

# Public Policy and Deep Reinforcement Learning on AWS

Emily Webber | Machine Learning Specialist at Amazon Web Services | To-be-open-sourced research project

# Public Policy Has Unique Challenges



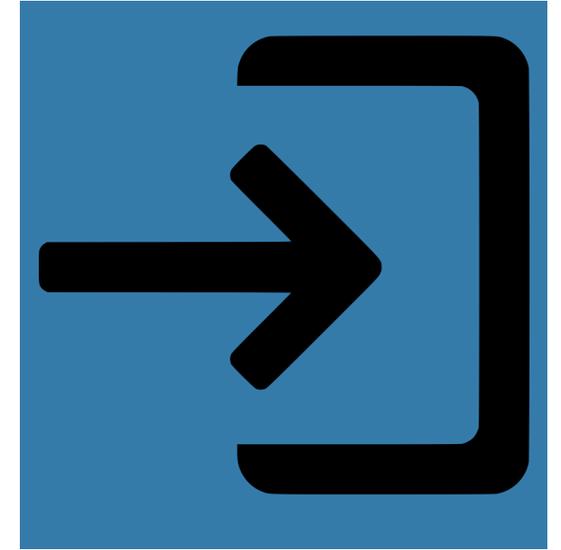
**Structural Inefficiency**



**Lack of single goal**



**Synthesize Information**



**Leadership Turnover**

# What if we used machine learning to optimize public policy?

---

Decades of  
Economic  
Data

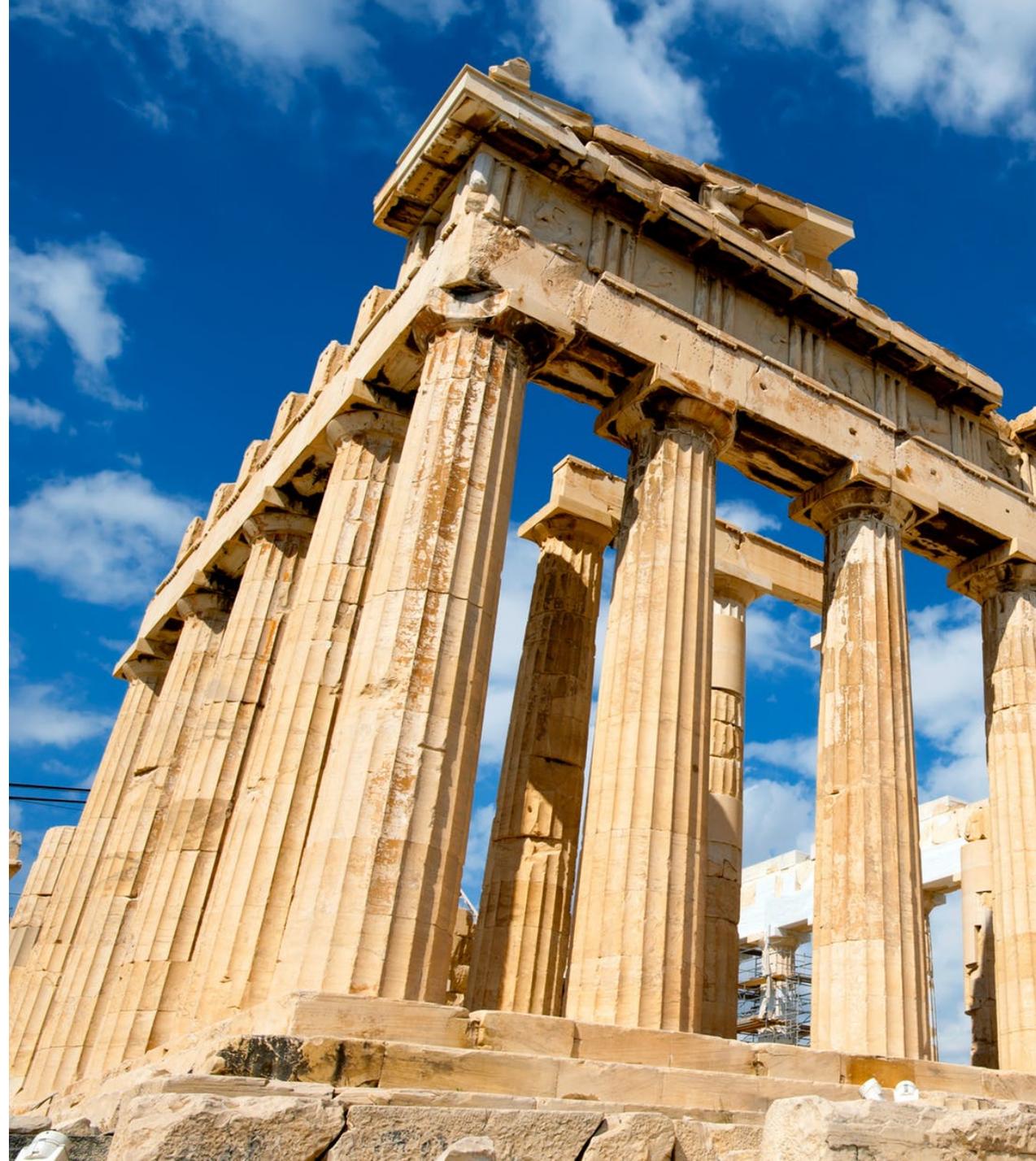
Normalized  
Policy Data

Personalized

Collaborative

Transparent

Reinforcement Learning



# Data-Driven Public Policy Analysis Is Not New

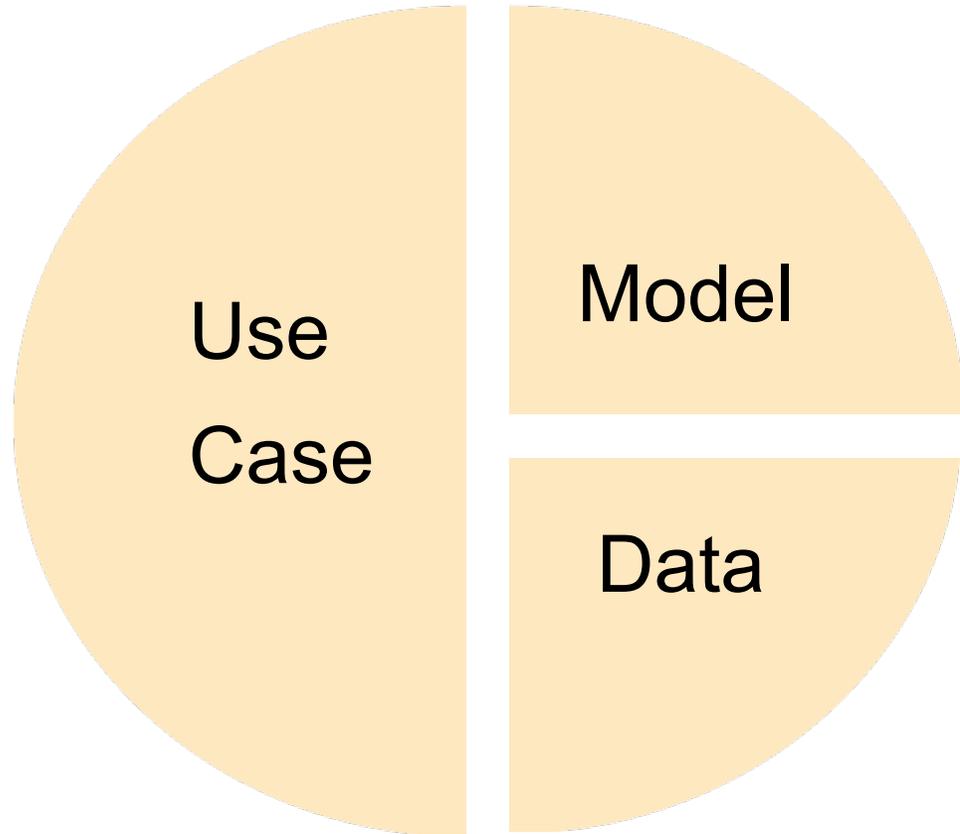
- Causal Inference
- Counterfactual analysis
- Intuitively, what would have happened if the policy (or, treatment) had not been applied?
- Can we convince ourselves that the two groups were nearly identical otherwise?

	Before	After
Treatment: Illinois	100	250
Control: New York	100	150

Did the treatment *cause* this difference?

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_T X_T + \dots + \epsilon$$

# Learning Theory Fundamentals



Machine Learning



Reinforcement Learning

# Mathematically speaking

## Bellman Equation for Reinforcement Learning

$$\underline{U(s)} = \underline{R(s)} + \underline{\gamma} \max_a \sum_{s'} \underline{T(s, a, s')} \underline{U(s')}$$

Utility per state,  
or value

Reward per state,  
a real number

Discount factor

For each possible  
adjacent state

Transition value

Recursive call  
on utility function

Available action  
Current state  
Adjacent state, iterable

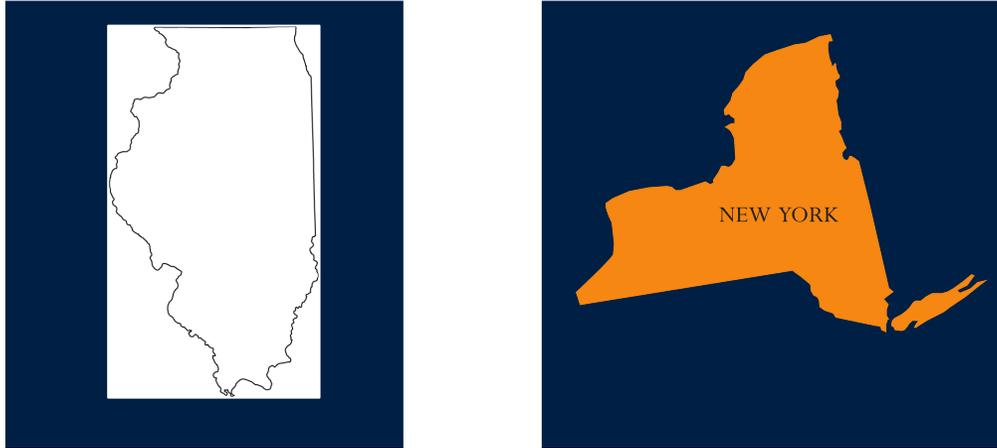
Available action

Current state

Adjacent state, iterable

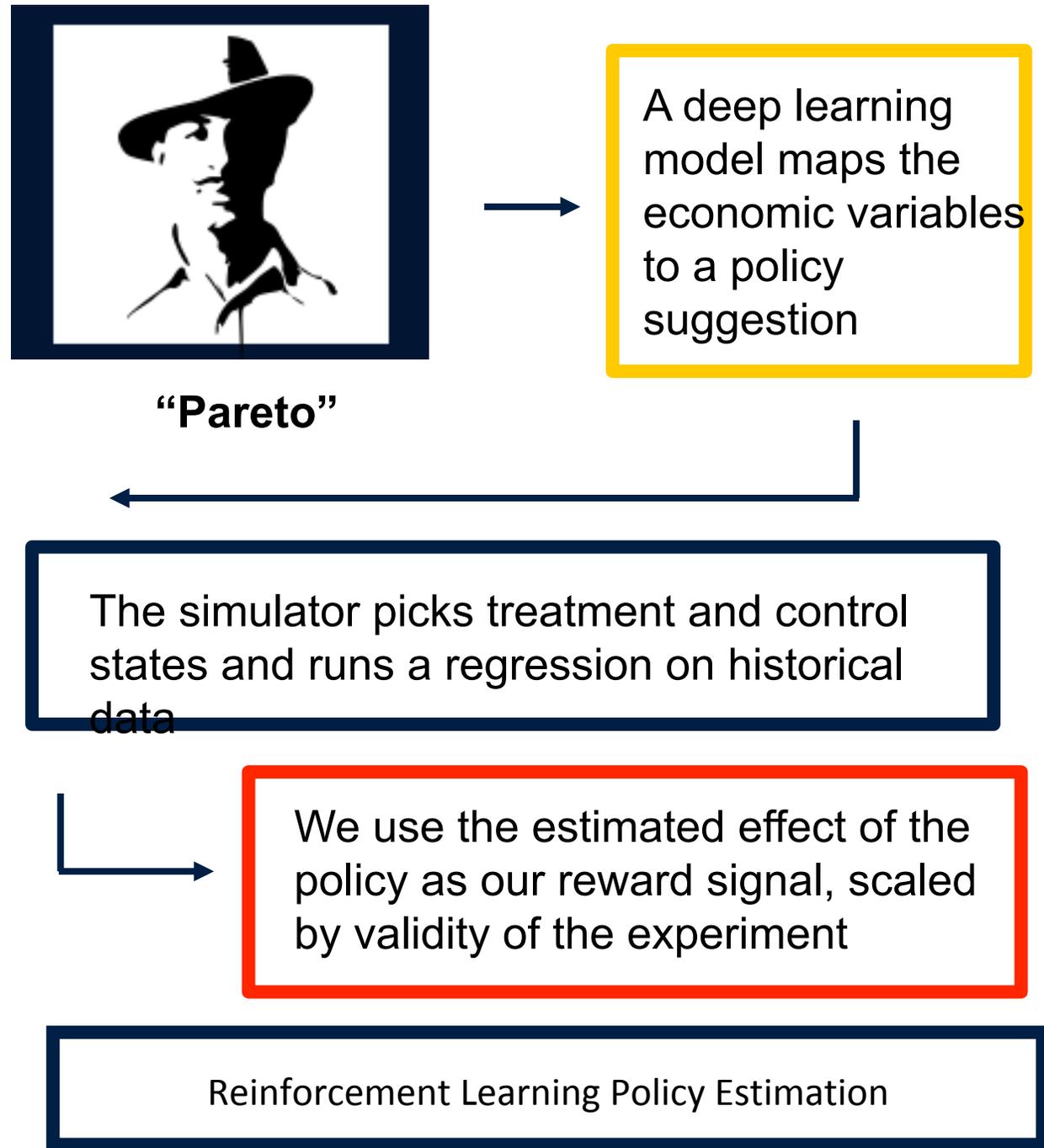


# Our reward function



- Ask, are they similar? T-test
- Use logical reasoning
- Eventually, scale with another ML model using data labelled by experts

Causal Inference



But how do we pick  
the right way to optimize?

# Philosophical Foundations

Utilitarianism

Personal Value

Egalitarianism

Equality

Kantian Rights

Universal  
Rights

Libertarianism

Freedom

## Pareto Improvements

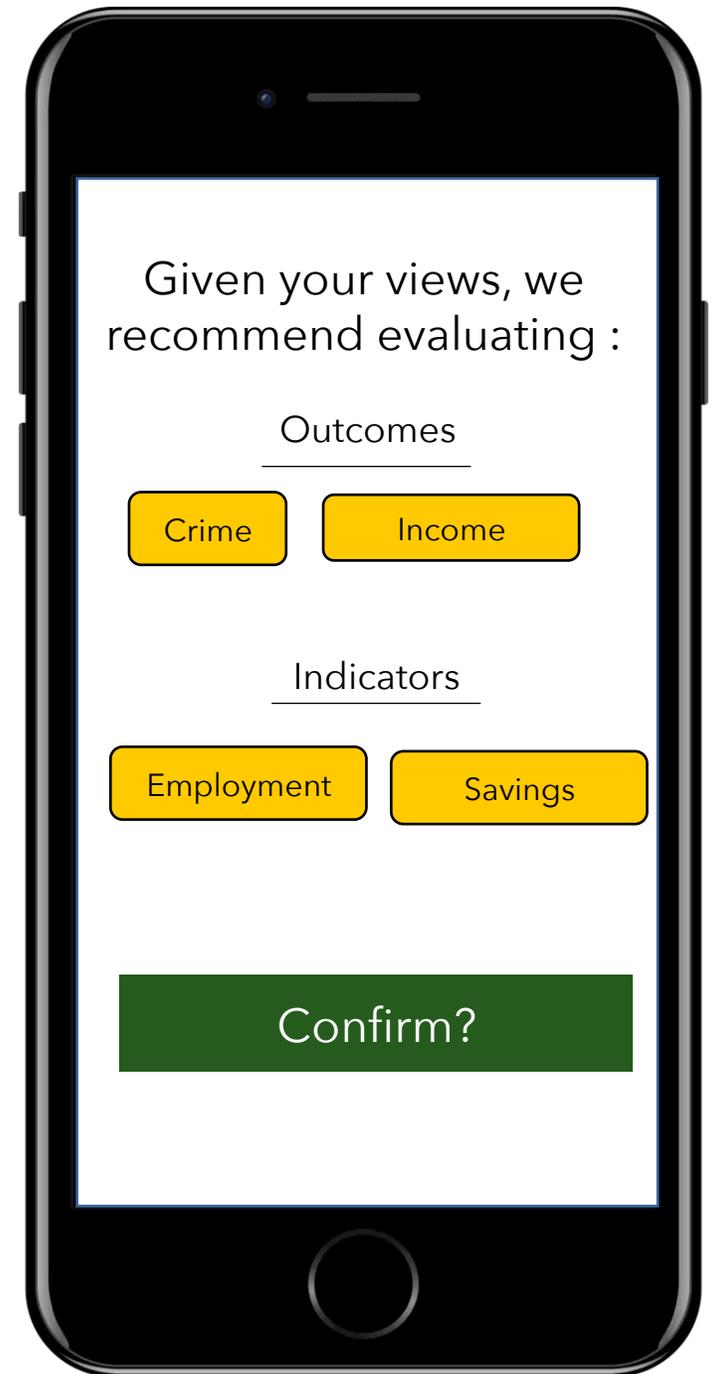
Improve at least one person,  
without making anyone worse off

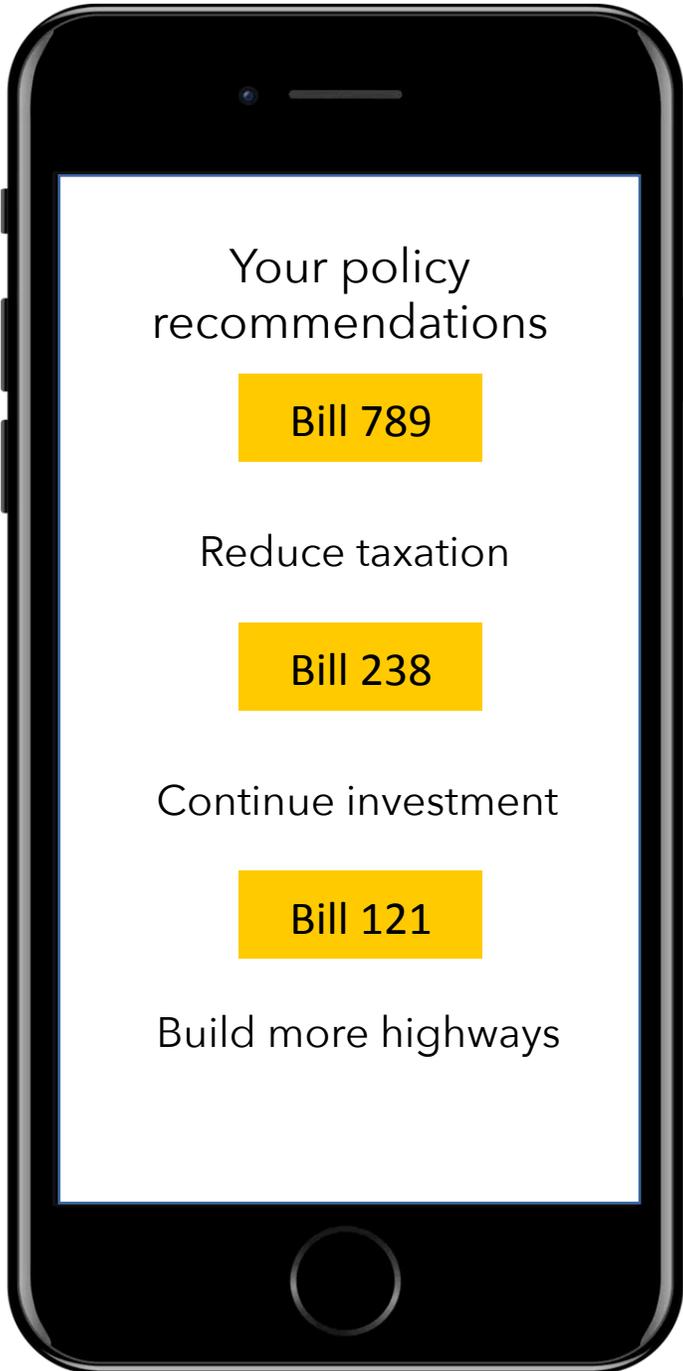
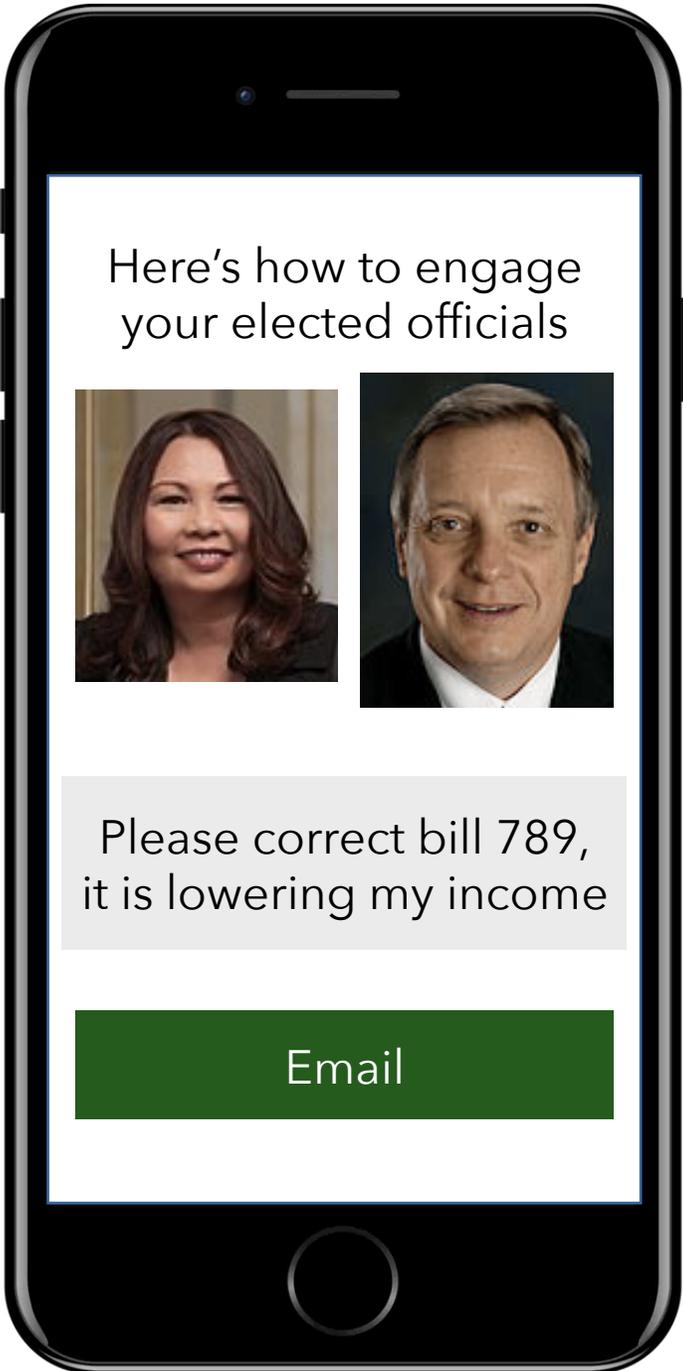
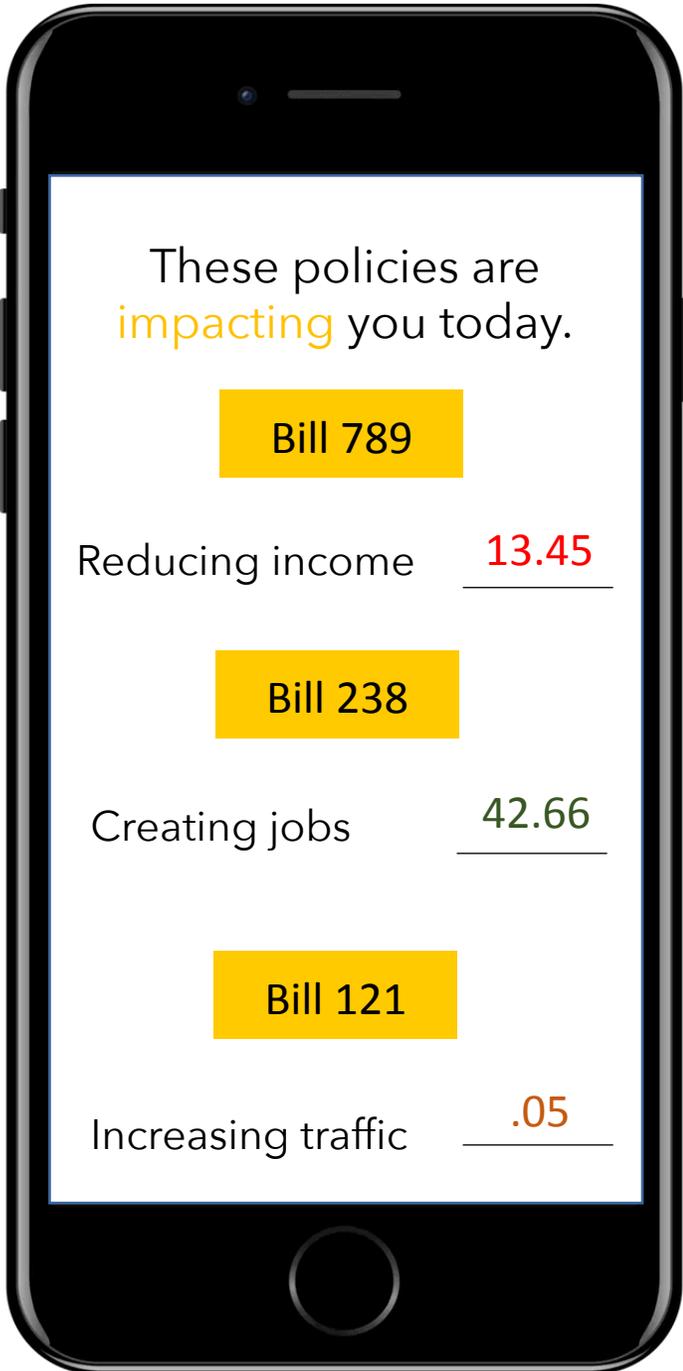


# **There is no single best optimization strategy**

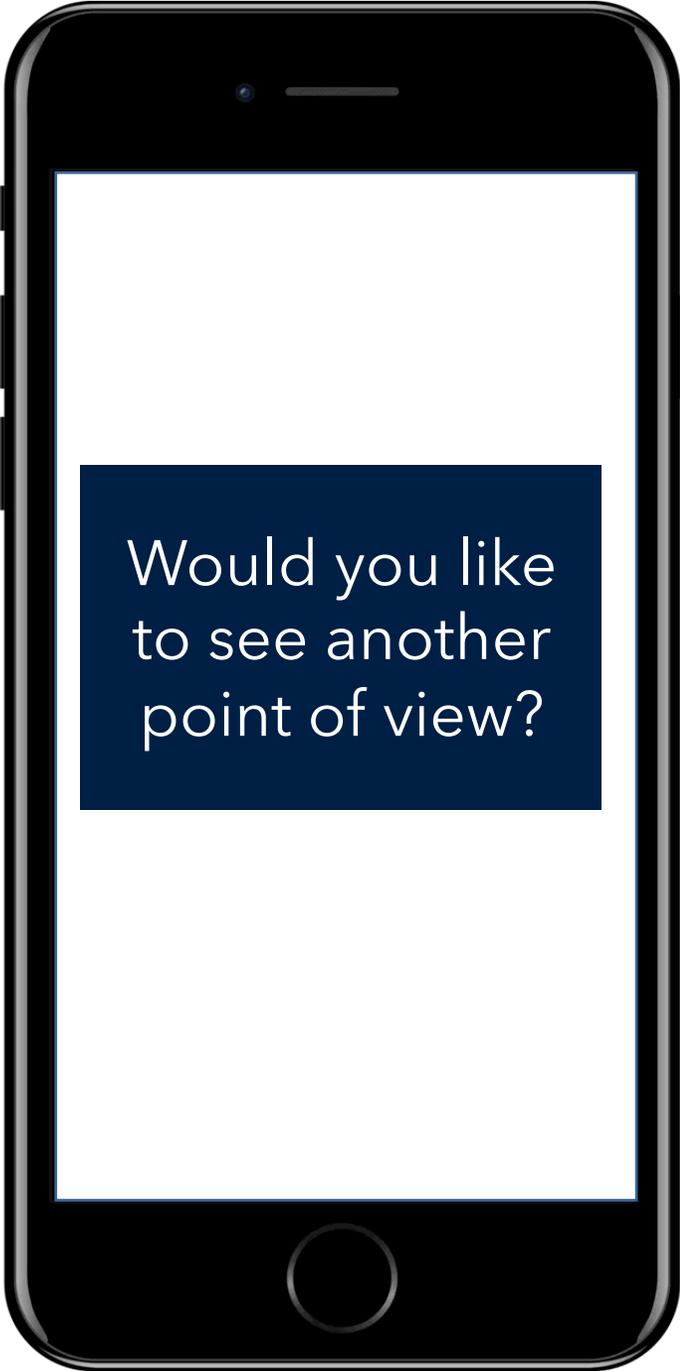
---

**What we can do is use data  
to automatically suggest policies  
based on user-defined preferences**

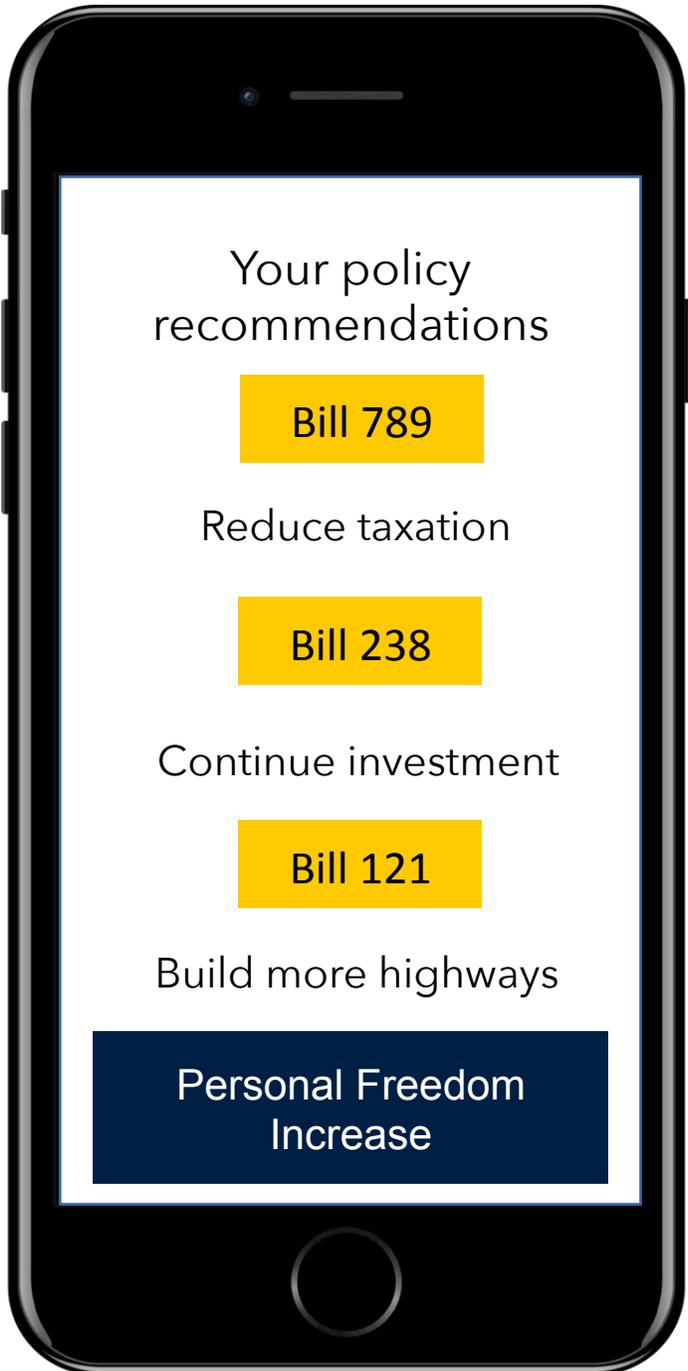




What if we could step into  
someone else's shoes?

A black smartphone is shown vertically. The screen is white and features a dark blue rectangular box in the center. Inside the box, the text "Would you like to see another point of view?" is written in white. The phone has a circular home button at the bottom and a small sensor at the top.

Would you like  
to see another  
point of view?



Your policy recommendations

Bill 789

Reduce taxation

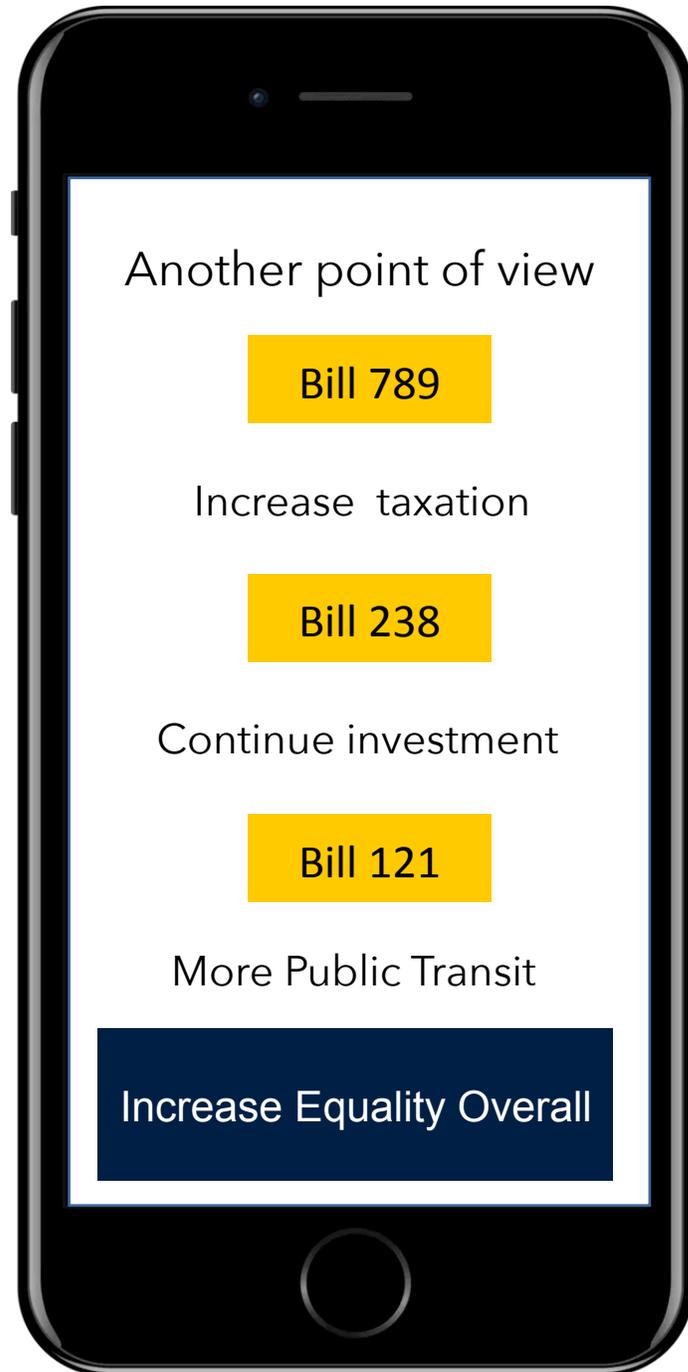
Bill 238

Continue investment

Bill 121

Build more highways

Personal Freedom  
Increase



Another point of view

Bill 789

Increase taxation

Bill 238

Continue investment

Bill 121

More Public Transit

Increase Equality Overall

Technically speaking:

```
for ism in philosophical_frameworks:
```

```
    utility = define_utility(ism)
```

```
    data = update_data/utility)
```

```
    model = get_pareto(data)
```

How should we handle air traffic delays?



## Utilitarianism

Do whatever increases overall utility

- Different people value timeliness differently
- Need multiple ways of defining utility for diverse stakeholders
- Use testing and surveys to get a numerical estimate for how different people value certain outcomes

## Egalitarianism

Do what increases overall equality

- Don't prioritize airline status travelers
- Don't let people pay more for perks
- Don't do special favors
- Treat each traveler, airliner, and airport the same

## Kantian Rights

Uphold human rights

- Uphold the human sanctity of travelers
- Provide food, lodging, respectful notice
- Make reasonable attempts to avoid delays

## Libertarianism

Preserve Freedom

- Let people pick for themselves
- Don't automatically make decisions for travelers
- Let travelers switch across airlines
- Ensure freedom of airlines and airports

There is fundamental overlap between  
the philosophical frameworks.

This overlap can be scaled by reward functions

# There is no single right answer

## We need a computational system that can:

- Synthesize different points of view
- Weight these based on criteria, like population size
- Be transparent, collaborative, timely
- Change with the times

To efficiently support existing governing bodies

An aerial photograph of a mountain range with a valley and a lake. The mountains are covered in green vegetation, and the valley floor is a mix of brown and green. A small lake is visible in the center of the valley. The sky is a clear, pale blue.

# Thank you!

Emily Webber | Amazon Web Services | LinkedIn  
[effective-policies@amazon.com](mailto:effective-policies@amazon.com) ← email me to collaborate!

*Disclaimer!*