Better Hedging Through Deep Learning

Using Machine Learning to Rebuild Derivative Pricing

Mark Higgins
COO, Beacon Platform Inc

Introduction

- We'll look at the original "Deep Hedging" work by JPMorgan's Quantitative Research Team ("Deep Hedging: Hans Buehler, Lukas Gonon, Josef Teichmann, and Ben Wood")
  - https://arxiv.org/abs/1802.03042

- Then extend to hedging the right to use a natural gas storage facility

- Joint work by Kinetic Mind and Beacon Platform
  - Kinetic Mind: Dr Kay Pilz – expert consultants in the application of machine learning to finance
  - Beacon Platform: Mark Higgins, Neil Palmer, Ben Pryke, and Stephen Dacek – enterprise innovation platform vendor

Classical Dynamic Hedging: from Theory to Practice

Black Scholes gave us a strong <u>but idealistic</u> foundation for dynamic hedging:

→ Perfect freedom to trade the underlying hedge instruments without transaction costs, and continuous rebalancing
→ Defines a "risk neutral" measure which transforms stochastic processes to reproduce market forward prices
→ Perfect hedging possible, so post-hedge PNL distribution is a delta function

But what about…

→ Trading costs (high bid-ask spreads, or varying bid-ask spreads across related markets)
→ Hedging constraints (frequency, notional limitations)
→ Unhedgeable risks (incomplete markets, non-market risks)

And how to balance cost vs residual risk…

→ Do I hedge with this instrument or that one?
→ What are my hedging objectives when PNL variance is non-zero?

Classical Hedging → Deep Hedging

- Many finance academics and practitioners have focused on these problems in the last few decades
- Some progress in separate areas but hard to bring it all together
- Problem: huge computational complexity with real markets & complex portfolios
  → Markets move fast, traders need real-time guidance
- But now:

COMPUTE + **ANALYTICS** + **MACHINE LEARNING** = **BETTER RISK MANAGEMENT**

COMPUTE	ANALYTICS	MACHINE LEARNING
Cloud-based Practically unlimited Pay-as-you-go	*Flexible platforms linked to the business*	*Developments in approaches to neural networks*

Introducing Deep Hedging

- First practical case (that I know of!) where neural networks can be applied to front office derivative pricing problems
- Feb 2018 paper by JP Morgan Quantitative Research team: Deep Hedging (https://arxiv.org/abs/1802.03042)
- Assumes there's a nonlinear function which, given the market conditions and existing hedge portfolio, tells you new hedge trades to do
- Look at distribution of post-hedges lifetime PNL of the position
- Choose a convex risk measure of that PNL distribution (eg expected shortfall) to minimize
  - Use a neural network for the nonlinear function
  - Train it by minimizing that convex risk measure

The Benefits of Deep Hedging

- Some interesting consequences:
  - Don't need to use risk neutral paths – can be real world measure
  - Reproduces standard Black-Scholes-like results (including extensions to stochastic vol etc) in the limit of continuous hedging, no transaction costs or size constraints, and no unhedgeable risks
  - Can use any realistic market model you want: transaction costs, unhedgeable risks, infrequent hedging, hedge size constraints, etc

- JPM QR's Deep Hedging paper applied this to a Heston stochastic volatility world and equity option pricing

- Much wider applicability, however, especially for hedging problems where hedging constraints are material
  - Commodity players managing physical assets (gas-fired power plant, natural gas storage, etc)
  - Life insurers hedging variable annuity product issuance (long dated, illiquid market exposure, unhedgeable risks)

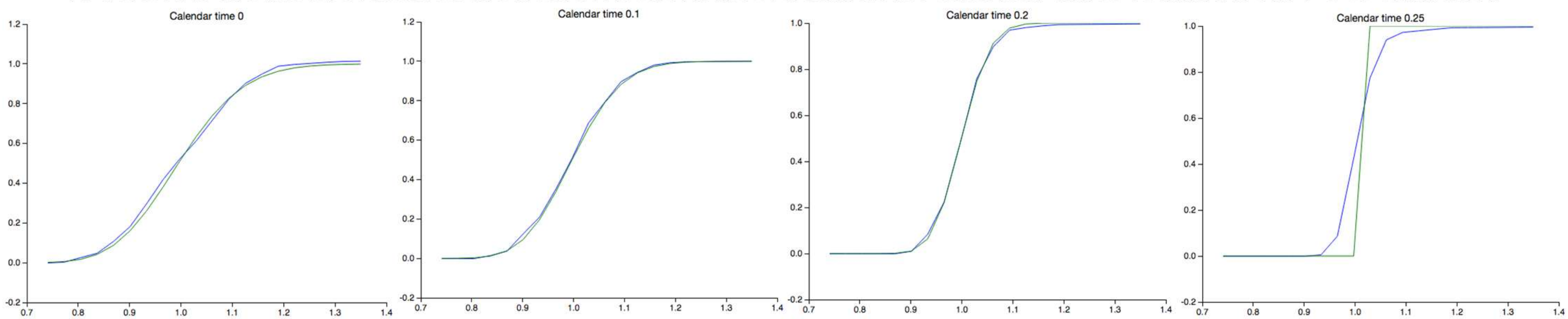
Deep Hedging – Black-Scholes Limit

- Let's start with a case we know: delta hedging a short European call option with the underlying asset, where the asset price follows a geometric Brownian motion with constant drift and volatility
  - All risks are hedgeable, and we're allowing for frequent hedging and no transaction costs, so should reduce to Black-Scholes
  - Note that the drift of the process isn't the risk neutral drift – it's the real world drift
  - Base case: 20% volatility, initial spot = 1, zero rates, zero real world drift, 0.25y to expiration, strike = 1

- How it works (using Google TensorFlow for neural network functionality, two hidden layers with 50 nodes each):
  - Neural network has three inputs: current spot, current time, and current delta hedge notional
  - Neural network has one output: incremental notional of delta to trade for the new hedge
  - Generate a subset of Monte Carlo paths (used 100 paths); each path has a number of time steps (used 100)
  - Run through those paths, generating hedges from the neural network at each point, and end up with a post-hedges PNL distribution for those paths (also generate post-hedges PNL distribution using Black-Scholes exact hedges to compare)
  - Calculate 70%-ile expected shortfall from those PNLs
  - Repeat, training network weights to minimize expected shortfall – used 1M total Monte Carlo paths, so 10k training steps

Deep Hedging – Black-Scholes Limit Results

- As expected we end up with the neural network spitting out hedge notionals that match the Black-Scholes deltas very closely
- Delta shown on y-axis, spot on x-axis; blue lines are neural network deltas, green lines are exact Black-Scholes deltas
- To generate the plots, evaluated the neural network (post training) for calendar times running from t=0 to expiration



Deep Hedging – Black-Scholes Limit Results

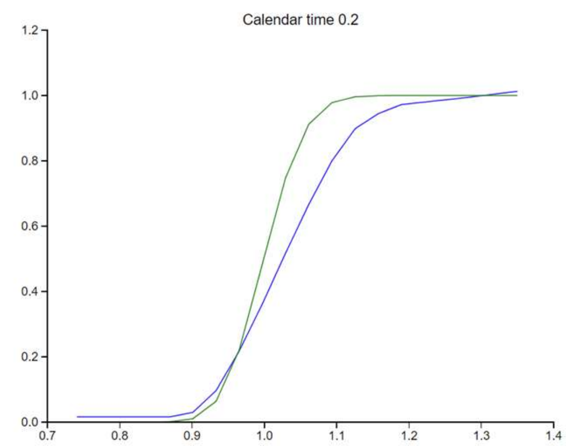
- Tried it for a few different cases:
  - Base: 20% vol, zero rates, zero drift, 70%-ile expected shortfall, 100 time steps
  - High Drift: 20% vol, zero rates, 10% (real-world) drift, 70%-ile expected shortfall, 100 time steps
  - Low %-ile: 20% vol, zero rates, zero drift, 25%-ile expected shortfall, 100 time steps
  - Infreq Hedge: 20% vol, zero rates, zero drift, 70%-ile expected shortfall, 30 time steps

Params	Deep Hedge Price	BS Hedge Price	BS Price	Mean MC Payoff	Mean PNL	PNL Std Dev
Base	0.0444	0.0437	0.0399	0.0399	-0.0399	0.0043
High Drift	0.0446	0.0437	0.0399	0.0543	-0.0399	0.0046
Low %-ile	0.0417	0.0413	0.0399	0.0399	-0.0399	0.0047
Infreq Hedge	0.0489	0.0468	0.0399	0.0399	-0.0399	0.0080

Deep Hedging – Black-Scholes with Transaction Costs

- Now we can take the framework and extend it to interesting stuff that traditional derivative pricing has trouble with
- For example: adding transaction costs
- Simple bid/ask model where spread is constant as a % of price, independent of notional (3bp TC means 6bp bid/ask spread)
- Chart shows the delta using the neural network (blue) compared to the Black-Scholes delta with 0.05y left to expiration

Params	Deep Hedge Price	BS Hedge Price	BS Price	Mean MC Payoff	Mean PNL	PNL Std Dev	Mean/SD Trans Cost	Mean/SD BS Trans Cost
0 TC	0.0447	0.0437	0.0399	0.0399	-0.0399	0.0045	0/0	0/0
3bp TC	0.0483	0.0455	0.0399	0.0398	-0.0409	0.0060	0.0010/0.0003	0.0016/0.0005
10bp TC	0.0498	0.0501	0.0399	0.0398	-0.0430	0.0063	0.0031/0.0007	0.0059/0.0012
30bp TC	0.0528	0.0636	0.0399	0.0398	-0.0465	0.0060	0.0066/0.0013	0.0168/0.0042
75bp TC	0.0691	0.1549	0.0399	0.0398	-0.0519	0.0132	0.0121/0.0022	0.0777/0.0309
1% TC	0.0836	0.3989	0.0399	0.0398	-0.0455	0.0302	0.0056/0.0000	0.2588/0.0810



Deep Hedging – Natural Gas Storage – Intro 1/2

- Now let's stretch our legs a bit and look at a more interesting example where this technique can really add business value: hedging the right to use a natural gas storage facility

- Natural gas storage: a hole in the ground where you can pump in gas when prices are low, leave it, and withdraw it (then sell it) when prices are high
  - Useful because of "seasonality" in the forward curve: since gas is used to generate electricity, prices tend to be higher in winters and lower in summers
  - Optionality since that spread can sometimes change sign, and the optimal months to inject/withdraw change as the forward curve moves and bends
  - Typically buy the right to use a storage facility for some relatively long period, like 5-10y, then hedge with short-dated forwards to dynamically manage the deltas
  - Bid/ask spreads in gas can be substantial, especially at the specific physical delivery location where a storage facility exists; though spreads are typically lower at more liquid "hub" delivery locations that are connected to other locations through pipelines

Deep Hedging – Natural Gas Storage – Intro 2/2

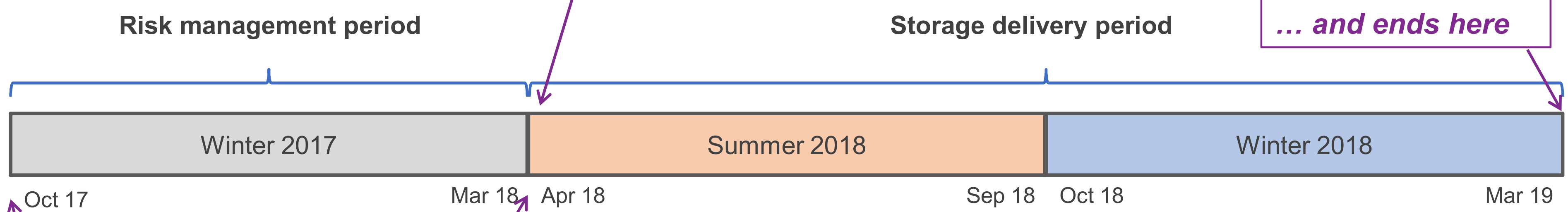
- The model framework is materially more complex than in the European option example because natural gas is not a standard financial asset, and so different points on the forward curve move with different volatilities and with less than 100% correlation
    - Need a "forward curve model", which defines how every point on the forward curve moves, rather than the more familiar "spot models" of financial assets
    - Also need to model the seasonal spreads, because those drive the value of the storage facility
    - We allow for two delivery locations for gas: one where the storage facility is located, which has high bid/ask spreads; and one "hub" location which trades as a basis to the storage facility's location, but has lower bid/ask spreads; we need to model that price basis
    - We use a 2-factor model for the forward curve as proposed in Andersen (Markov Models for Commodity Futures, https://ssrn.com/abstract=1138782), and a 1-factor model for the basis, where all forward prices are lognormally distributed

Deep Hedging – Natural Gas Storage – Hedging Window & Physical Storage Window

- Storage must start empty and end empty
- Strip of 365 daily options to inject or withdraw
- Injection rate: 100 days from empty to max
- Withdrawal rate: 33 days from max to empty

Storage contract window to inject or withdraw starts here …

… and ends here



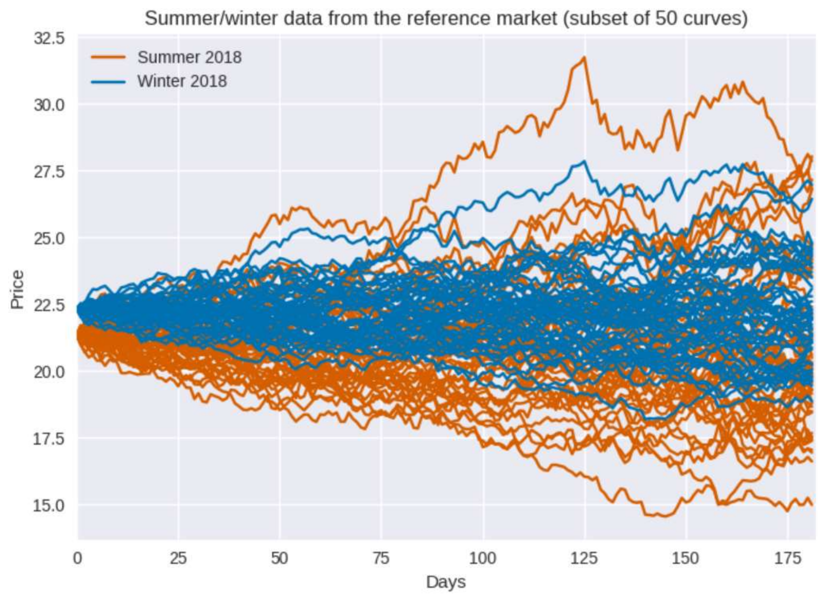
Forward hedging of storage contract starts here …

… and ends here

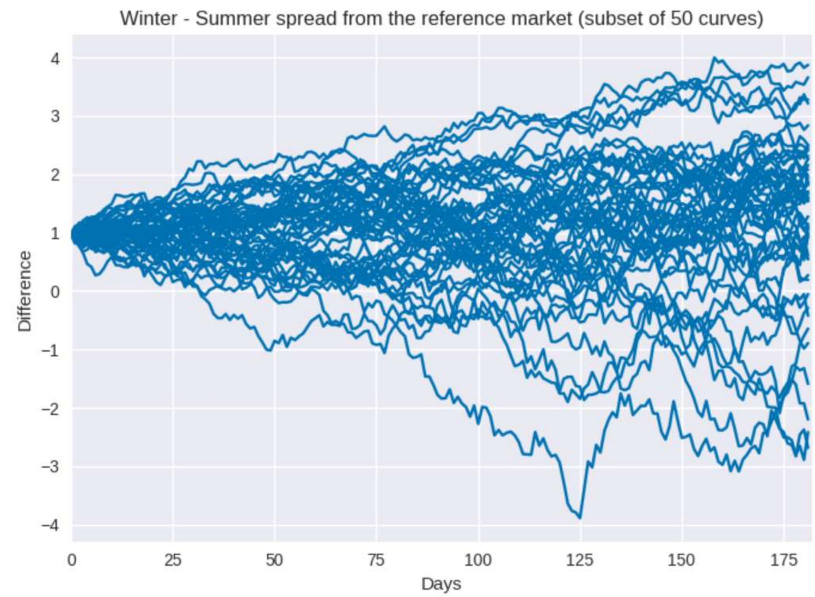
- Typical state of gas market: Winter gas price > Summer gas price
- Expected behaviour is "inject in Summer, withdraw in Winter" (if economic to do so)
- Hence buy gas from market in Summer and sell back to market in Winter
- Hence value of seasonal storage is linked to the "Winter – Summer spread"
- Summer Contract and Winter Contract swaps exist for hedging purposes

Deep Hedging – Natural Gas Storage – What do the market paths look like?

- Outright simulated Winter and Summer contract prices during the hedging window prior to injection/withdrawal period (EUR/MWh)



- Here's the same data but showing how "Winter – Summer" spread diffuses (main driver of storage value)



Beacon Platform & Kinetic Mind

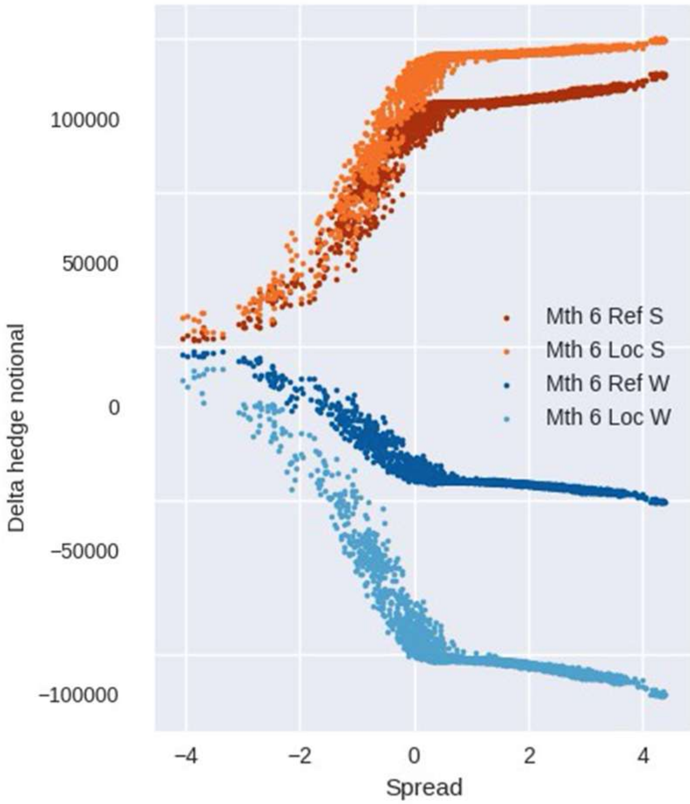
beacon
ACCELERATE TECHNOLOGY

Deep Hedging – Natural Gas Storage - Results

- High level results are what you would expect qualitatively
  - As transaction costs at the storage location increase, the fraction of hedge done at the liquid hub location increases, but the total long/short notionals summed across locations are roughly constant
  - As basis volatility increases, the fraction of hedge done at the storage location increases
  - As both transaction costs and basis volatility increase, expected shortfall increases

- The advantage of this method is that it tells you quantitatively both how you should distribute your delta hedge across the two locations, and that it tells you what a realistic post-hedges PNL distribution looks like (which helps determine what you are willing to pay for the storage rights)
  - Cannot get that without a lot of simplifying assumptions using the standard Black-Scholes-style real option pricing methods

Deep Hedging – Natural Gas Storage – Results on Training of Network



Scatter plot shows network can learn appropriate hedging behavior.

- Output of network is the recommended hedge position
- need to be long Summer – orange/red
- need to be short Winter – light blue/dark blue

- Main driver of storage value is Winter-Summer spread (x-axis)

- Positive spread → storage option is ITM, hedges of high magnitude

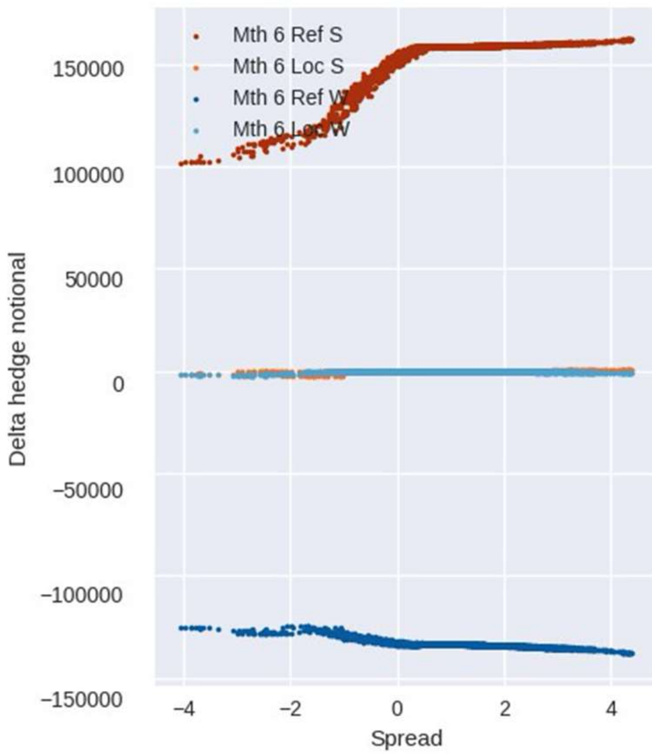
- Zero or negative spread → storage is ATM/OTM, hedges of smaller magnitude

Deep Hedging – Natural Gas Storage – Results show Basis Volatility can overcome Costs

LEFT PLOT:

- Costs at hub = 1%
- Costs at storage loc = 10%
- Zero basis vol

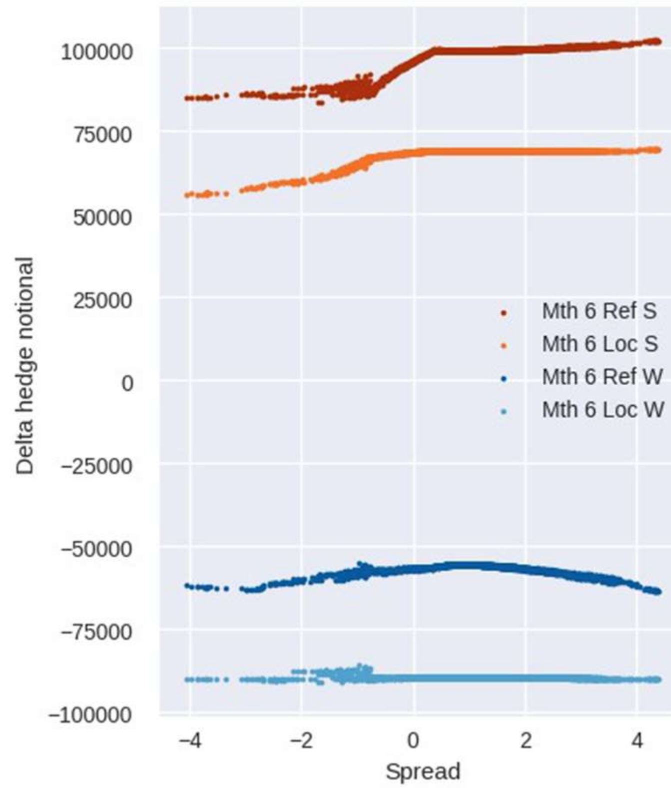
- OUTCOME: All hedging activity done at the hub



RIGHT PLOT:

- Costs at hub = 1%
- Costs at storage loc = 10%
- High basis vol (30%)

- OUTCOME: Hedging activity now split between hub & storage



Deep Hedging – Natural Gas Storage – Future Work and Conclusions

- We haven't modeled the basis between illiquid long-dated forward prices – which real long-dated storage contracts are exposed to – and the more liquid short-dated prices
  - Effectively the long-dated prices are a kind of unhedgeable risk, which the deep hedging framework handles consistently
- We haven't fully modeled the storage facility itself self-consistently; we simulate only to the first injection date of the storage facility, and price it at intrinsic value at that point
  - Better would be to run the simulation through the life of the storage facility and self-consistently find the value from the full life PNL distribution
- We haven't modeled the full set of tradable hedge instruments – really just the summer and winter prices as blocks
  - This ignores extra optionality from trading between nearby individual months on the forward curve as small changes in forward curve shape change the optimal injection and withdrawal dates

- What we have done is show that, for a relatively simple toy model, hedging using the deep hedging framework gives materially better results: lower transaction costs and a smaller PNL standard deviation
  - Should be able to extend this to real storage facilities and more complex forward curve models and really improve outcomes for commodity traders