Lecture 1

Markets, Mechanisms and Machines

David Evans and Denis Nekipelov

What this course is about

- Interaction between market design and algorithms

What this course is about

- Market design with data
  - Learning about participating agents
  - Track changing agent behavior
  - Optimize market (revenue, welfare, etc.)
    - Optimization of existing mechanisms
    - Development of new mechanisms

Market design in practice

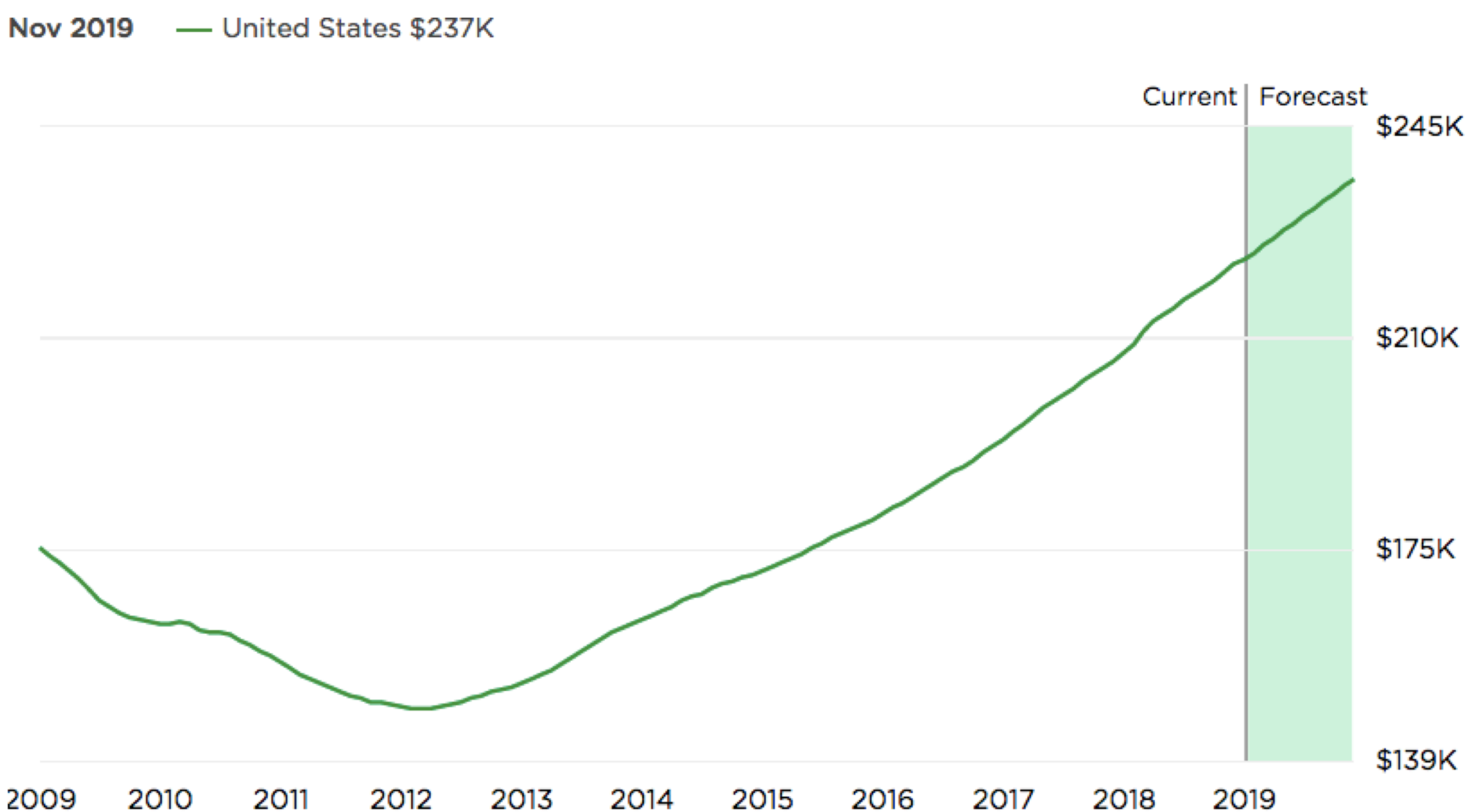
- Establish goal (typically, optimal allocation)
  - Which user sees the ad?
  - Which hotel shows up first in search?
  - Whose code is executed first?
  - Which customer gets what price for an airline ticket?
- Have model of agent behavior
  - Predict how agent responds to mechanism
  - Predict how agent responds to other agents

Market design in practice

- Why having model is important?
  - Not sufficient to observe how agents behave currently
  - Mechanism designer needs to anticipate how agents behave in the settings that have not been yet observed
- Typical task of mechanism designer: observe how agents behave under mechanism A and use this observation to predict how they will behave if instead mechanism B was used

Assistance to mortgage borrowers

- Financial crisis in 2008 resulted in drastic drop in real estate prices (source: Zillow.com/research/data)



Assistance to mortgage borrowers

- Price drop resulted in cascading effect
  - Mortgage-paying households observe dropping real estate price
  - The value of their mortgage becomes close or below the market price of the house
  - When monthly cost of mortgage combined with depreciation of property values exceeds alternative cost, optimal choice is to walk out
- Collectively, this individual optimizing behavior drives real estate prices further down

Assistance to mortgage borrowers

- Mortgage defaults were catalyst of crisis
- Goal of the Federal Reserve: construct measures to slow down defaults
  - Tightening of underwriting standards (reduce defaults of new mortgages)
  - Assistance to current borrowers
    - Mortgage principal write-down
    - Mortgage interest subsidy
    - “Payment holiday” for underwater mortgages

Assistance to mortgage borrowers

- How can we predict which policies are more effective?
  - Cannot use the past: similar crisis never happened
  - Most proposed policies have never been tried
- Need a model of borrower’s behavior
  - Borrowers have preferences (utility) over owning the house vs. renting
  - They engage in “selfish” dynamic optimization and compare expected value of future flow of utility from owning vs. renting
- Resulting prediction is called counterfactual

Assistance to mortgage borrowers

Probability of “eventual default” for subprime mortgage borrowers (FICO <600) under various policies

Scenario	Counterfactual		Baseline	
	Default	Prepay	Default	Prepay
Scenario 1: Home Price Decline	0.1129	0.4717	0.0983	0.4944
Scenario 2: Home Price Increase	0.2204	0.2635	0.2696	0.2177
Scenario 3: Higher Credit Quality	0.2329	0.2232	0.2696	0.2177
Scenario 4: 10% Principal Writedown	0.1167	0.5783	0.1482	0.5536
Scenario 5: 20% Principal Writedown	0.0914	0.5938	0.1482	0.5536
Scenario 6: LTV Cap at 0.8	0.1316	0.5692	0.1482	0.5536
Scenario 7: LTV Cap at 0.9	0.1458	0.556	0.1482	0.5536

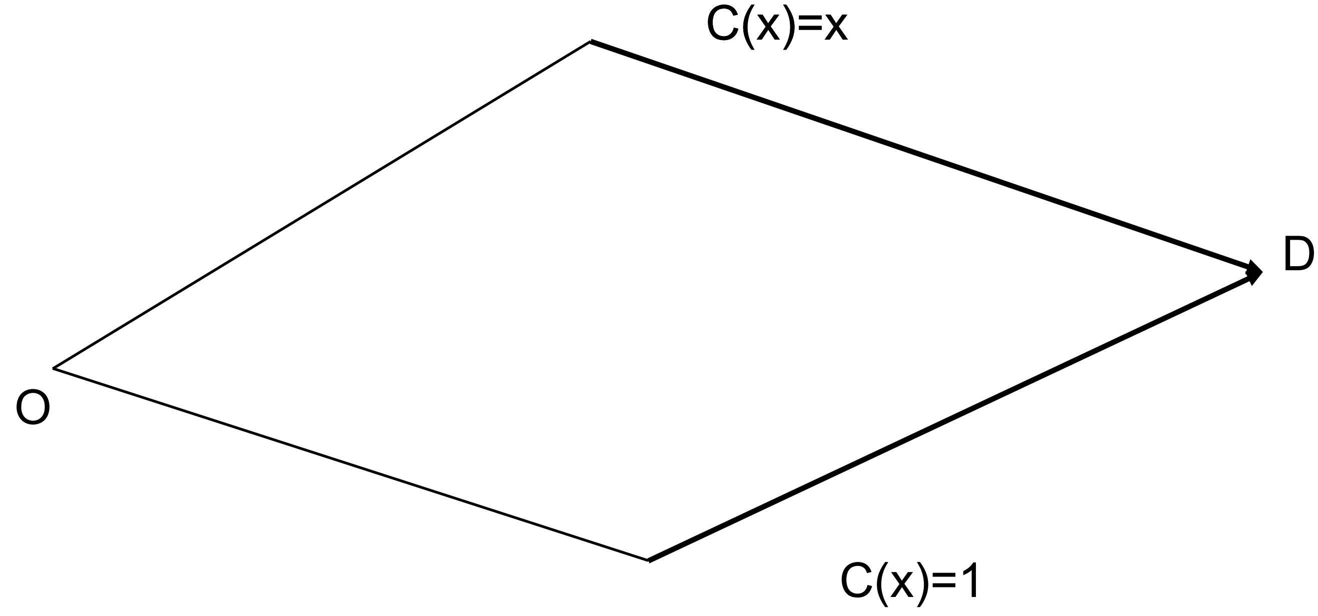
The first (second) column reports predicted probability of eventual default or prepay by December 2009 under the specified counterfactual (baseline) scenario. The baseline default and prepayment probabilities differ across scenarios since different scenarios examine different subgroups. For instance, Scenarios 4-7 examine all loans while Scenario 1 examines loans originated in 2004. For Scenario 1, the cumulative housing price growth rate over the sample period is 11.8% under the baseline case while it is only 5.6% under the counterfactual case. For Scenario 2, the cumulative housing price growth rate over the sample period is 4.5% under the baseline case while it is 20.9% under the counterfactual scenario.

Market design in practice

- In many environments agents compete for the same resources and goal of mechanism designer is to allocate them "fairly"
  - Allocate items to agents most needing them
  - … so that no one is willing to "switch places" with any other agent
  - Can be a difficult task
    - Unlike physical systems, agents respond both to changes in mechanism and to each other's behavior
    - Changes in market mechanism can affect **market equilibrium**

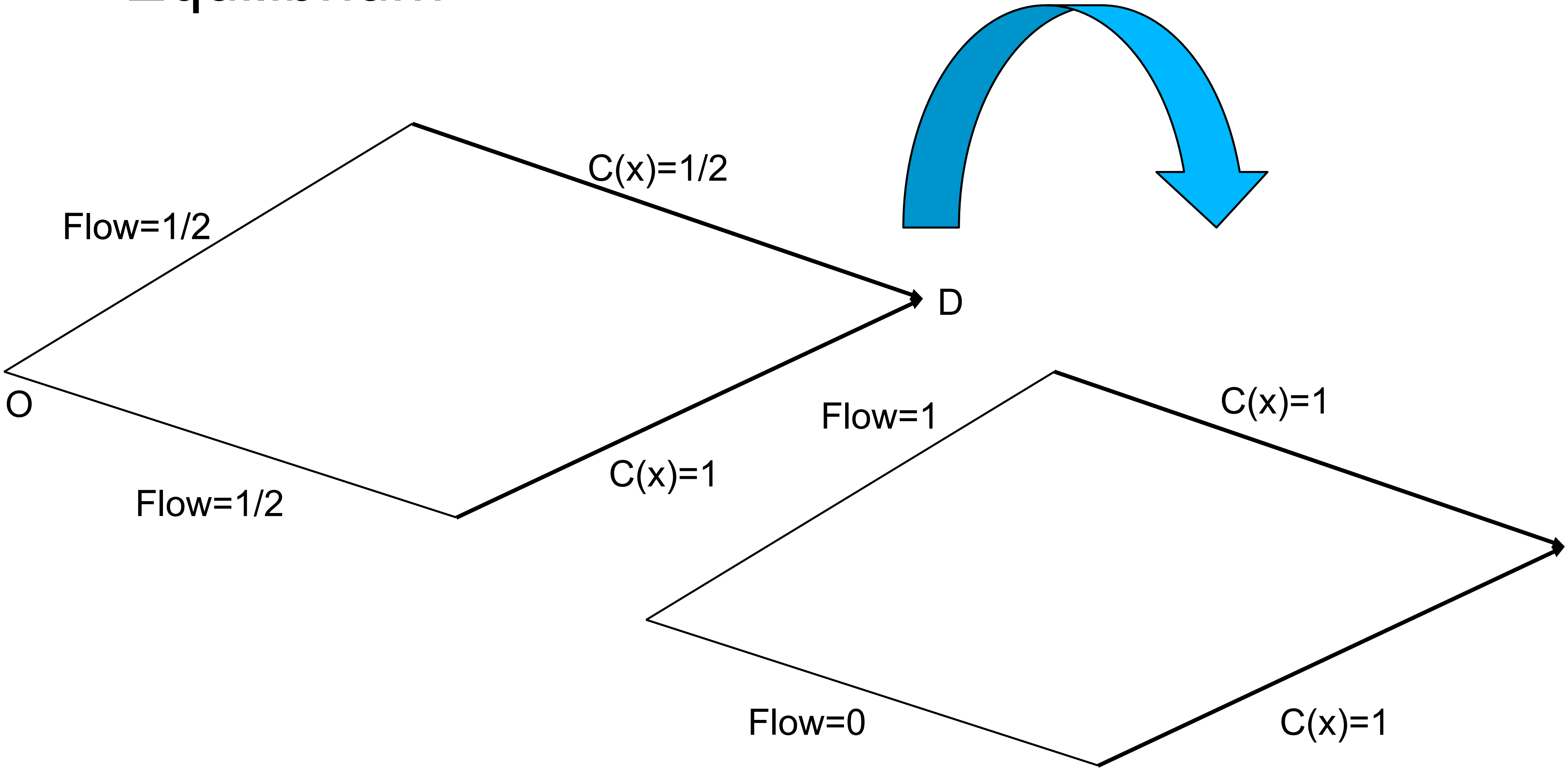
Selfish routing

- “Small” agents (relative to overall volume)
- Each agent wants to optimize path
- Traffic flow is at Nash equilibrium if it based on minimum cost paths
- Total cost of flow is equal to the sum of costs of all agents

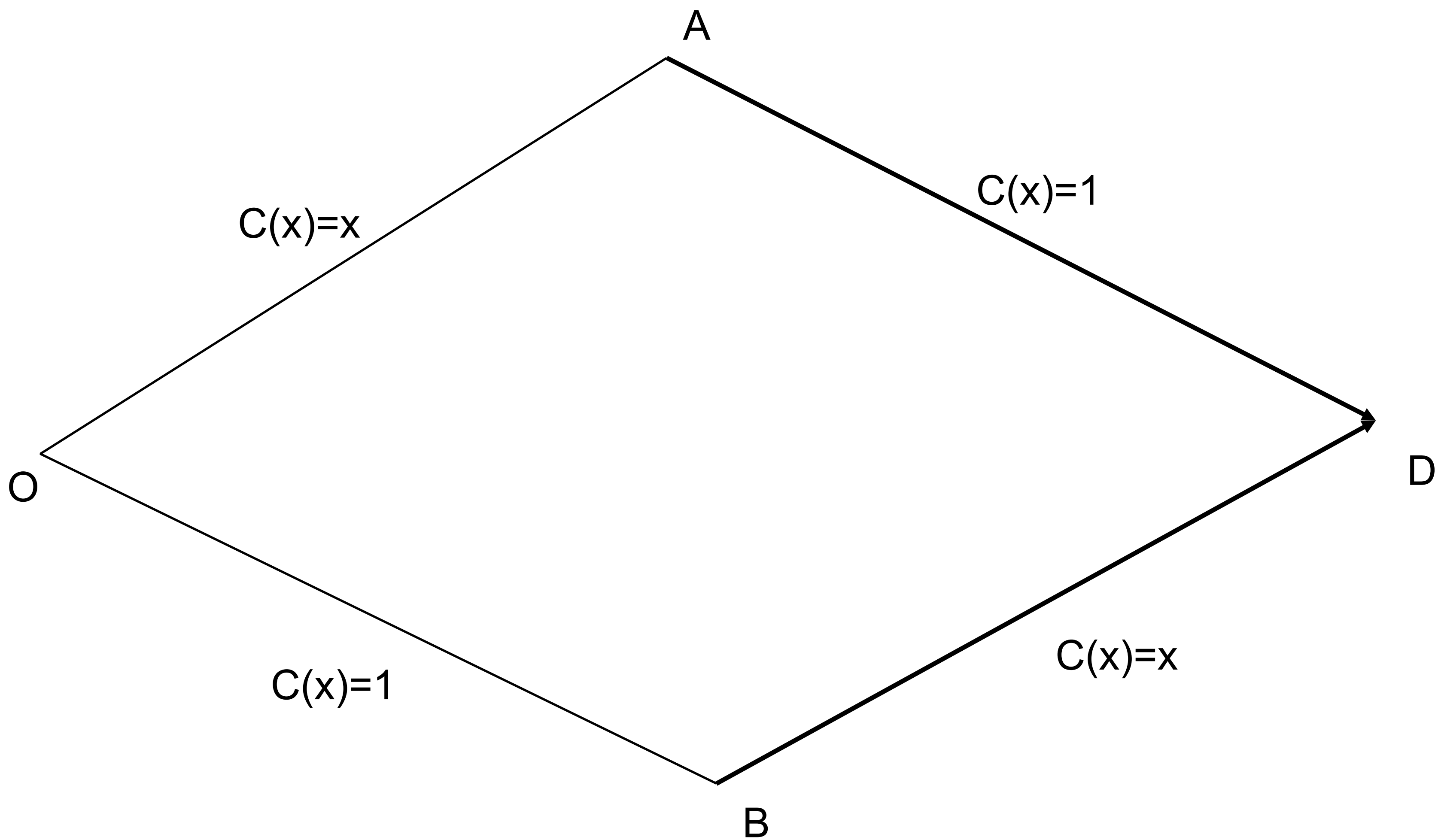


Selfish routing

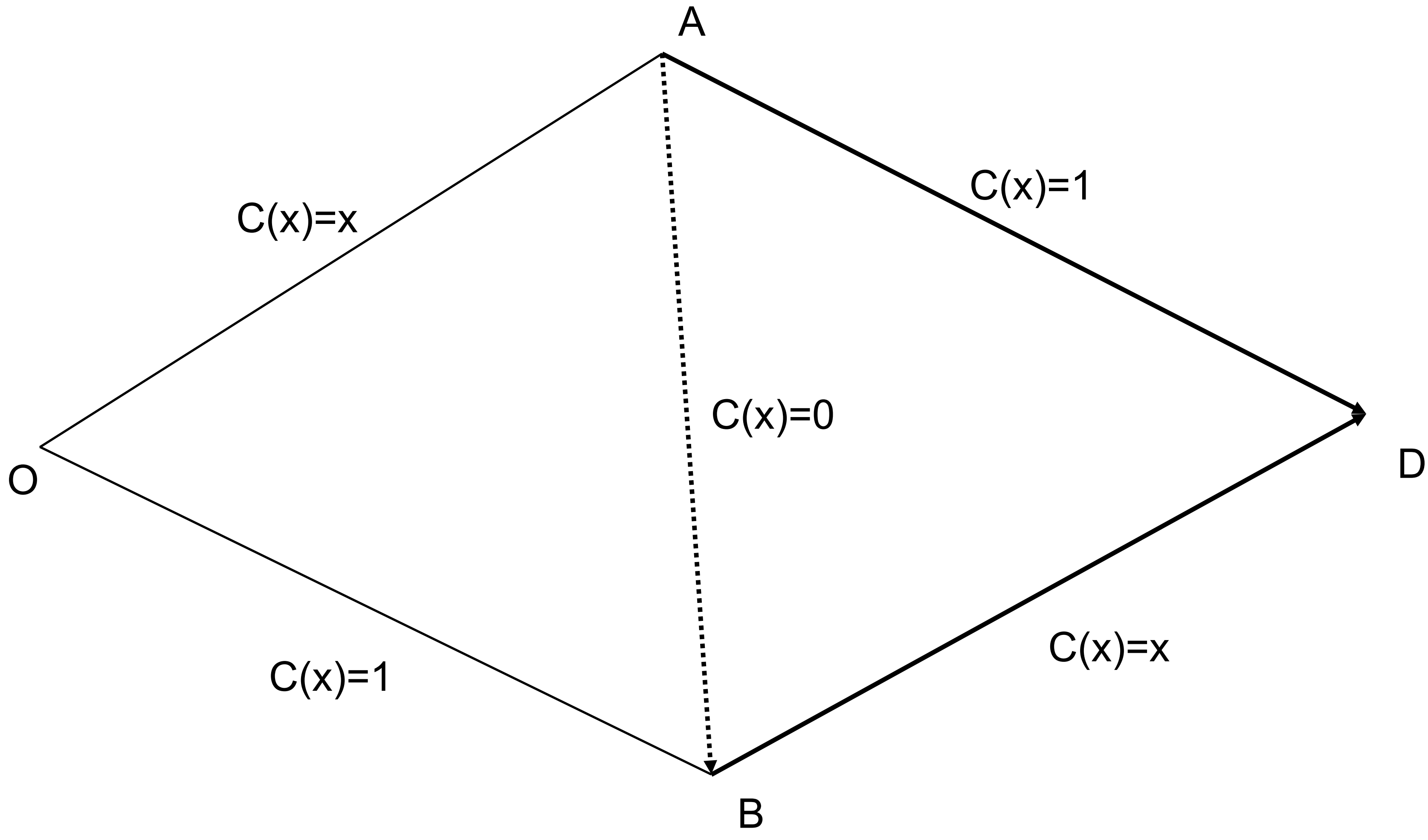
- Equilibrium



Selfish routing



Selfish routing



Selfish routing

Modeling and causal inference

- Not sufficient to know how aggregate traffic flow responds to changes in cost of routes
  - Need a model for agents' behavior
  - Need to quantify how costs affect behavior
- This requires to quantify the **causal** impact of changes in the mechanism on agents

Causation vs. Correlation

"Do you think all these film crews brought on global warming or did global warming bring on all these film crews?"

Causation vs. Correlation

- Does a change in X really cause a change in Y? Or do they just co-vary?
- To evaluate policies and test theories, we need to establish causation.
- But in the real world, correlation and causation are often very difficult to separate.
  - Does drinking red wine reduce the risk of a heart attack?
  - Does watching reality TV cause stress?
  - Does smoking cigarettes cause cancer?

Causation vs. Correlation

- We observe a positive relationship between crime and the number of police officers
  - Is it because police officers create crime?
  - Or (more likely!) is it because more police officers are assigned to more troublesome neighborhoods?
- We observe that unemployed people who attend a job training program wait for shorter periods before finding a job.
  - Is it because the program helped them, or is it because those who joined the program are the most skilled/motivated ones, so that they would have waited less anyway?
- So how do we establish causation?

Causation vs. Correlation

- Ideally, we would like to have **experimental data**: for example, two identical plots of land, where the same crop is cultivated, using the same techniques, **but** where different fertilizers are used.
- Then if the outcome of interest (yield per acre, for example) is different between the two plots, we can safely infer that a different fertilizer *causes* differences in the average yield
- Classic example: clinical trials (compliance?).
- **BUT**, in most cases, all what we have is **observational data** (you can't have the **same** person with **and** without college, or the **same** economy with **and** without a tax cut...).
- In general, to find answers, we need a model & Econometric/Machine learning techniques

Modeling vs. Prediction

- Model explains how vector of inputs X (features or regressors) results in outcome Y
  - Typical mathematical model: joint probability distribution F(Y,X)
    - Y and X are **assumed** to be random variables
    - Joint distribution of Y and X is induced by agents' behavior (e.g. Y is the index of path chosen and X is its cost)
  - Estimation of the model requires us to take a sample of data $\{Y_t, X_t\}_{t=1,\ldots,T}$ and construct distribution G(Y,X) using this sample, such that with high probability G and F are close
  - Constructing such G can be (very) difficult

Modeling vs. Prediction

- Instead of fully modeling relationship between Y and X we be OK with presenting a “surrogate” $Y^*(X)$ that is “close” to Y
- $Y^*(X)$ is our prediction for Y
  - Need precise definition what “close” means
  - Formalized using *loss function* that penalizes prediction errors
  - One common loss function is square loss:

$$(Y^*(X) - Y)^2$$

  - It is nice, because it is smooth and convex for many choices for $Y^*(X)$

Modeling vs. Prediction

- Goal: for randomly drawn example (Y,X) find $Y^*(X)$ that minimizes *expected loss*
- Mathematically we need to minimize a functional

$$E[(Y^*(X) - Y)^2]$$

with respect to all possible functions $Y^*(X)$

- Note that

$$E[(Y^*(X) - Y)^2] = E[(Y-E[Y|X])^2]+E[(Y^*(X) - E[Y|X])^2]$$
$$= Var(Y) + E[(Y^*(X) - E[Y|X])^2]$$

Modeling vs. Prediction

- Since variance of Y is a universal constant and

  $E[(Y^*(X) - E[Y|X])^2] \geq 0$,

  expected loss is minimized at

  $Y^*(X) = E[Y|X]$
- Function $E[Y|X]$ (conditional expectation of the outcome given regressors/features) is called a regression function
- Estimation of the regression function leads to prediction minimizing the squared prediction error

Regression

- A special important case is when the search for predictors $Y^*(X)$ is restricted to linear functions of regressors:

  $$Y^*(X)=a_0+a_1X_1+\ldots+a_kX_k$$

- Minimization of mean-squared loss

  $$E[(Y- a_0 - a_1X_1-\ldots-a_kX_k)^2]$$

  produces best linear predictor for that loss

- Important advantage: it is smooth and globally convex; has analytic solution for minimum

Regression

- Can regression have causal interpretation?
- Yes, if the true model is a regression model
- Suppose that joint distribution of Y and X is generated by model

$$Y= a_0+a_1X_1+\ldots+a_kX_k+\varepsilon, \quad E[\varepsilon|X]=0$$

- Then

$$E[Y|X]= a_0+a_1X_1+\ldots+a_kX_k$$

- I.e. finding best linear predictor for the mean square loss also finds the true model

Regression

- For model

  $Y= a_0+a_1X_1+\ldots+a_kX_k+\varepsilon,\ \ E[\varepsilon|X]=0$

  linear regression recovers causal impact of X
- Coefficient a_p has interpretation of an impact of change in regressor X_p by 1 holding all other regressors fixed
- I.e. a_p is the causal impact of X_p on the expected value of the outcome variable

Regression vs. Causation

- We can use linear regression methods for causal purposes as long as the true data generating model is linear
- This is referred to as the assumption of *correct specification* of the model
- Correct specification requires that
  1. The true model generating Y is linear in X
  2. X observed in the data includes all factors impacting Y

Regression vs. Causation

- There is no guarantee for correct specification
  - Can be verified in out of sample tests: randomly split your data into two subsets and use one to estimate the model and another one to compute prediction error
  - Sometimes provided by formal theory
- Without correct specification results do not have causal interpretation
  - Still produce best linear prediction and my still be useful

Regression vs. Causation

- With data, regression minimizes empirical mean-squared error

  $$\sum_{t=1,\ldots,T}(Y_t - b_0 - b_1X_{1t} - \ldots - b_kX_{kt})^2$$

- First order condition for b's leads to a system of linear equations
- Implemented in all existing statistical packages
- Requires additional technical restriction of no perfect linear dependence between X's (otherwise there are multiple solutions and optimization will be "stuck")

Properties of linear regression

- If the linear regression model is correctly specified, and distributions of X's are "well-behaved"
  - Each b_k is close to a_k with high probability as the sample size T gets large
  - The distribution of $\sqrt{T}(b_k - a_k)$ is approximated by normal distribution centered at zero, i.e. we can compute the probability which a_k are likely.
  - Therefore, we can formally test if specific value a_k is compatible with observed data. E.g. can $a_k = 0$, i.e. regressor X_k has no causal impact on Y?