



**What We Can Learn from the Film Industry  
about Group Wisdom and the Winner's Curse**

**Pierre Delfiner**

AAPG International Conference and Exhibition, Paris, France - September 11-14, 2005

Good afternoon!

Seeing the title of my talk some of you may wonder if they are at the right conference! :-) What does the film industry have to do with Petroleum Geology?

Well it's just that I long had a few questions in the back of my head, and was waiting to find dataset to answer them. One such question is about the explanation of the so-called "winner's curse", which I will recall in a minute.

## Outline

---

### MOTIVATION FOR THIS STUDY

### THE BET CLUB

- Principles of the game
- How estimates are formed

### ANALOGY WITH PROSPECT EVALUATION

### DATABASE ANALYZED

### RESULTS

- Relative to committee evaluations
- Relative to bidding

## The winner's curse

The **winner's curse** is the notion that

*“In competitive bidding, the winner tends to be the player who most overestimates the true tract value”*

(Capen et al., *Competitive Bidding in High-Risk Situations- JPT, June 1971*)

The reasoning behind this statement is as follows:

1. The mean of all competitors' bids is equal to the true value, or close
2. The winning bid, being the maximum, is higher than the mean
3. Therefore the winning bid is higher than the true value

**Conclusion: the winner overpays!**



The first time I came across the notion of 'winner's curse' was at a talk given by Ed Capen in Paris, an excellent talk. I then read a paper he published in 1971 on the theory of competitive bidding, which is also excellent.

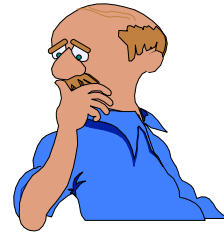
The explanation of the curse is explained in the slide.

Capen's recommendation to protect ourselves from the winner's curse is to lower the bid and he specifically proposes to apply a discount factor of 35%. In other words, we should bid 35% of the estimated value of the prospect. This strategy will protect us from the winner's curse and probably also from winning any bid. :-) Here I will take the opposite approach and rather try to estimate by how much we have to multiply our estimate to win the bid.

## Is it a fact?

### mean of estimates = true value ?

- With petroleum prospects the true value is never known
- We would like to check this on real data



We have access to a large dataset from the film industry: the Bet Club

This dataset is used as an analog for prospect evaluation

movie = prospect

player = geologist

Many prospects, many estimates, the true values are known

The proof of the winner's curse is based on the assumption that the mean of all estimates is equal to the true value.

This notion is very intuitive and very common. It is the idea that the errors of the different estimates cancel out by averaging, and the larger the group of people that make an estimate the more accurate the average.

But is this an established fact?

The problem with prospect evaluation is that we never really know the answer (that may be a saving grace for us, geoscientists!) :-) We could work with simulations, but they lack the flavor of reality.

We will use the database from the film industry as an analog and see what we can learn from it.

## Questions considered

---

### Group wisdom

- Should we tap into the collective wisdom of all our geologists?
- Should we trust our best expert?
- Is there a better way?

### Bidding

- Is the winner's curse inevitable?
- Can we predict how much the winner overpays?

We will address two types of questions

- Questions about group wisdom

For a given prospect should we ask the opinion of 200 geologists or should we ask the guru, for example the turbidite guru if the field is a turbidite?

- Questions about bidding

If the estimates are a basis for bids, does the maximum estimate consistently exceed the true value? And if so can we predict by how much?

## Principles of the Bet Club

---

- In France new movies are released every week on Wednesday.
- On Tuesdays, professionals of the film industry log on a website and record their forecast of the number of admissions for these new movies.
- These forecasts are estimates of the number of admissions on the first day of release in Île de France only (the greater Paris area).
- When the actual number of admissions becomes known estimates are compared with true values and a ranking of the players is established.
- This is only a game. No financial stakes involved.

(Thanks to Boris Pugnet!)



## Information used for predicting admissions (1/2)

### Analogs

- Results of similar films in Ile de France (accessible on paying sites or proprietary databases)
- Results of the same film abroad (especially in the US, accessible on the Internet)

### Print and advertising (budget and quality)

- Number of screens in France (accessible to all), and in Ile de France (accessible to only few players)
- Film marketing and promotion (number of TV programs, of articles, etc.)



Estimation of the number of admissions is a complex process. To give you an idea of this process here is a list of pieces of information used to predict admissions. Not all points have a direct analogy with prospect evaluation.

Analogs are the main source of information to establish a prediction. For each movie the players try to find what they call a '**reference film**'. Sometimes the reference is obvious

sequels: Star Wars3 => Star Wars 2  
Ocean 12 => Ocean 11

last film from the same director: Clint Eastwood, Scorsese

Sometimes the analog is more fuzzy. It can be for example the style of film (e.g., French comedy).

The other most important piece of information is the **number of screens**. Obviously if the movie is shown in 900 theaters (Spider-Man) it has more chances to attract viewers tickets than if there are only 5 screens. The maximum for Ile de France alone is 52 screens.

The number of a screens do not have an analogy for exploration.

## Information used for predicting admissions (2/2)

---

### Film credits and quality

- Film director, casting, producer
- Film type and targeted audience (teen age movie, art-house movie,...)
- Film notoriety (e.g., nominated in Cannes, adaptation from a famous novel, remake, sequel, etc.)
- Press coverage, press reviews, viewers' reviews
- Quality of preview (trailer usually available on the web)

### Competition

- Competition from other movies
- Other factors: weather, holidays, etc

Here are other factors. The casting is quite important. Some people like Bruce Willis and they go and see any movie with Bruce Willis.

Competition from other movies, or other activities, is also important.

## Analogy with prospect evaluation

### Similarities

- Film industry extremely risky
- Each film is different
- Large dispersion of estimates
- Existence of “nuggets” (Star Wars 3 = Kashagan)
- Level of information varies among players
- Analogs are used
- Estimation process is both analytic and intuitive

### Differences

- The Bet Club has no financial stake
- No incentive other than giving the best estimate
- The value of a prospect is not the same for all companies
- Film attendance involves human component



- There are major flops. A recent example is ‘The Alamo’. It was released in the US in 2004 for a total cost of 137 M\$ (production: 107 M\$, marketing: 30 M\$) and generated an estimated loss of 125 M\$! It was such a flop in the US that it was not released in the rest of the world. On the other hand ‘La marche de l’empereur’ (March of the Penguins) was a major unexpected success: Warner Independent bought the distribution rights for less than 1 M\$ and grossed over 70 M\$ in the US (source: [www.boxofficemojo.com](http://www.boxofficemojo.com)) .
- The film industry is called a prototype industry. **Each oil field is special.**
- Large dispersion of estimates **Same for prospects (e.g., North Sea rounds)**
- Company makers **in both cases**
- The level of information varies among players: part of the information is accessible to all players, other parts is accessible to only a few  
**Some companies know more (e.g. the Operator)**
- Analogs are used **in both cases**
- Estimation process is both analytic and intuitive **in both cases**

### DIFFERENCES

- This is just a game and people tend to focus better on a decision when there are financial rewards attached to it. However the players are professionals, and their reputation is at stake which provides incentive to do serious work.
- This is a strong point. In prospect evaluation there may be strategy involved, with films the only goal is to come up with the best estimate.
- A prospect may be worth more for a company that already has facilities in the area than for one that doesn’t.
- The attendance to movies involves human behavior .  
**This component is absent with petroleum prospect**

## Database analyzed

---

- 125 movies (from 24/11/04 to 25/05/05)
- 42 players initially.
- Problem with missing data: all players did not participate every week.
- Only players who contributed to at least 80% of the forecasts were retained (33 players).
- Missing data replaced by the mean estimate for that movie.

## Database view

### PLAYERS →

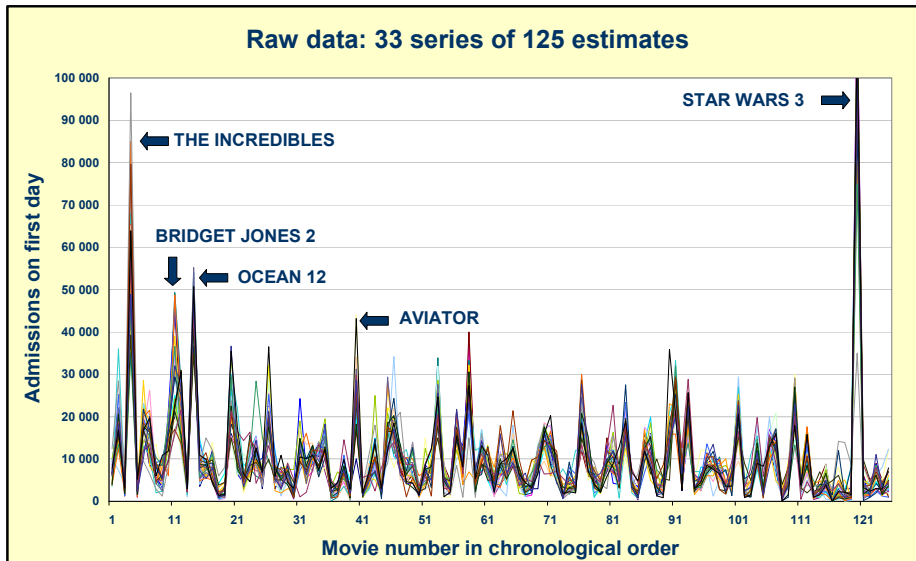
	Prévisions	Allan the player	JP	Lionel	Eric	Lo	Rémi	Seb Yoda
241104	Holy Lola	7500	6500	6031	3822	6333	4320	6500
	36, Quai des Orfèvres	22000	16000	15336	18222	21900	18327	18000
	l'enlèvement	1800	2000	2032	1222	2006	1200	3000
	les Indestructibles	67000	60000	61236	36222	68000	65427	55000
	Si seulement	2320	2700	3954	3222	3003	953	2200
011204	Narco	19700	17500	15516	17222	23110	16323	16000
	Pole express	12500	8500	14756	14222	15980	8500	11500
	Mémoire effacée	7500	8000	6954	5222	7220	5555	4200
	Ghost in the shell 2 : Innocence	8500	3500	3401	4222	8008	3350	6200
081204	Blade : Trinity	14000	16000	17021	15222	16961	17348	14000
	Bridget Jones 2	23000	20000	31804	24222	25052	36510	45621
	Les Dalton	19000	20000	22221	19222	15051	15230	18000
	Les 3 Rois mages	3500	2000	4802	2422	8008	4500	4000
151204	Ocean 12	37500	40000	45632	42222	38650	48012	50100
	Le dernier Trappeur	10000	3500	5302	9222	14566	5600	10000
	Les temps qui changent	6800	7000	7034	7222	8620	8011	10000
	Un petit jeu sans conséquences	11000	12000	7352	8222	6100	4600	9000
	Famille à louer	13	900	946	1922	4204	500	2500
	Noel de foie	2900	2200	945	2222	6632	1800	2500
221204	Benjamin Gates et le trésor des templiers	25000	16000	26456	22222	21001	15300	21320
	Les désastreuses aventures des orphelins Baudelaires	9500	12000	11441	6222	13333	12200	14321
	Coup de foudre à Bollywood	6000	4000	2631	5222	8412	5600	4820
	Les soeurs fâchées	9500	6000	12458	6822	9820	7800	8500
291204	The Grudge	16000	12000	13652	8222	11276	12400	7213
	Irresistible Alfie	3200	3000	5587	6822	6597	4000	7009
050105	Alexandre	27000	16000	16859	32222	21012	17000	22412

MOVIES

You are not expected to read this. It is just to show the structure of the database: an Excel file with 125 rows, the movies, and 33 columns, the players.

The green and blue represent the weeks, and the grey cells are missing data replaced by the means.

## Raw data



AAPG, Paris, France.  
September 11-14, 2005.

11



Each curve in this graph represents the estimates of a player.

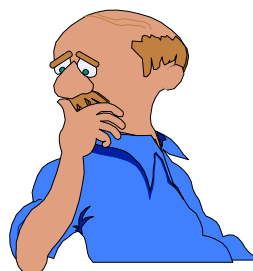
There are 33 series over 125 movies.

Notice the large dispersion in the size of movies (we call "size" the number of admissions). For Star wars 3 the scale goes beyond 100,000, while there are some movies that sell less than 50 tickets.

## Group wisdom

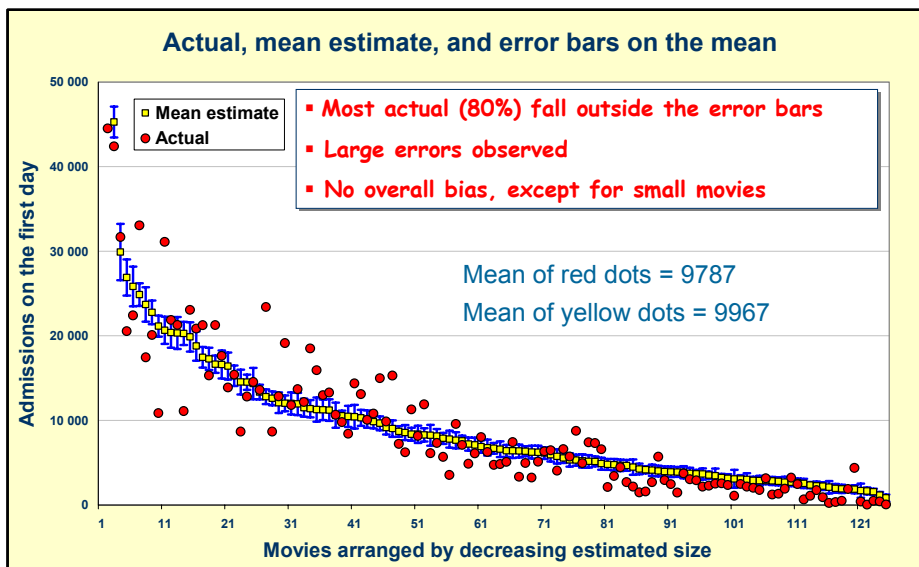
---

**Is the mean of all the players' value estimates statistically equal to the true value?**



So we now look at our first question.

## Mean estimates and actual values



AAPG, Paris, France.  
September 11-14, 2005.

13



This is a key graph. Movies are arranged by decreasing size, size being the mean of all estimates. On the x-axis is the movie number and on the y-axis the number of admissions on the first day, where the scale has been truncated to 50,000 (Star Wars 3 goes up to 120,000).

The yellow dots represent mean estimated sizes. Errors bars in blue are confidence intervals for the mean ( $2 \text{ standard deviation} / \sqrt{33}$ ).

The red dots are the actual values. Clearly they generally fall outside the error bars. For each movie there is a bias.

But overall the red dots jump back and forth around the yellow dots. No systematic pattern is observed, except for small movies which tend to be overestimated. However, over all movies there is no bias. [ANIMATION] The mean of all estimates, the yellow dots, comes within 2% of the mean of all actual values.

We can derive two important conclusions from this graph:

- 1. For each movie there is a bias.** Where does it come from? We will discuss this in the next slide.
- 2. The bias is random from movie to movie.** If we make an analogy with an exploration portfolio the consequences are the following: **Over a portfolio of prospects if we promise to deliver the mean we will deliver the mean.**

## Explanation of the bias

### ❑ Overestimation of small movies

Most players work for majors and are unfamiliar with small movies

### ❑ Insufficient diversity among players

All from same age group and similar socio-economic background

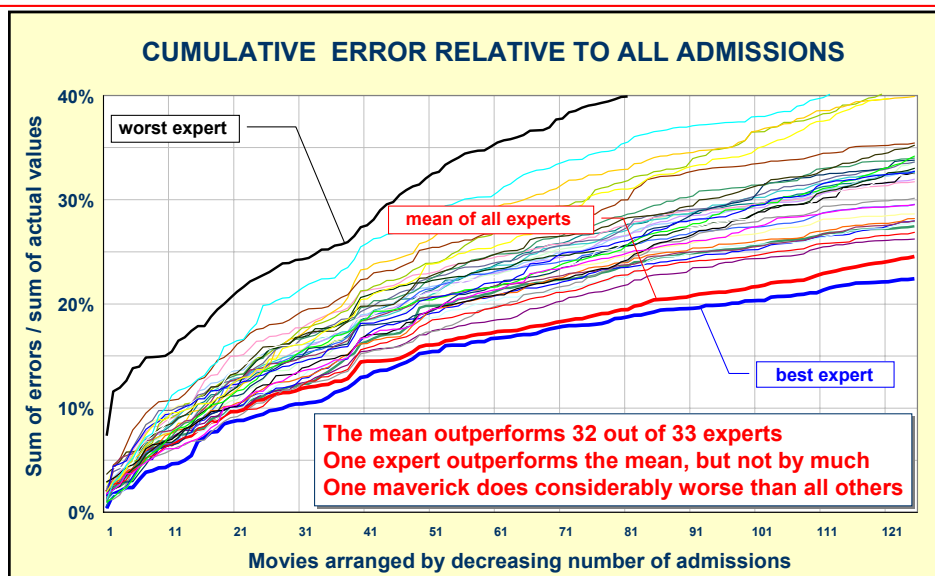
The same reference movies are used

Star Wars 3	=>	Star Wars 2
Ocean 12	=>	Ocean 11
Aviator	=>	Gangs of New York (Scorsese)
Robots	=>	Ice age (cartoon)
Million \$ baby	=>	Mystic River (Clint Eastwood)

In the oil industry we also observe a tendency to overestimate small fields (due to an economic cutoff effect?)

One condition for the canceling out of errors is the independence of estimates. If they all use the same reference, that reference will pull all the estimates in the same directions and create a bias.

## Performance ranking



AAPG, Paris, France.  
September 11-14, 2005.

15



Now let us look at the performance of the players of the Bet Club. These curves show the CUMULATIVE ERROR the various movies sorted by decreasing size. These errors are added in absolute value and are scaled by the total number of actual admissions. So they represent an error as a percentage of the total number of tickets sold. The best expert has the smallest error and is represented by the blue curve. The worst expert has the largest error and is represented by the black curve.

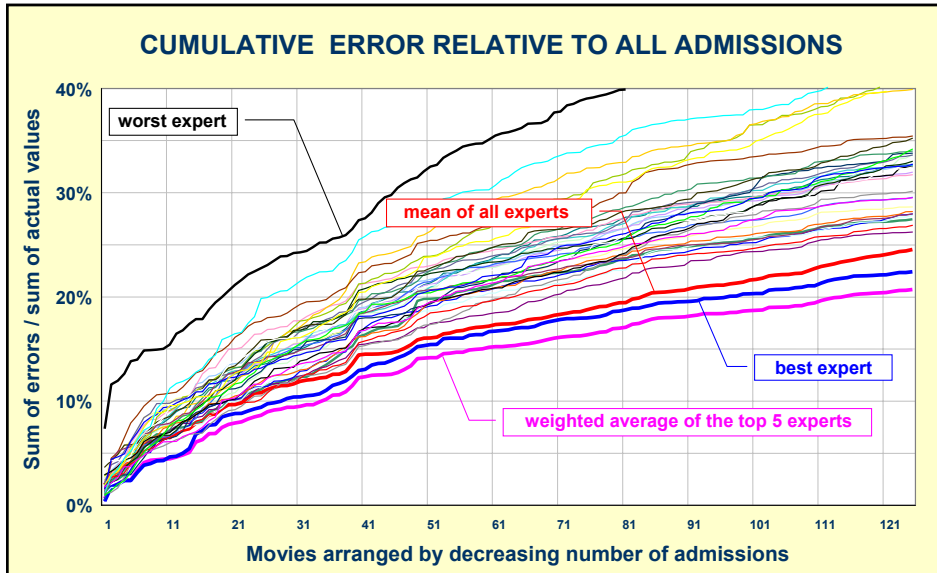
We observe a large dispersion of estimates. One expert is completely off. We also see that in bids, some estimates are completely off.

Notice also that the best expert outperforms the second best by an appreciable margin.

Now let us introduce a new player, Mister MEAN. How does he perform? **[ANIMATION]** This is shown by the red curve. Again, this curve represents a score, the cumulative error, and this is why the red curve does not plot in the middle of the other curves. Our new expert, Mister MEAN, has a rather low error, and does better than all but one expert.

**[ANIMATION]** Conclusions

## Performance ranking



AAPG, Paris, France.  
September 11-14, 2005.

16



Is it possible to do still better than the mean?

Here is the purple curve. **[ANIMATION]** It is 16% better than the mean. Can anyone guess what it is?

Answer: the weighted mean of the top 5 players.

## Mixture of experts

---

Select top 5 experts and compute a weighted average.

- 53% of Lionel
- 17% of Boris
- 12% of Eric
- 11% of Thomas
- 10% of Alex

**Weights represent  
credibility**

This estimate does 16% better than the mean (20.7% vs 24.5%).

Above weights were adjusted on the database but are not essential. The mean of the 5 experts gives about the same performance (21.3%).

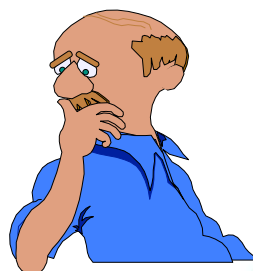
A new estimate is formed by computed a weighted average of the estimates provide by the best 5 players.

The weights can be interpreted as representing credibility. In fact they are not really essential, the straight average of the 5 best players would have about the same performance. What is key is the selection of the top 5.

## Bidding

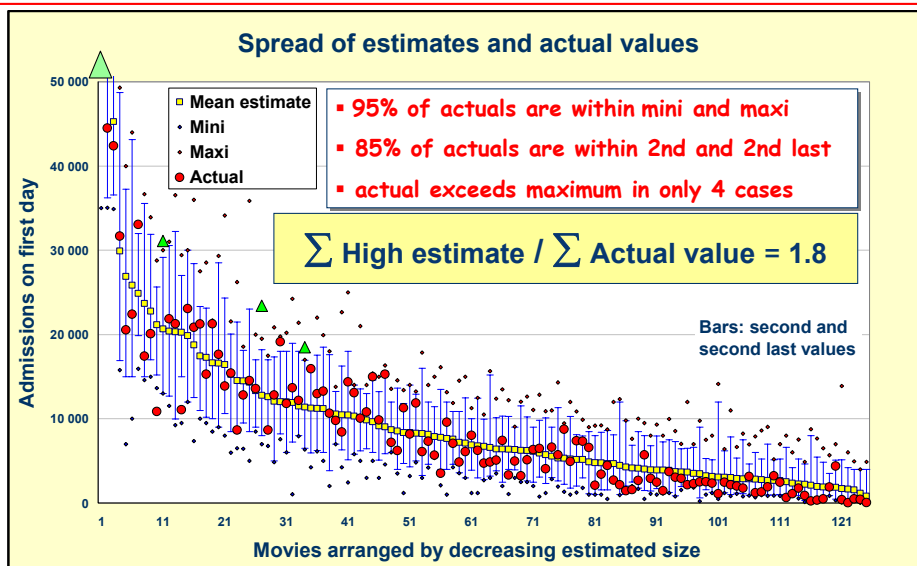
---

**Does the maximum estimate consistently exceed the true value?**



We suppose here that our estimate of the value of a prospect is used as our bid. Does the winner overpay?

## Mini, maxi estimates and actual values



AAPG, Paris, France.  
September 11-14, 2005.

19



This graph shows the (mini, maxi) range. These are the blue and red diamonds. The blue bars represent the intervals between the second and second to last values; this is to avoid the influence of extreme estimates.

We see that most points fall within the (mini, maxi) ranges.

### [ANIMATION]

The actual value exceeds the maximum estimate in only 4 cases, marked with a green triangle.

On the average the maximum estimate is about 80% higher than the true value. To win we have to pay 80% more than the true value.

### Conclusions

- Collectively the players do a good job at predicting the range.
- Although the mean estimate is not equal to the true value, the winner's curse does take place anyway

## Multiplicative factor

**Is it possible to estimate the multiplicative factor to apply to our estimate to match the winning bid?**

### Assumptions

1. All players bid independently according to the same lognormal distribution
2. My bid is the P50 of this distribution

$$E(\text{factor}) \cong \exp\left[\sigma \Phi^{-1}\left(\frac{n}{n+1}\right)\right] \quad \Phi: \text{standard normal cdf}$$

It depends on

- the number of competitors  $n$  (including us)
- the dispersion of estimates  $\sigma$  (logarithmic standard deviation)

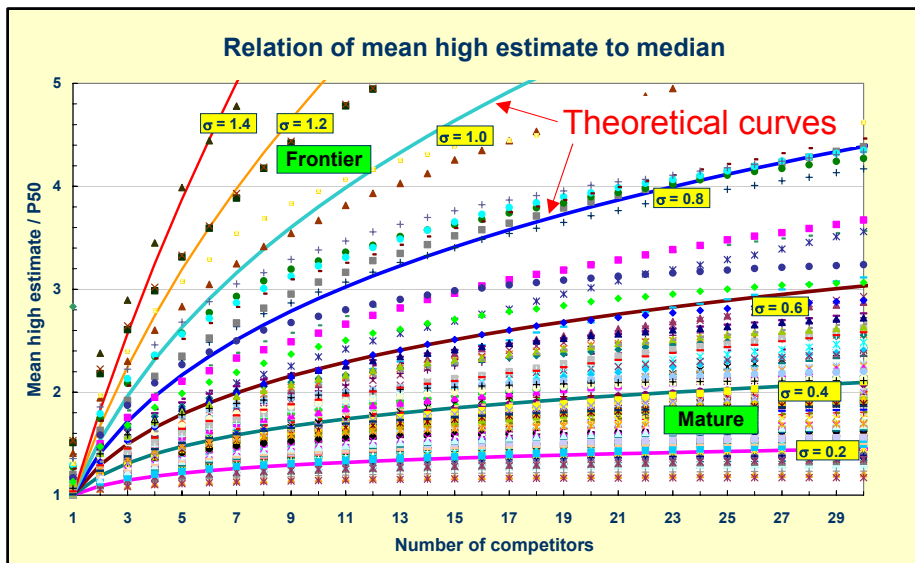


It is possible to estimate the multiplicative factor if we make some simple assumptions about the distribution of the bids. We assume that all estimates are independent and follow the same lognormal distribution with logarithmic standard deviation  $\sigma$ .

This parameter  $\sigma$  captures the uncertainty of estimates, it is large in frontier areas and small in mature areas. Then the expected value of the multiplicative factor is given by the above formula. It depends on

- the number of competitors: obviously the more players compete the higher the maximum will be
- the uncertainty about the prospect (whether it is frontier or mature)

## Multiplicative factor: actual and lognormal model



AAPG, Paris, France.  
September 11-14, 2005.

21



When we apply the above model to the film data we get this picture.

The points are the mean high estimates obtained by sub-sampling the database. For example we take a subset of 5 values, compute their maximum, then take another subset of 5 value and compute the maximum, repeat this and in the end average all these maxima.

The continuous curves are computed using the lognormal formula of the previous slide. The **amazing** result is that the curve match the points, very well for small sigmas and acceptably for large sigmas.

This indicates that our simple lognormal model is reasonably accurate and enables prediction of the multiplicative factor.

## What have we learned from the film data?

### Group wisdom

- On a case by case basis there is a bias
- This bias may be caused by lack of sufficient diversity
- Over a large enough portfolio the mean performs quite well
- A top expert can outperform the mean
- A weighted average of top experts does even better
- Committees should preserve the diversity and independence of its members
- Multiple independent assessments are a benefit of Joint Ventures

### Bidding

- The winner's curse exists even if mean estimate  $\neq$  true value
- It is possible to calculate the multiplicative factor to apply to our estimate to match the winning bid



It is now time to wrap up what we have learned from the film data that can be of value for prospect evaluation. Obviously we have to be careful when formulating conclusions since the analogy has its limits.

For each movie we have seen that there is often a bias. This bias can be explained by the lack of diversity among players who use the same reference. The same phenomenon may well be at work with geologists. If they share the same cultural background, the same exploration experience, they may all be driven by the same analog. This creates the risk of a bias. In order for errors to cancel out it is crucial to ensure enough diversity and independence of opinions.

Over a large number of movies there is no bias. The mean of all estimates falls within 2% of the mean of actual values. This matches what we observe in oil exploration, if we promise the mean we deliver the mean.

Can we improve on the mean? The film data say "Yes". One player in our panel outperforms the mean, although not by much. However, a weighted average, or simply an average, of the top 5 players improves our estimate by about 15%.

Can these observations help improve the prospect evaluation process? I believe it can by using as a leading thread the fundamental benefit of diversity and independence. When recommendations are made by committees there is a risk of group behavior, for example members may align with the most senior person. How to implement a decision process within a company needs further consideration. One idea, for example, is to ask committee members to make their assessments independently before they meet, at least evaluate the various chance factors (evaluating volumes is more work).

One instance where we have currently have multiple independent assessments of the same quantity **joint ventures**. Beside mitigating risk, joint ventures generally improve the quality of the work by providing a diversity of viewpoints.

A brief conclusion on bidding. We have seen that the winner's curse can take place despite the fact that the mean estimate is not equal to the true value. The winner's curse seems to be an inescapable reality. Rather than try to avoid it seems more interesting to try estimate its magnitude.

**The Wisdom of Crowds**  
**James Surowiecki**  
**2004, Doubleday**

“The more influence a group’s members exert on each other ... the less likely it is that the group’s decisions will be wise ones.”