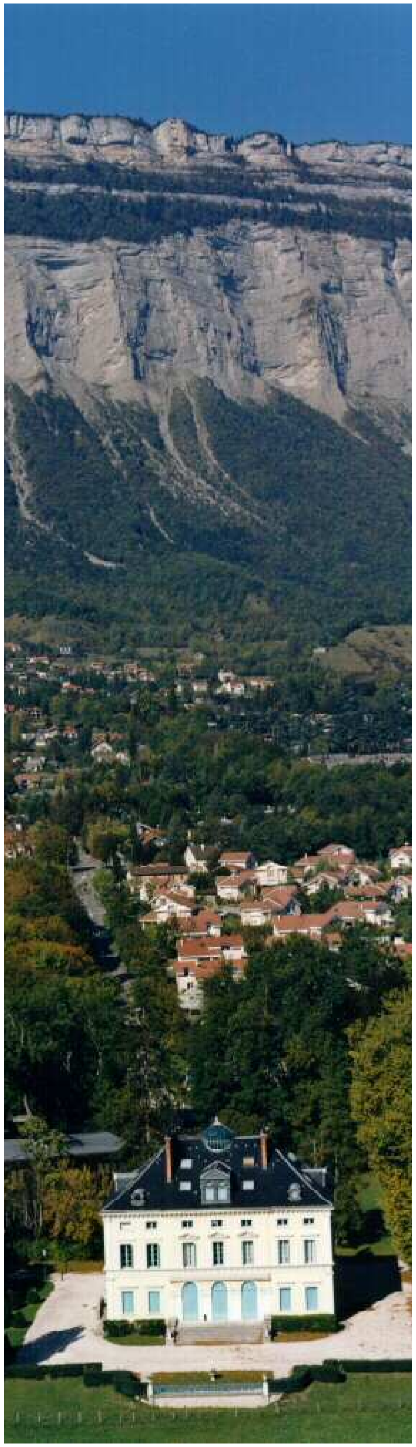
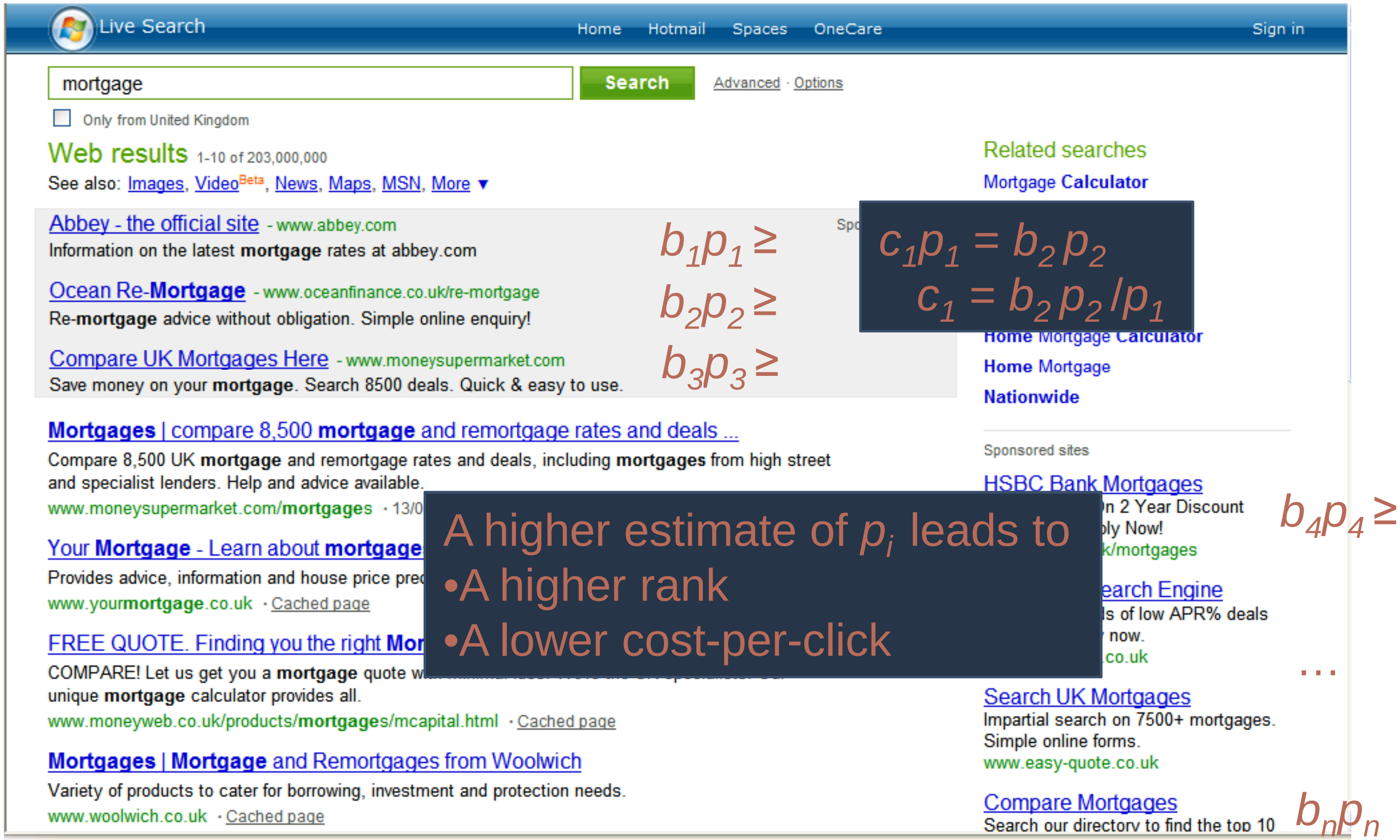
Learning optimally from self-interested data sources in on-line ad auctions

Onno Zoeter* and **Chris Dance**
Xerox Research Centre Europe

* = large part of this work done while at
Microsoft Research Cambridge



Paid search and the GSP auction



$b_1 p_1 \geq$

$b_2 p_2 \geq$

$b_3 p_3 \geq$

$b_4 p_4 \geq$

...

$b_n p_n$

$c_1 p_1 = b_2 p_2$

$c_1 = b_2 p_2 / p_1$

A higher estimate of p_i leads to

- A higher rank
- A lower cost-per-click

Machine learning and the GSP auction

Bids b_i are observed, click probabilities p_i are not, and need to be learned.

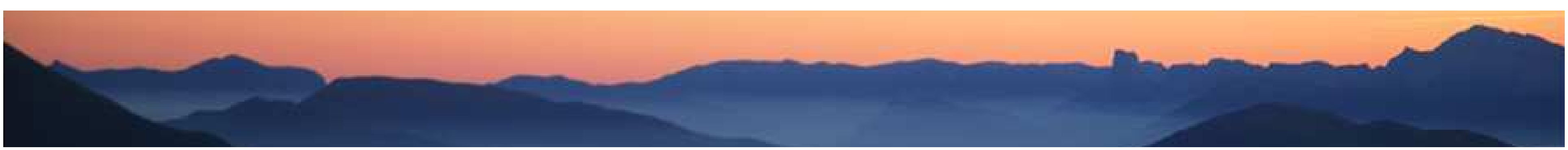
Treating them as known or as point estimates leads to serious sub-optimalities.

Single round approximation in GSP

When p_i are unknown optimal ad placement involves exploring new ads.

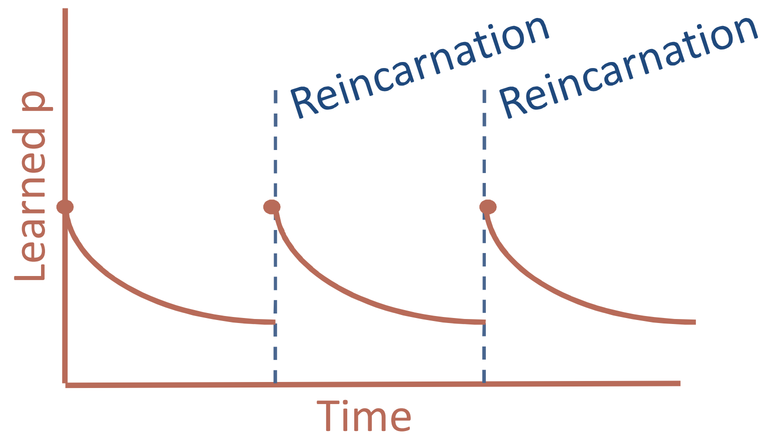
Example: current winner has b_1=1, p_1=.06. Shouldn't we try a new ad with b_2=1, p_2=.05?

Problem 1: GSP only cares for revenue in ***this*** round. It is not optimal for multiple rounds.

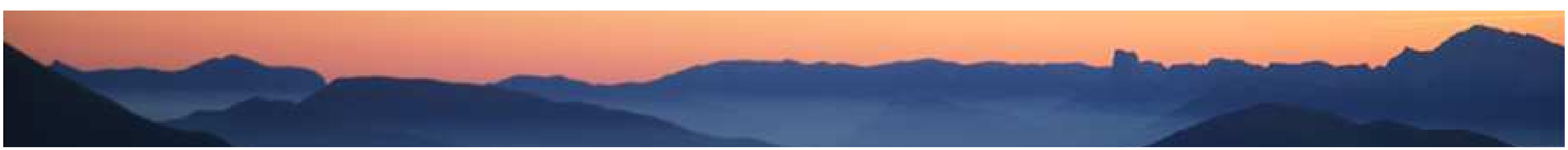


Advertiser cheating in GSP auctions

Problem 2: once estimate $\widehat{p}_i$ falls below initial estimate for a new ad, it is optimal for advertiser *i* to reincarnate.



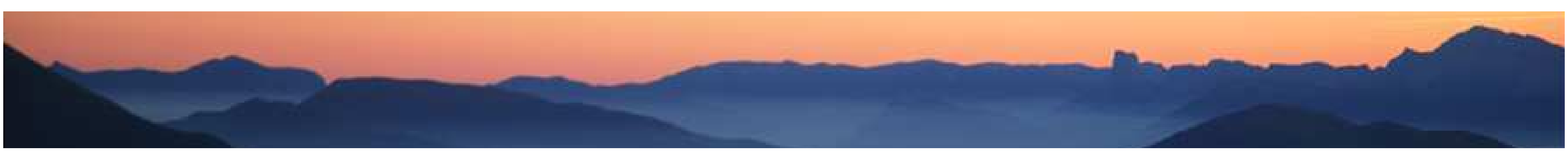
For now this needs to be policed.



An alternative paid search auction

Question: is there an alternative paid search auction where reincarnating is not beneficial? Where ads are optimally explored?

This talk: **yes!** We discuss an application of the dynamic-VCG mechanism where advertisers have to submit a bid and a belief.

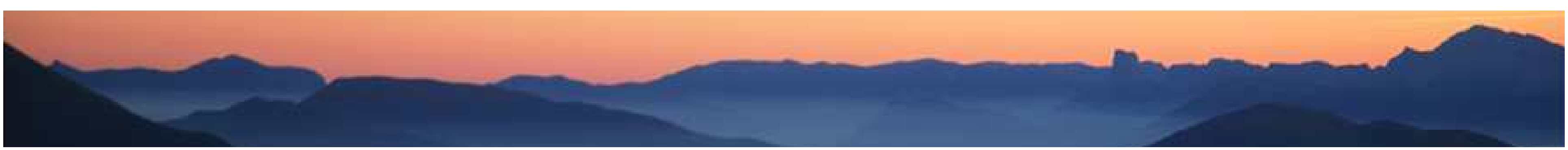


Optimal placements with uncertain p_i: multi-armed bandit problems

Optimal ad placement involves exploring new ads.

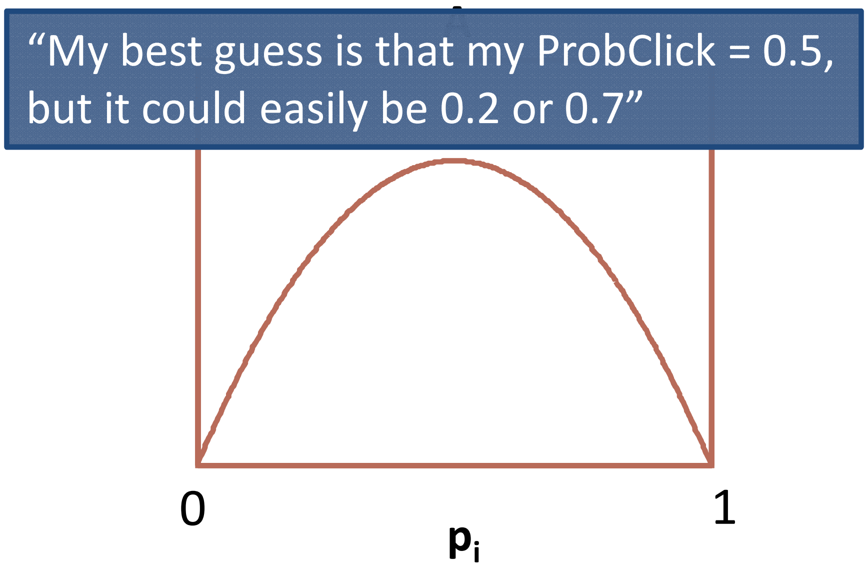
We can interpret it as a reinforcement learning problem.

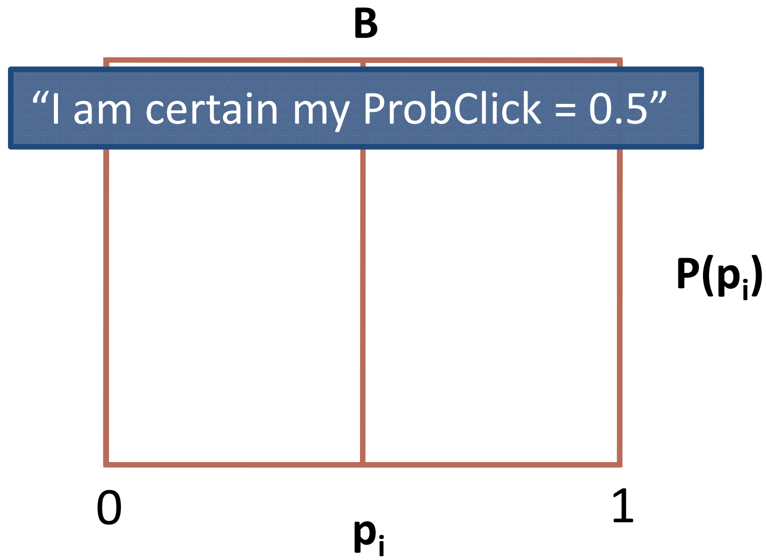
For a single slot, this is a classic multi-armed bandit problem.



A multi-armed bandit example

Two ads, same $E[p_i]$, different $Var[p_i]$.





Not only best guess for p_i matters. Need to maintain belief.

You get two rounds, do you place A or B?

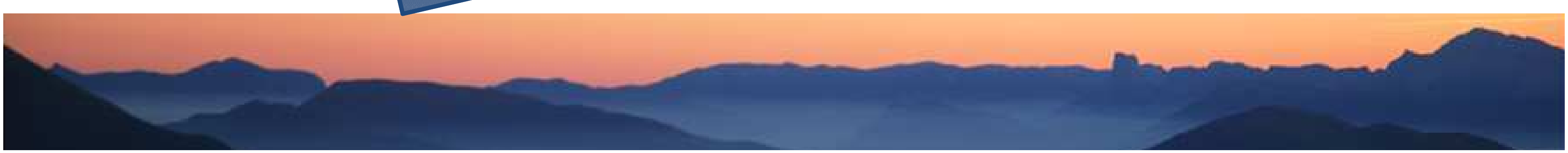
You get one round, do you place A or B?

Bergemann and Välimäki's dynamic-VCG mechanism

At each round

1. Agents report their private state.
2. The centre acts optimally according to the reported state.
3. Agent *a* pays for this round the reduction in utility for the other agents that his presence in this round implies (*externality*).

The multi-round extension of the Vickrey-Clarke-Groves mechanism



Properties of dynamic-VCG

Important properties of dynamic-VCG:

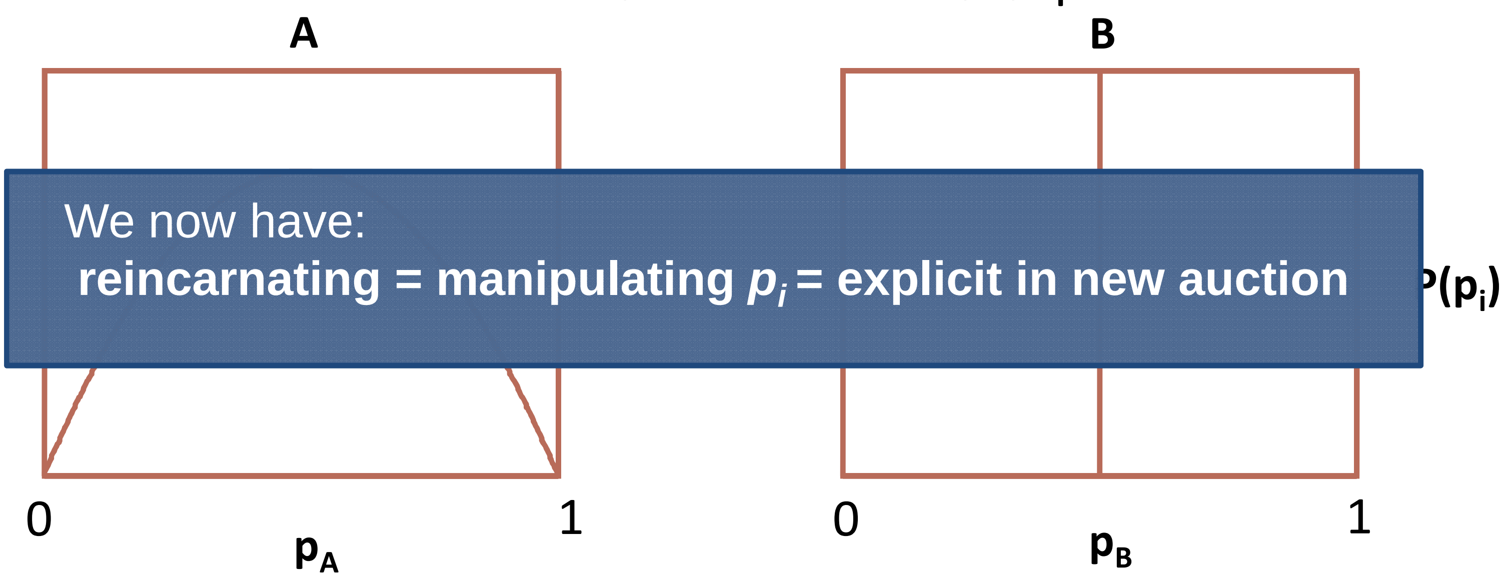
Efficiency: the assignment maximizes expected multi-round total utility.

Truth-telling: the situation where all agents report true values for a click and true beliefs is an equilibrium

Making reincarnation a legitimate part of the auction

We require advertisers to submit

- A bid
- A belief over their probability p_i



We now have:
reincarnating = manipulating p_i = explicit in new auction

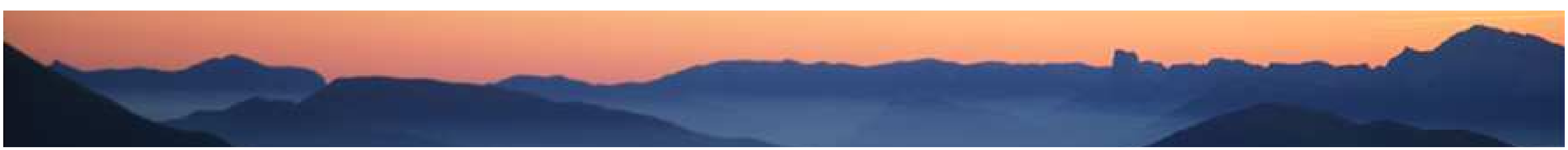
A new single slot ad auction

- At t=1 centre provides default CTR beliefs
- At each round t=1,…
  1. Advertisers submit bid and optionally override CTR beliefs.
  2. Centre places ad that maximizes multi-round return to advertisers.
  3. Advertiser pays externality for this round (i.e. even if not clicked)
  4. Centre uses Bayes rule to update beliefs

Properties of the new ad auction

1. The p_i's are learned at the optimal rate.
2. The placement is optimal over all rounds.
3. Reincarnation is not beneficial.
   - Advertiser can report belief in every round.
   - In particular belief associated with fresh ad.
   - But reporting truth is optimal.

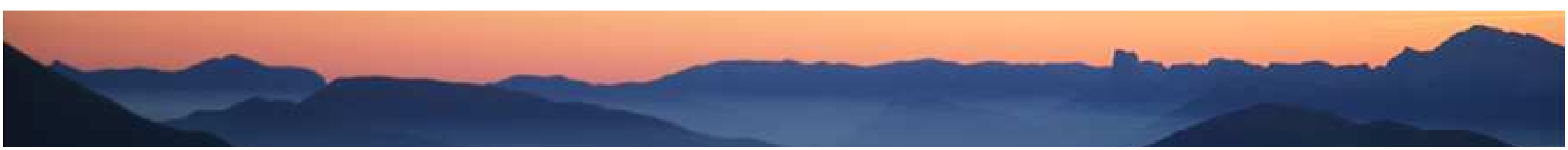
This auction gives "loyalty benefits" at exactly the right rate



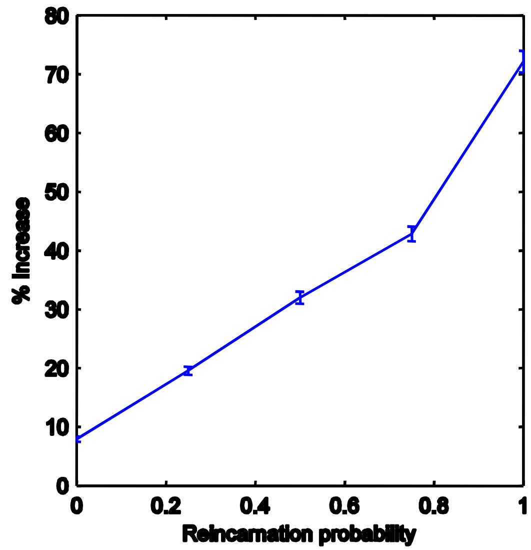
Truth-telling intuition

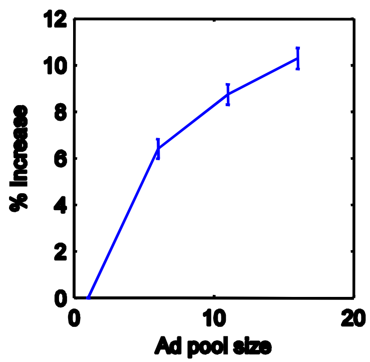
Externality pricing rule → pay-per-impression

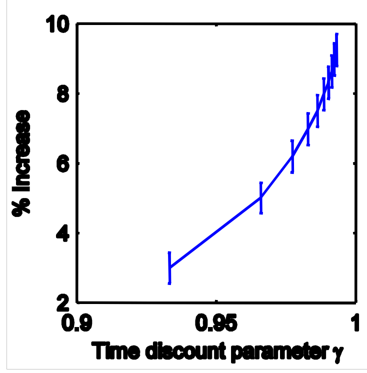
- Optimistic p_i: too many costs-per-impressions.
- Pessimistic p_i: lost opportunities.
- Similarly sub-optimal to be too confident/uncertain about beliefs.

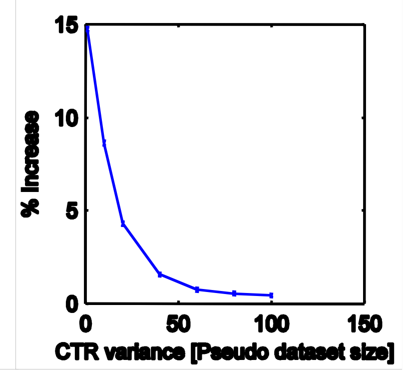


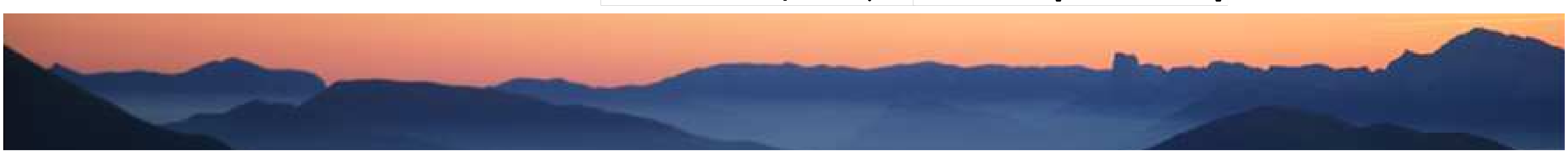
Experiments











Machine Learning and Incentives

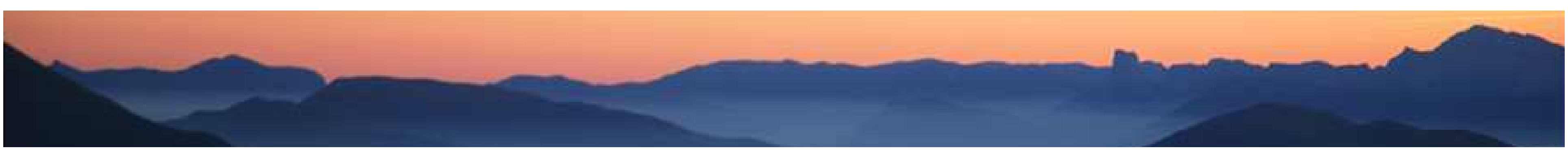
Learning a model in biology, medicine, etc.

≠

Learning a model of humans on the web.

Incentives form an important dimension often overlooked in Machine Learning.

Trick here is relatively widely applicable.



Questions?