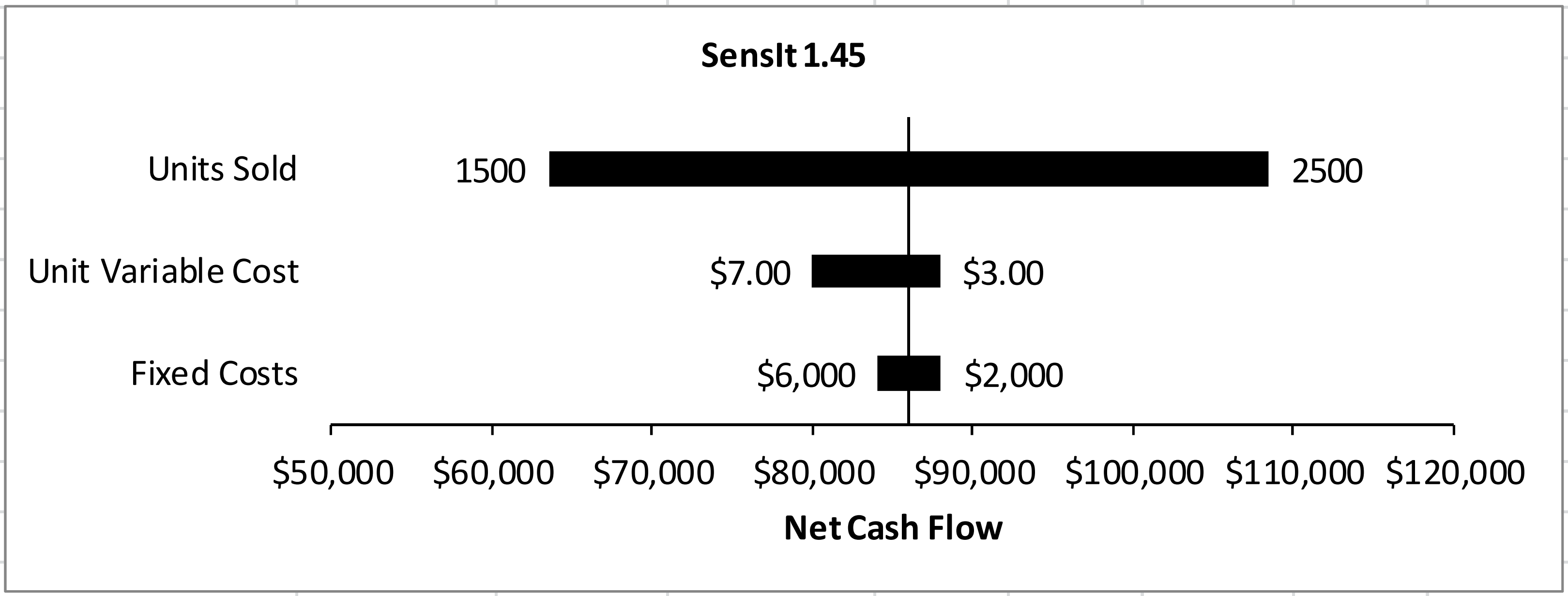
	A	B	C	D	E	F	G	H	I
2		Input Cells		Input Assumptions				Scenarios	
3	Unit Price	$49		Minimum	Base Case	Maximum		Worst	Best
4	Units Sold	2000		1500	2000	2500		1500	2500
5	Unit Variable Cost	$4.00		$3.00	$4.00	$7.00		$7.00	$3.00
6	Fixed Costs	$4,000		$2,000	$4,000	$6,000		$6,000	$2,000
7									
8	Net Cash Flow	$86,000						$57,000	$113,000

	A	B	C	D	E	F	G	H	I
2		Input Cells						Scenarios	
3	Unit Price	49		Minimum	Base Case	Maximum		Worst	Best
4	Units Sold	2000		1500	2000	2500		1500	2500
5	Unit Variable Cost	4		3	4	7		7	3
6	Fixed Costs	4000		2000	4000	6000		6000	2000
7									
8	Net Cash Flow	=B4*(B3-B5)-B6						=H4*(B3-H5)-H6	=I4*(B3-I5)-I6

Tornado chart for simple spreadsheet model

Single-factor sensitivity analysis

	Corresponding Input Value			Net Cash Flow Output Value				Percent
Input Variable	Low Output	Base Case	High Output	Low	Base	High	Swing	Swing^2
Units Sold	1500	2000	2500	$63,500	$86,000	$108,500	$45,000	96.2%
Unit Variable Cost	$7.00	$4.00	$3.00	$80,000	$86,000	$88,000	$8,000	3.0%
Fixed Costs	$6,000	$4,000	$2,000	$84,000	$86,000	$88,000	$4,000	0.8%



Simple spreadsheet model with random inputs

	A	B	C	D	E	F	G	H	I
3	Unit Price	$49		RAND()		Type			
4	Units Sold	1798		0.156835		Normal	Mean=2,000	StDev=200	
5	Unit Variable Cost	$6.06		0.926868		Triangular	Min=$3	Mode=$4	Max=$7
6	Fixed Costs	$3,000		0.190587		Discrete	Lookup	Value	Probability
7							0.0	$2,000	0.1
8	Net Cash Flow	$74,206					0.1	$3,000	0.2
9							0.3	$4,000	0.4
10							0.7	$5,000	0.2
11							0.9	$6,000	0.1

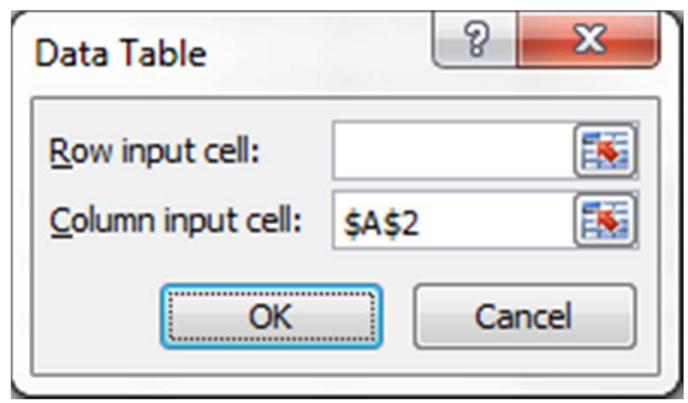
	A	B	C	D
3	Unit Price	49		RAND()
4	Units Sold	=ROUND(NORMINV(D4,2000,200),0)		=RAND()
5	Unit Variable Cost	=ROUND(IF(D5<(4-3)/(7-3),3+(D5*(7-3)*(4-3))^0.5,7-((1-D5)*(7-3)*(7-4))^0.5),2)		=RAND()
6	Fixed Costs	=VLOOKUP(D6,G7:H11,2,TRUE)		=RAND()
7				
8	Net Cash Flow	=ROUND(B4*(B3-B5)-B6,0)		

Setting up an Excel data table for Monte Carlo simulation

	A	B	C	D	E	F
1	Trial	Unit Price	Units Sold	Unit Variable Cost	Fixed Costs	Net Cash Flow
2		='RAND Model'!B3	='RAND Model'!B4	='RAND Model'!B5	='RAND Model'!B6	='RAND Model'!B8
3	1					
4	2					
5	3					

	A	B	C	D	E	F
1	Trial	Unit Price	Units Sold	Unit Variable Cost	Fixed Costs	Net Cash Flow
2		$49	1985	$5.75	$4,000	$81,851
3	1					
4	2					
5	3					

For 10,000 trials,
select A2:F10002.
In Excel 2007 & 2010
choose Data >
What-If Analysis >
Data Table



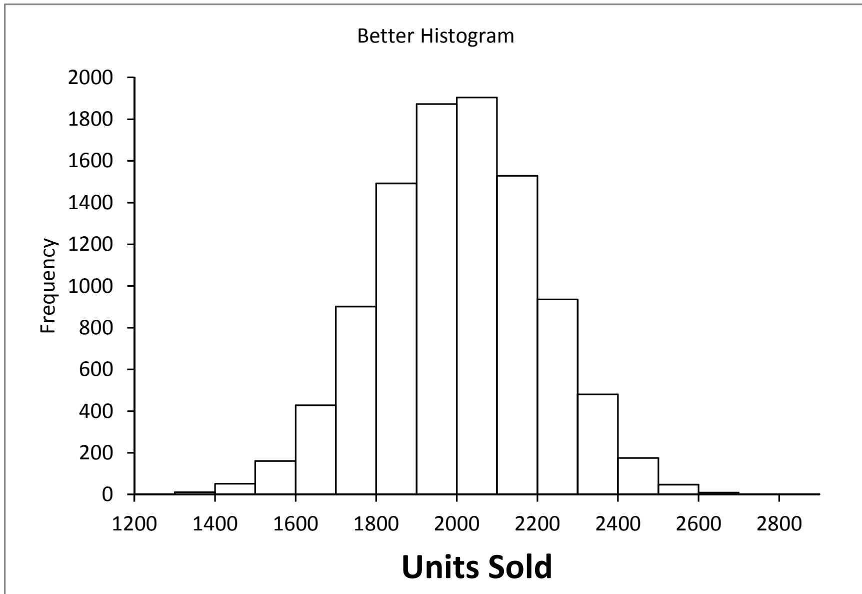
Results of an Excel data table for Monte Carlo simulation

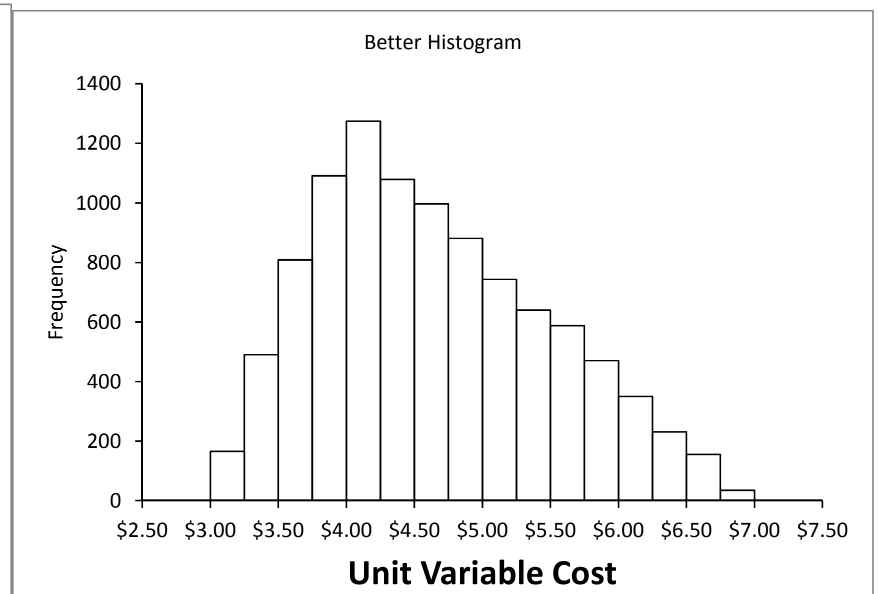
	A	B	C	D	E	F
1	Trial	Unit Price	Units Sold	Unit Variable Cost	Fixed Costs	Net Cash Flow
2		='RAND Model'!B3	='RAND Model'!B4	='RAND Model'!B5	='RAND Model'!B6	='RAND Model'!B8
3	1	=TABLE(,A2)	=TABLE(,A2)	=TABLE(,A2)	=TABLE(,A2)	=TABLE(,A2)
4	2	=TABLE(,A2)	=TABLE(,A2)	=TABLE(,A2)	=TABLE(,A2)	=TABLE(,A2)
5	3	=TABLE(,A2)	=TABLE(,A2)	=TABLE(,A2)	=TABLE(,A2)	=TABLE(,A2)

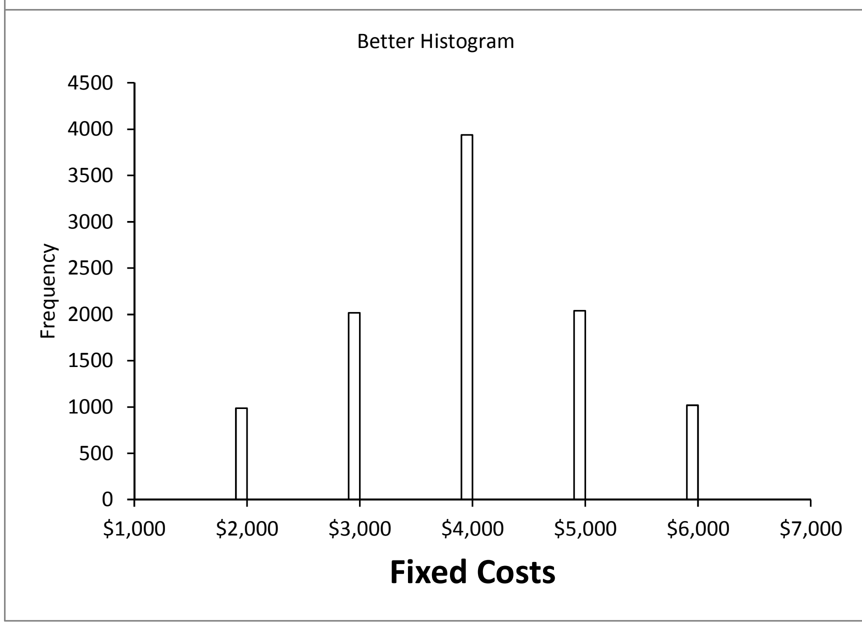
	A	B	C	D	E	F
1	Trial	Unit Price	Units Sold	Unit Variable Cost	Fixed Costs	Net Cash Flow
2		$49	2096	$5.49	$2,000	$89,197
3	1	$49	2152	$4.13	$3,000	$93,560
4	2	$49	2039	$3.91	$3,000	$88,939
5	3	$49	2077	$5.12	$6,000	$85,139

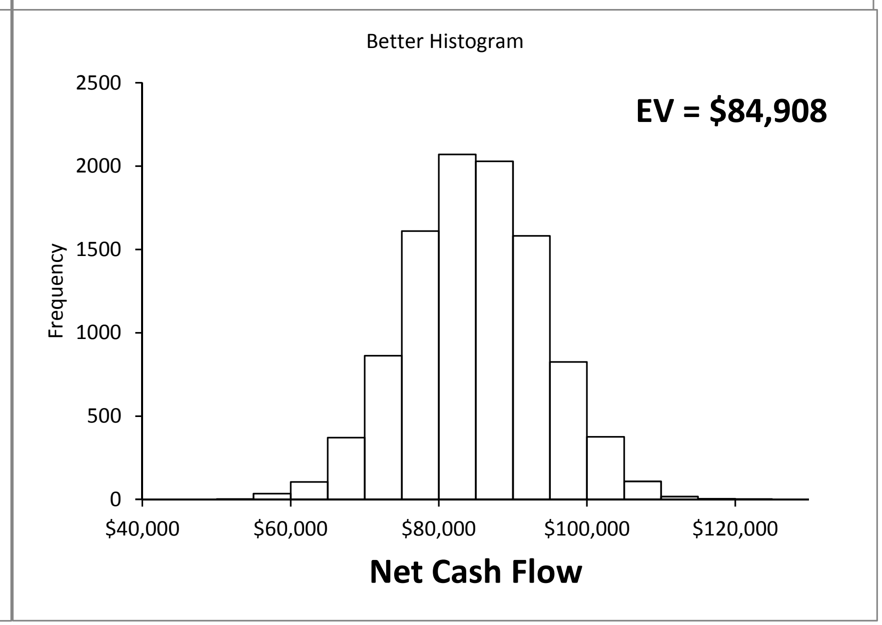
Data table is volatile.

For subsequent analysis, copy and paste special values.

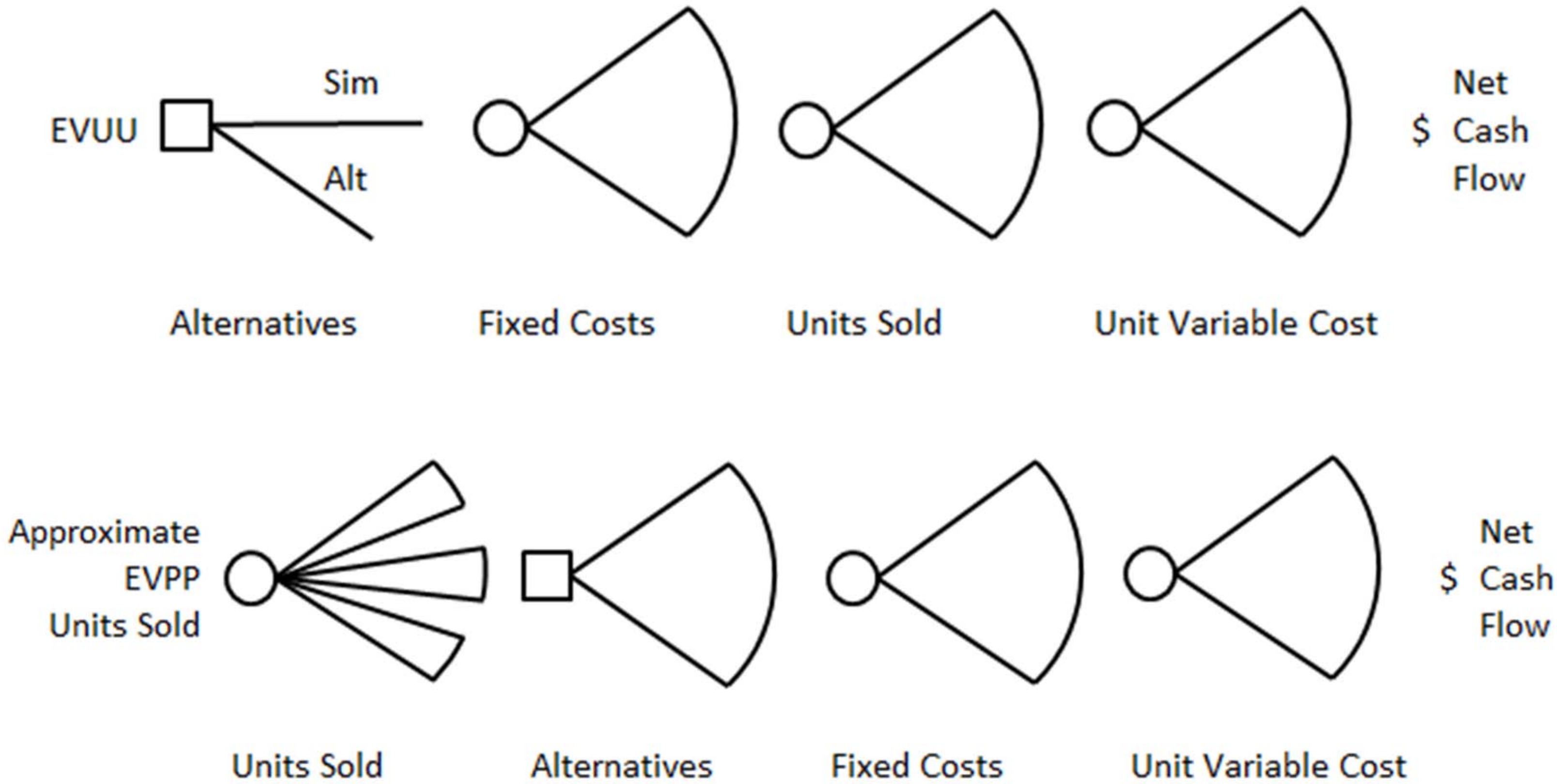

Better Histogram
Frequency
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1200
1400
1600
1800
2000
2200
2400
2600
2800
Units Sold


Better Histogram
Frequency
1400
1200
1000
800
600
400
200
0
$2.50
$3.00
$3.50
$4.00
$4.50
$5.00
$5.50
$6.00
$6.50
$7.00
$7.50
Unit Variable Cost


Better Histogram
Frequency
4500
4000
3500
3000
2500
2000
1500
1000
500
0
$1,000
$2,000
$3,000
$4,000
$5,000
$6,000
$7,000
Fixed Costs


Better Histogram
EV = $84,908
Frequency
2500
2000
1500
1000
500
0
$40,000
$60,000
$80,000
$100,000
$120,000
Net Cash Flow

Decision tree analogy for Monte Carlo simulation EVPI



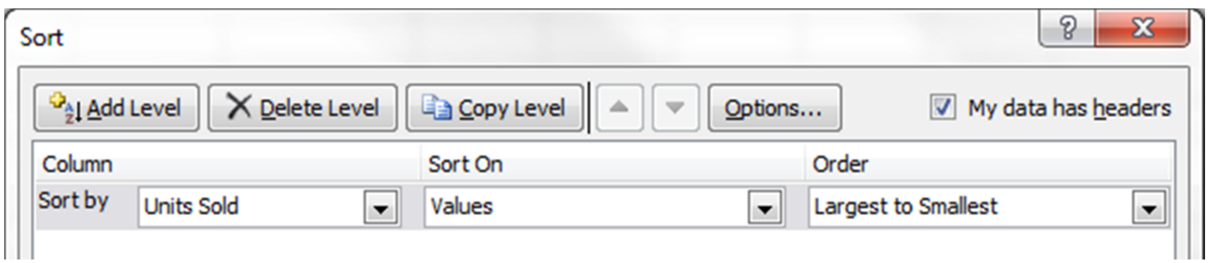
Approx. EVPI(Units Sold) = Approx. EVPP(Units Sold) - EVUU

To determine EVPP(Units Sold), first …

Copy the Units Sold and Net Cash Flow data to another sheet

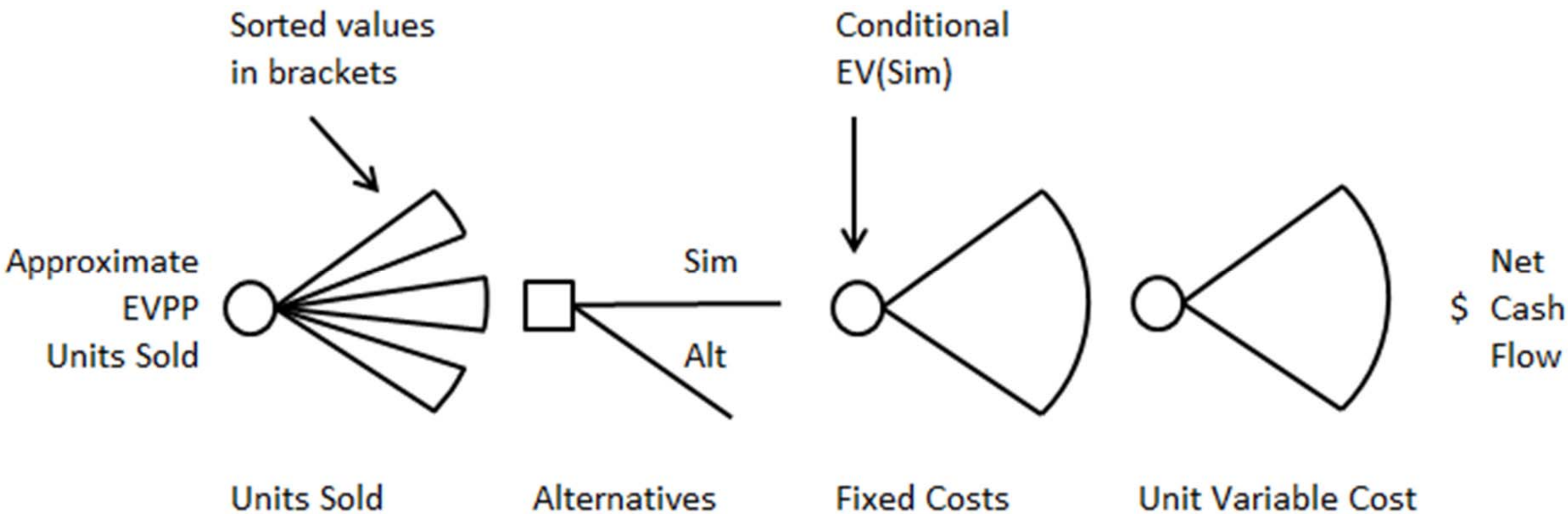
	A	B	C
1	Trial	Units Sold	Net Cash Flow
2	1	2152	$93,560
3	2	2039	$88,939
4	3	2077	$85,139

Sort the simulation results with Units Sold as key



	A	B	C
1	Trial	Units Sold	Net Cash Flow
2	3726	2810	$122,539
3	3910	2718	$120,071
4	528	2698	$118,381

Compute Conditional EV(Sim) for each bracket

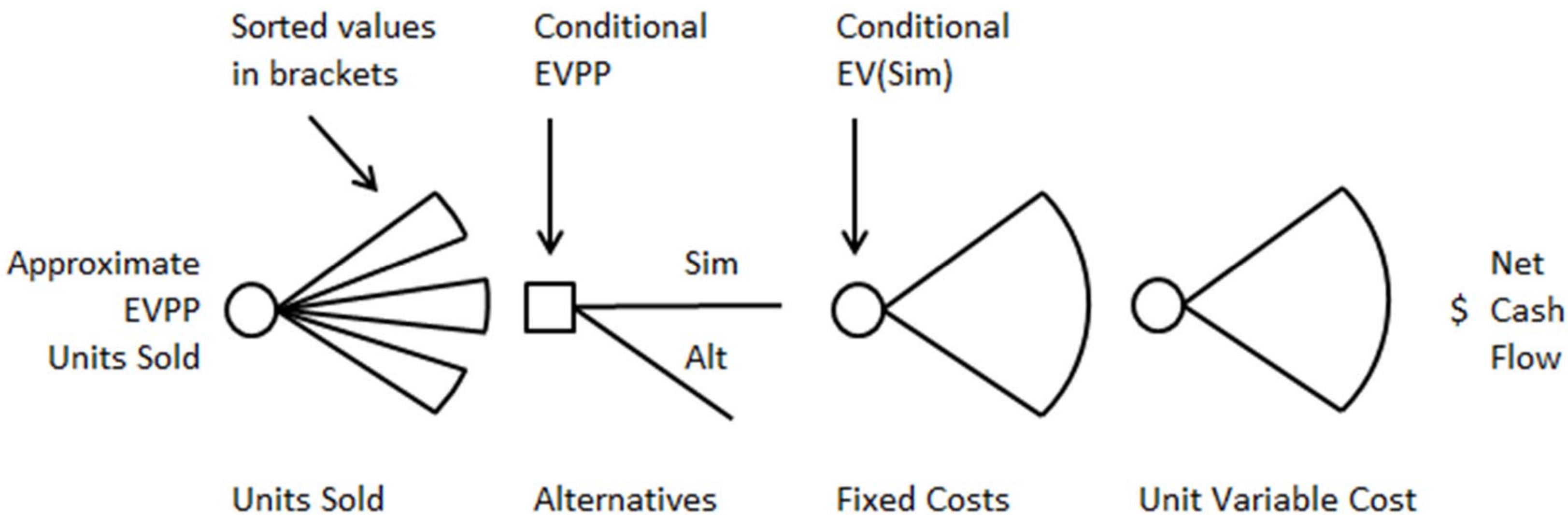


	A	B	C	D	E	F
1	Trial	Units Sold	Net Cash Flow		Bracket	CondEV(Sim)
2	3726	2810	$122,539		1	=AVERAGE(OFFSET(C2,(E2-1)*100,0,100))
3	3910	2718	$120,071		2	$103,815
4	528	2698	$118,381		3	$102,241

Syntax: OFFSET(reference, rows, cols, [height], [width])

100 brackets * 100 trials each = 10,000 trials total

Compute Conditional EVPP for each bracket



	A	B	C	D	E	F	G	H	I	J
1	Trial	Units Sold	Net Cash Flow		Bracket	CondEV(Sim)	CondEVPP		EV(Sim)	$84,908
2	3726	2810	$122,539		1	$108,259	=MAX(F2,J3)			
3	3910	2718	$120,071		2	$103,815	$103,815		EV(Alt)	$80,000
4	528	2698	$118,381		3	$102,241	$102,241			

Compute EVPP and EVPI

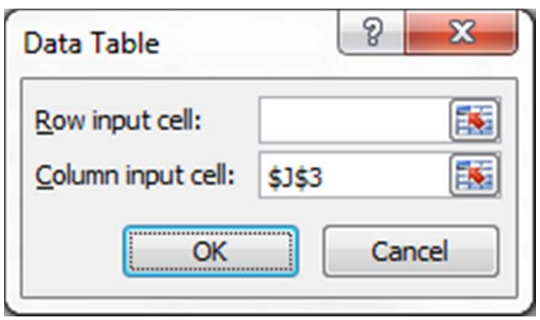
	A	B	C	D	E	F	G	H	I	J
1	Trial	Units Sold	Net Cash Flow		Bracket	CondEV(Sim)	CondEVPP		EV(Sim)	$84,908
2	3726	2810	$122,539		1	$108,259	$108,259			
3	3910	2718	$120,071		2	$103,815	$103,815		EV(Alt)	$80,000
4	528	2698	$118,381		3	$102,241	$102,241			
5	873	2685	$111,938		4	$101,098	$101,098		EVUU	$84,908
6	7438	2683	$112,040		5	$100,219	$100,219			
7	5970	2654	$116,385		6	$99,271	$99,271		EVPP	$86,538
8	8095	2649	$114,258		7	$98,372	$98,372			
9	3665	2632	$113,519		8	$97,846	$97,846		EVPI	$1,630
10	8989	2625	$107,565		9	$97,243	$97,243			

	A	B	C	D	E	F	G	H	I	J
1	Trial	Units Sold	Net Cash Flow		Bracket	CondEV(Sim)	CondEVPP		EV(Sim)	=AVERAGE(C2:C10001)
2	3726	2810	$122,539		1	$108,259	$108,259			
3	3910	2718	$120,071		2	$103,815	$103,815		EV(Alt)	$80,000
4	528	2698	$118,381		3	$102,241	$102,241			
5	873	2685	$111,938		4	$101,098	$101,098		EVUU	=MAX(J1,J3)
6	7438	2683	$112,040		5	$100,219	$100,219			
7	5970	2654	$116,385		6	$99,271	$99,271		EVPP	=AVERAGE(G2:G101)
8	8095	2649	$114,258		7	$98,372	$98,372			
9	3665	2632	$113,519		8	$97,846	$97,846		EVPI	=J7-J5
10	8989	2625	$107,565		9	$97,243	$97,243			

EVPI(Units Sold) for values of the other alternative

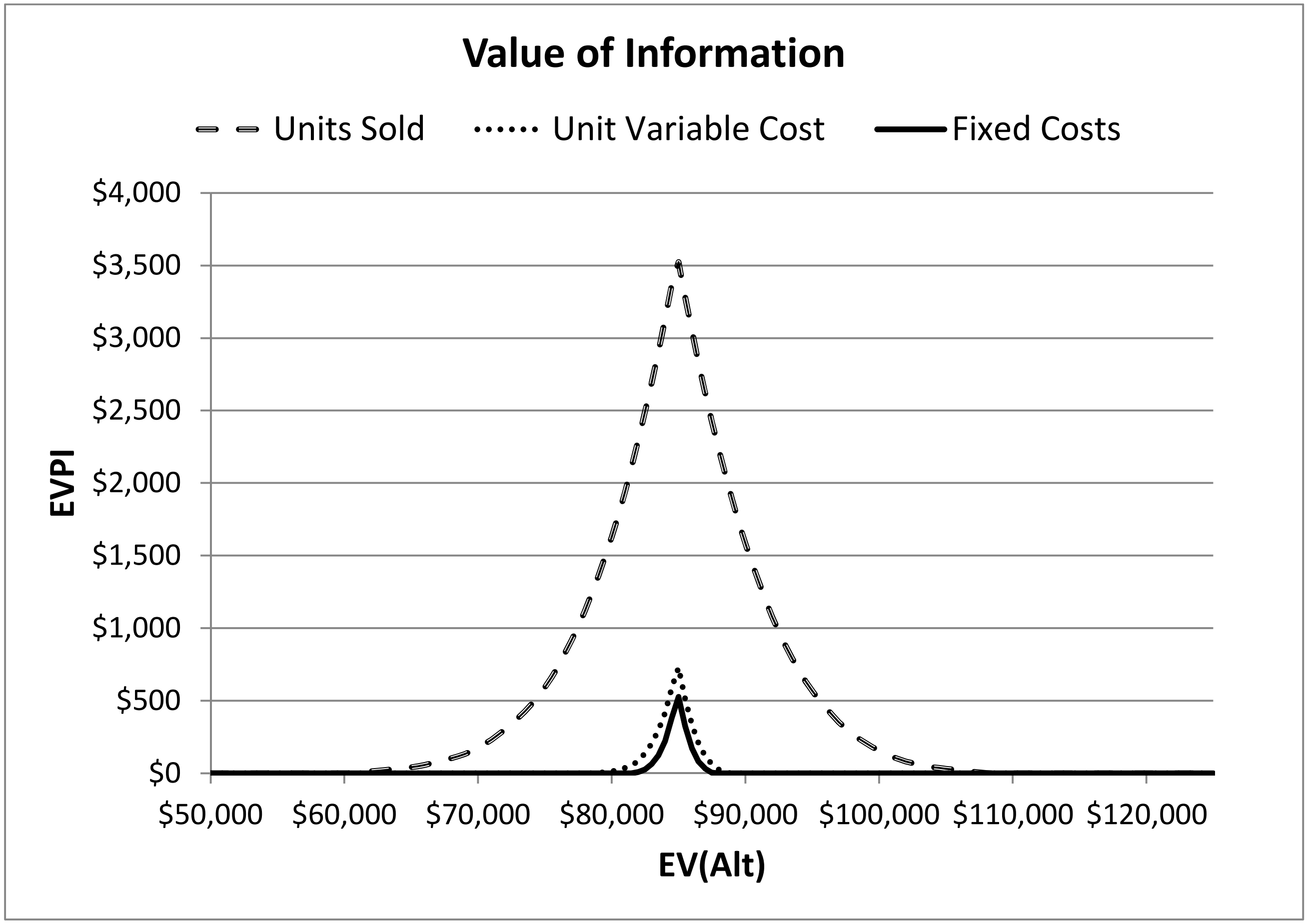
Set up an Excel Data Table
$60K to $110K, step $5K
Formula for output: =J9
Select L2:M13
Data > What-If Analysis
  > Data Table

	I	J	K	L	M
1	EV(Sim)	$84,908		EV(Alt)	EVPI
2					=J9
3	EV(Alt)	$80,000		$60,000	
4				$65,000	
5	EVUU	$84,908		$70,000	
6				$75,000	
7	EVPP	$86,538		$80,000	
8				$85,000	
9	EVPI	$1,630		$90,000	
10				$95,000	
11				$100,000	
12				$105,000	
13				$110,000	



	I	J	K	L	M
1	EV(Sim)	$84,908		EV(Alt)	EVPI
2					$1,630
3	EV(Alt)	$80,000		$60,000	$0
4				$65,000	$43
5	EVUU	$84,908		$70,000	$178
6				$75,000	$594
7	EVPP	$86,538		$80,000	$1,630
8				$85,000	$3,524
9	EVPI	$1,630		$90,000	$1,584
10				$95,000	$569
11				$100,000	$156
12				$105,000	$33
13				$110,000	$0

Three EVPIs for values of the other alternative



Excel features used in this project

- Worksheet functions in what-if model
  - RAND, NORMINV, IF, VLOOKUP, ROUND
- Monte Carlo simulation
  - Data Table
- EVPI calculations
  - Data Sort, AVERAGE, OFFSET, MAX
- Diagrams and charts
  - Drawing tools, XY charts
  - Better Histogram add-in
    http://www.treeplan.com/better.htm

References

- Felli, J.C., Hazen, G.B. 1999. Do sensitivity analyses really capture problem sensitivity? An empirical analysis based on information value. Risk, Decision and Policy 4(2) 79-98. Accessed June 28, 2010, http://users.iems.northwestern.edu/~hazen/Do_SAs_Capture_Problem_Sensitivity.pdf

- Hubbard, D., Samuelson, D.A. 2009. Modeling Without Measurements: How the decision analysis culture's lack of empiricism reduces its effectiveness. OR/MS Today 36(5) 26-31. Accessed June 28, 2010, http://www.lionhrtpub.com/orms/orms-10-09/risk.html

- Middleton, M.R. 1982. Risk Analysis and the Expected Value of Perfect Information. Proceedings, American Institute for Decision Sciences, Western Regional Conference. Accessed June 28, 2010, http://www.mikemiddleton.com/RiskAnalysisEVPI_1982.pdf

Value of Information
in
Spreadsheet Monte Carlo Simulation Models

INFORMS 2010 Austin

Michael R. Middleton, Ph.D.
Decision Toolworks
Mike@DecisionToolworks.com
415.310.7190

PowerPoint Slides, Slides PDF File, and Excel Workbook

http://www.DecisionToolworks.com/SimulationEVPI2010.pptx
http://www.DecisionToolworks.com/SimulationEVPI2010.pdf
http://www.DecisionToolworks.com/SimulationEVPI2010.xlsx