Harmonizing, aggregating, and comparing results from expert elicitations

Max Henrion, Evan Sherwin
Lumina Decision Systems,
Los Gatos, California

Erin Baker
University of Massachusetts
at Amherst

Energy Modeling Forum,
Snowmass, July 2012

Lumina DECISION SYSTEMS

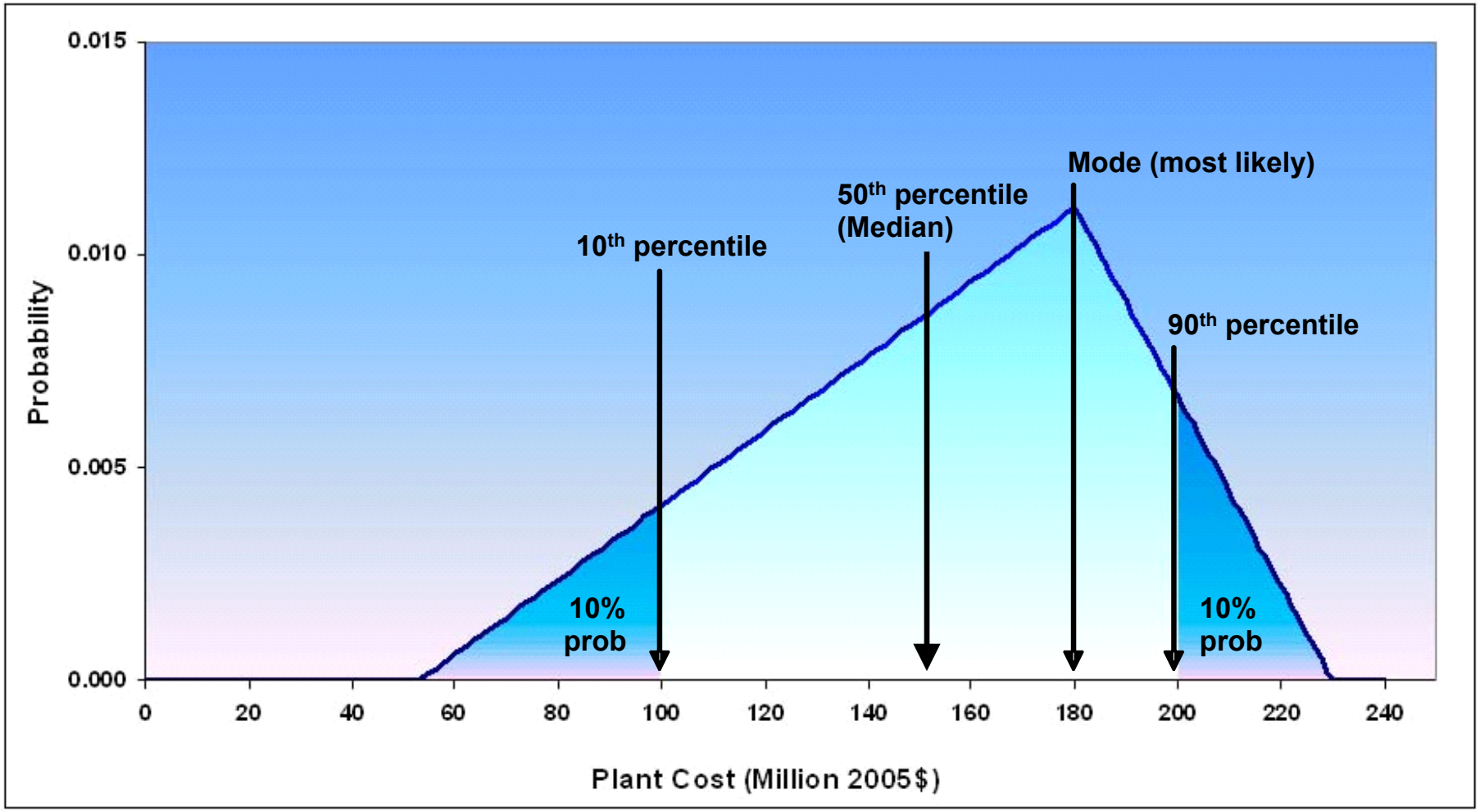
TEaM

Introduction

- TEAM goal is to harmonize, aggregate, and compare expert elicitation results from Harvard, FEEM and U Mass.
- Outline
  - Methods to fit distributions.
  - Methods to aggregating distributions over experts and over TEAM contributors
  - Comparing results: Uncertainty and learning
- *These are early draft results for review and discussion.*

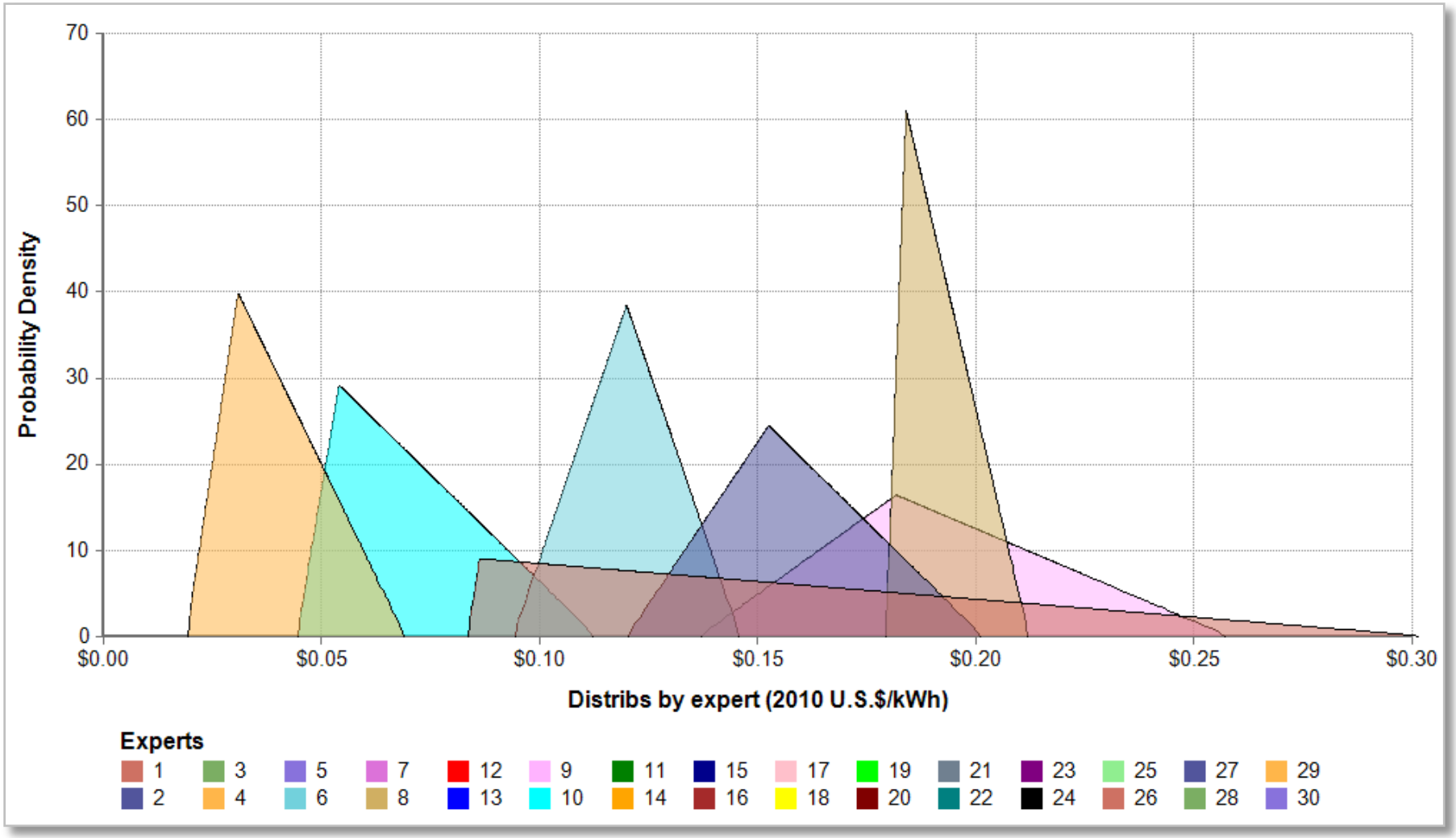
Lumina DECISION SYSTEMS

Fitting a Triangular Distribution to 10th, 50th, 90th percentiles



Lumina DECISION SYSTEMS

Sample results by expert:
Harvard Solar Levelized Cost 2030, High R&D



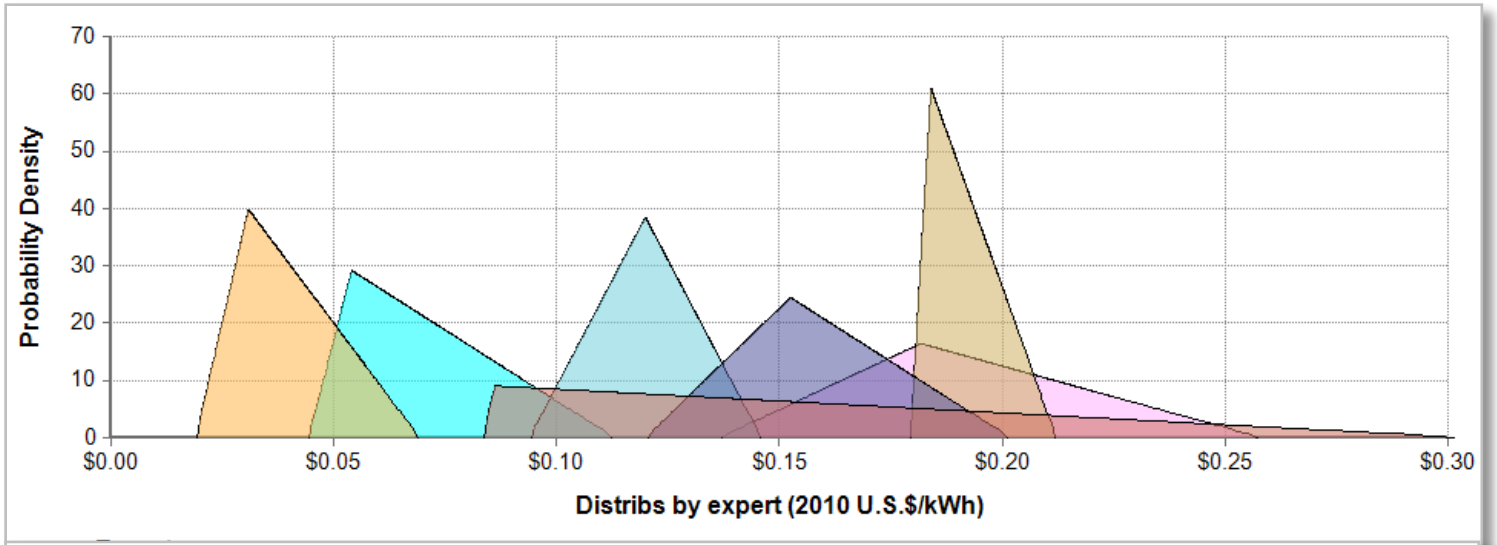
Lumina DECISION SYSTEMS

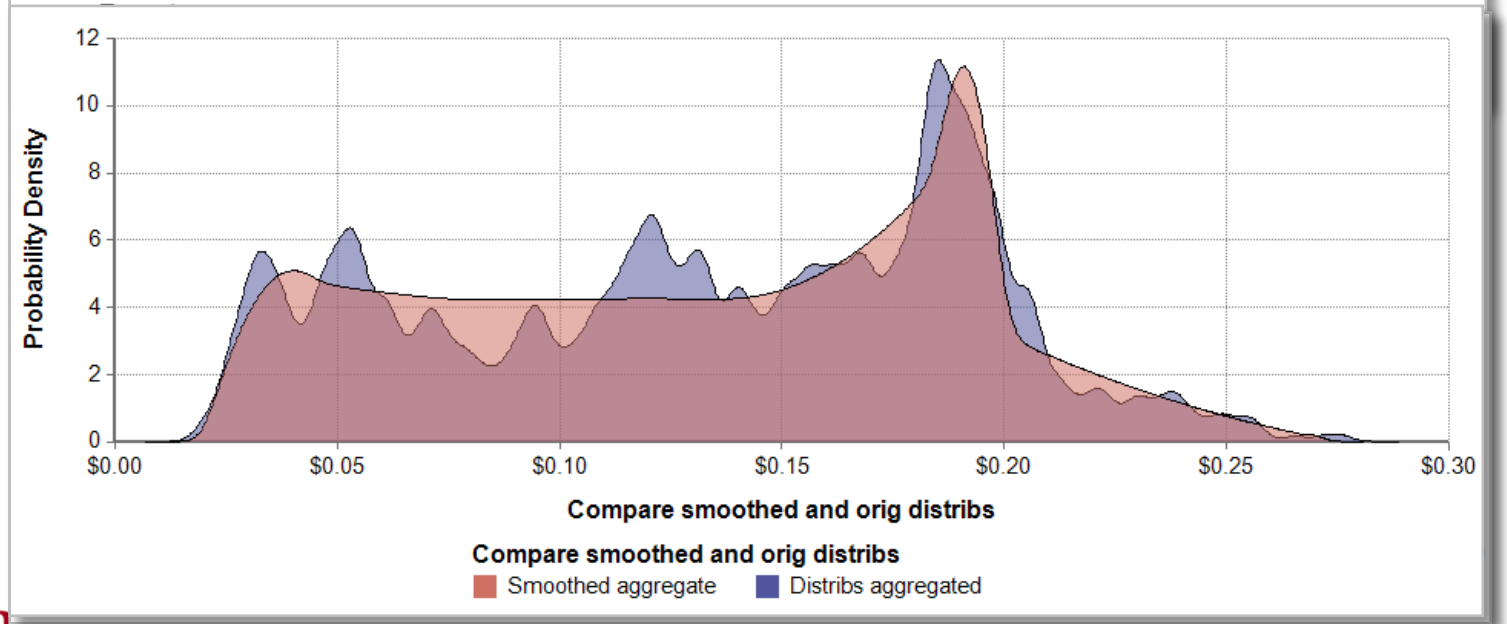
Aggregating distributions over experts

- Most common and easiest method is *Laplacean mixing*: Select from each distribution at random
- For now, we use equal weighting
  - Could be weighted by expertise rated by self or others.
  - Is width of distribution an indicator of expertise?

Pierre-Simon Laplace
(1749-1827)
Théorie analytique des probabilités, 1812

Aggregating: Laplacean mixing with smoothing:
Harvard Solar Levelized Cost 2030, High R&D





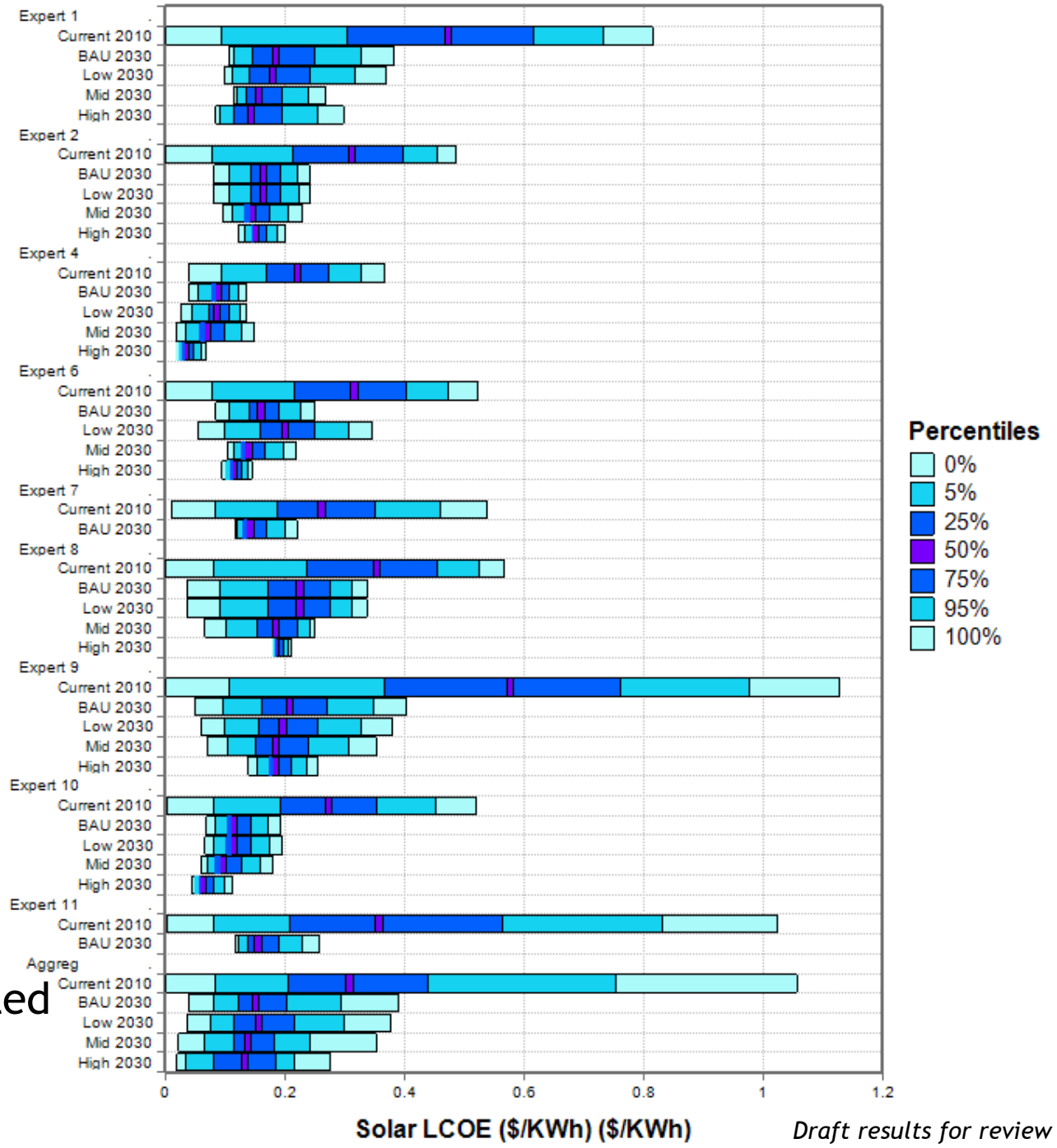
Lumina DECISION SYSTEMS

College scarf probability box plot:

Harvard Solar Levelized Cost by year & R&D level

Experts 1 to 11

Aggregated



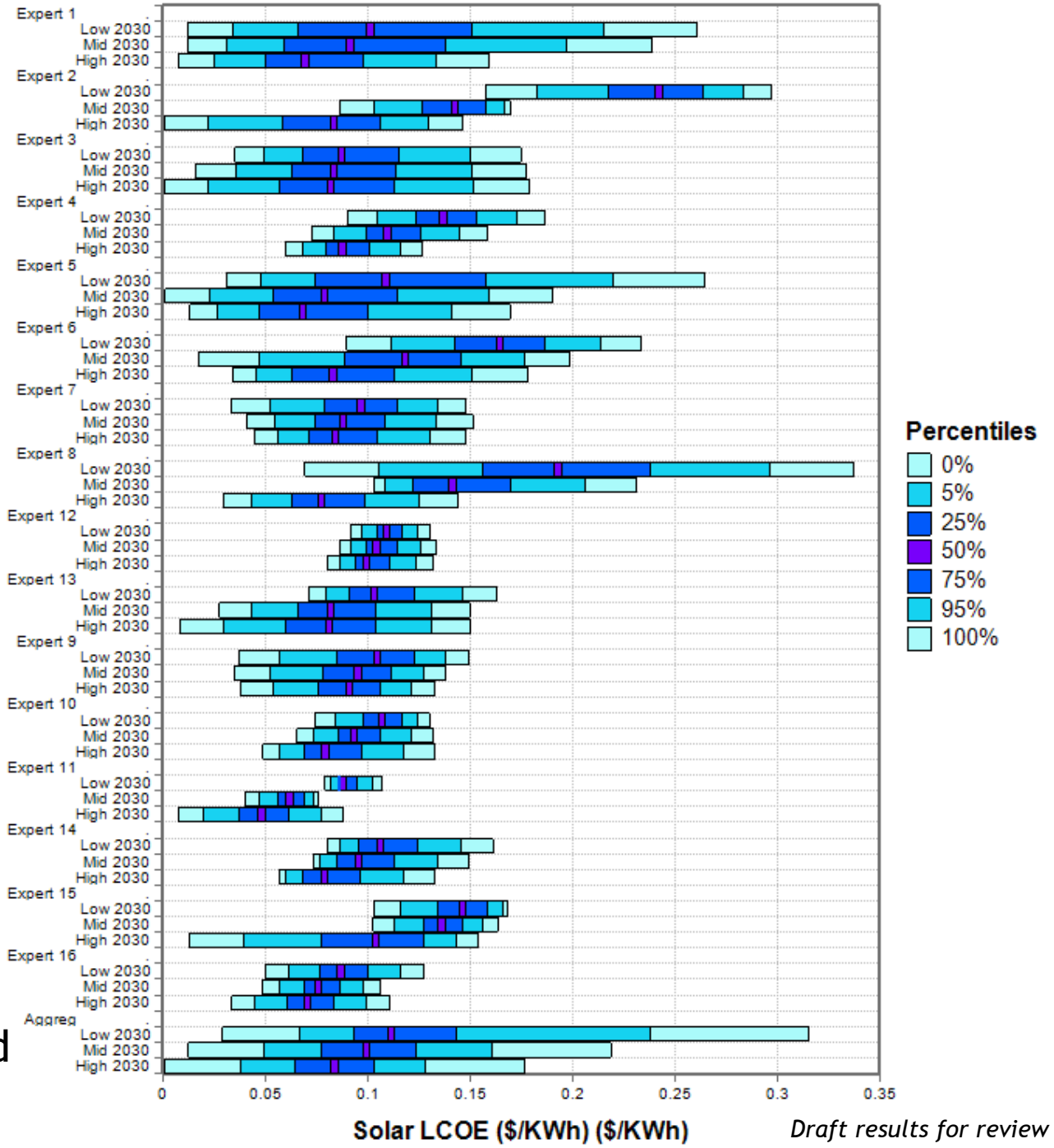
Draft results for review

College scarf probability box plot:

FEEM Solar Levelized Cost by year & R&D level

Experts 1 to 16

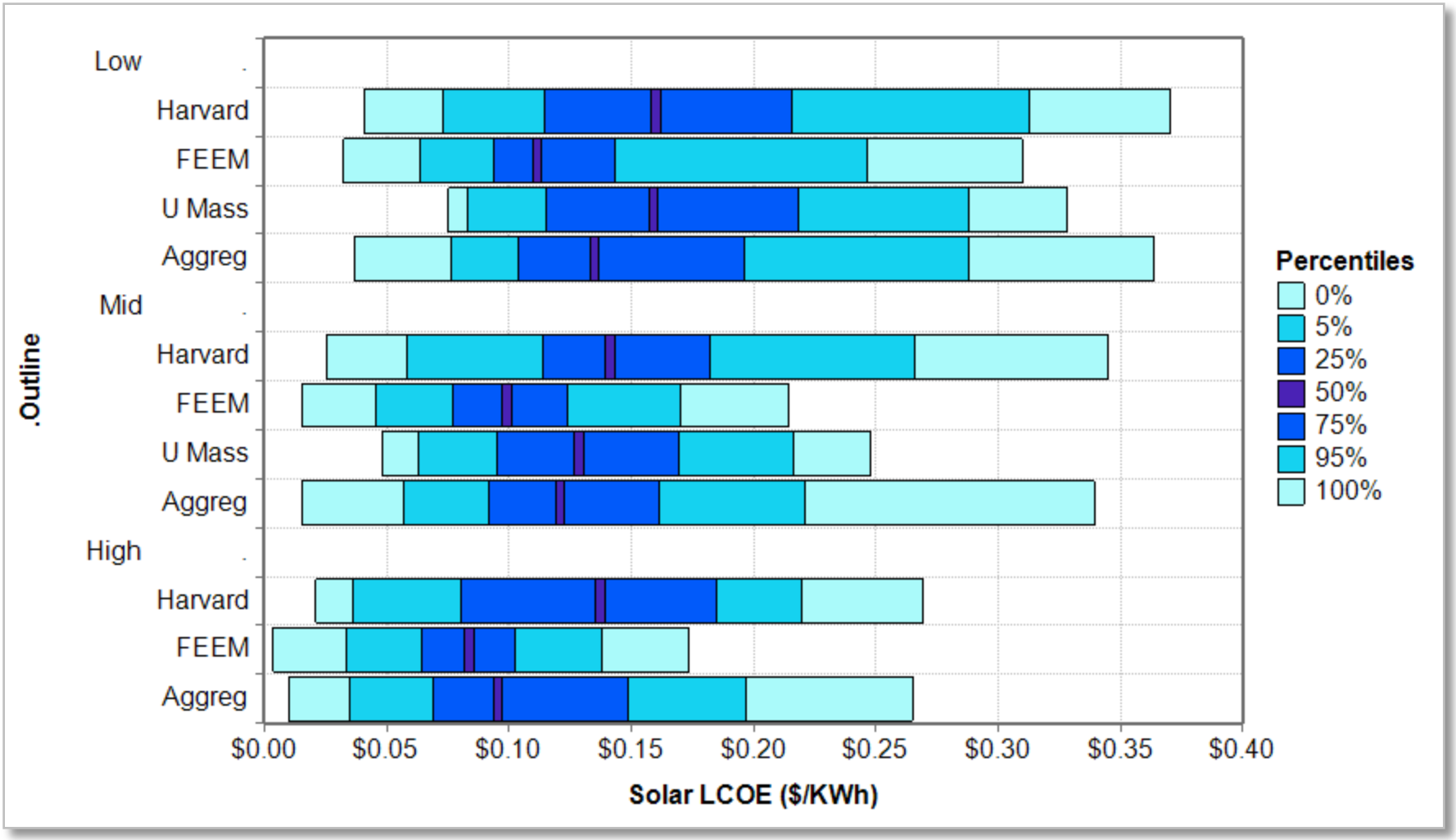
Aggregated



Lumina DECISION SYSTEMS

Draft results for review

Solar levelized cost of electricity: By R&D level, contributor and aggregated



Lumina DECISION SYSTEMS

Draft results for review

How many sample runs (scenarios) do we need?

For reasonable convergence		**1000**
Compare results from three Contributors plus one aggregation	x 4 =	4000
Three methods of fitting distributions	x 3 =	12,000
Three ways to smooth distributions	x 3 =	36,000
Six technologies each with three levels of R&D. Each combination could be an R&D portfolio.	$3^6 = 729$ =	26,244,000
Or use a single set of covering distributions and apply importance sampling to estimate effects of all these changes to the PDF.		5000

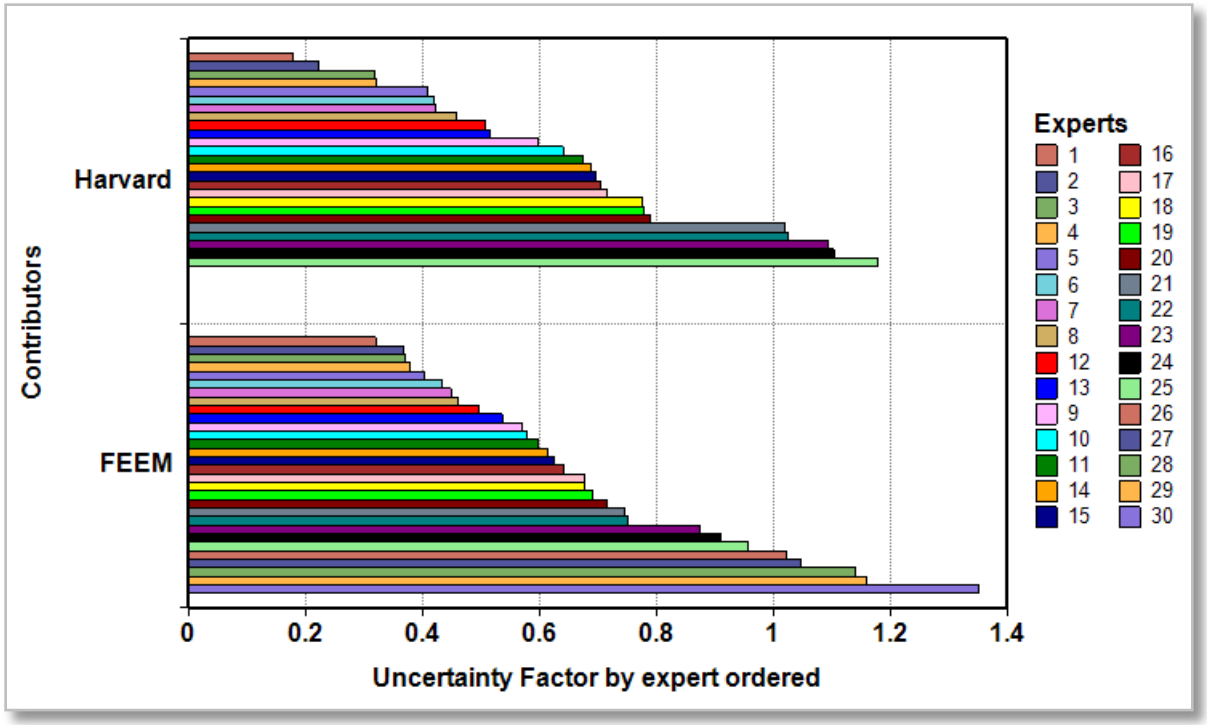
Lumina DECISION SYSTEMS

Covering distributions for use in models

- A ***covering distribution*** is at least as wide as the distributions it covers (for all R&D levels and contributors).
  - We used a range that is 5% smaller than the min and larger than the maximum of all elicited distributions.
- We use a uniform distribution for efficiencies and log-uniform for costs.
  - Log-uniform gives more coverage at low costs: More likely to lead to high technology adoption and effect on emissions – most interesting model behavior.
- Input to the models: 10,000 samples from a covering distribution for each eight quantities.
- Importance weights applied to model results let one investigate the effects of any set of distributions and R&D funding portfolio without having to re-run the model.
- Dramatically reduces number of model runs needed.
- It also means we can continue to revise distributions after the models have been run.

Lumina DECISION SYSTEMS

Variability in uncertainty factors: By experts and contributors



Uncertainty Factor is defined as the range P_{10} to P_{90} divided by P_{50}

$$=(P_{90} - P_{10})/P_{50}$$

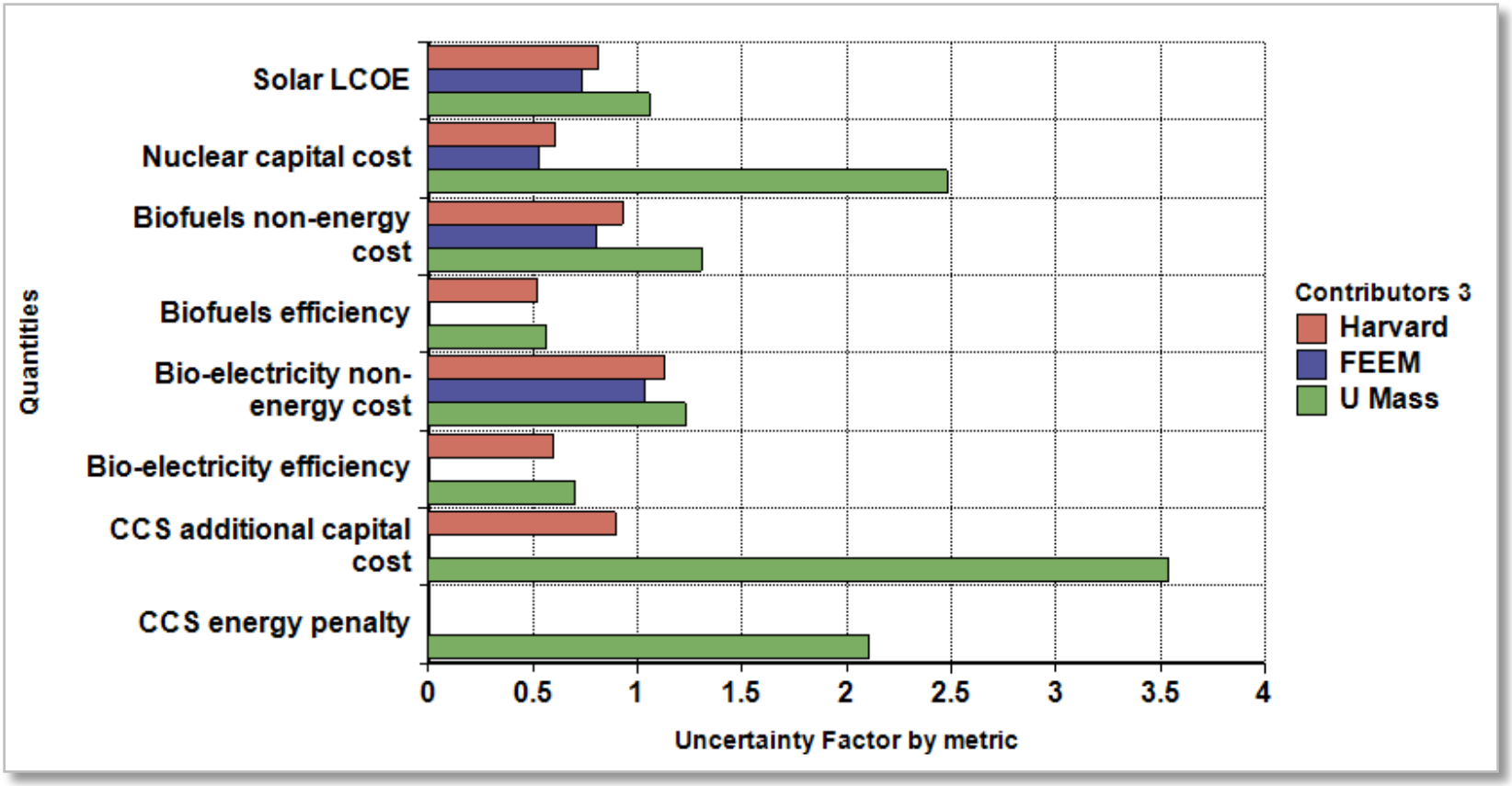
In this chart, they are sorted by average uncertainty factor for each expert to show variability among experts.

- Similar average UF by contributor: Harvard: 0.74, FEEM: 0.72
- Huge range of UF over experts: from 0.18 to 1.35.
- Suggests that some experts are far too overconfident (range too narrow)
  - And/or some are underconfident

Lumina DECISION SYSTEMS

Uncertainty factors:

By metric and year and contributor



Uncertainty Factor is defined as the range P_{10} to P_{90} divided by P_{50}

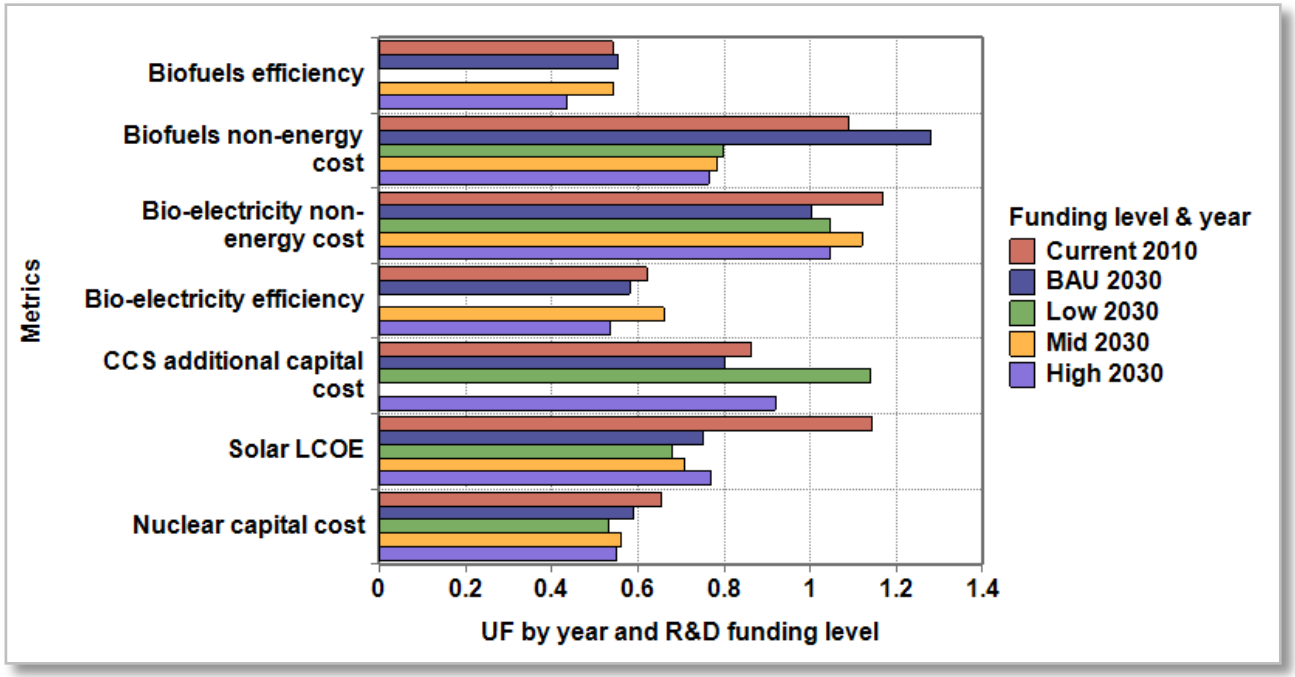
$=(P_{90} - P_{10})/P_{50}$

In this chart, they are averaged over metrics and R&D levels for each contributor.

- FEEM has slightly lower UFs than Harvard, and U Mass some what higher.
- Efficiencies have smaller UF than costs, as expected.
- Nuclear cost is an exception with a low UF, for Harvard and FEEM.

Lumina DECISION SYSTEMS

Uncertainty factors:
By metric and year and R&D level



Uncertainty Factor is defined as the range P_{10} to P_{90} divided by P_{50}

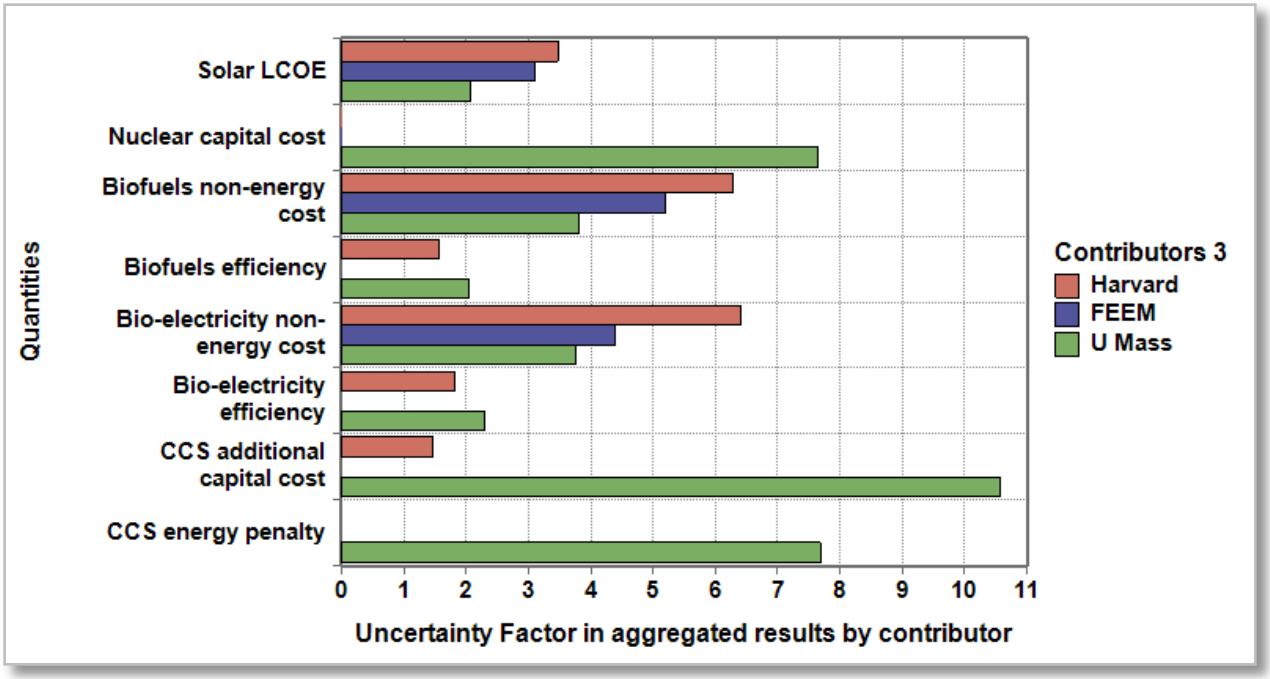
$$=(P_{90} - P_{10})/P_{50}$$

In this chart, they are averaged over experts and contributors for each metric, funding level and year.

- Efficiencies have smaller UF than costs, as before.
  - Nuclear cost is an exception with a low UF.
- Uncertainty factor does not increase over time (as one might expect) and often decreases – e.g. bio-electricity cost, solar, and nuclear.

Lumina DECISION SYSTEMS

Uncertainty factors for aggregates
By metric and contributor



Uncertainty Factor is defined as the range P_{10} to P_{90} divided by P_{50}

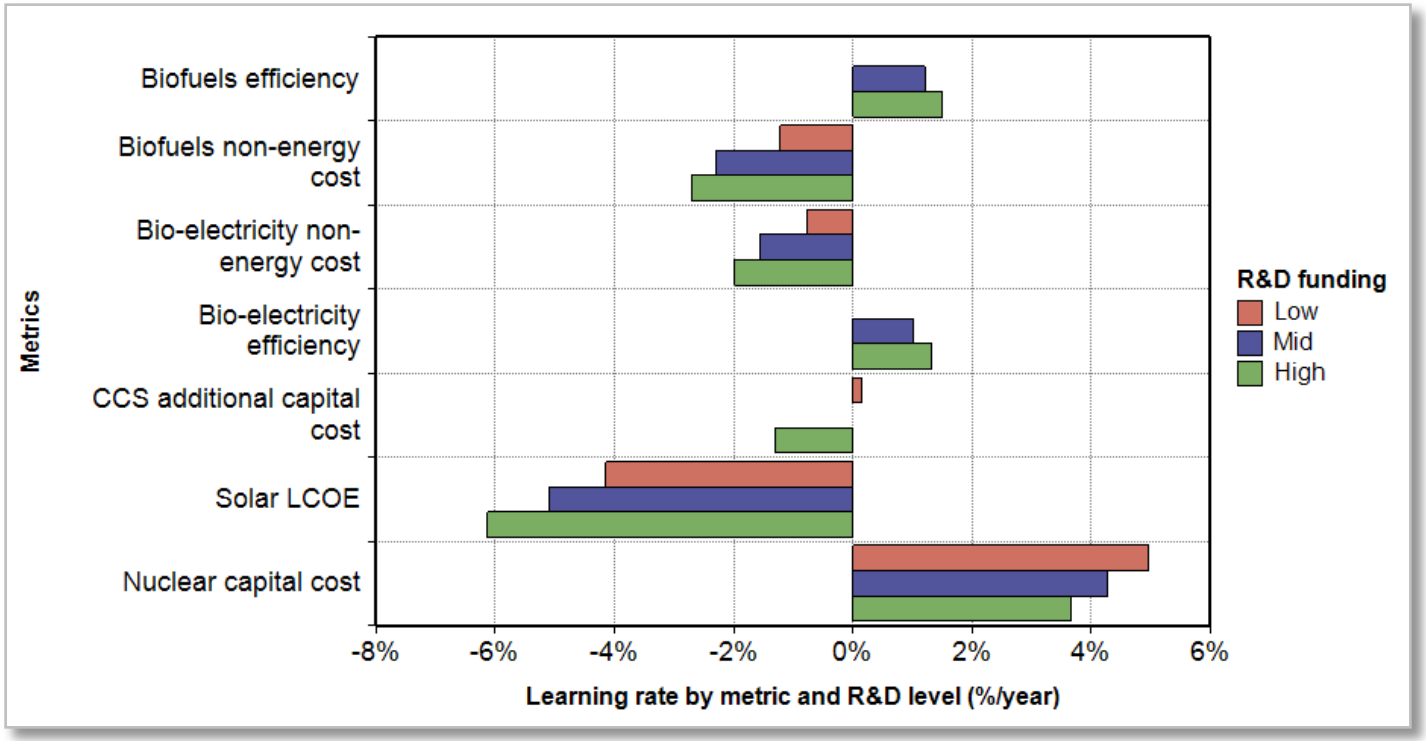
$$=(P_{90} - P_{10})/P_{50}$$

These are Ufs for aggregate distributions averaged over R&D level for each Contributor.

- Uncertainty factors are much larger for aggregated distributions than for individual experts.

Lumina DECISION SYSTEMS

Learning rate per year 2010-30:
By metric and R&D funding level



Learning rate is here defined as the compound annual growth rate (CAGR) from 2010 to 2030 for median estimates, averaged over experts.*

$$=(P_{50}[2030]/P_{50}[2010])^{(1/20)} - 1$$

*Harvard only since others did not provide 2010 estimates.

- Efficiency increases, and costs decrease over time
  - except Nuclear and CCS low R&D
- R&D funding level improves learning
  - and reduces increase rate for Nuclear
- Solar costs decrease most rapidly

Lumina DECISION SYSTEMS

Issues in harmonizing EE results

Issue	What we did	What else might we do?
Varying units, e.g. $/liter gasoline equivalent to $/GGE	Convert to US units	
Varying currency and nominal currency years. Should we convert from Euros to USD for 2005 and then from nominal 2005 to 2010, or vice versa?	Convert to USD first	
Some assessed only PV, some specific types of PV, some utility, commercial, and residential, some PV and CSP, some the best of PV and CSP.	Treat them all as "Solar"	
Different assumptions about latitude and climate, which affects insolation levels.	Ignore	
Differing definitions of Nuclear	Ignore	
Varying R&D funding amounts by contributor	Classify into low, mid, and high, ignoring amounts.	

Lumina DECISION SYSTEMS

Issues in fitting distributions

Issue	What we did	What else might we do?
Some experts assumed no change with R&D funding level	That may be a valid opinion. Accept as given.	
Some estimates imply that increased R&D funding could increase technology costs.	Accept as given.	
Harvard estimated biofuel cost for gasoline, diesel, and jet-fuel substitutes	Use gasoline and ignore the others.	
Some distributions are implausibly narrow, suggesting overconfidence. e.g. Harvard Ex8 Mid R&D 2030 biofuel cost [$0.52,0.57, 0.60] & biofuel efficiency [80%, 84%, 88%].	Accept as given.	
Some distributions have implausibly low or high values – e.g. bio-fuels production efficiency from 38% to 99%	Accept as given.	
Triangular distribution has the advantage of looking arbitrary – which is also a disadvantage.	Use triangular anyway.	
Sometimes fitting a triangular to given P10, 50, 90 gives a tail with negative values or efficiency above 100%	Truncate distribution below zero and, for efficiency, above 100%.	

Lumina DECISION SYSTEMS

Issues in aggregating distributions

Issue	What we did	What else might we do?
Aggregating over experts		
Some experts may be more expert than others.	Use equal weighting for all experts.	
Laplacean mixture gives odd multimodal distributions	Smooth result using piecewise cubic on CDF. But still often multi-modes	
Aggregating over contributors (Harvard, FEEM, U Mass)		
Contributors had different numbers of experts.	Use equal weighting	
Contributors had different elicitation processes	Use equal weighting	
…		

Lumina DECISION SYSTEMS

Lumina
DECISION SYSTEMS