

# Revealing the Oil Majors' Adaptive Capacity to the Energy Transition with Deep Multi-Agent Reinforcement Learning

Dylan Radovic  
Lucas Kruitwagen  
Christian Schroeder de Witt



# Outline

Introduction

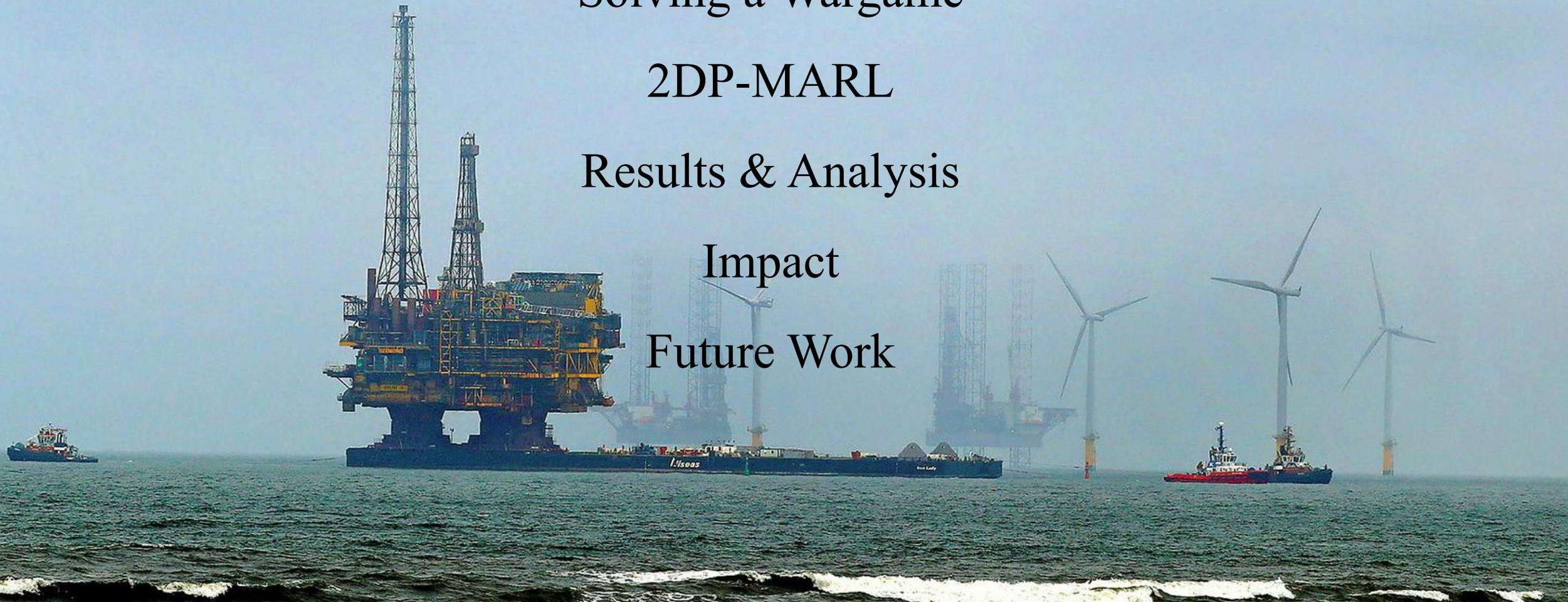
Solving a Wargame

2DP-MARL

Results & Analysis

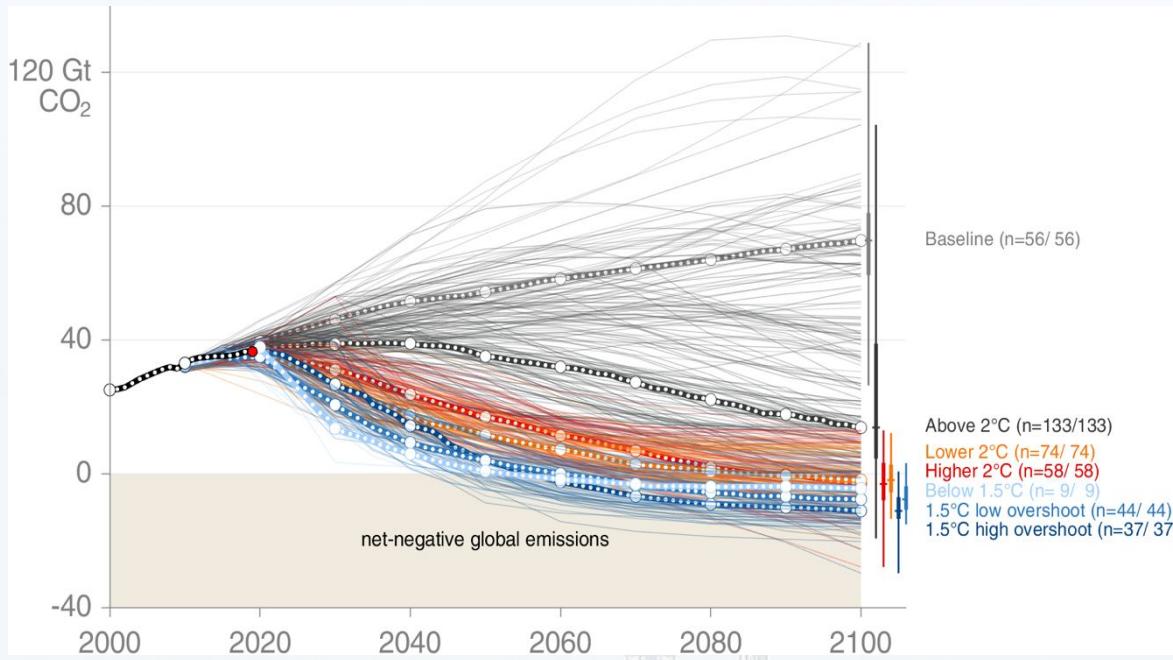
Impact

Future Work



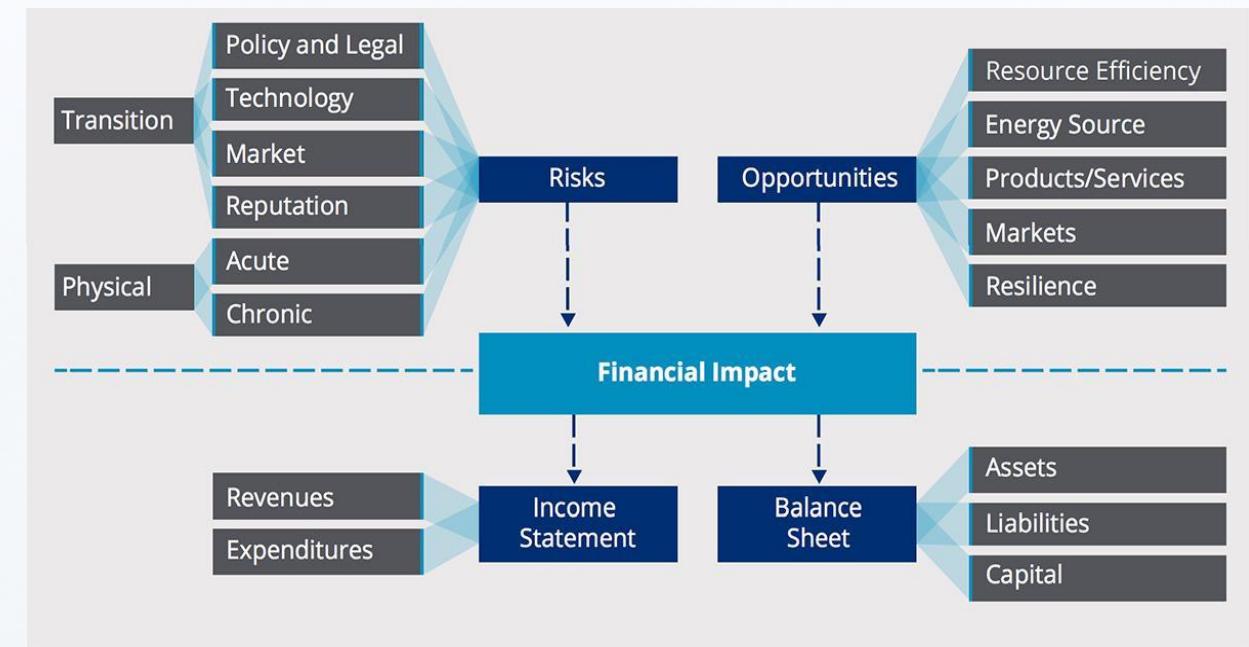
# Introduction





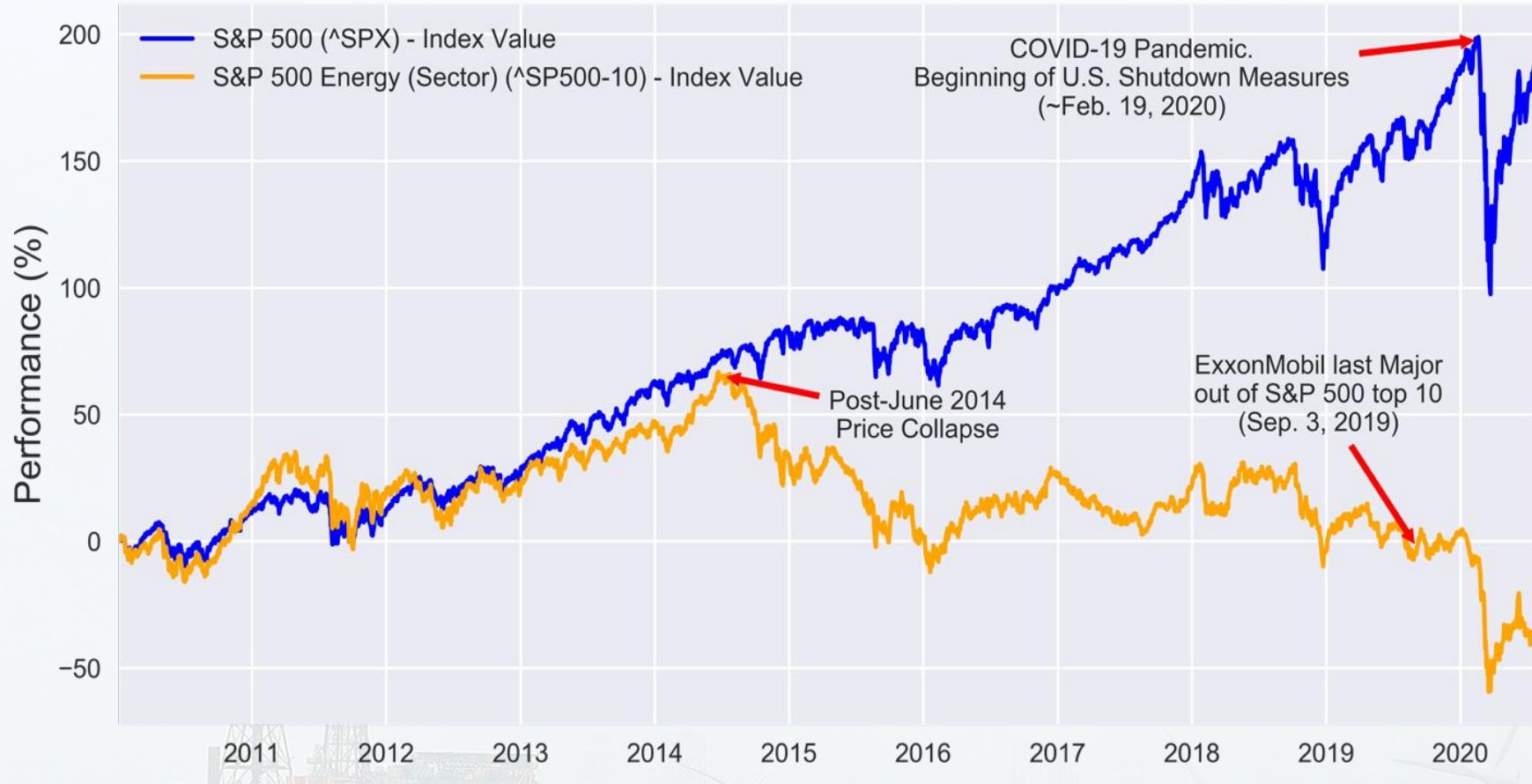
## Energy transition manifesting in response to climate change

- Though pathways vary, all scenarios predicated on low-carbon energy production
- Deep decarbonization requires massive reallocation of capital



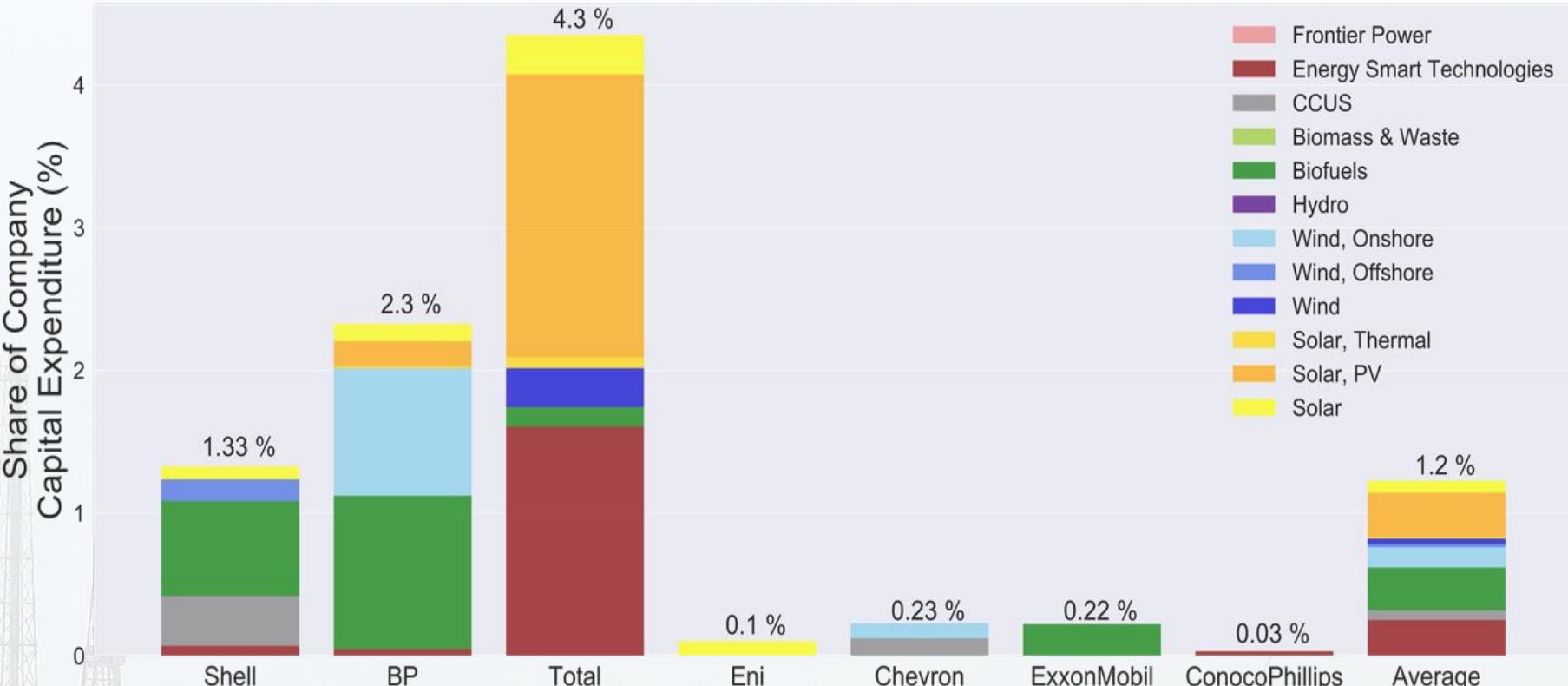
## Transition risks will arise, however, so will opportunities

- Rapid low-carbon transition poses an existential threat for the fossil fuel industry, particularly the oil Majors
- Adaptation may prove financially favorable



**The Majors have been in decline since 2014, COVID-19 has accelerated it**

- Oil & gas companies have consistently underperformed despite economic growth
- Financial recovery post-COVID through BAU becomes increasingly unlikely as the impeding energy transition unfolds



## Low-carbon actions speak louder than net-zero words

- Majors have set out plans for a full-scale decarbonization of their business models
- Throughout the previous decade, the Majors' low-carbon efforts proved minimal
- Majors have yet to make significant moves into low-carbon business models as the upside and downside risks of doing so remain unclear

# Aim

Provide tangible insights into the Majors' adaptive capacity to the energy transition by

*Exploring upside and downside risks of a first low-carbon mover*



*Solving for robust business pathways amidst energy transition uncertainty*



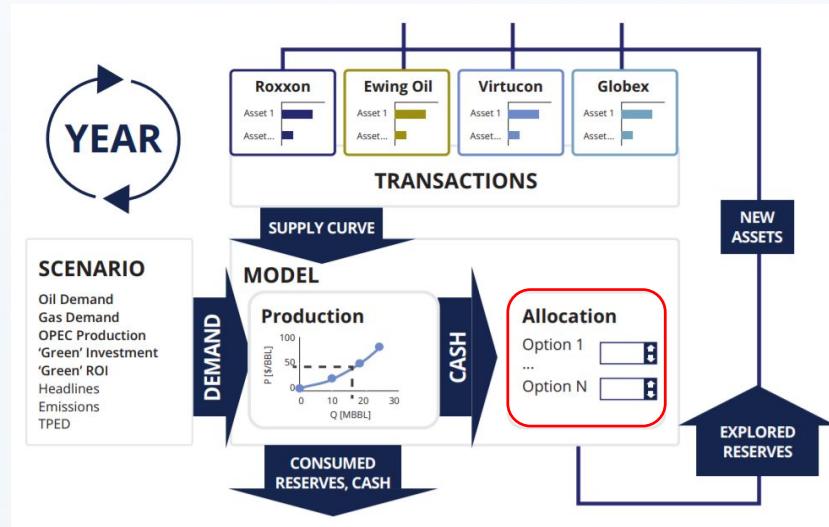
# Solving a Wargame



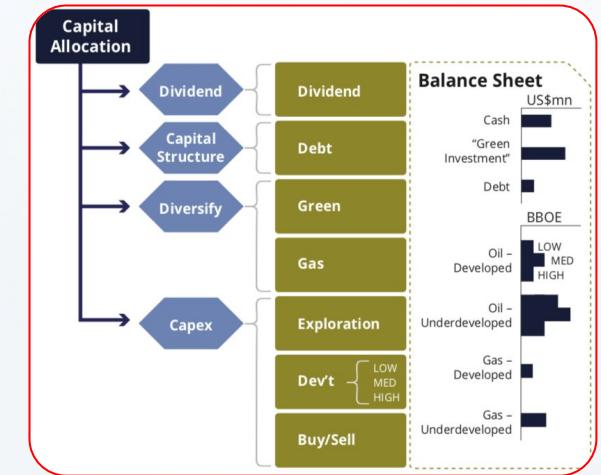
MARCH 2018

# CRUDE AWAKENING: MAKING OIL MAJOR BUSINESS MODELS CLIMATE-COMPATIBLE

BEN CALDECOTT, INGRID HOLMES, LUCAS KRUITWAGEN,  
DILEIMY OROZCO AND SHANE TOMLINSON



2DP Wargame Schematic



Player Allocation Options

## 2 Degrees Pathways Wargame – testbed for oil & gas transition

- The 2 Degrees Pathways (2DP) wargame created to

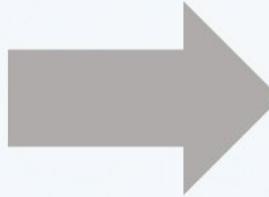
*"To help inform company, investor, government, and civil society thinking around pathways the oil and gas majors can take to become 1.5C/2C-compatible"*

- Originally played with human players, revealing more about human bias than robust, climate-compatible pathways

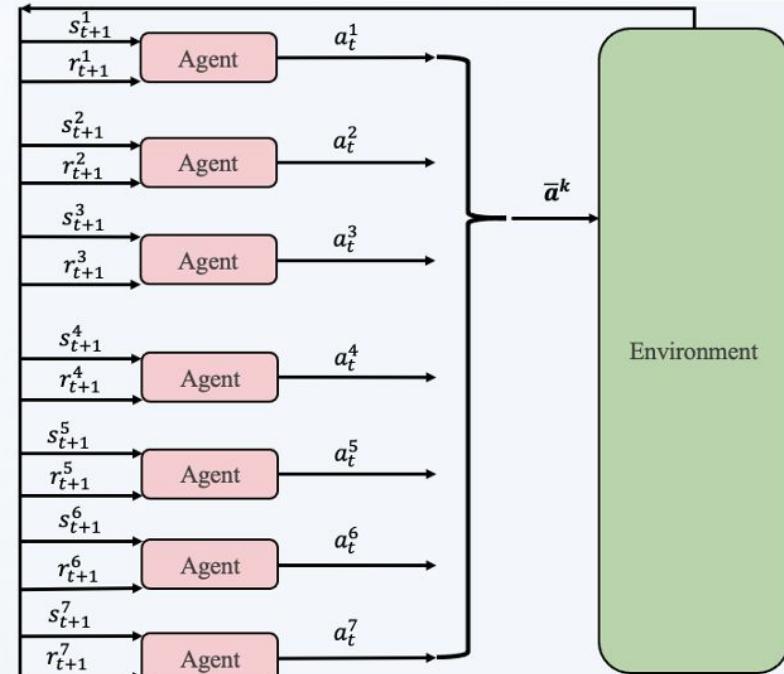
## Partially Observable Stochastic Game

$$\langle \mathcal{I}, \mathcal{S}, \{b^0\}, \{A_i\}, \{\mathcal{O}_i\}, \mathcal{P}, \{\mathcal{R}_i\} \rangle$$

Notation	Description
$\mathcal{I}$	Finite set of agents indexed $1, \dots, n$
$\mathcal{S}$	Finite set of states $\{s^1, \dots, s^N\}$
$b^0 \in \Delta(\mathcal{S})$	Initial state distribution
$A_i$	Finite set of actions available to agent $i$ and $\vec{A} = \times_{i \in \mathcal{I}} A_i$ is the set of joint actions (i.e. action profiles), where $\vec{a} = \langle a_1, \dots, a_n \rangle$ denotes the joint action
$\mathcal{O}_i$	Finite set of observations for agent $i$ and $\vec{\mathcal{O}} = \times_{i \in \mathcal{I}} \mathcal{O}_i$ is the set of joint observations where $\vec{o} = \langle o_1, \dots, o_n \rangle$ denotes the joint observation
$\mathcal{P}$	Set of Markovian state transition and observation probabilities, where $\mathcal{P}(s', \vec{o} s, \vec{a})$ denotes the probability that taking joint action $\vec{a}$ in state $s$ results in a transition to state $s'$ and joint observation $\vec{o}$
$\mathcal{R}_i: \mathcal{S} \times \vec{\mathcal{A}}$	Reward function for agent $i$



## Multi-Agent Reinforcement Learning



## Framing 2DP as a Partially Observable Stochastic Game

- A continuous control problem with multiple competing entities, 2DP is best characterized as a Partially Observable Stochastic Game (POSG)—a Markovian framework resemblant of real-life market competition

## Solving with Deep Multi-Agent Reinforcement Learning

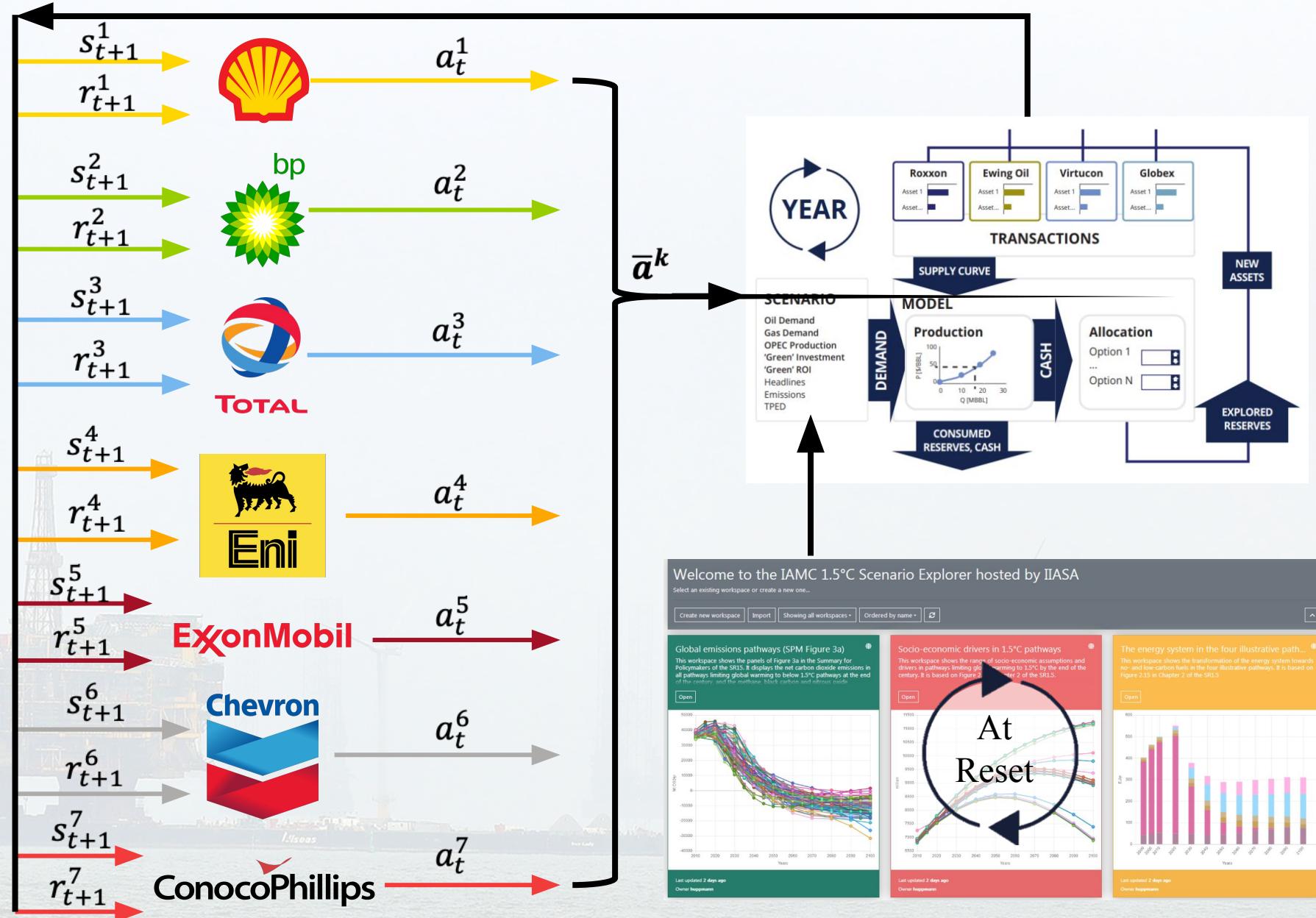
- Advances in Deep Multi-Agent Reinforcement Learning (MARL) have achieved superhuman-level performance in POSGs of high-dimensional settings

# 2DP-MARL



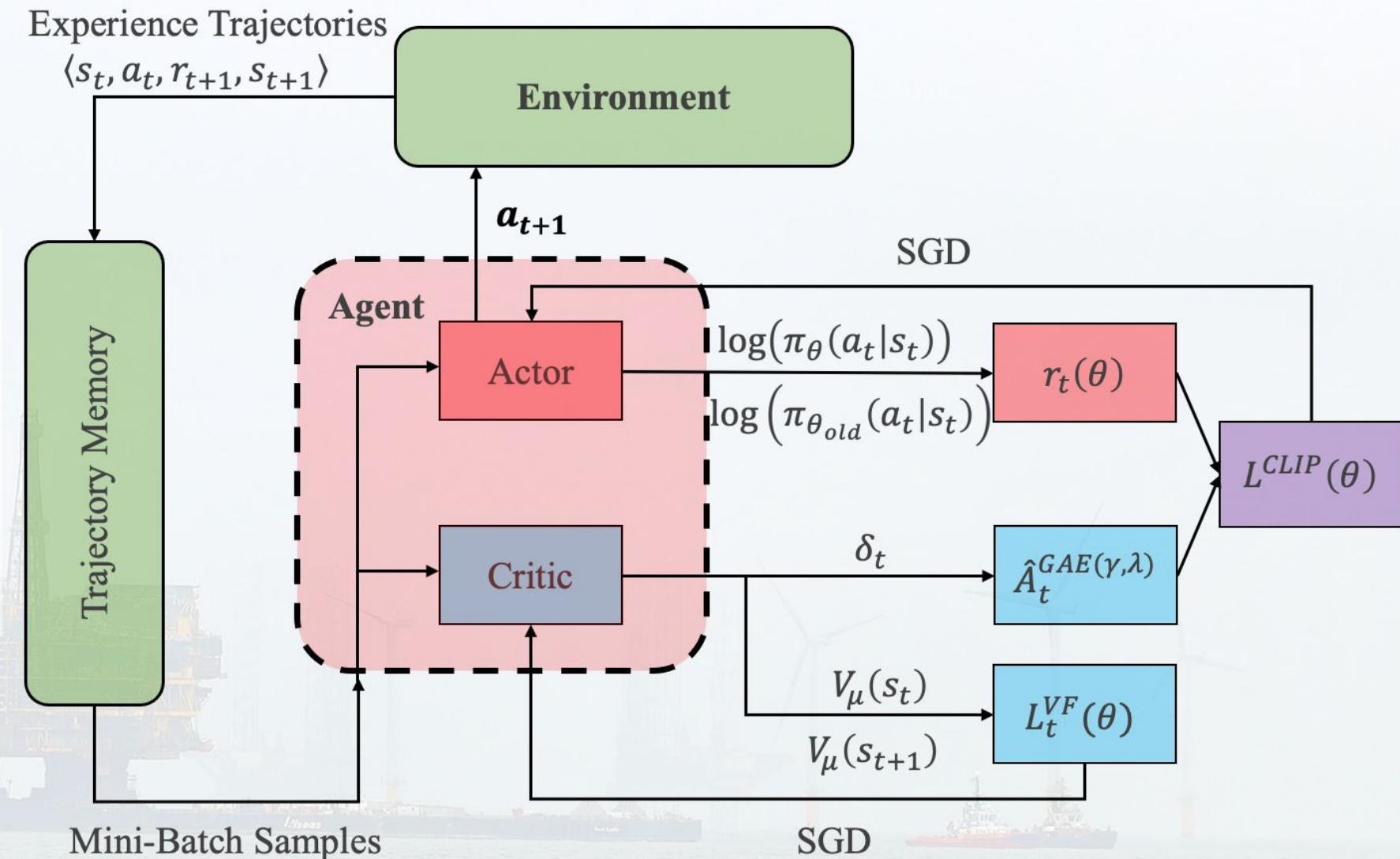
# A novel microeconomics model relevant to the energy transition discussion

- 2DP-MARL combines wargaming and advances in Deep MARL to help guide the Majors' stakeholders in assessing climate-compatible business models



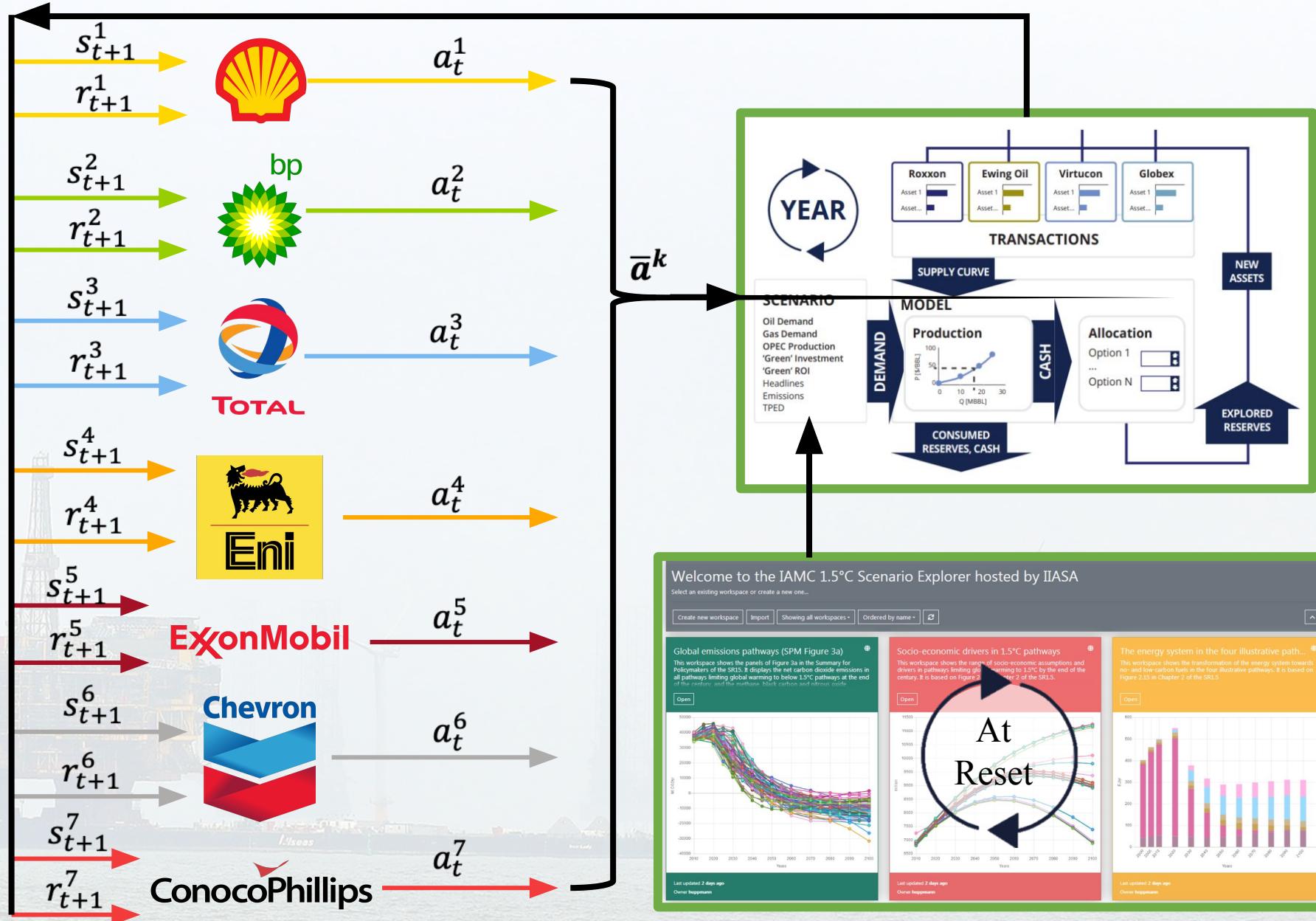
## A2C/PPO as foundational algorithms

- Algorithms employed take from Advantage Actor-Critic and Proximal Policy Optimization to solve a game in a high-dimensional, continuous space



## Environment

- Original 2DP Wargame, 2020-2040
- Climate-compatible scenario data generated from Integrated Assessment Modeling Consortium (IAMC) and International Institute for Applied Energy Systems (IIASA) ensemble

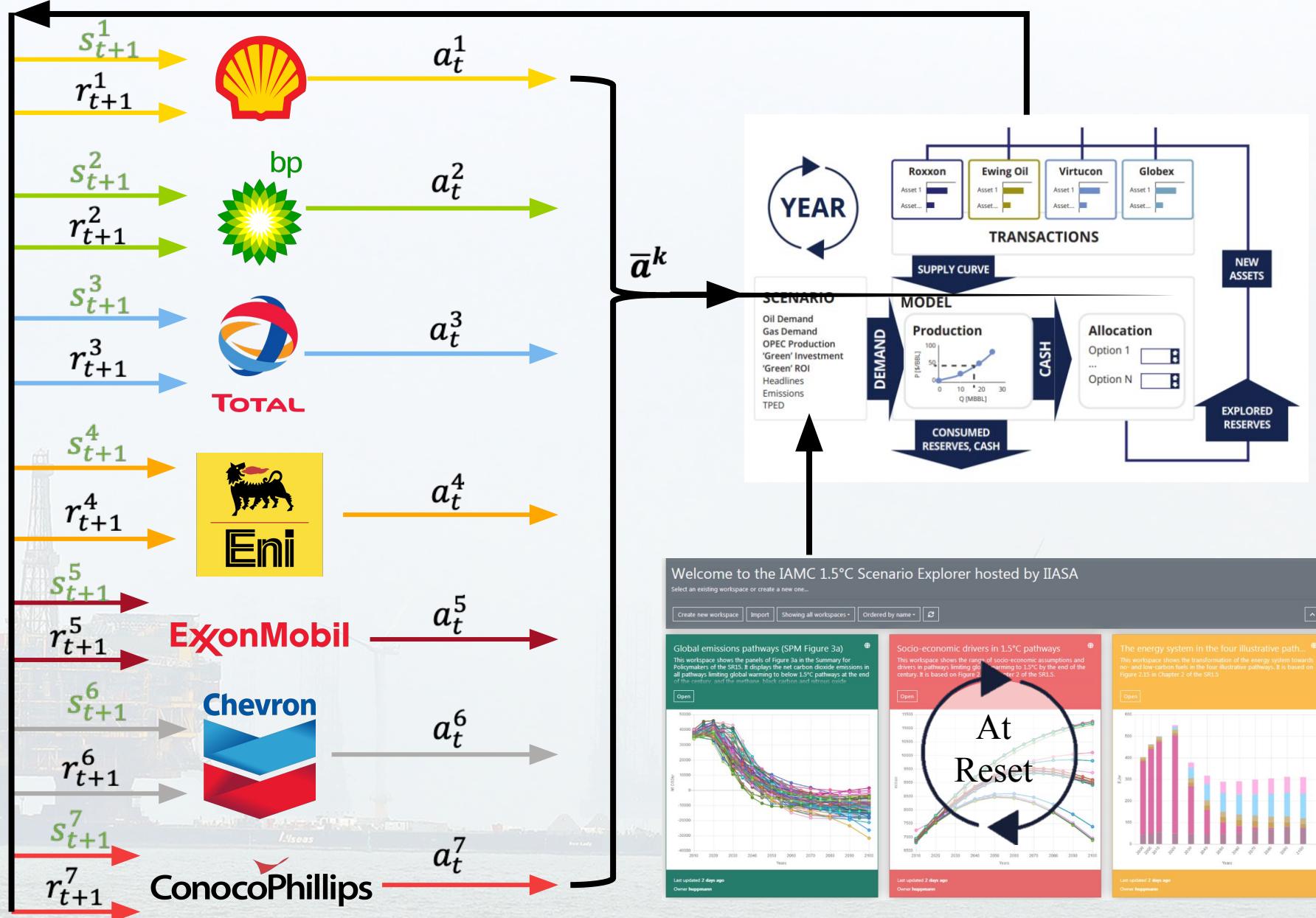


## Environment

- Original 2DP Wargame, 2020 – 2040
- Climate-compatible scenario data generated from IAMC/IIASA ensemble

## Observations

- Incomplete information to mimic real-life competition (i.e. players may only view some of their adversaries' assets)



## Environment

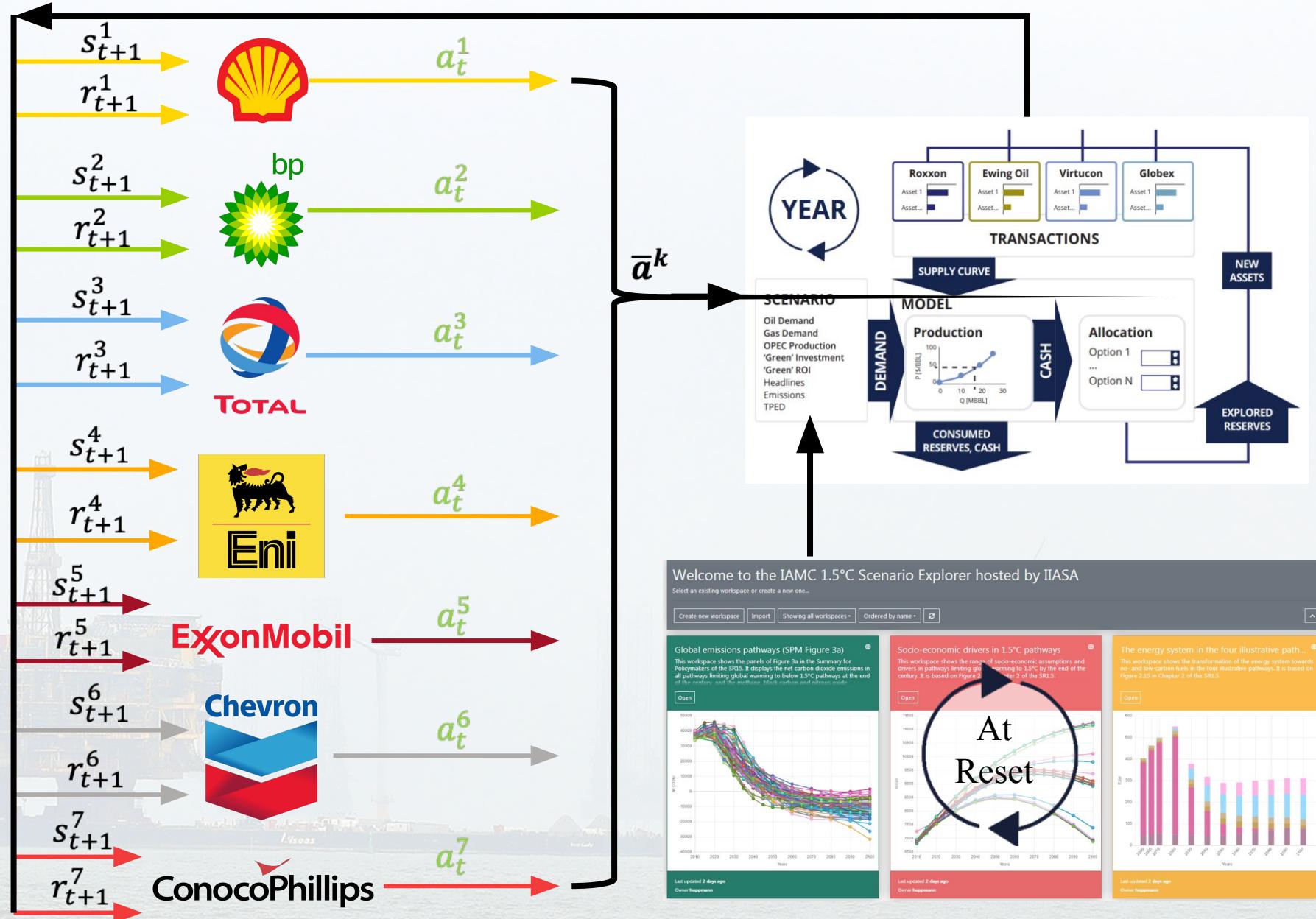
- Original 2DP Wargame, 2020 – 2040
- Climate-compatible scenario data generated from IAMC/IIASA ensemble

## Observations

- Incomplete information to mimic real-life competition (i.e. players may only view some of their adversaries' assets)

## Actions

- Similar to 2DP wargame capital allocation options, includes player-to-player trading and 'green' auction house



# Environment

- Original 2DP Wargame, 2020 – 2040
- Climate-compatible scenario data generated from IAMC/IIASA ensemble

# Observations

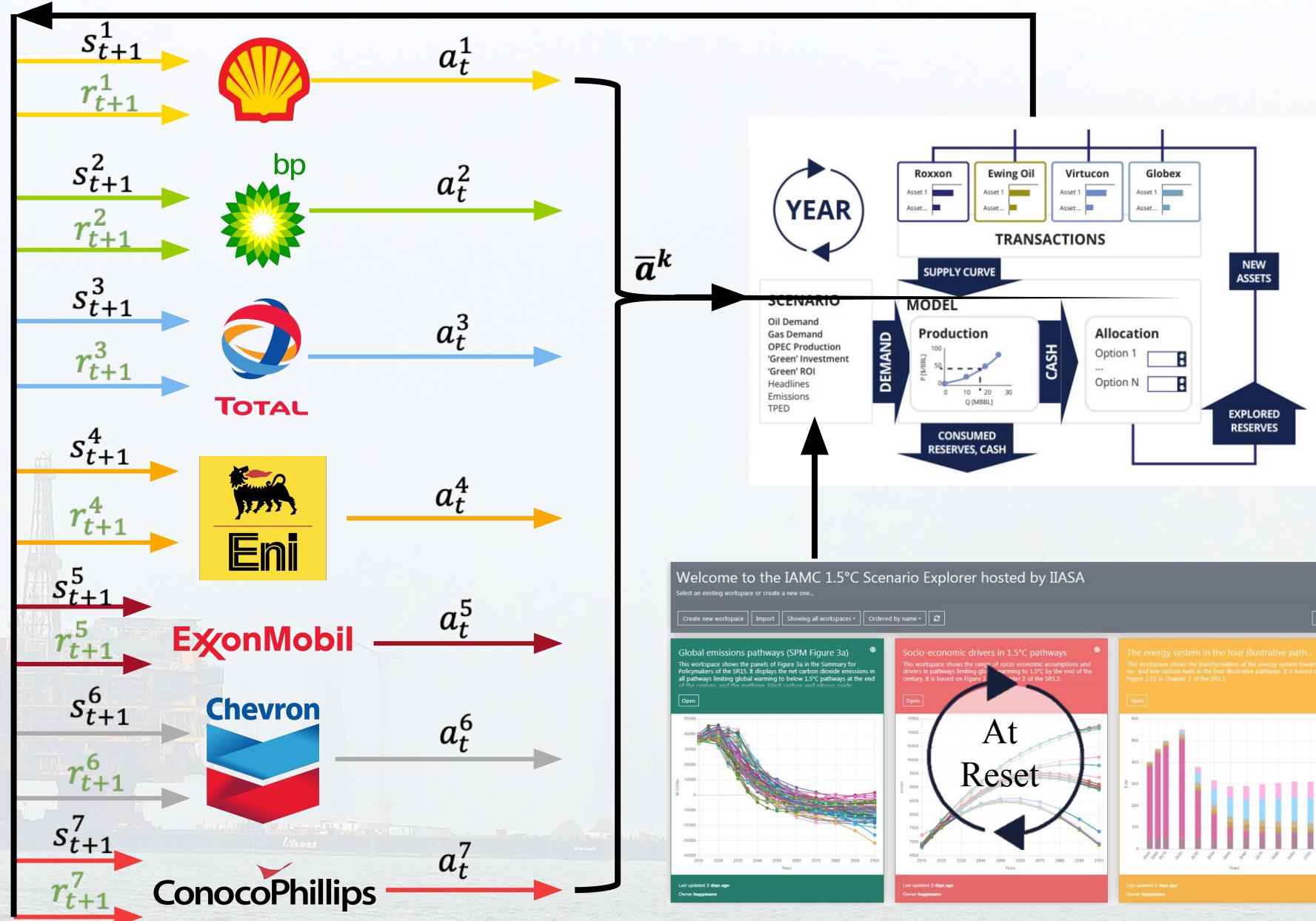
- Incomplete information to mimic real-life competition (i.e. players may only view some of their adversaries' assets)

# Actions

- Similar to 2DP wargame capital allocation options, includes player-to-player trading and 'green' auction house

# Rewards

- Focus on maximizing shareholder value through dividends
- Negative rewards to enforce realistic behavior



# Reason for Negative Rewards:

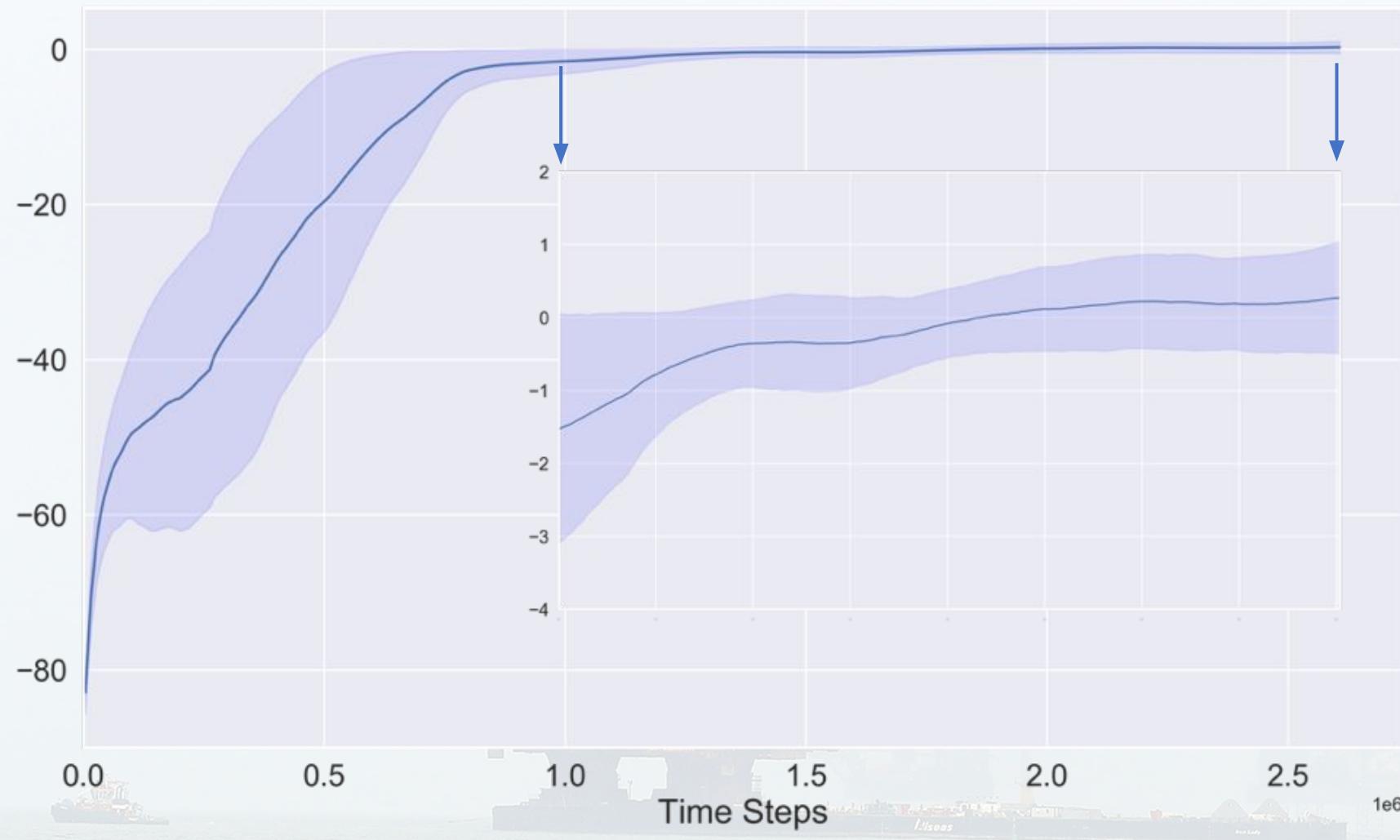
# Intelligent agents love to find loopholes!

Year	Game	Info	Policies						Spending Habits						Financial and Industry Metrics									
			O_Pri	Brrw	P_Div	P_Debt	Upg_L	Upg_M	Upg_H	Exp_O	Σ?	Div(\$)	Dbt_(\$)	Dev_(\$)	M(\$)	Dev_H(\$)	Expo(\$)	Σ(\$)	TEV(\$)	ROI(%)	ROE(%)	D/E(%)	R/P	Cost of Capital (%)
2019	2019	14.0	26	33.04	31.74	4.7	9.4	14.09	7.03	1000	351209	337390	2050	24680	74677	800176	3530542.52	83	0	16.34	2.27	0.053		
2020	2020	14.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	70328	249827	1025	5125	12300	61570	703285.16	0	0	67.31	0.0	0.094		
2021	2021	14.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	238264	846386	1025	5125	12300	208593	1311694	2382643.65	0	0	67.33	0.0	0.094	
2022	2022	14.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	909812	321925	4148	20743	49785	7965130	501930	9098121.71	110	0	66.77	13.84	0.093	
2023	2023	14.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	3296811	11711260	3420	17102	41047	2886267	17955910	32968111.32	110	0	66.86	24.29	0.093	
2024	2024	14.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	11692371	415434806	11588	57942	139062	10236351	63672123	116923716.66	110	0	66.99	58.59	0.094	
2025	2025	14.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	21116360	14605763	44250	531810	35996247	2396674	221254	0	67.06	98.98	0.094			
2026	2026	14.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	143646880	510276782	160348	881741	19241794	15758917	275285887	1436468897.78	110	0	67.11	242.62	0.094	
2027	2027	6.71	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	49934367	1775916149	568686	284340	682423	437678792	2723765874	4999583476.96	31	0	67.14	541.04	0.094	
2028	2028	29.3	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	17380840	561310380	199971	999895	3799498	352162453	4969790754	1738092725	0	0	67.15	1884.89	0.094	
2029	2029	9.6	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	17380840	561310380	199971	999895	3799498	352162453	4969790754	1738092725	3	0	67.16	6559.69	0.094	
2030	2030	3.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	20992158327	174572006	743153588	121577442	16375651	16378319145	176559085	20992158327	0	0	67.16	22810.72	0.094	
2031	2031	0.5	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	7295179822	25101664025	84534669	273348	1015437308	3014537312	397817363843	7295179739.09	0	0	67.16	79292.54	0.094	
2032	2032	0.2	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	253512649498	990558283	189038424	169670172	3525614056	22193376687	13821973958	2535167631929.83	0	0	67.16	275580.95	0.094	
2033	2033	0.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	8899712557	31297479593	181202881	125234674	77126209280	408008762055	8811137931236.22	0	0	67.16	957787.11	0.094		
2034	2034	0.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	306142231652	1087590125428	354813833	1774019675	42091675	26801919024	16680921572106	306221056969	0	0	67.16	3328174.68	0.094	
2035	2035	0.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	10638605397243	377195184113	23130187569	16150937968	14762551212	913308840793	57965875055442	106460598054745.27	0	0	67.16	11565572.19	0.094	
2036	2036	0.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	369697174024	221679394308	4284612257	1420611339	514177462173	323659718113	2014345189801	367969531132427.25	0	0	67.16	40191075.53	0.094	
2037	2037	0.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	128471686134444	4563699316468688	148899550238	764497751170	7879646028801	1124734422679	6999595231312314	1284716861589922.2	0	0	67.16	139666219.72	0.094	
2038	2038	0.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	446444580288805	152743850124	2587160972146	890851075130	14833850140	14853850140	64535082365.26	446535082365.26	0	0	67.16	485347562.66	0.094	
2039	2039	0.0	51	5.25	18.65	25.0	21.16	29.35	4.6	1000	1551422440356350	55111724644352	17981056337	890851781679	215733714762095	13582752137157	845313550267470	1.551726519807627e+16	0	0	67.16	1686608279.81	0.094	

In 21 years, Shell receives a reward of 2178246

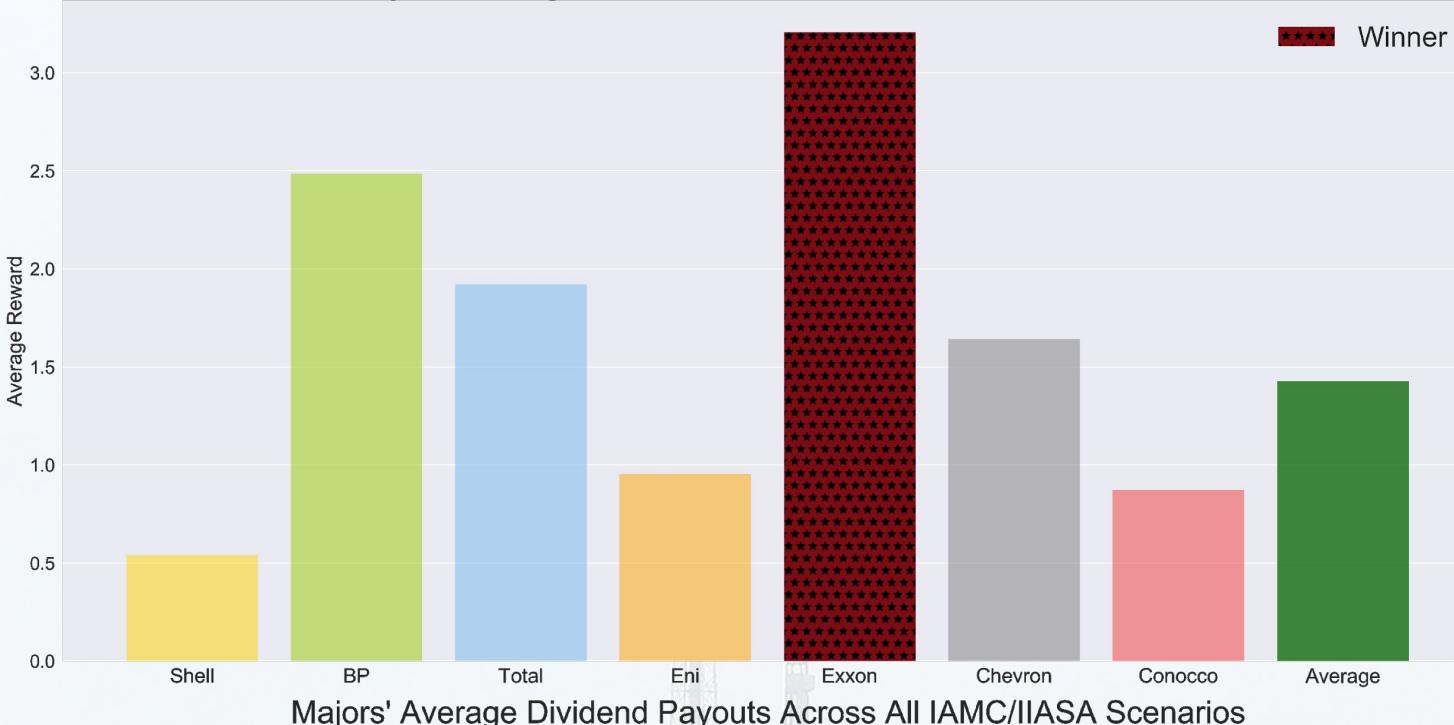
# Results & Analysis





- Convergence towards a scenario-robust solution is reached after  $\sim 2.6M$  timesteps, or 320 pass throughs each scenario ( $\sim 130,000$  games played)
- Convergence towards a positive reward indicates agents achieved realistic, robust strategies to maximize dividend payouts
- Oil and gas, ‘green’ and debt holdings are explored to evaluate the Majors’ business models

## Majors' Average Reward Across All IAMC/IIASA Scenarios



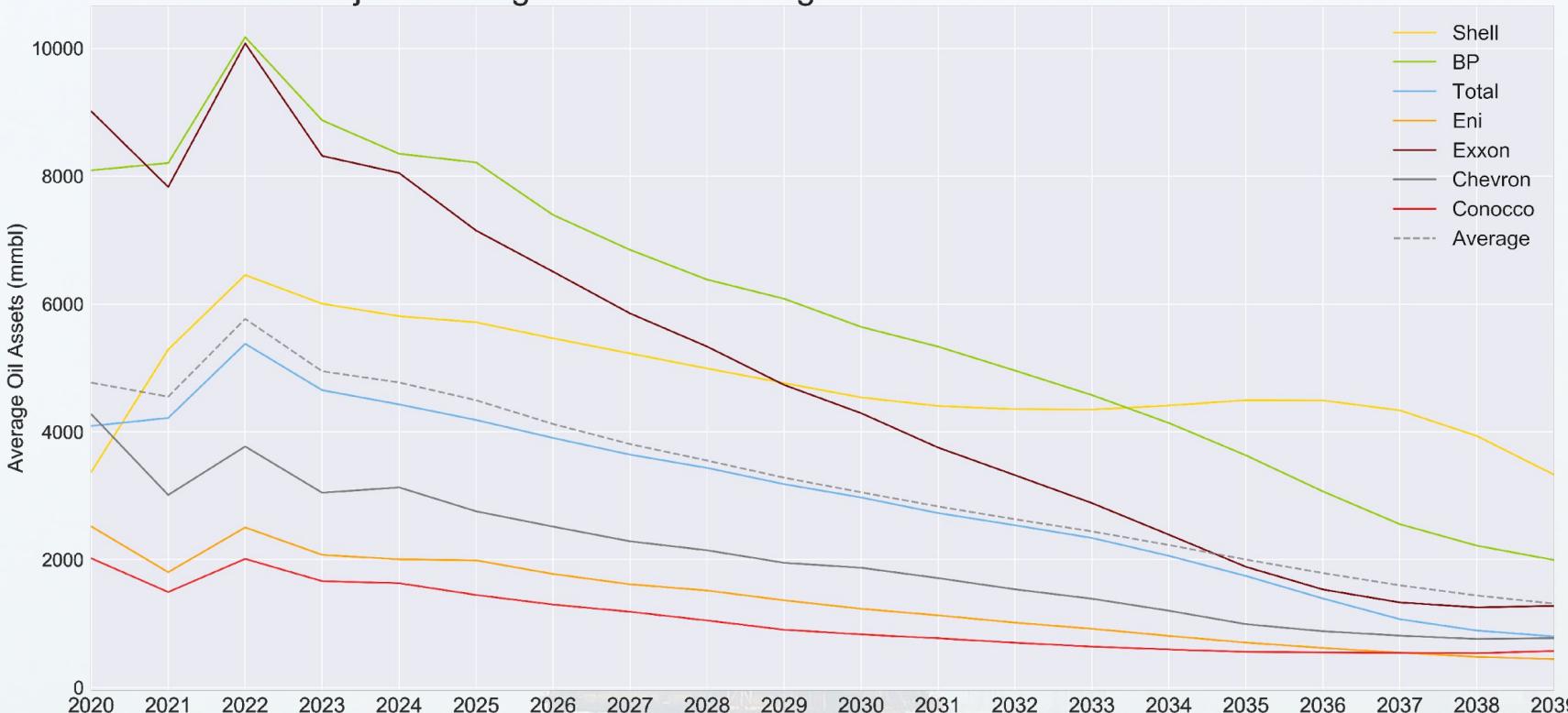
- Across all scenarios, Exxon pays out the most dividends while maintaining a realistic strategy pathway

## Majors' Average Dividend Payouts Across All IAMC/IIASA Scenarios



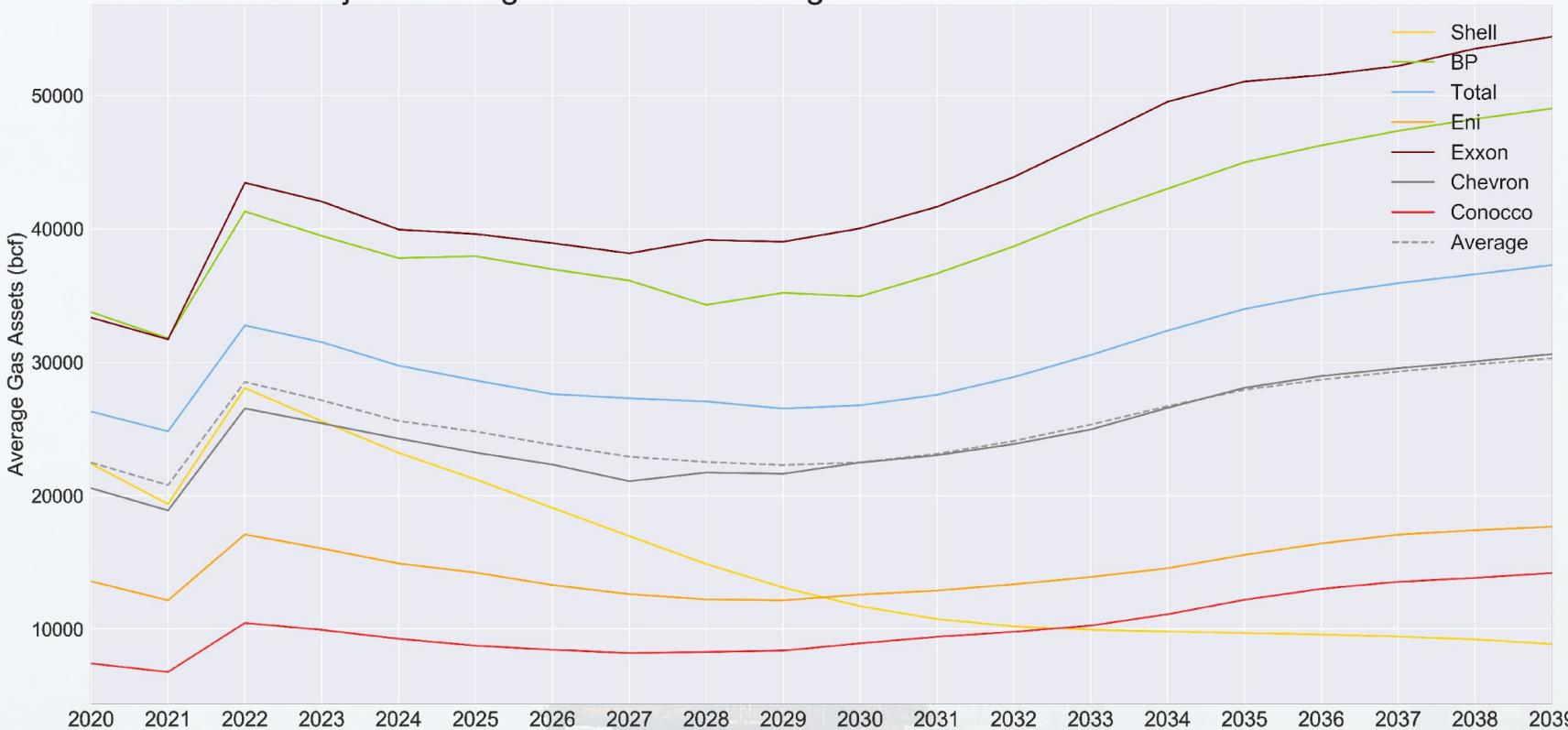
- Average yearly dividend payouts take an initial dip before noticeably increasing in the latter half of the game

### Majors' Average Oil Asset Holdings Across All IAMC/IIASA Scenarios



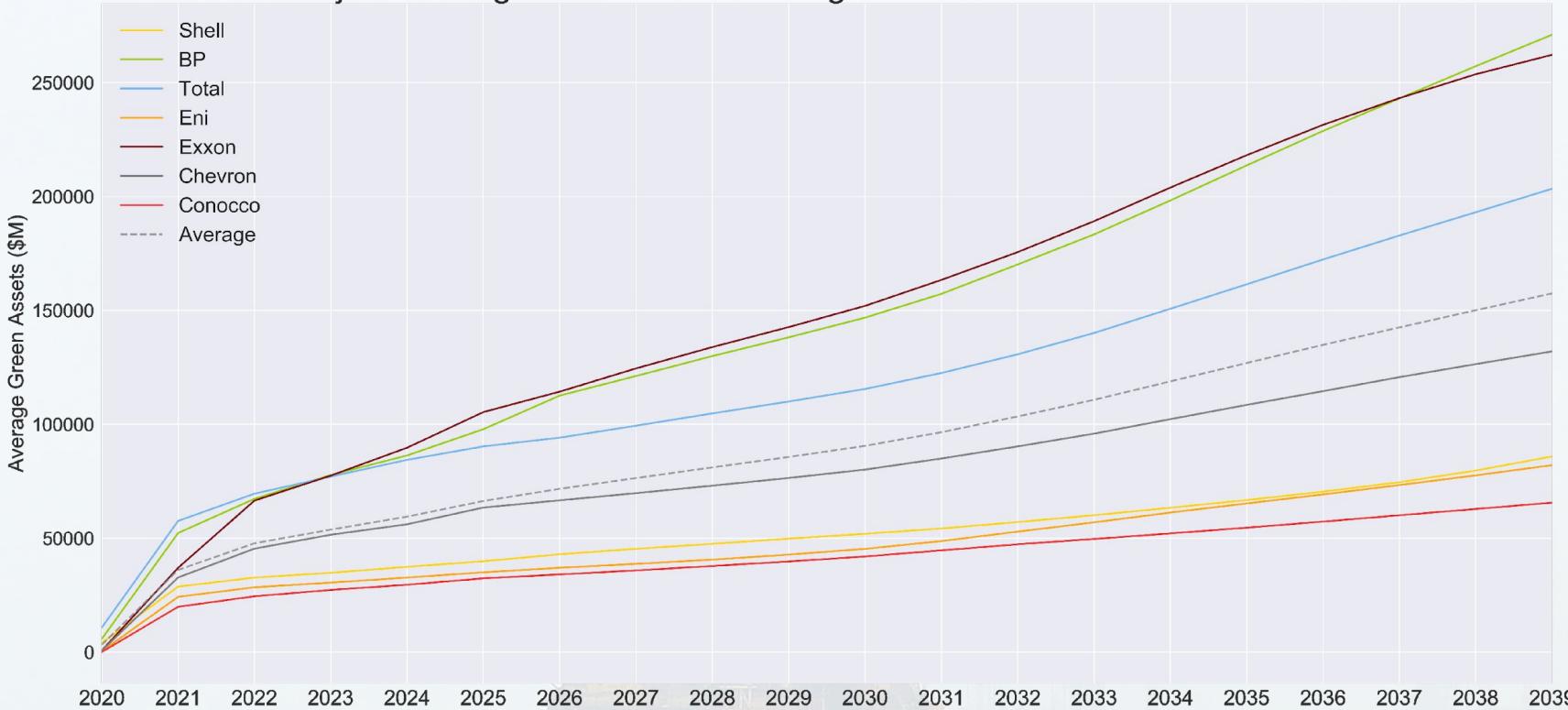
- Majors heavily divest from oil assets early, continue divestment throughout
- Shell and BP maintain the greatest oil market share at the end game

### Majors' Average Gas Asset Holdings Across All IAMC/IIASA Scenarios



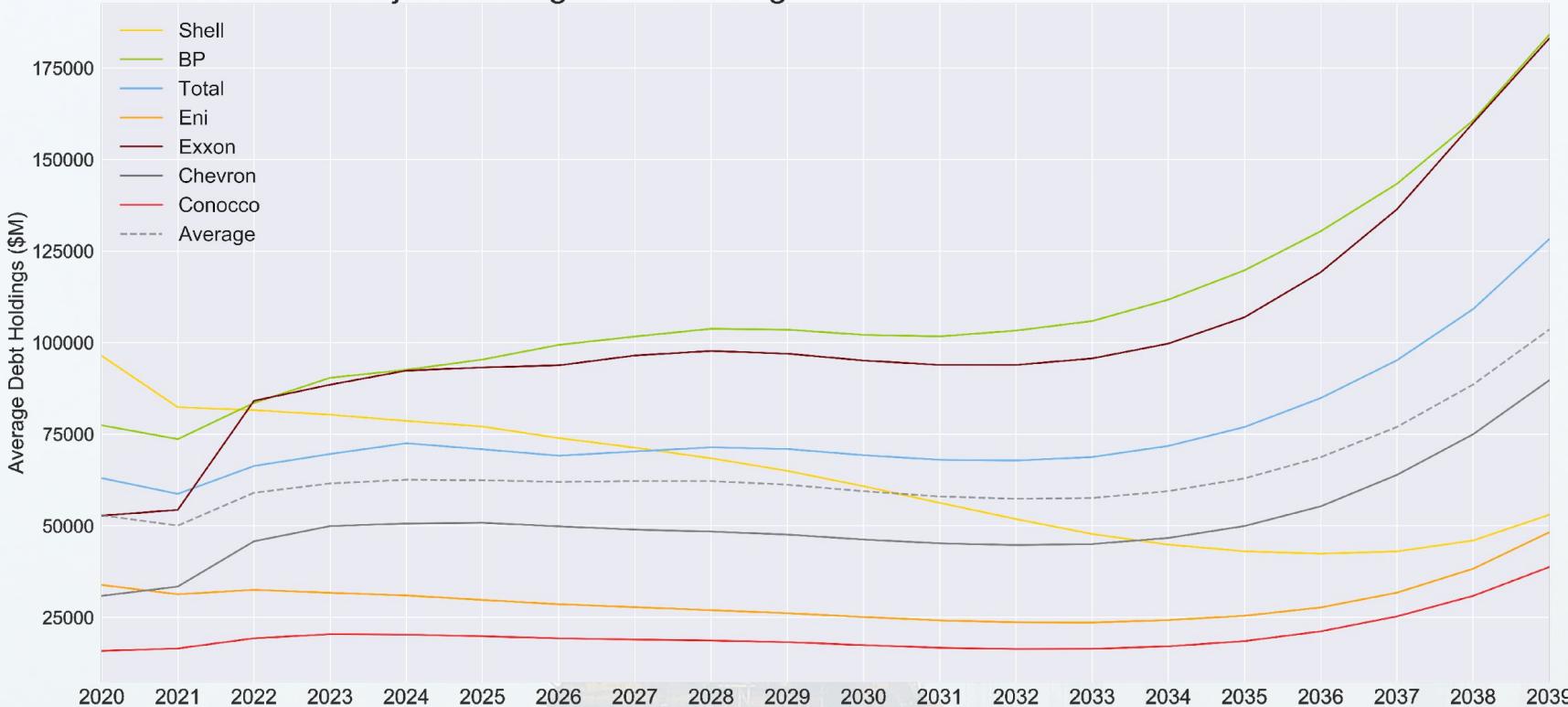
- Majors gradually increase gas asset holdings, suggesting the carbon asset's net income stability
- Gas market shares remain similar, however, Shell tends to let its market share slip

## Majors' Average Green Asset Holdings Across All IAMC/IIASA Scenarios



- Majors heavily invest in ‘green’ assets; display similar ‘first-mover’ behavior and long-term trends
- Exxon and BP act as leaders of the ‘green’ movement primarily due to their large, initial balance sheets

## Majors' Average Debt Holdings Across All IAMC/IIASA Scenarios



- Hints of a leverage transition (i.e. borrowing cash to buy ‘green’ assets) undertaken by BP and Total, Exxon and Chevron.
- Levels of debt increase towards the end game due to Majors’ new debt resilience

# Impact



## **Moving First, Benefits Outweigh Costs**

Moving first into 'green' assets allows Majors to accumulate higher returns long-term and maintain stable levels of debt



## **Decarbonization of Emission Leaders**

Providing a financial case for the decarbonization of the Oil Majors helps stakeholders understand the potential these carbon emission leaders have to become net-zero vanguards

## **A Robust, 'Green' Strategy**

Going 'green' while diminishing net income reliance on oil assets proves a robust response to uncertainty in a 1.5C world

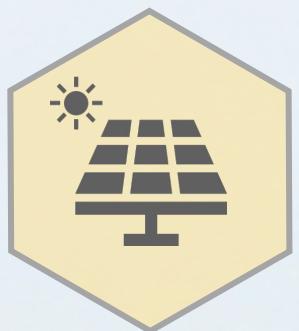
# Future Work





Capital cost  
sensitivity  
analysis

Modeling  
National  
Oil Companies



Expanding  
'green' asset  
allocation



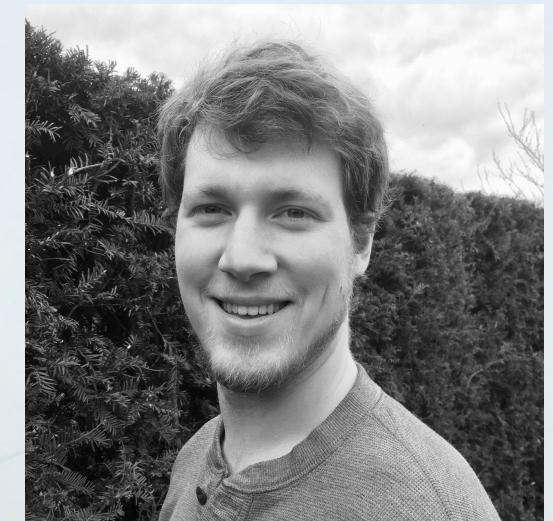
Focus on corporate  
governance  
questions



Mass rollout  
as 1-player  
game



# Thank You!



**Lucas Kruitwagen**

University of Oxford

[lucas.kruitwagen@smithschool.ox.ac.uk](mailto:lucas.kruitwagen@smithschool.ox.ac.uk)

**Dylan Radovic**

Imperial College London

[dylan.radovic19@imperial.ac.uk](mailto:dylan.radovic19@imperial.ac.uk)

**Christian Schroeder de Witt**

University of Oxford

[christian.schroeder@stcatz.ox.ac.uk](mailto:christian.schroeder@stcatz.ox.ac.uk)