# IV Use of probabilities

## 1. Decision making from probabilities

# IV.1.1. A story from 1930's California

**In the 1930's Irving Krick, a meteorologist from Cal Tech, established the first private weather forecast firm in in the USA in competition with US Weather Bureau (USWB).**

**Assume we are in a region with adverse weather 30% of the time: 9 days/month or 122 days/year.**

**This is not quite true for sunny California, but it will make the story more easy to tell and understand**

# But first some theory:

## What to do when probability **p** is issued?

1. If you do nothing there is a chance **p** to lose **L**.
2. On average the loss will be **p·L**
3. If you take protective action it will cost **c**
4. Only if **p·L > c** is it worth while to take action
5. The “break even” point is **p = c/L**

Assume that adverse weather will cause a loss **L = €100** per day

For a certain occupation the cost of protection per day may range from c = **€0** to c = **€100** **(the same as the loss)**

We can now calculate the average Expected Monetary Value per day, i.e. the average cost and loss per day if there is no forecast information

## IV.1.2. The local weather forecasters at the USWB make very good forecasts with 80% being correct.

| | Obs rain | Obs dry |
|---|---|---|
| Fc rain | 20 | 10 |
| Fc dry | 10 | 60 |

**All forecasts were well tuned:**

**The number of rain forecasts (30) over 100 days matches the number of observed rain days (30)**

# Expected Monetary Value (EMV)

| Fore casts | Obs rain | Obs dry |
|---|---|---|
| Fc rain | Hit | False alarm |
| Fc dry | Miss | |

| Fore casts | Obs rain | Obs dry |
|---|---|---|
| Fc rain | Cost of | protection |
| Fc dry | Loss | |

$$\text{EMV} = \underbrace{\text{Cost of protection} \cdot (\text{Hits} + \text{false alarms})}_{\text{Protective action taken}} + \underbrace{\text{Loss} \cdot \text{Misses}}_{\text{Protective action \underline{not} taken}}$$

This matrix also reflects the actions and their consequences

| | Obs rain | Obs dry |
|---|---|---|
| Fc rain | **20** | **10** |
| Fc dry | **10** | **60** |

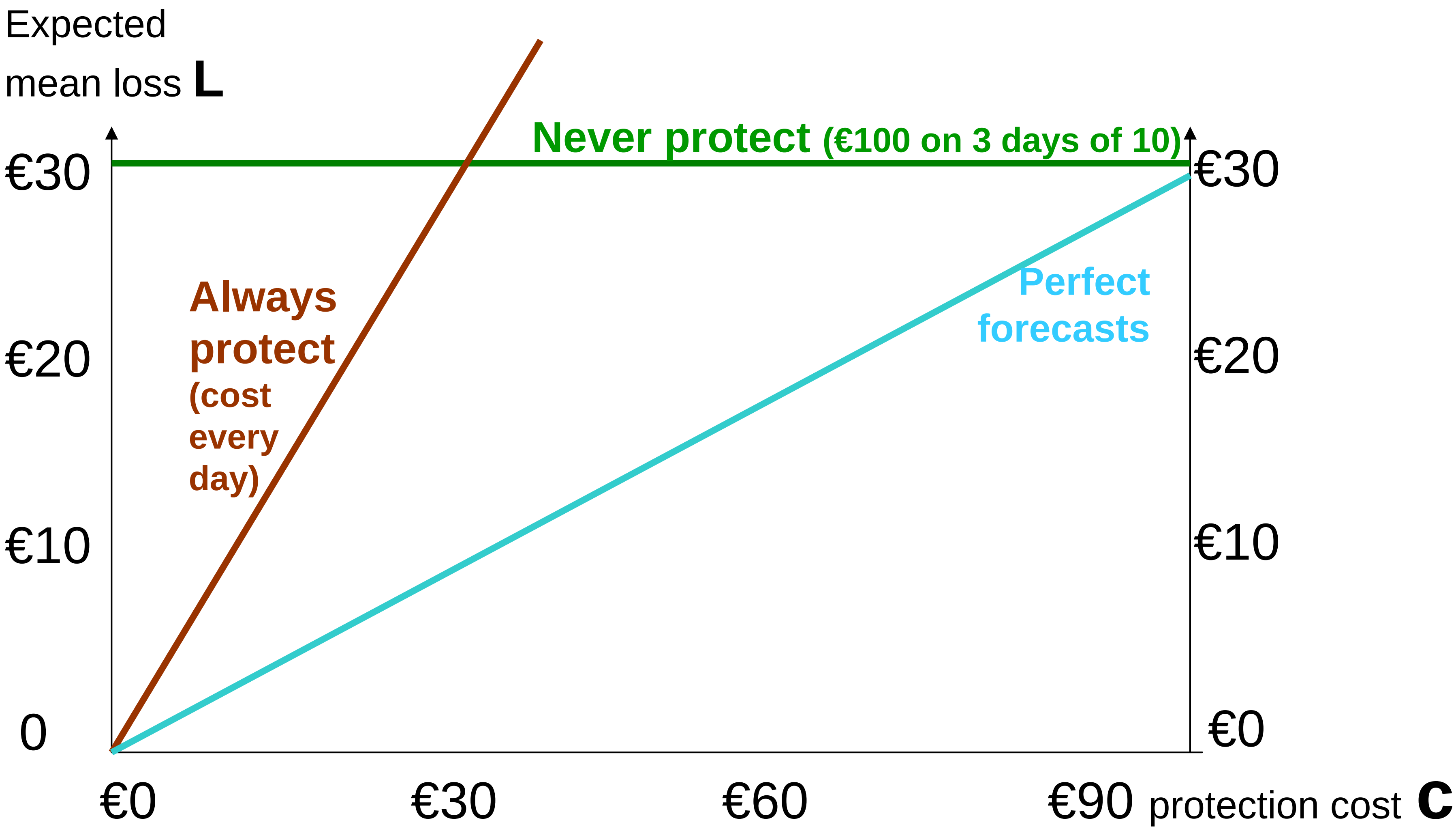
Actions were taken

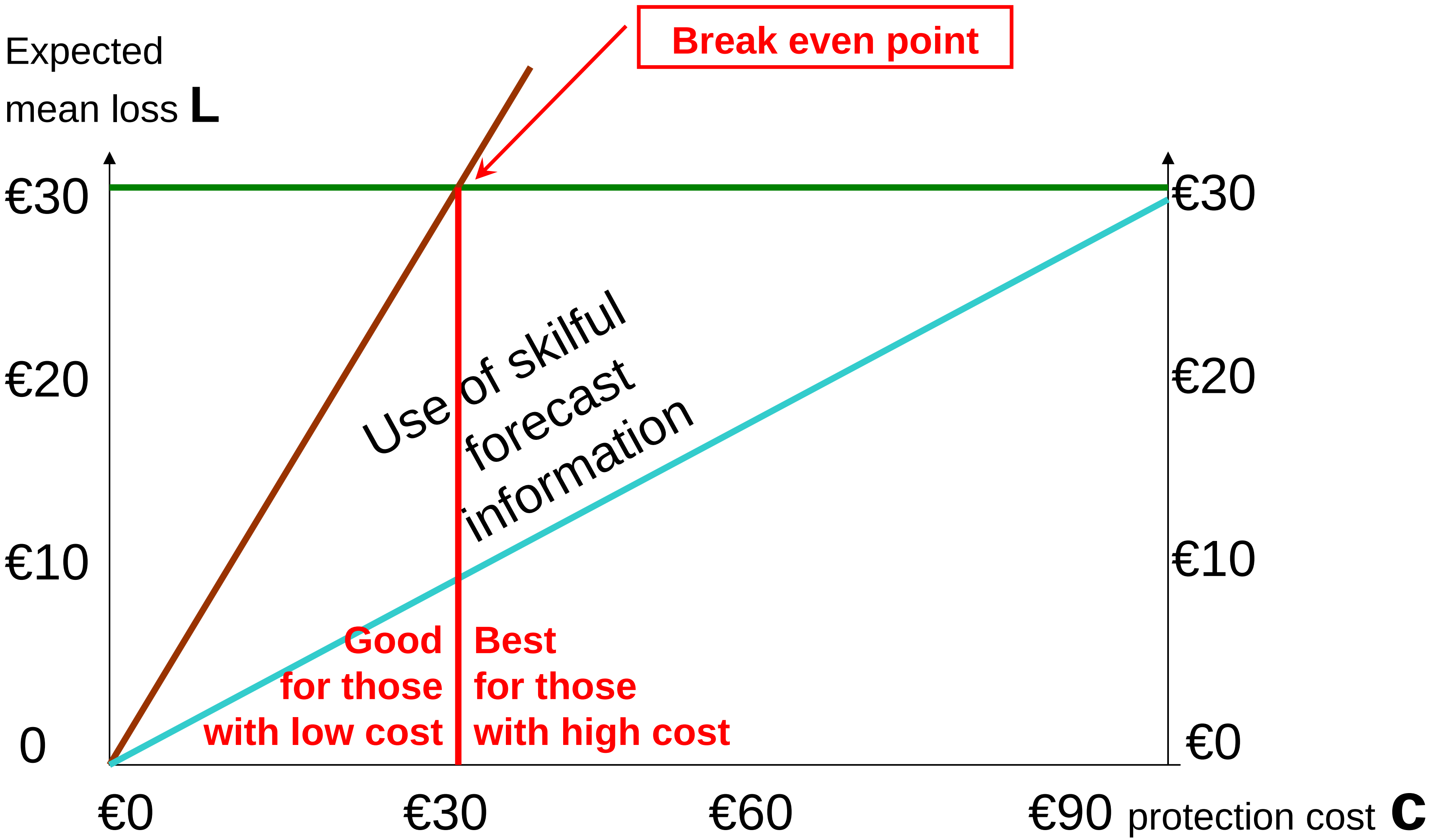
No actions were taken

Losses

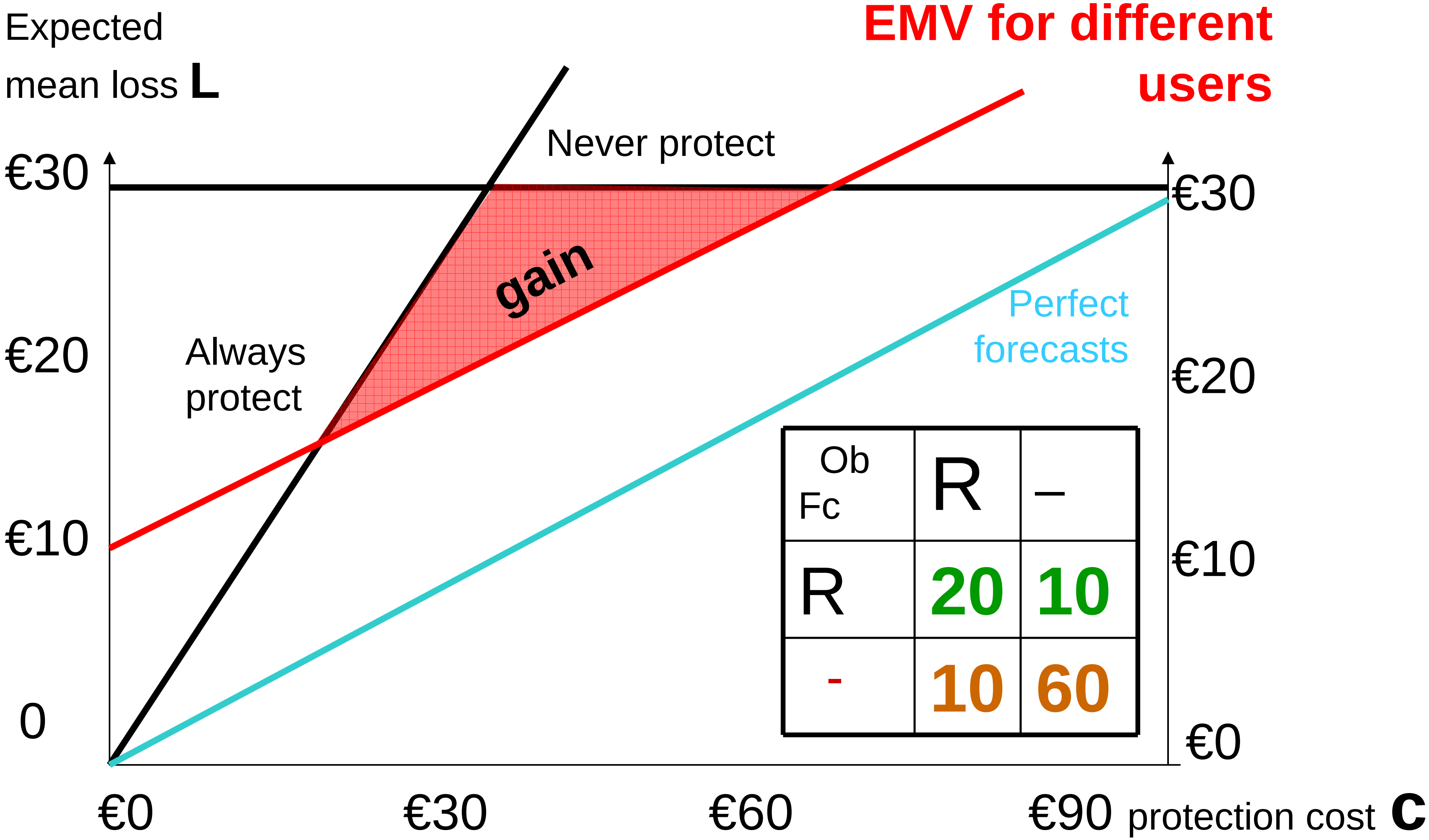
From this it is possible to calculate the Expected Monetary Value (EMV)

# With no forecast information you can chose to a) protect every day or b) never protect

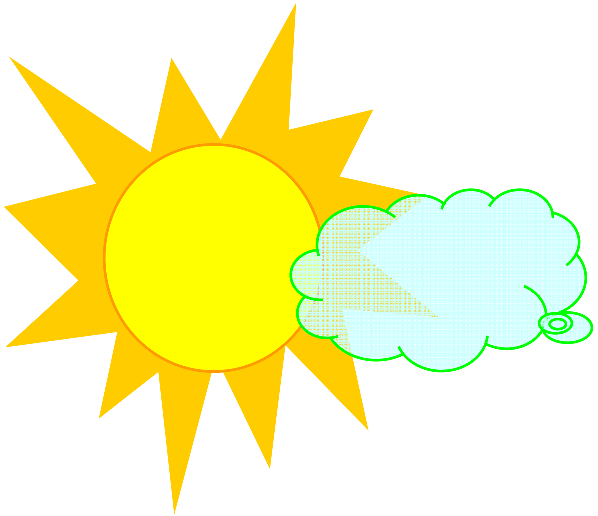



Break even point
Expected
mean loss L
€30
€20
€10
0
€30
€20
€10
€0
Use of skilful
forecast
information
Good
for those
with low cost
Best
for those
with high cost
€0
€30
€60
€90
protection cost c

# The expected loss per day for different protection costs c



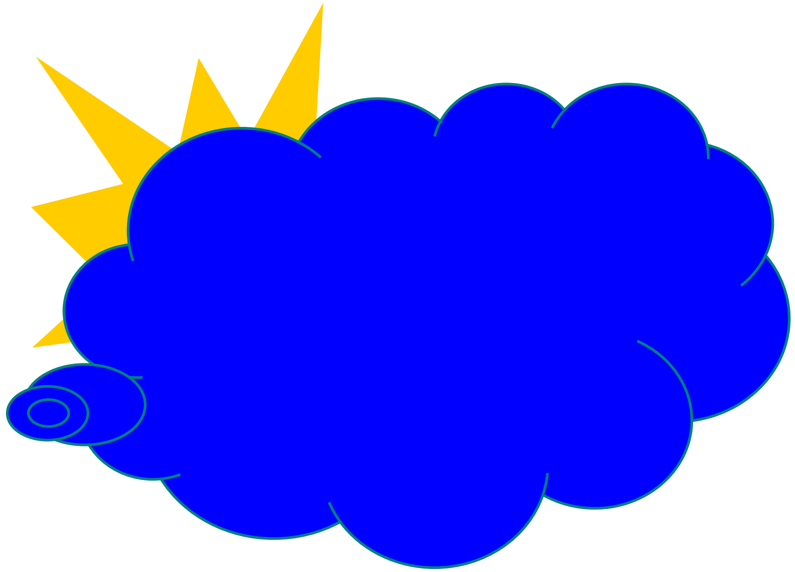
# Irvin Krick’s privately made forecasts were very *bad*



When the Weather Bureau promised the public **sunny and mostly dry. . .**

..the Irving Krick forecast to some of his clients said:

**Probably rain**



When the Weather Bureau warned the public about **probable rain. . .**

..the Irving Krick forecast to some of his clients said:

**Probably dry**

# Verifications showed that Irvin Krick's privately made forecasts were very *bad*

| Fore casts A | Obs rain | Obs dry |
|---|---|---|
| Fc rain | 30 | 30 |
| Fc dry | 0 | 40 |

**Over-forecasting** rain
(60 days vs 30)

| Fore casts B | Obs rain | Obs dry |
|---|---|---|
| Fc rain | 5 | 0 |
| Fc dry | 25 | 70 |

**Under-forecasting** rain
(5r days vs 30)

**Still Krick's private weather firm earned him millions**

**Why?**

## A: The rain was *over-forecast* for the Hollywood studios because of their low c/L ratio.

**Low cost:** Staying at home and risk missing a fine day.

**High loss:** To have the stars and equipment unnecessarily taken out on the prairie in case of unpredicted rain.

## B: The rain was *under-forecast* for the water authorities because of their high c/L ratio.

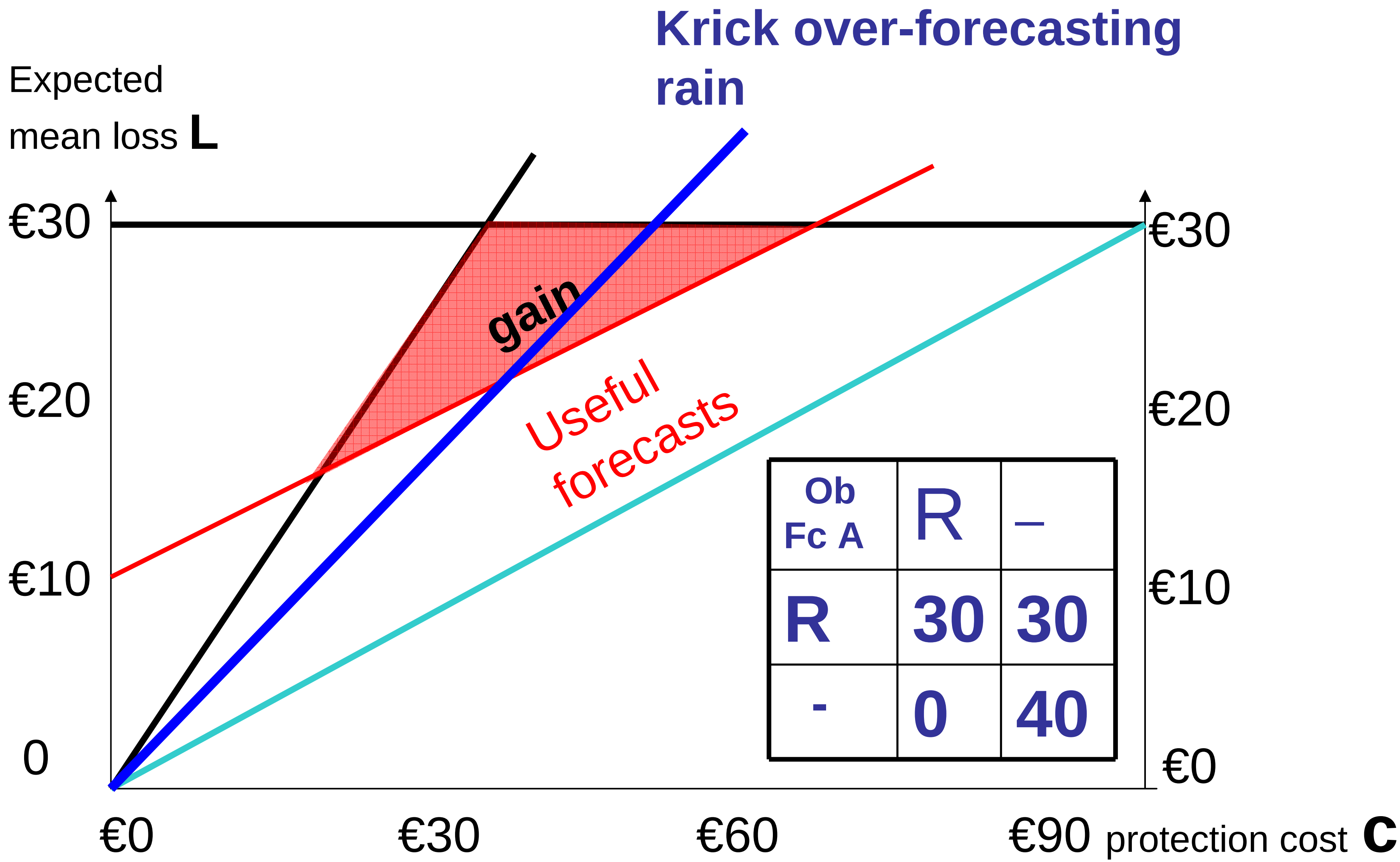
**High cost:** Spilling expensive water to lower the water levels to avoid over-filling or ability to adjust the prices.

**High loss:** Unplanned water spill or risk of damaging the dam in case of unpredicted rain.

| **Hollywood** Low cost/Loss | **Rain occurred** | **Staying dry** |
|---|---|---|
| **Rain forecast** Action: staying at home (cheap) | Staying at home (minor cost) | Missing a shooting (minor cost) |
| **Dry forecast** Action: take out expensive stars and equipment to the praire | No shooting (great economic loss) | |

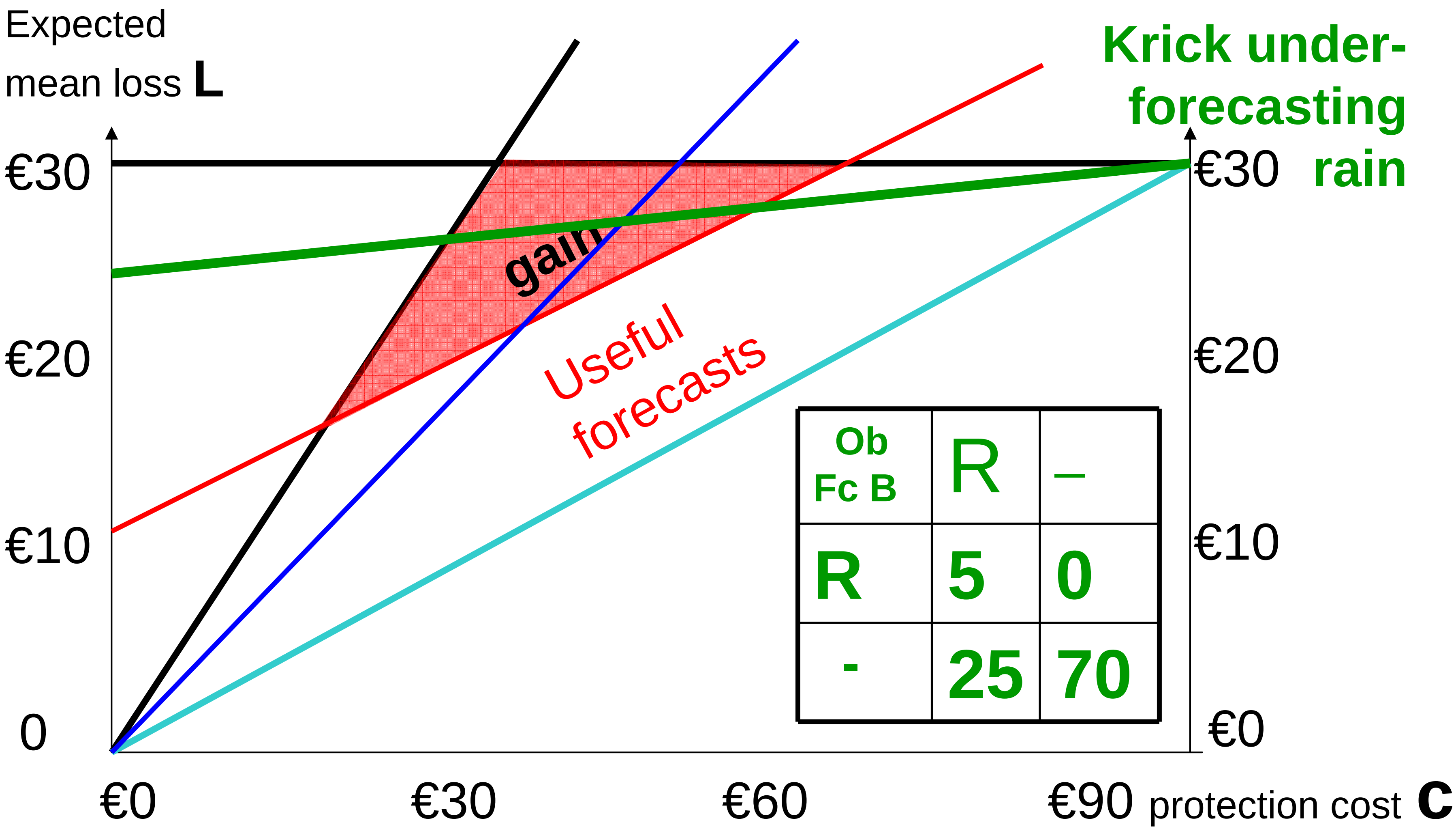
| Water indus.<br>High cost/Loss | Rain occurred | Staying dry |
|---|---|---|
| Rain forecast<br>Action: spilling expensive water | Not enough rain might fall | Unnecessary spill of expensive water |
| Dry forecast<br>Action: not spilling expensive water | Unforeseen damage | |

# The expected loss per day when Krick over-forecast rain



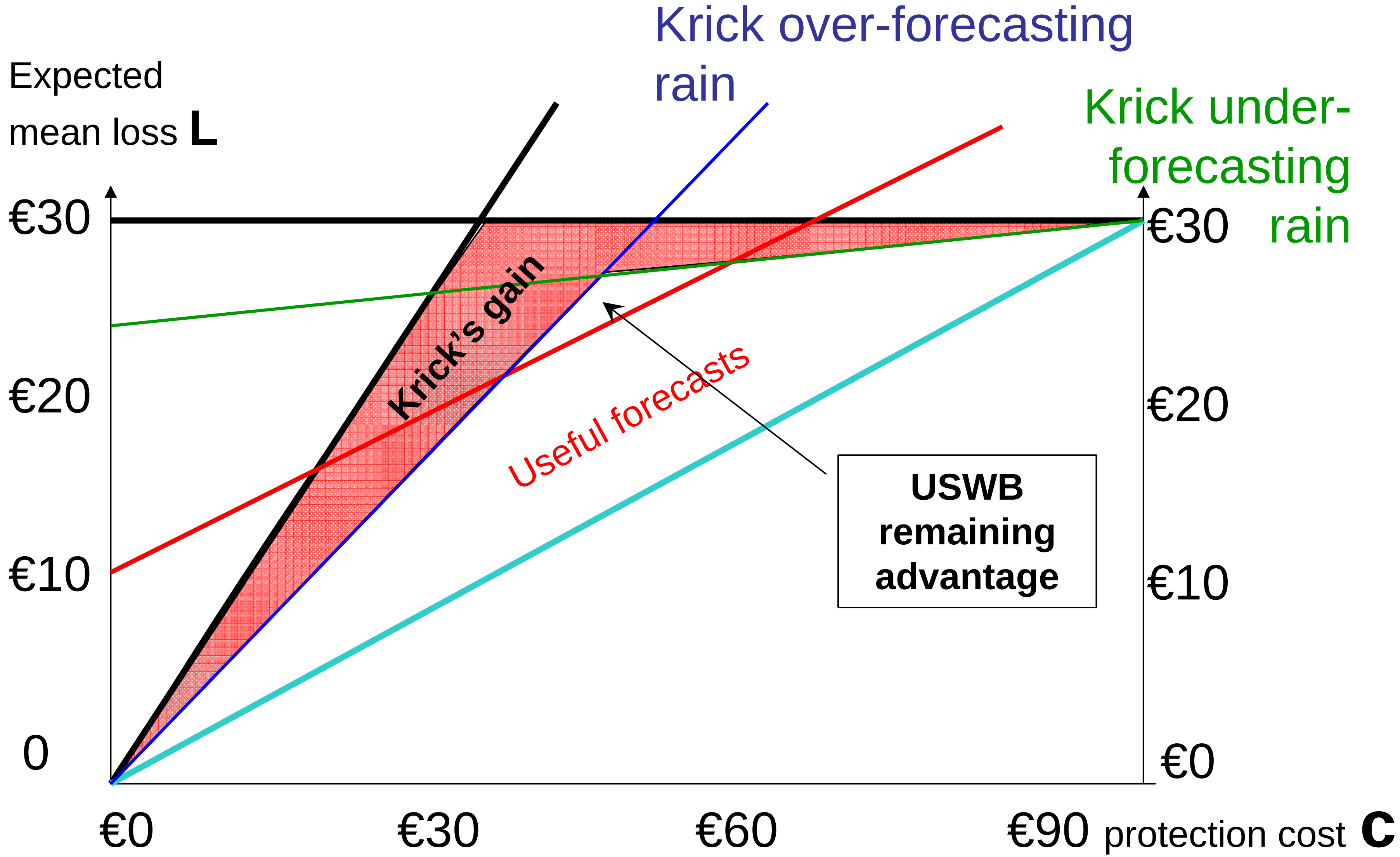

| Ob<br>Fc A | R | – |
|---|---|---|
| R | 30 | 30 |
| - | 0 | 40 |

# The expected loss per day when Krick under-forecast rain




| Ob<br>Fc B | R | – |
|---|---|---|
| R | 5 | 0 |
| - | 25 | 70 |

# The expected loss per day for different protection costs c



# IV.1.3. How the US Weather Bureau could have fought back using probabilities

# If the US Weather Bureau had chosen to become less categorical it could also have served *both* low and high cost-loss customers

| USWB | Obs rain | Obs dry |
|---|---|---|
| Fc rain | 20 | 10 |
| Fc dry | 10 | 60 |

| USWB | Obs rain | Obs dry | |
|---|---|---|---|
| Fc rain | 10 | 0 | |
| ??? | 20 | 20 | 50-50% |
| Fc dry | 0 | 50 | |

| Ob / Fc | R | – |
|---|---|---|
| R | 15 | 15 |
| - | 15 | 55 |

| Ob / Fc | R | – |
|---|---|---|
| R | 10 | 20 |
| - | 20 | 50 |

| Ob / Fc | R | – |
|---|---|---|
| R | 15 | 35 |
| - | 15 | 35 |

**Persistence**

| Ob / Fc | R | – |
|---|---|---|
| R | 7 | 0 |
| ?? | **23** | **40** |
| - | 0 | 30 |

**Climatology**

| Ob / Fc | R | – |
|---|---|---|
| R | 5 | 0 |
| ?? | **25** | **60** |
| - | 0 | 10 |

**Pure guess (always uncertain)**

| Ob / Fc | R | – |
|---|---|---|
| R | 0 | 0 |
| ?? | **30** | **70** |
| - | 0 | 00 |

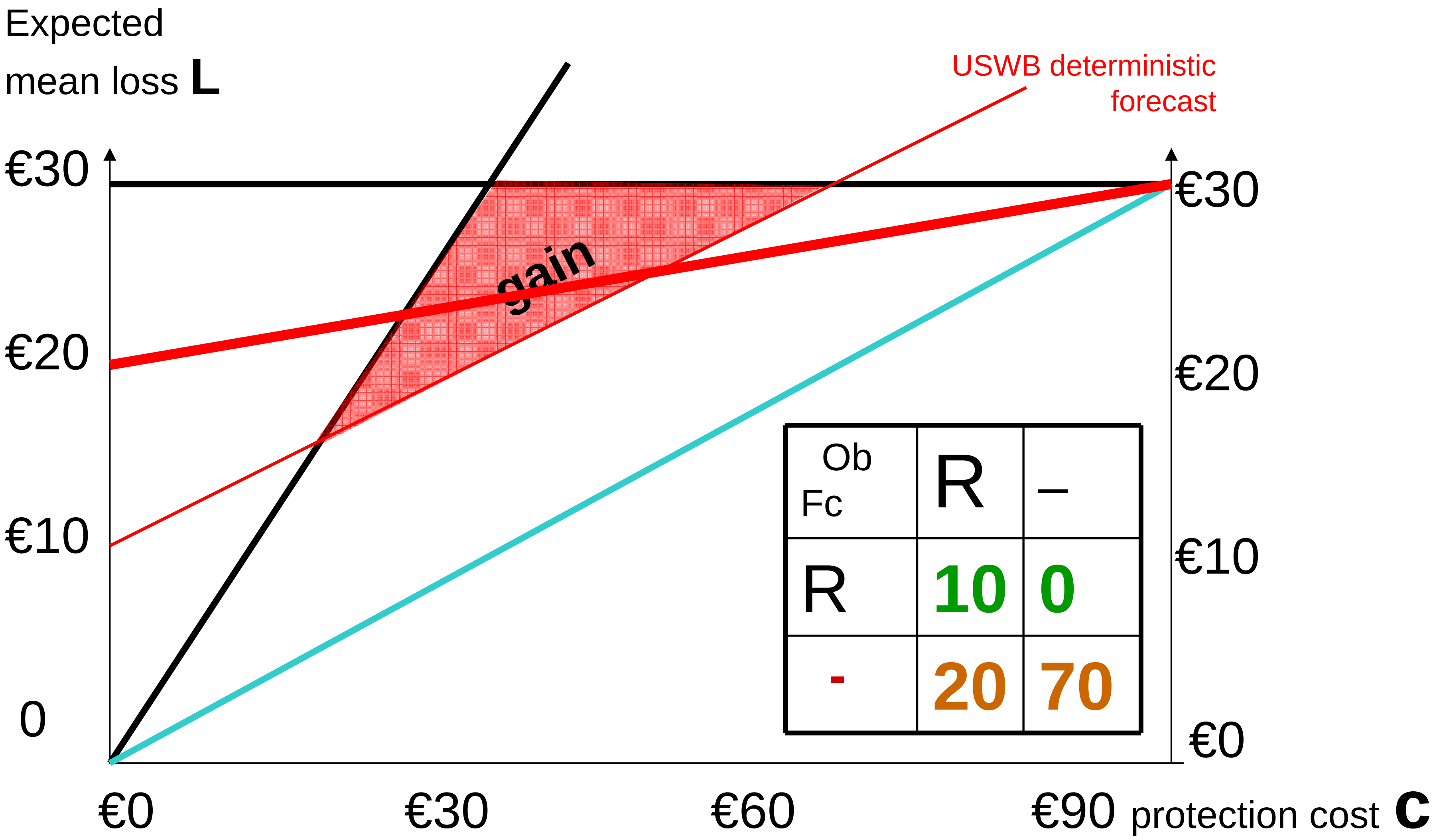
## It allows those who are <u>not</u> sensitive to rain to interpret the ??? as "it might <u>not</u> rain"

| USWB | Obs rain | Obs dry |
|---|---|---|
| Fc rain | 10 | 0 |
| ??? | 20 | 20 |
| Fc dry | 0 | 50 |

→

| USWB | Obs rain | Obs dry |
|---|---|---|
| Fc rain | 10 | 0 |
| Fc dry | 20 | 70 |

**These are the EMV (total cost) for those who interpreted ??? as “it might not rain”**



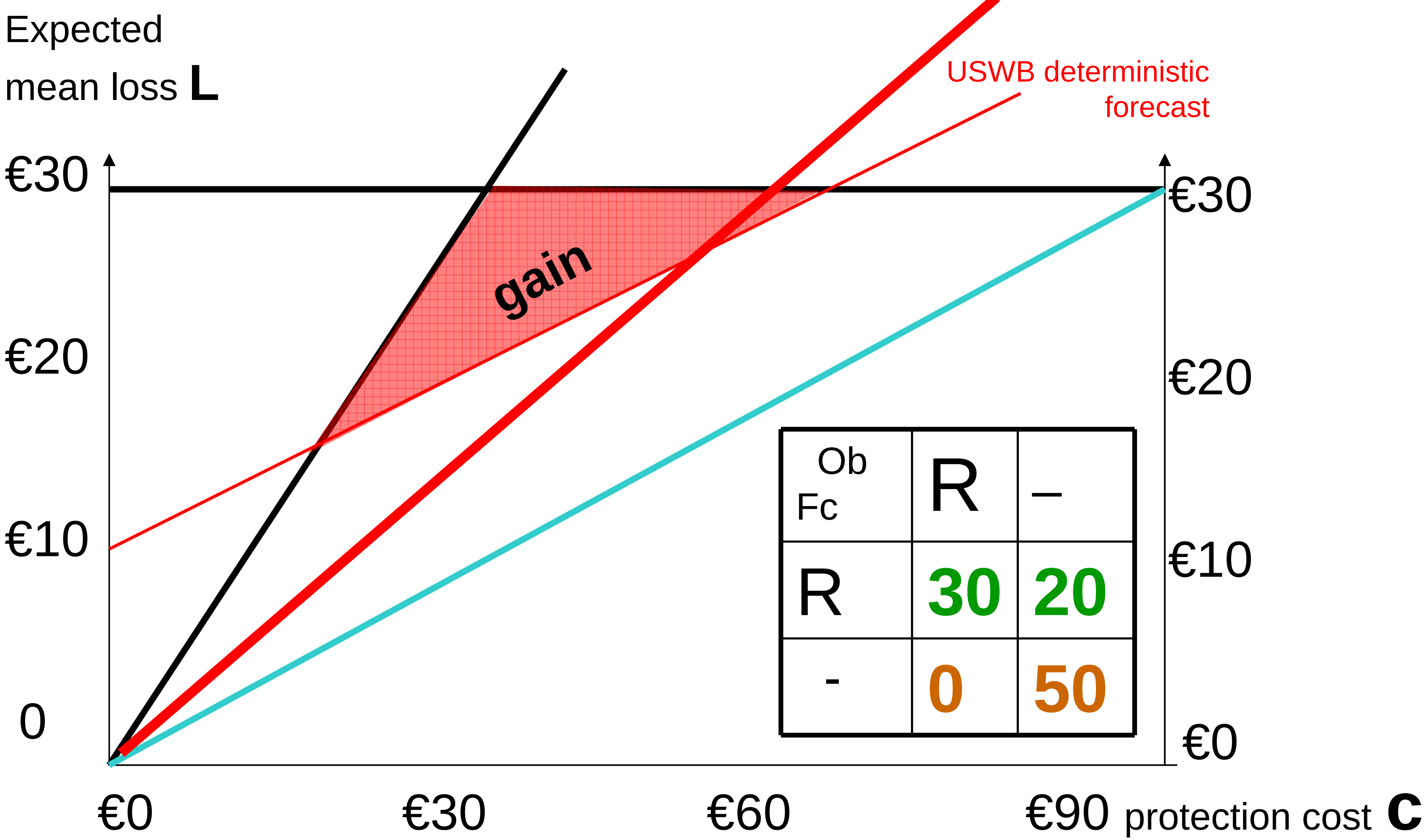
# It allows those who are sensitive to rain to interpret the ??? as "it might rain"

| USWB | Obs rain | Obs dry |
|---|---|---|
| Fc rain | 10 | 0 |
| ??? | 20 | 20 |
| Fc dry | 0 | 50 |

→

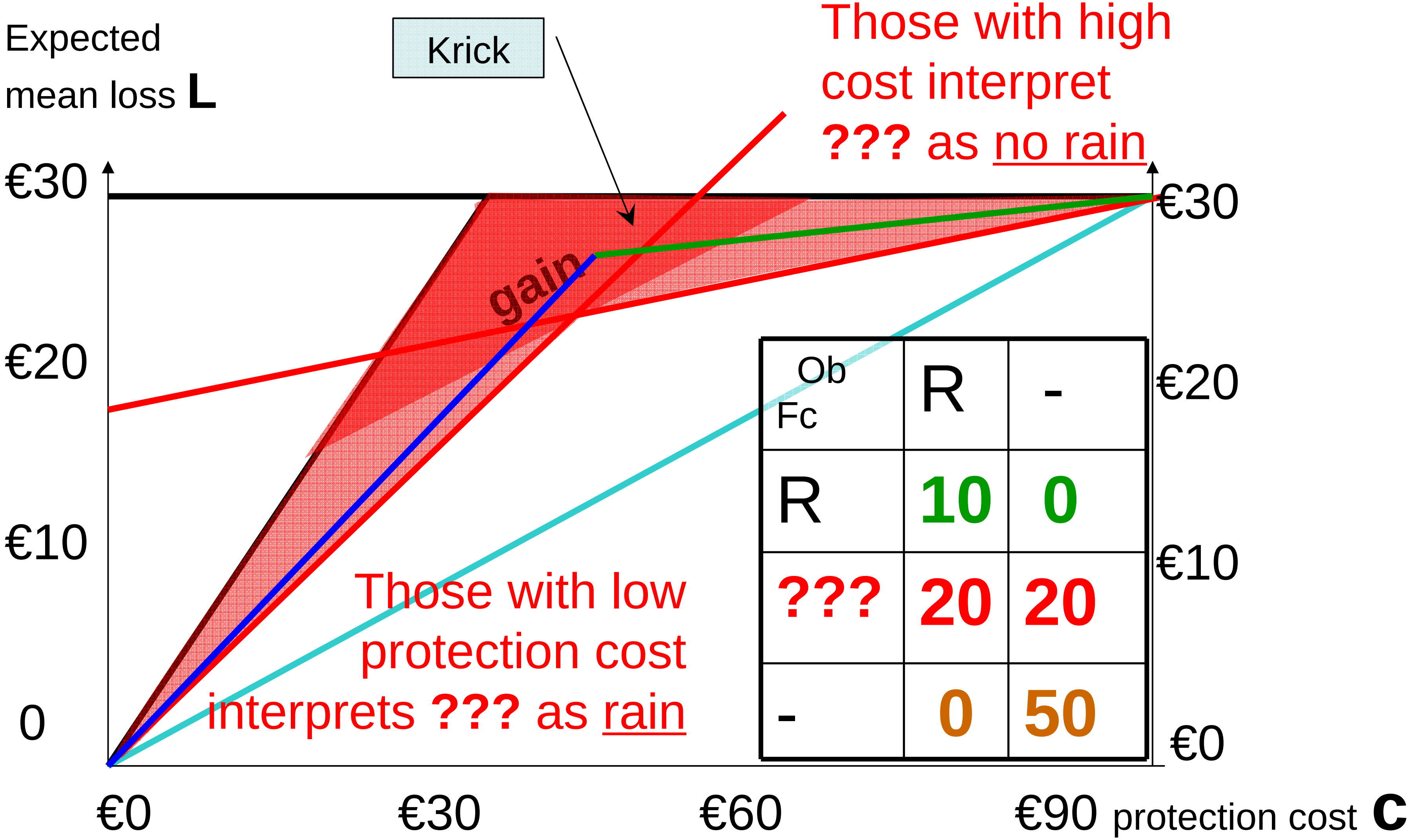
| USWB | Obs rain | Obs dry |
|---|---|---|
| Fc rain | 30 | 20 |
| Fc dry | 0 | 50 |

# These are the EMV (total cost) for those who interpreted ??? as "it might rain"




| Ob / Fc | R | – |
|---|---|---|
| R | 30 | 20 |
| - | 0 | 50 |

# And them put them together . . .




| Ob / Fc | R | - |
|---|---|---|
| R | 10 | 0 |
| ??? | 20 | 20 |
| - | 0 | 50 |

## I repeat:

Categorical

| Ob / Fc | R | – |
|---|---|---|
| R | 20 | 10 |
| - | 10 | 60 |

Non-categorical

| **Ob** / **Fc** | R | - |
|---|---|---|
| R | 10 | 0 |
| ?? | **20** | **20** |
| - | 0 | 50 |

This is the matrix for those

| Ob / Fc | R | – |
|---|---|---|
| R | 30 | 20 |
| - | 0 | 50 |

with <u>low</u> protection cost

This is the matrix for those

| Ob / Fc | R | – |
|---|---|---|
| R | 10 | 0 |
| - | 20 | 70 |

with <u>high</u> protection cost

# Which ones of the 40 forecasts are more or less certain or uncertain?

Categorical

| Obs / Fc | R | – |
|---|---|---|
| R | 20 | 10 |
| - | 10 | 60 |

Non-categorical

| Obs / Fc | R | - |
|---|---|---|
| R | 10 | 0 |
| ??? | 20 | 20 |
| - | 0 | 50 |

Probabilistic

| Obs / Confidence | R | – |
|---|---|---|
| certain | 10 | 0 |
| almost certain | 8 | 2 |
| rather certain | 6 | 4 |
| rather uncertain | 4 | 6 |
| very uncertain | 2 | 8 |
| certain | 0 | 50 |

**Can we quantify that uncertainty?**

# Or with probability numbers

Categorical

| Obs<br>Fc | R | – |
|---|---|---|
| R | 20 | 10 |
| - | 10 | 60 |

Non-categorical

| Obs<br>Fc | R | - |
|---|---|---|
| R | 10 | 0 |
| ??? | 20 | 20 |
| - | 0 | 50 |

Probabilistic

| Obs<br>Prob% | R | – |
|---|---|---|
| 100 | 10 | 0 |
| 80 | 8 | 2 |
| 60 | 6 | 4 |
| 40 | 4 | 6 |
| 20 | 2 | 8 |
| 0 | 0 | 50 |

# IV.1.4. And now to the practise

## How the USWB could gave swept the floor with Krick's private weather service if they had realised to potential of probabilistic forecasts

# Decision matrix for different people when P=100%

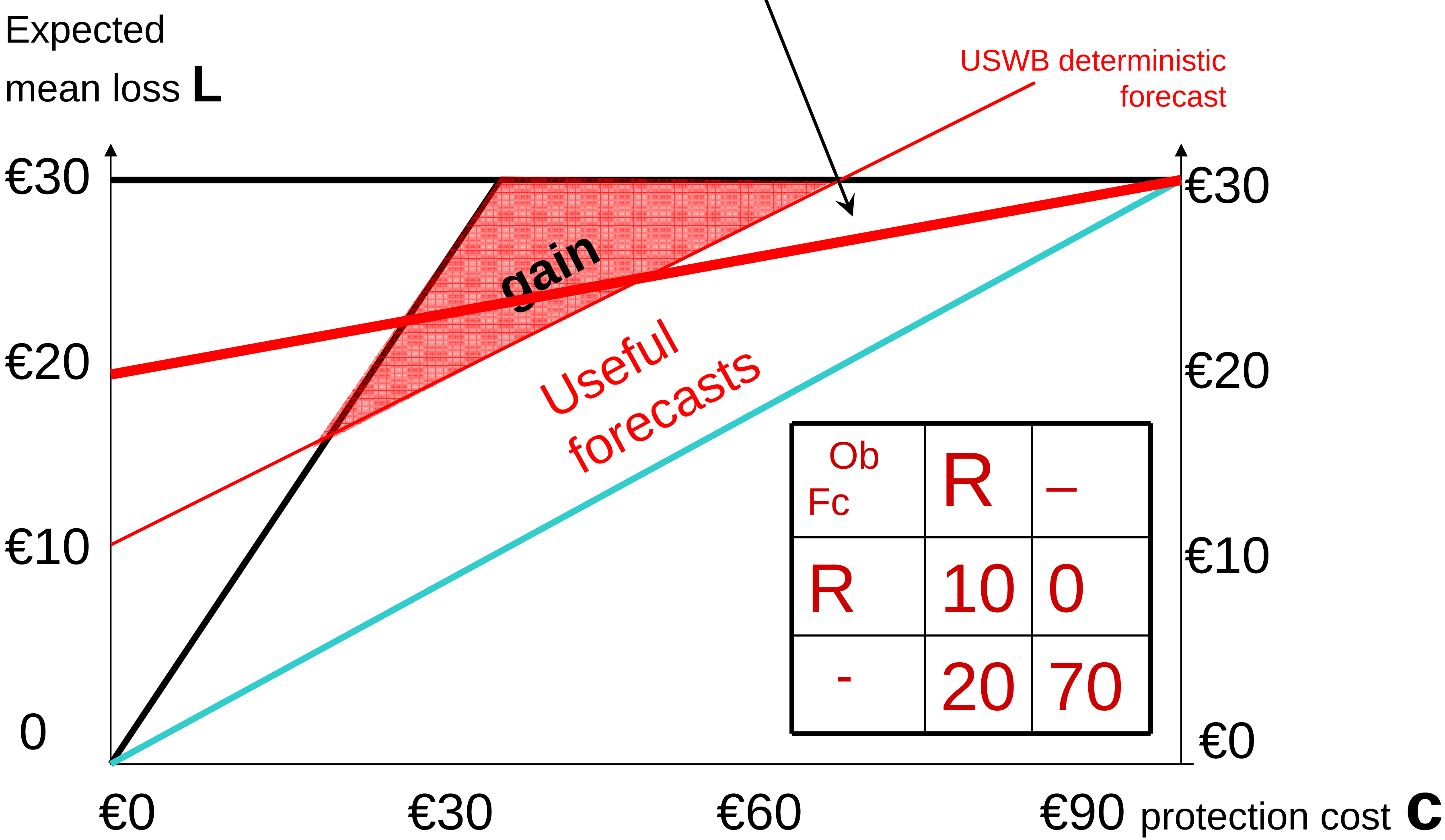
| Ob<br>Prob | R | – |
|---|---|---|
| 100 | 10 | 0 |
| 80 | 8 | 2 |
| 60 | 6 | 4 |
| 40 | 4 | 6 |
| 20 | 2 | 8 |
| 0 | 0 | 50 |

| Ob<br>Fc | R | – |
|---|---|---|
| R | 10 | 0 |
| - | 20 | 70 |

Decision matrix

Probability matrix

# Gains for people with c/L almost 100%



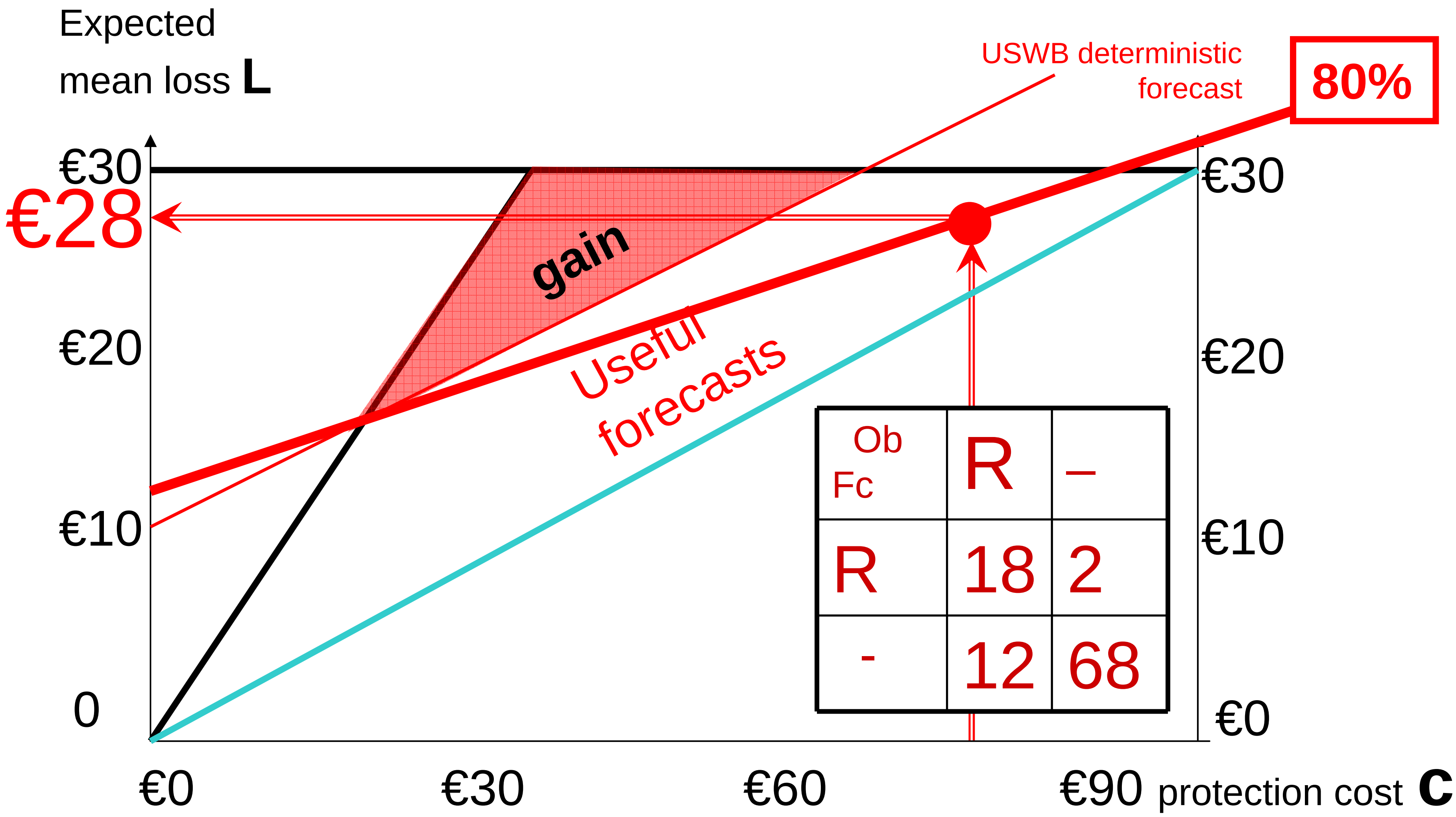
# Decision matrix for people with c/L around 80%

| Ob Prob | R | – |
|---|---|---|
| 100 | 10 | 0 |
| 80 | 8 | 2 |
| 60 | 6 | 4 |
| 40 | 4 | 6 |
| 20 | 2 | 8 |
| 0 | 0 | 50 |

→

| Ob Fc | R | – |
|---|---|---|
| R | 18 | 2 |
| - | 12 | 68 |

# Gains for people with c/L around 80%




| Ob / Fc | R | – |
|---|---|---|
| R | 18 | 2 |
| - | 12 | 68 |

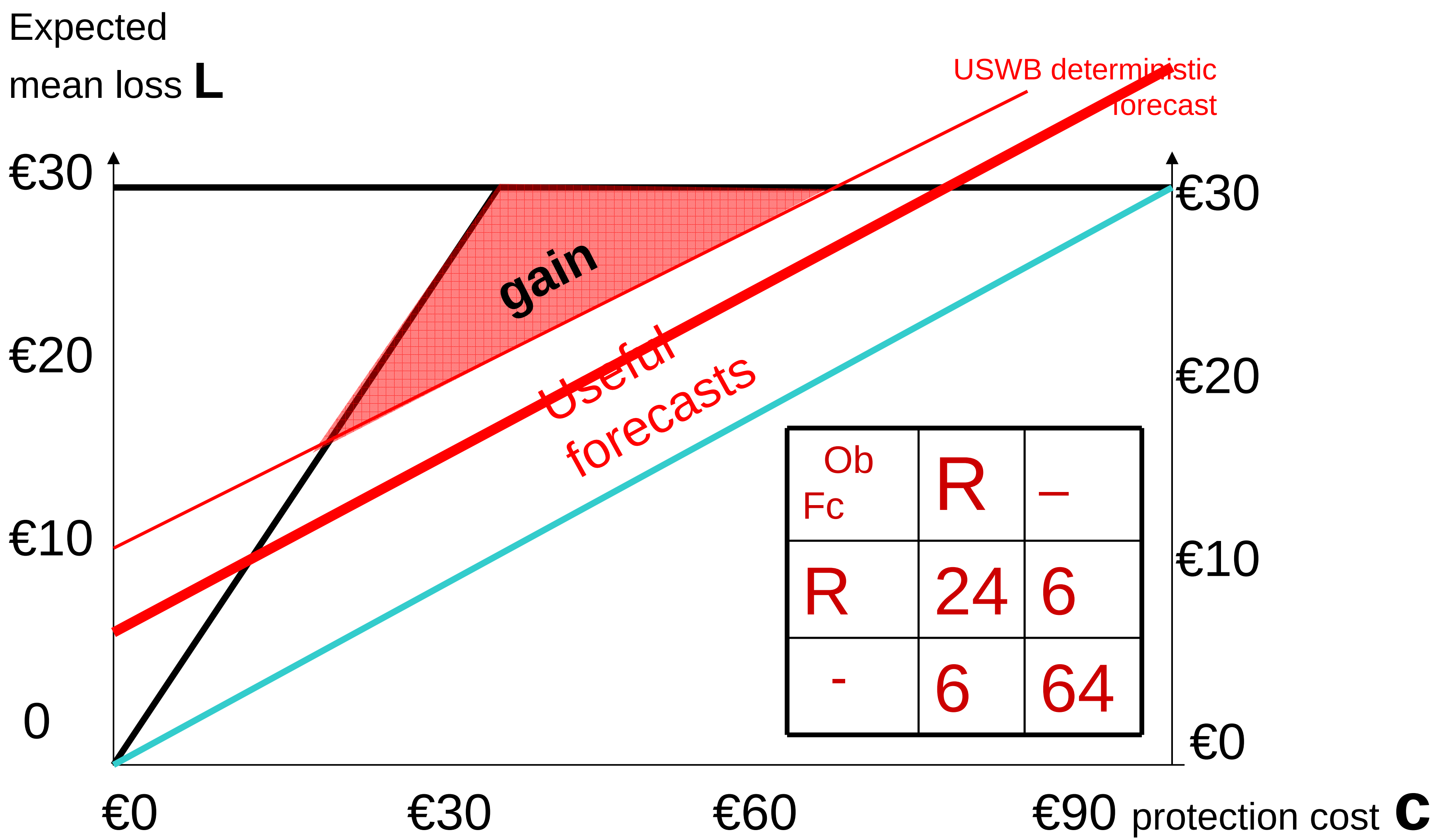
# Decision matrix for people with c/L around 60%

| Ob<br>Prob | R | – |
|---|---|---|
| 100 | 10 | 0 |
| 80 | 8 | 2 |
| 60 | 6 | 4 |
| 40 | 4 | 6 |
| 20 | 2 | 8 |
| 0 | 0 | 50 |

→

| Ob<br>Fc | R | – |
|---|---|---|
| R | 24 | 6 |
| - | 6 | 64 |

# Gains for different people when P = 60%



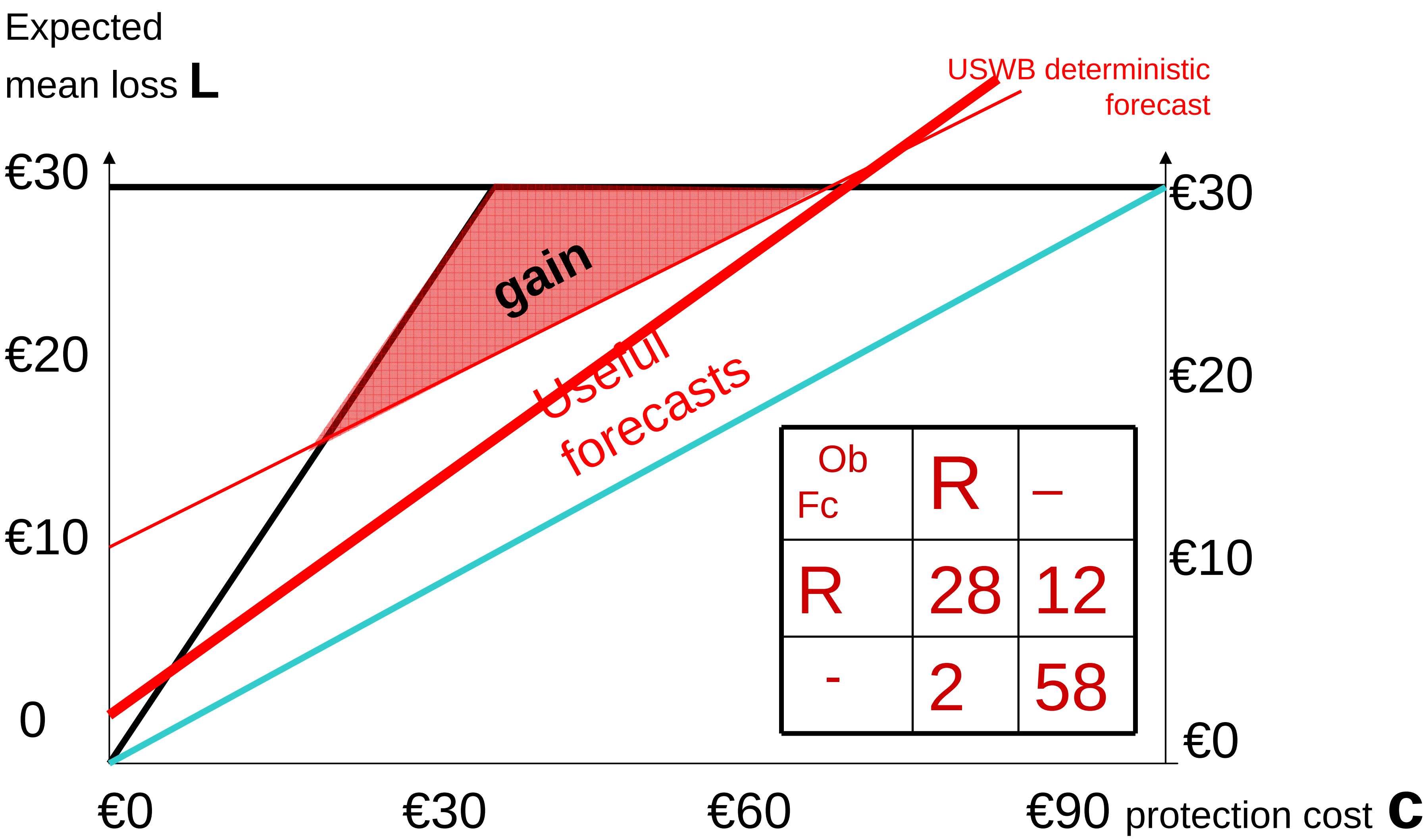
# Decision matrix for people with c/L around 40%

| Ob Prob | R | – |
|---|---|---|
| 100 | 10 | 0 |
| 80 | 8 | 2 |
| 60 | 6 | 4 |
| 40 | 4 | 6 |
| 20 | 2 | 8 |
| 0 | 0 | 50 |

→

| Ob Fc | R | – |
|---|---|---|
| R | 28 | 12 |
| - | 2 | 58 |

# Gains for people with c/L around 60%




| Ob / Fc | R | – |
|---|---|---|
| R | 28 | 12 |
| - | 2 | 58 |

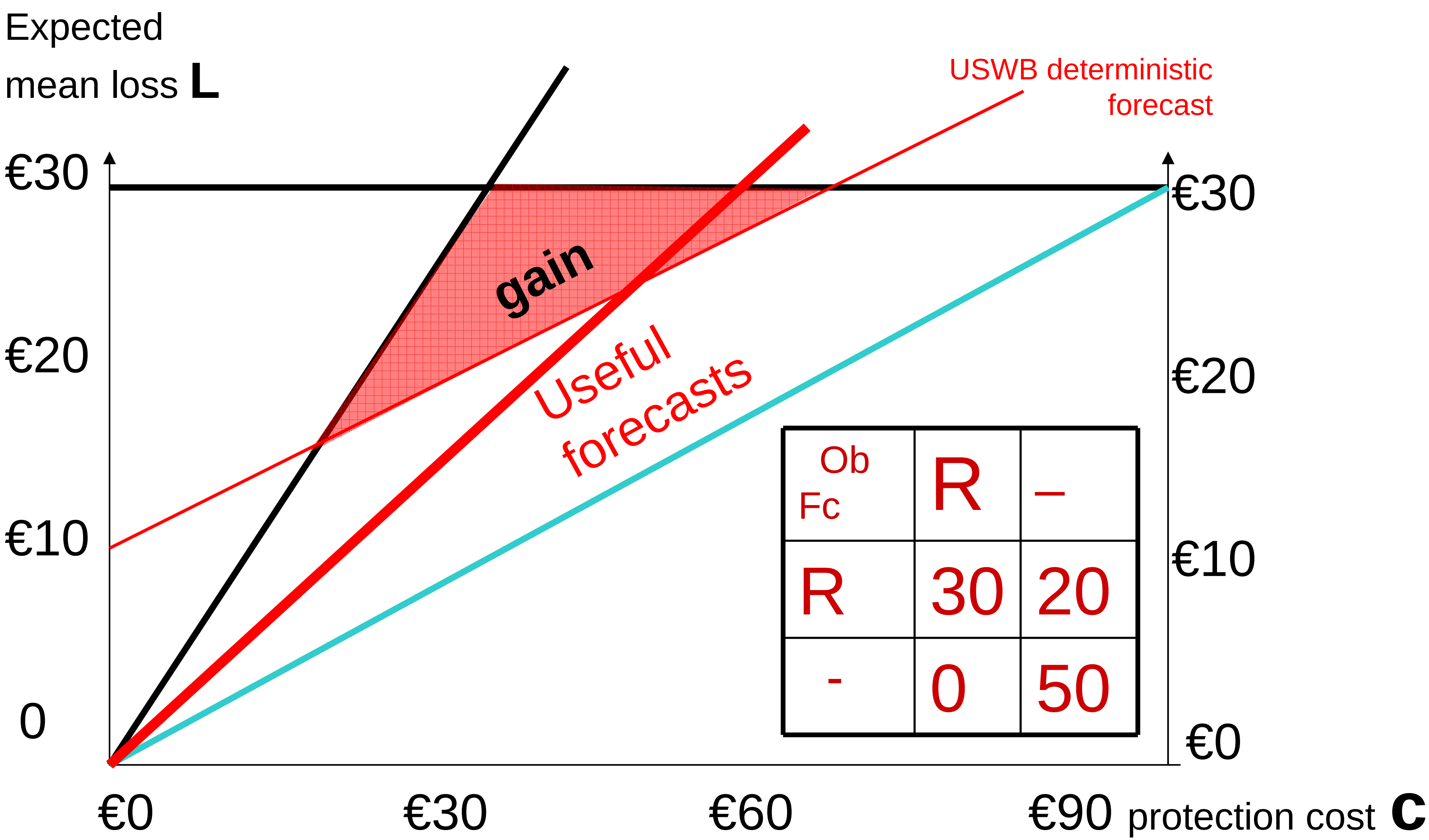
# Decision matrix for people with c/L around 20%

| Ob Prob | R | – |
|---|---|---|
| 100 | 10 | 0 |
| 80 | 8 | 2 |
| 60 | 6 | 4 |
| 40 | 4 | 6 |
| 20 | 2 | 8 |
| 0 | 0 | 50 |

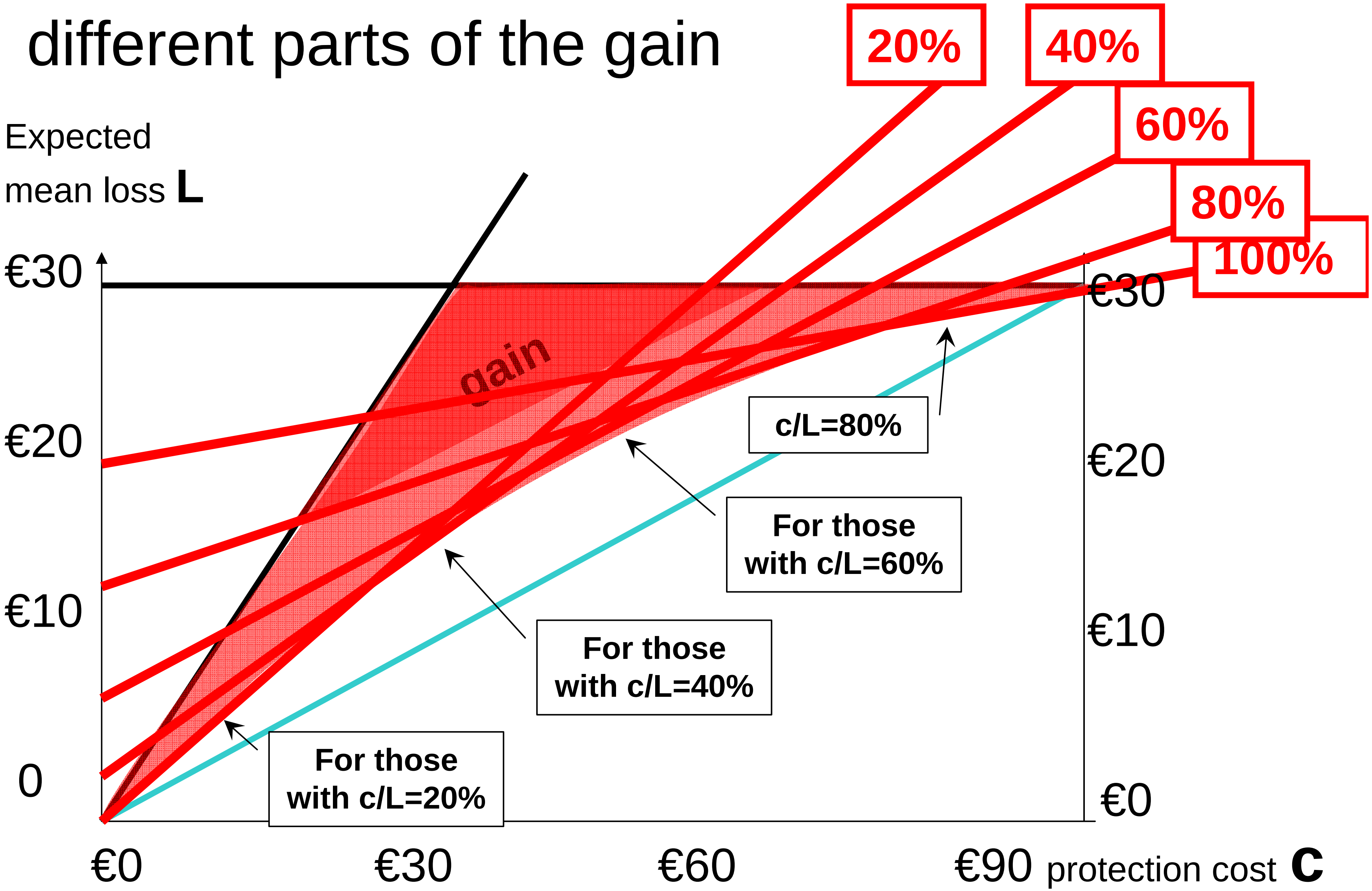
→

| Ob Fc | R | – |
|---|---|---|
| R | 30 | 20 |
| - | 0 | 50 |

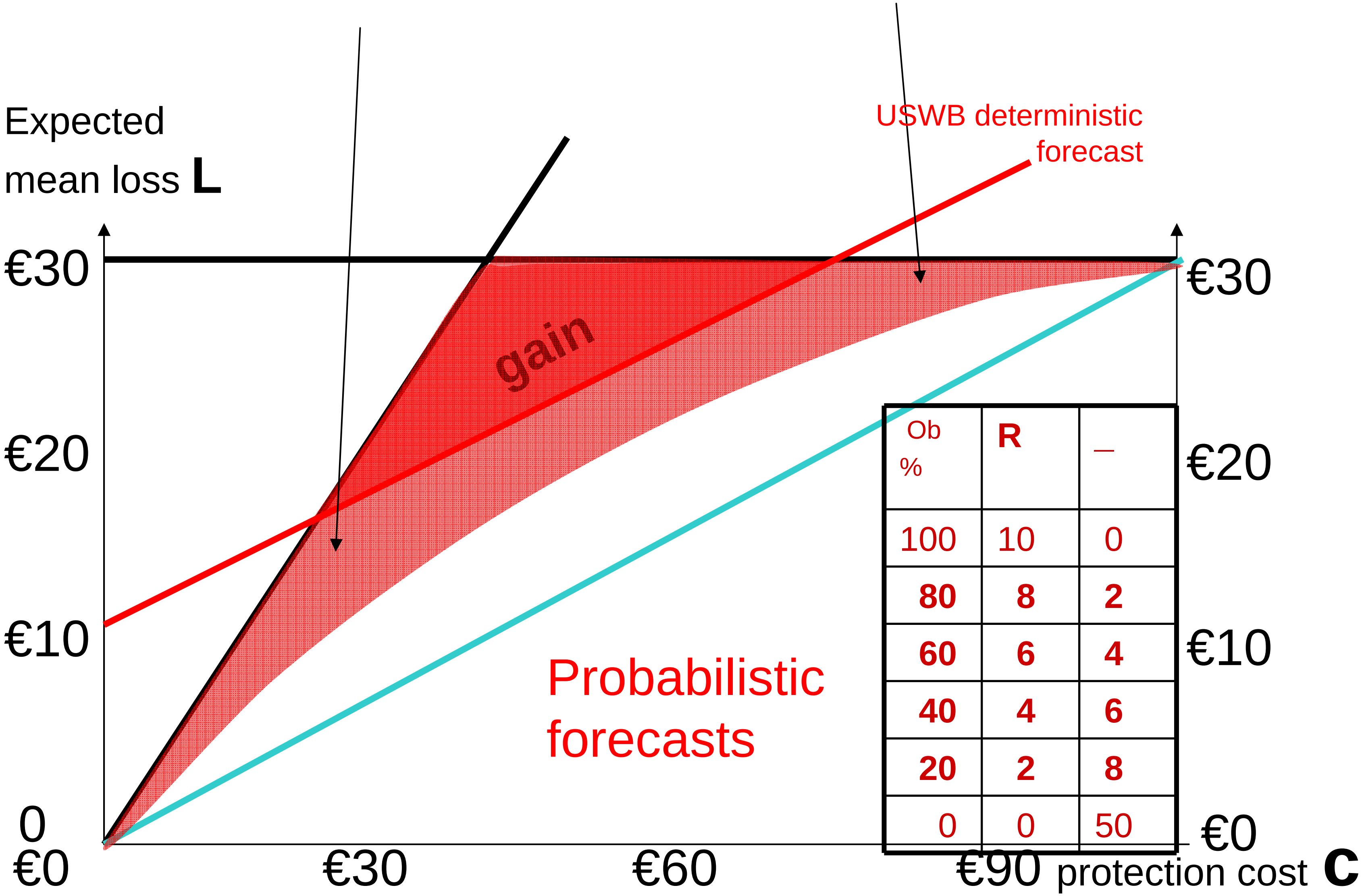
# Gains for people with c/L around 20%



# Different users benefit from different parts of the gain



# Probabilities yield gains for all possible protection costs




| Ob % | R | — |
|---|---|---|
| 100 | 10 | 0 |
| 80 | 8 | 2 |
| 60 | 6 | 4 |
| 40 | 4 | 6 |
| 20 | 2 | 8 |
| 0 | 0 | 50 |

END