Carnegie Mellon University

Heinz College

INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

CARNEGIE MELLON UNIVER... PITTSB... ...NA

Behaviors of People and Systems: Decision Problems And Data Driven Analysis

Ramayya Krishnan
Heinz College of Information Systems and Public Policy
Carnegie Mellon University
rk2x@cmu.edu

Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Outline

- 2 decision problems
  - Policy interventions, economic activity and health risks
    - SIR model, choice models to support counterfactual analysis
  - Supply chain shocks and shortages
    - Understanding and developing policy response to supply chain risks and failures

Opportunities for Collaboration

- Between AI, Data Science and the OR communities at three layers
  - Methods
    - ML, high dimensional statistics, causal inference and optimization
  - Data
    - Creating benchmark data sets (like the ARPA-E Power grid) for societal systems like critical supply chains will be hugely beneficial
  - Policy
    - Opportunity to work on problems that have direct policy impact

Data driven Covid Policy Analysis

With Beibei Li (CMU) and Zihao Yang (CMU)

Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Economic Activity and Health Risks

- Project designed to understand the interaction between economic activity and covid health risks at points of interest (restaurants, malls, …)

- Opportunity to design policy interventions that preserve economic activity where possible while reducing health risks

- Access to unique data sets on economic data, mobility behaviors and disease state by zip code
  - State-wide spending data across NAIC codes and different modes (online vs. offline) from a super regional bank by zip code
  - Mobility data from a GPS trace provider by zip code
  - Data from the State about covid cases by zip code

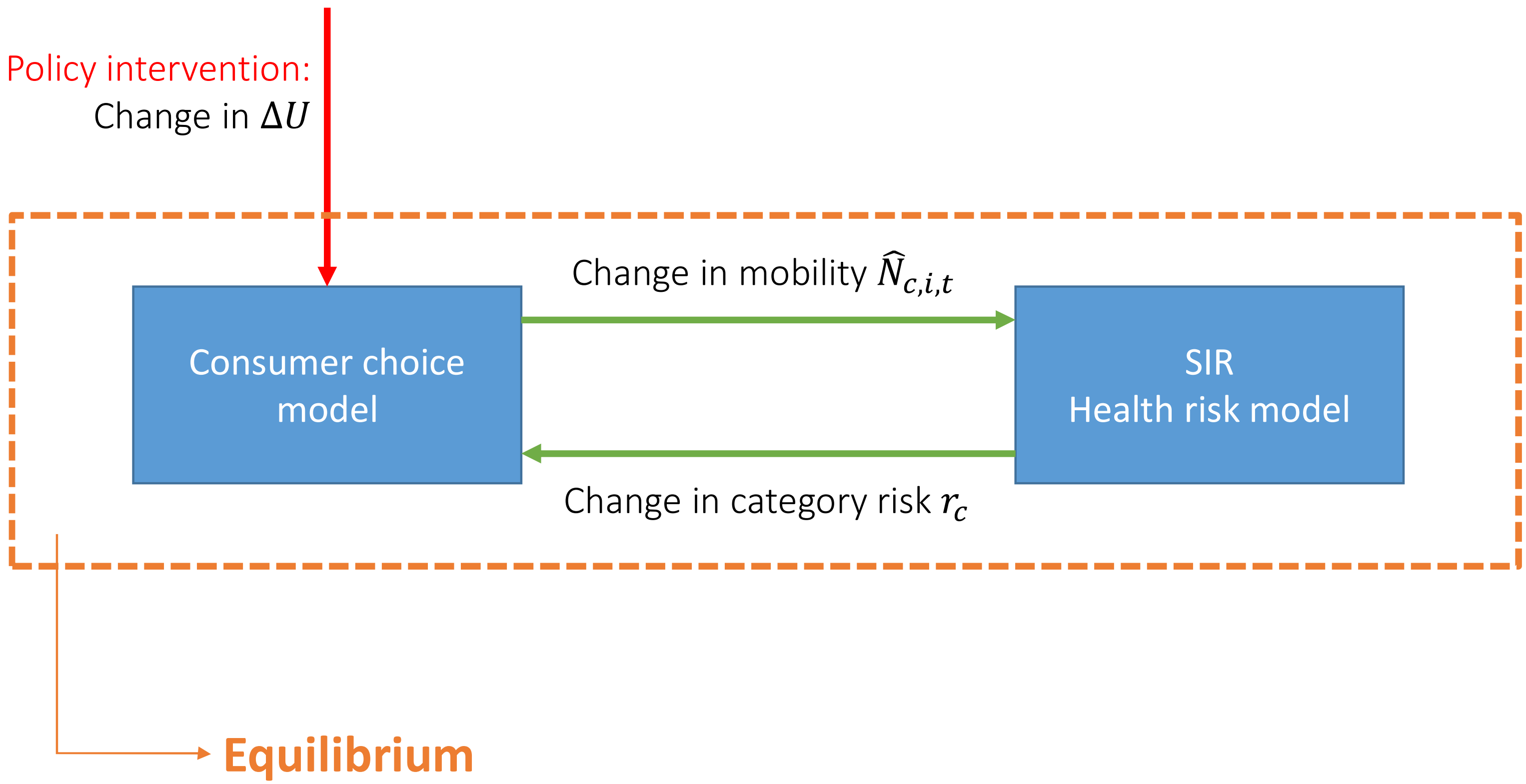
Questions of interest

- What is the differential risk of acquiring covid in a place of economic activity (e.g., restaurants vs. Grocery stores)?

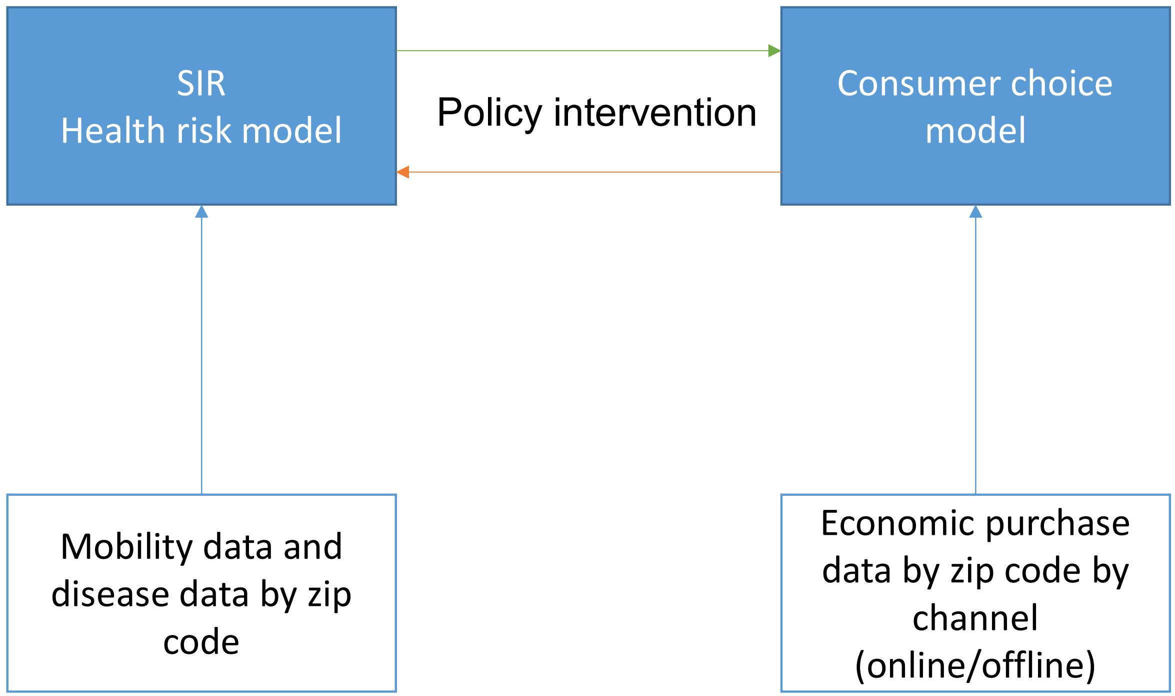
- How would perceived risk of exposure change shopping channel choices (offline vs. Online)?

- How can policy interventions introduced to minimize health risk while retaining economic activity be evaluated?

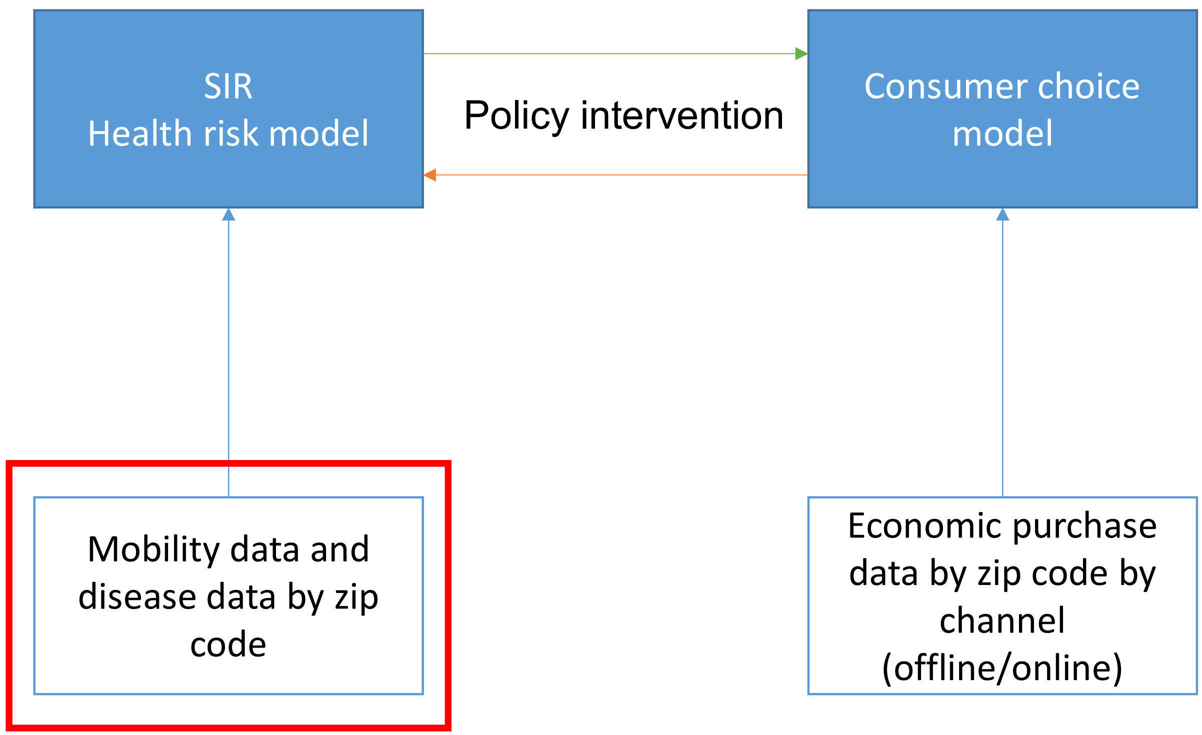
Counterfactual experiment

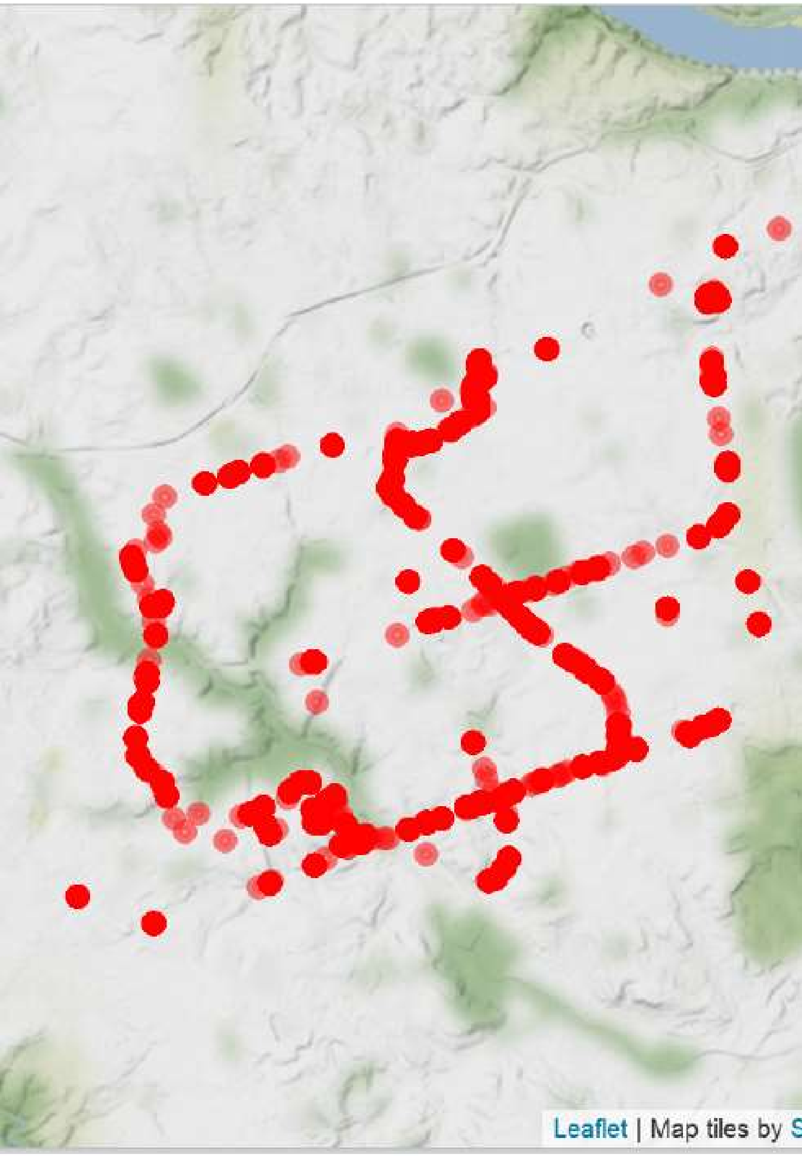


Interactions between consumer choice and health risk



Interactions between consumer choice and health risk





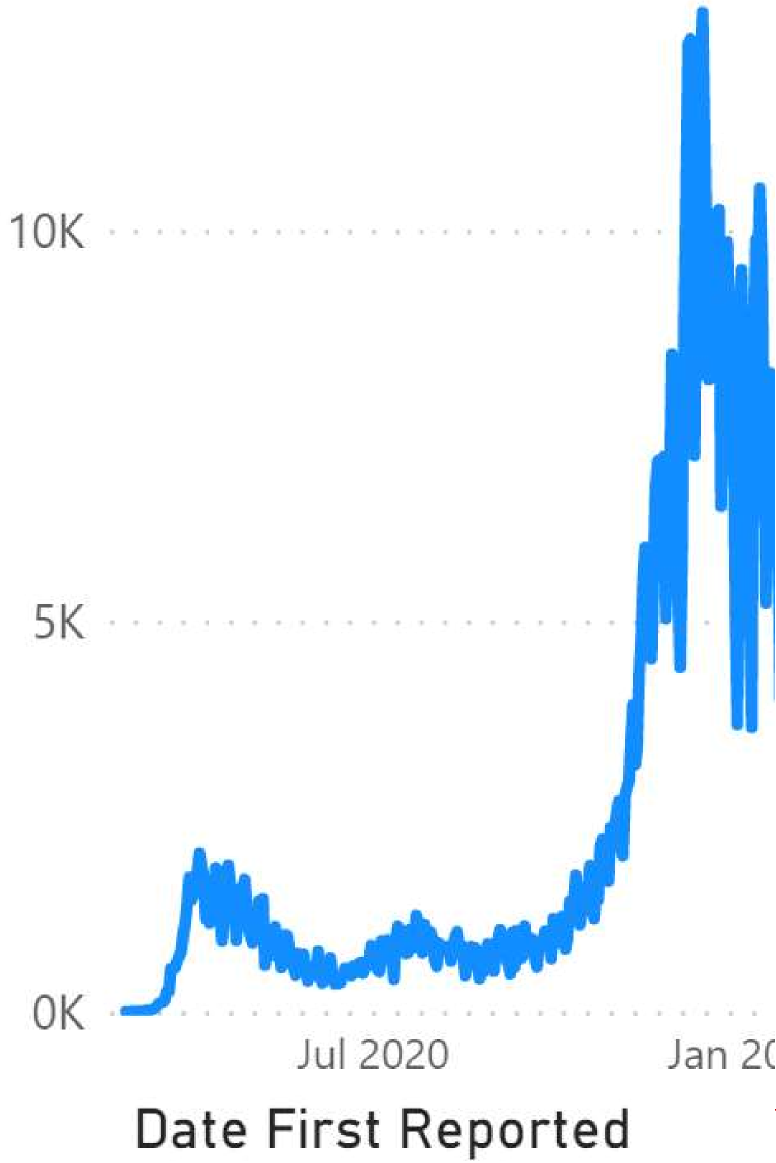
Dataset description

- **Mobility dataset: Safegraph/Mode data**
- Record: ‘Date’, ‘UserID’, ‘Latitude’, ‘Longitude’, …
- From Jan. 2020 to Nov. 2020
- 10B records in 2020 in PA, more than 300B records in the US

https://www.safegraph.com/blog/safegraphs-data-sourcing-process

Dataset description

- **Point of interest dataset: SafeGraph POI Dataset**
- Record: ‘Placename’, ‘Category (SIC code)’, ‘Latitude’, ‘Longitude’, …
- From Jan. 2020 to May 2021
- 170k POIs in Pennsylvania

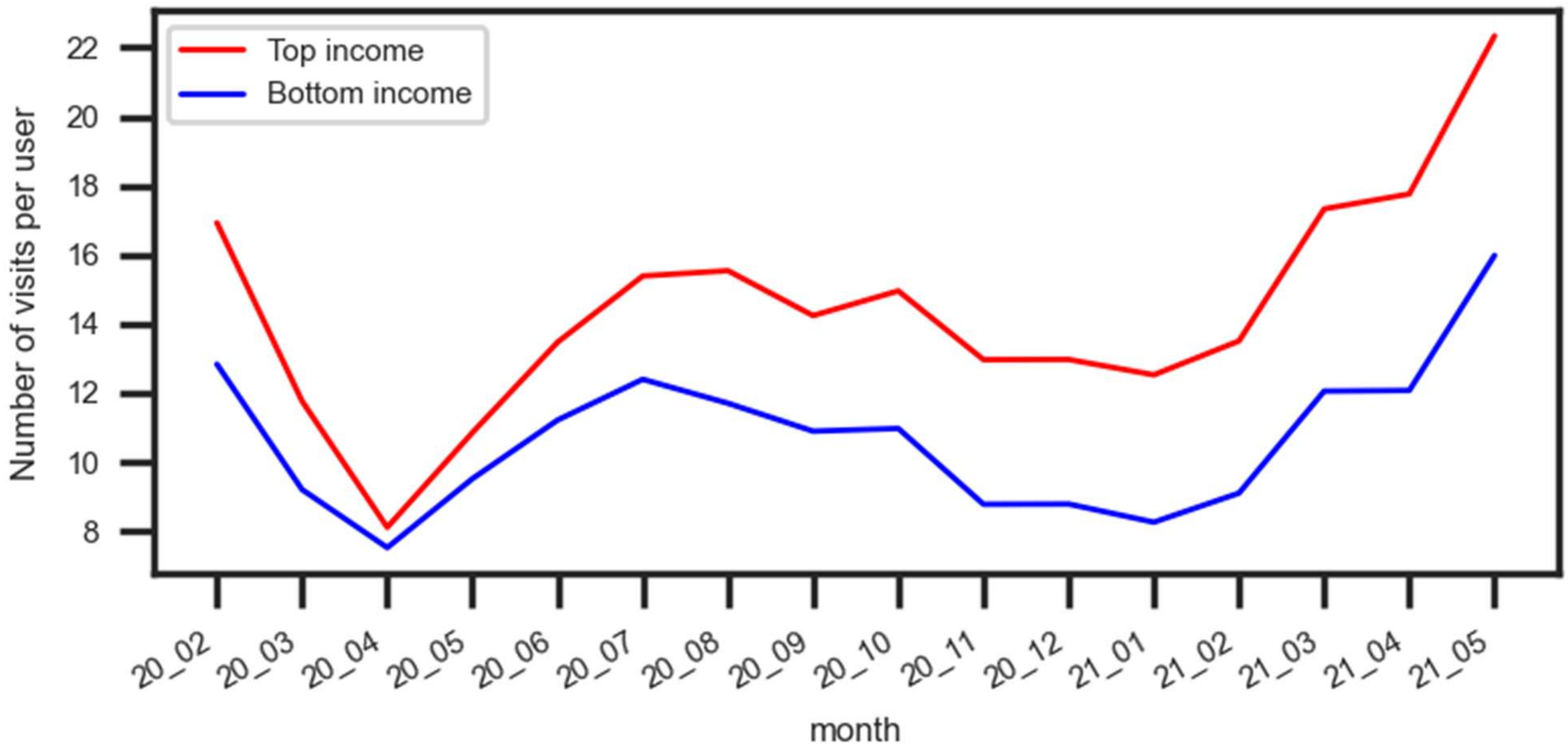


Dataset description

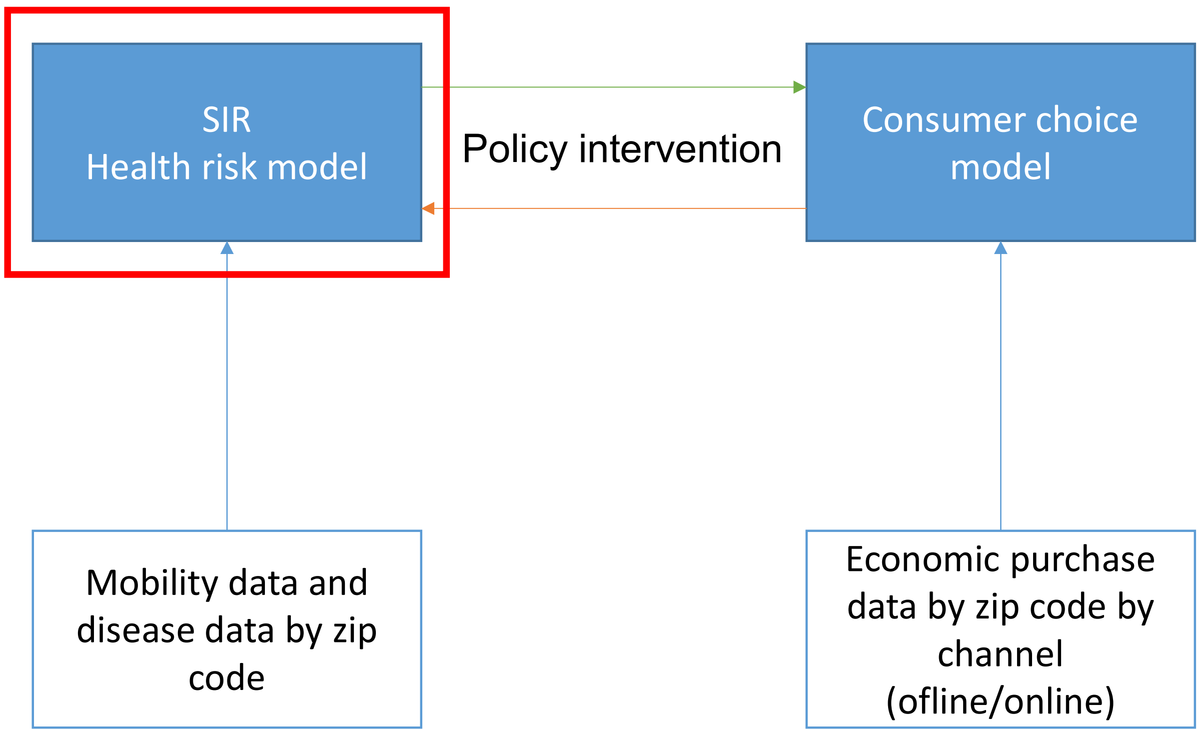
- **Covid-19 case dataset from the State by Zip Code:**
- Record: ‘Date’, ‘Zipcode’, ‘Number of case’
- From Mar. 2020 to May 2021

Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Mobility behavior of different income groups



Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT


SIR
Health risk model
Policy intervention
Consumer choice model
Mobility data and disease data by zip code
Economic purchase data by zip code by channel (ofline/online)

Individual SIR model

- Individuals have three states:
- S: Susceptible, people who can potentially contract the disease
- I: Infectious, people who have been infected and are capable of infecting susceptible individuals.
- R: Removed, people who have been infected and have recovered or died.

- We assume each individual has a set of probabilities p_s, p_i, p_r for each state.
- For a susceptible individual, each visit to a POI could result in contracting Covid-19 with some probability.
- For a susceptible individual, there are zip code/CBG level risks based on housing denisty etc. Of contracting associated Covid-19
- We assume events of infection are independent.

Individual SIR model

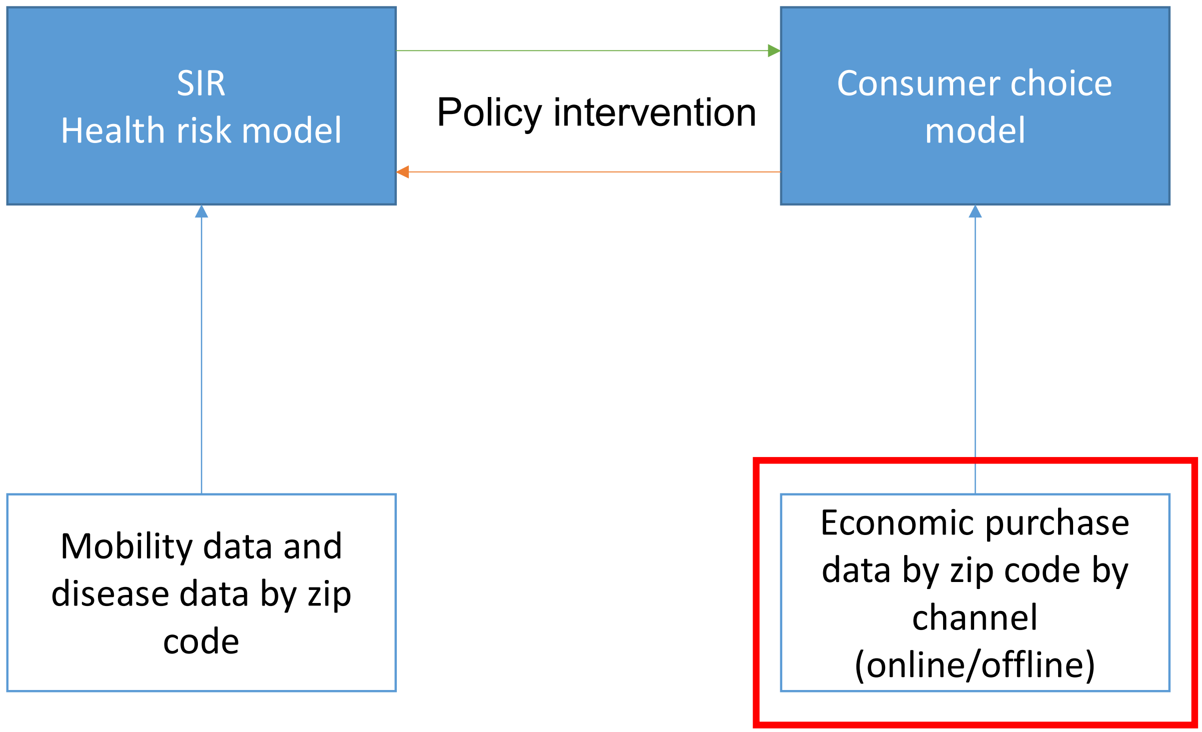
- The transition probabilities:
- $P_{S\to I}^{(t)} = 1 - (1 - \lambda_g)\prod_{i=1}^{n}\left(1 - \frac{\beta_c N_{o,I}^{(t)}}{a_o}\right)$
- $P_{I\to R}^{(t)} = 1/\delta_R$

λ_g	Probability of contracting Covid-19 by living in CBG g
β_c	Risk of category c
$N_{o,I}^{(t)}$	Number of infectious people co-locating at POI o on day t
a_o	The building area of POI o
δ_R	The average days of recovery

We are interested in β_c, the risk of category c.

Research question

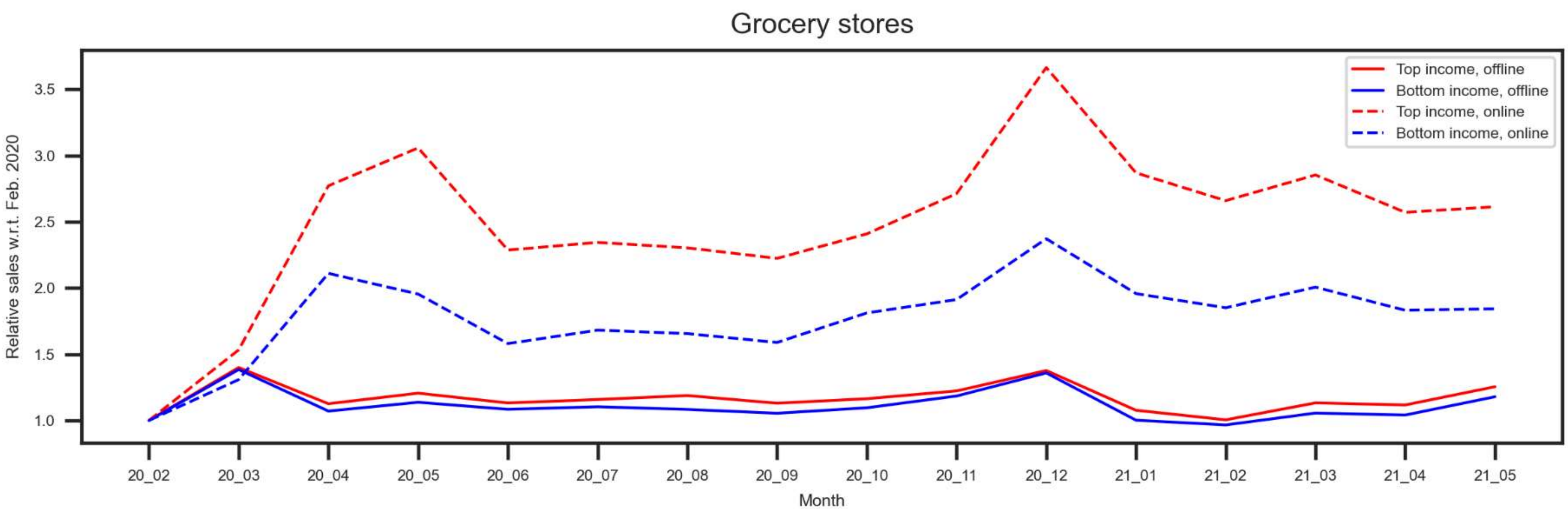
- What is the relationship between mobility and the spread of Covid-19? ☑
- How would perceived risk of exposure change consumer behaviors?
- What policies can policy makers adopt in order to control the spread of Covid-19 by changing mobility matterns tied to offline shopping?


SIR
Health risk model
Policy intervention
Consumer choice model
Mobility data and disease data by zip code
Economic purchase data by zip code by channel (online/offline)

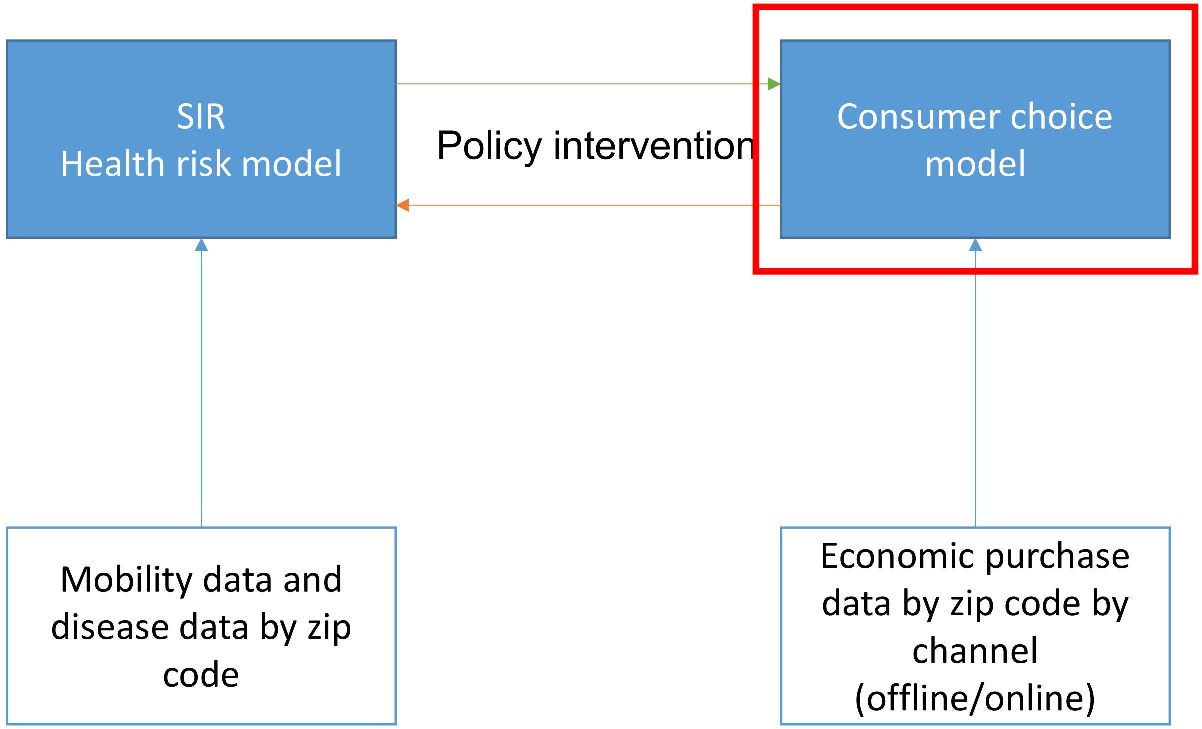
Dataset description

- **Consumer behavior dataset, from a large super-regional bank**
- Record: 'Date', 'Category (SIC Code)', 'Customer CBG', 'Transaction amount', 'Online/Offline' …
- From 02/01/2020 to 05/01/2021
- 1M aggregated records per month in PA
- Credit card transaction record with consumer geolocation, industry category and online/offline identification

Economic behavior of different income groups



Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT


SIR
Health risk model
Policy intervention
Consumer choice model
Mobility data and disease data by zip code
Economic purchase data by zip code by channel (offline/online)

Binary Choice Model (Logit):

$$p_{off} = \frac{e^{\Delta U}}{e^{\Delta U} + 1}, \qquad p_{on} = \frac{1}{e^{\Delta U} + 1}$$

$$\Delta U_i = U_{offline,i} - U_{online,i} = \beta_0 + \beta_1 b_c + \beta_2 d + \beta_3 b_c d + \beta_4 r_{c,i,t}$$

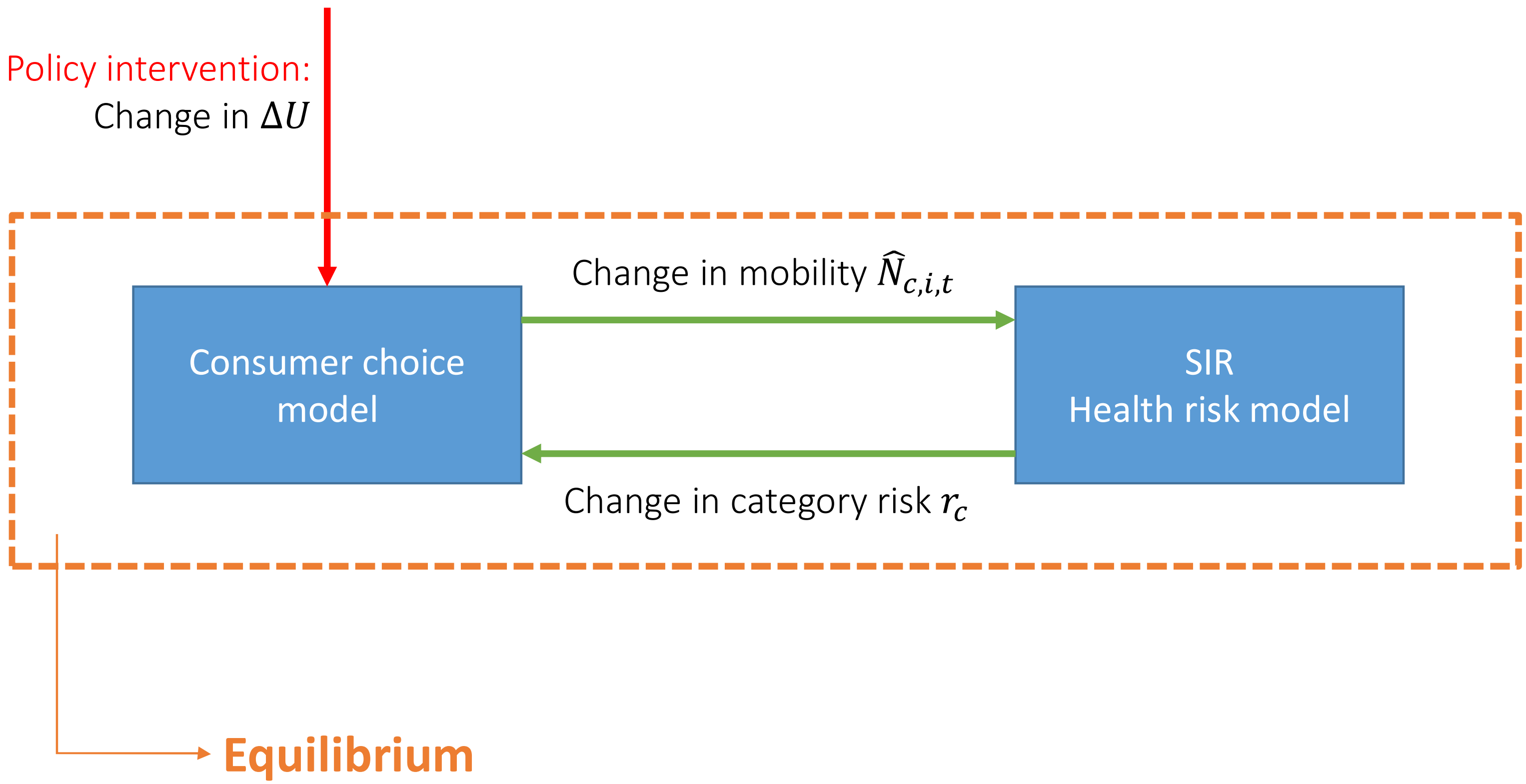
$r_{c,i,t} = r_c \times \widehat{N}_{c,i,t}$	Perceived risk of category c on day t
r_c	Risk of category c estimated from mobility data
$\widehat{N}_{c,i,t}$	Moving average of number of people visiting category c
b_c	Difference of going offline v.s. online in category c
$d_1 \sim d_7$	Dummy variable of day of week (Monday to Sunday)

Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Results

	Top 25% income zip codes	Bottom 25% income zip codes
Perceived risk	-0.0025 [0.001]**	-0.047 [0.002]**
Restaurants (Baseline: Grocery stores)	-1. 4246 [0.047]***	-1.0183 [0.057]***
Fastfood chains	-1.1605 [0.047]***	-0.6698 [0.057]***
Drug stores	-0.4173 [0.080]***	-0.7498 [0.085]***
Book stores	-6.8558 [0.089]***	-7.0184 [0.159]***
Tuesday (Baseline: Monday)	0.1046*	0.0204
…		
Saturday	0.3773***	0.2768***
Sunday	0.0910*	0.1024*
Book stores × Tuesday	0.5254***	0.5901**
…		

Counterfactual experiment



Key Ideas and Open Questions

Create platform for policy interventions and counterfactual analysis

Experiment with different structural models of consumer behavior as well as disease diffusion

Develop methods to address non-representativeness of data

Supply Chains, Resilience and Public Policy

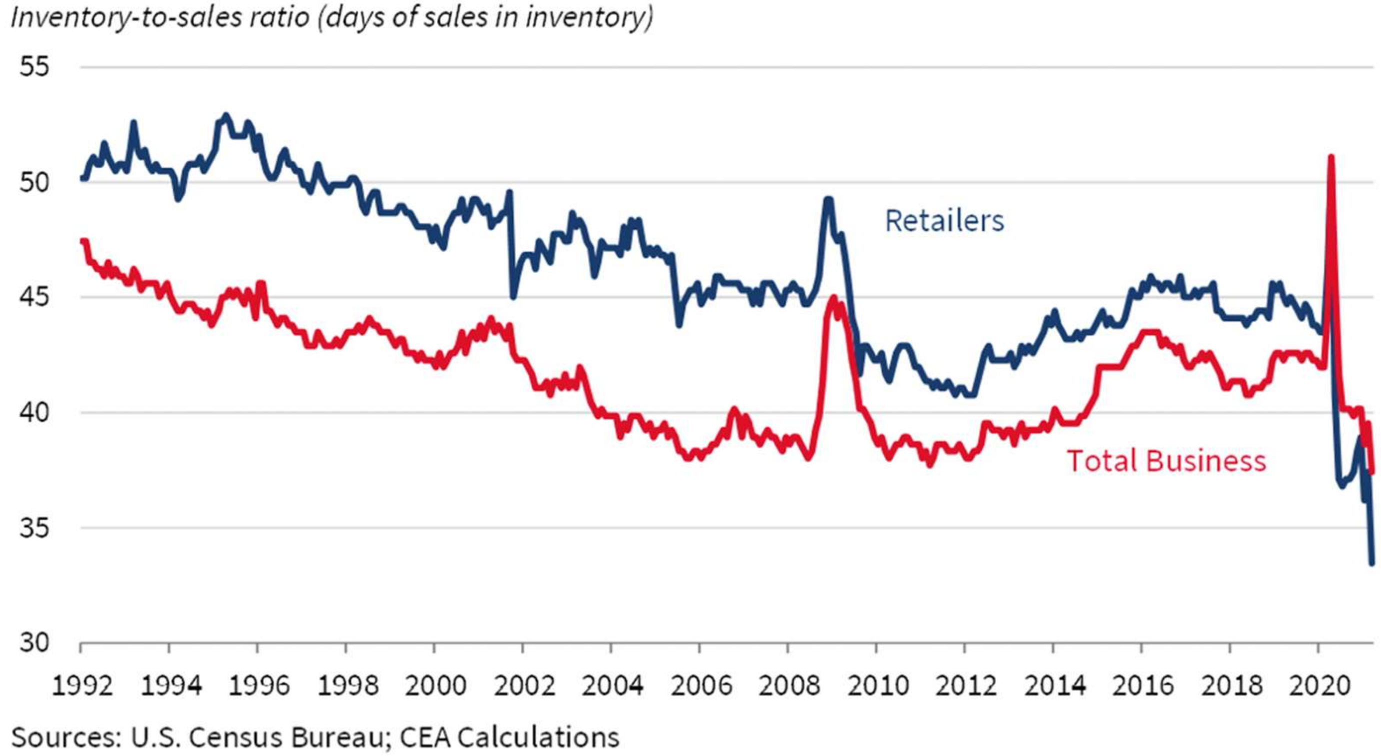
Pascal Van Hentenryck - Georgia Institute of Technology
Peter Zhang – Carnegie Mellon University

Carnegie Mellon University Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Opportunities for Collaboration

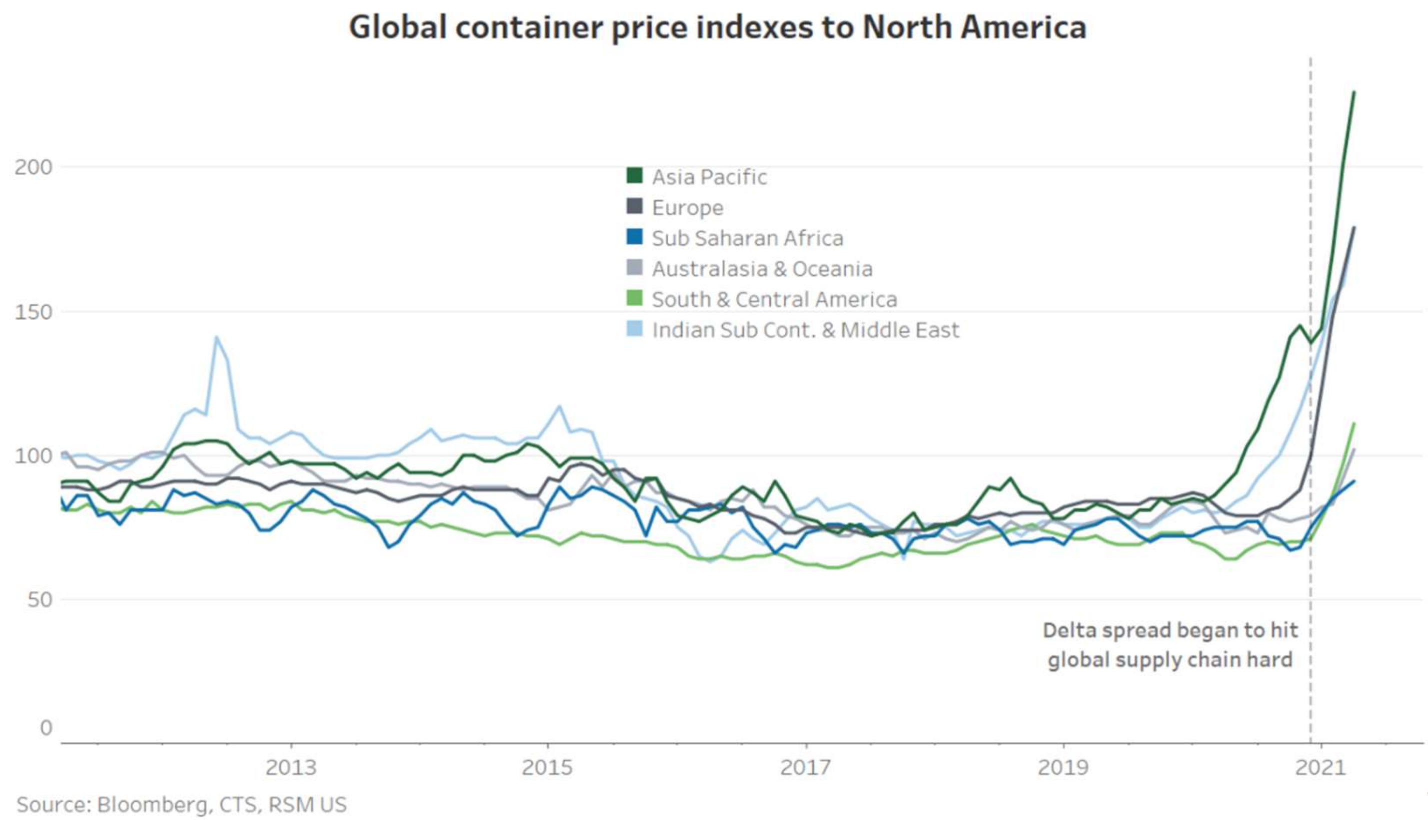
- Between AI, Data Science and the OR communities at three layers
  - Methods
    - ML, high dimensional statistics, causal inference and optimization
  - Data
    - Creating benchmark data sets (like the ARPA-E Power grid for societal systems like critical supply chains will be hugely beneficial
  - Policy
    - Opportunity to work on problems that have direct policy impact

Efficiency of Today's Supply Chains



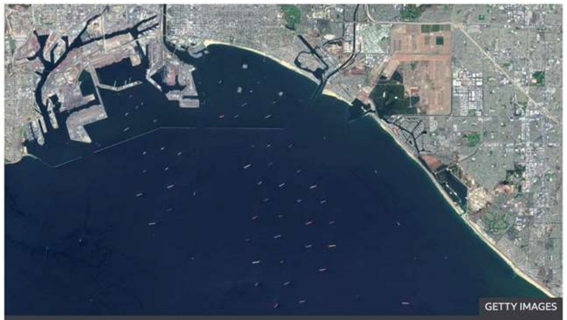
White House (2021): https://www.whitehouse.gov/cea/blog/2021/06/17/why-the-pandemic-has-disrupted-supply-chains/

Efficiency of Today's Supply Chains



Carnegie Mellon University Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Supply Chain Disruptions



Ships laden with cargo are waiting outside Los Angeles and Long Beach ports.

How bad is the congestion?

What Happened During Covid-19?

- E-commerce expansion
  - first and last mile issues
- Workforce shortage
  - at factories and in logistics
- Resilience issues
  - dependencies on suppliers
  - limitations of just-in-time systems

Real-World Supply Chain Resilience

JULY 29, 2021
By Rainer Schuster, Gaurav Nath, Pepe Rodriguez, Chrissy O'Brien, Ben Aylor, Boris Sidopoulos, Daniel Weise, and Bitan Datta

The COVID-19 crisis has been a wake-up call for supply chain managers. For years, companies have focused on eliminating redundancy in sourcing to reduce fixed costs and promote efficiency. Greater efficiency, however, came at the expense of diminished flexibility and effectiveness—a tradeoff the pandemic-induced supply chain disruptions have made painfully clear.

Complexity in Today's Supply Chains

Visualization of an automaker's (partial) supply chain

Carnegie Mellon University HeinzCollege
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Risks in Today's Supply Chains

- Significant increase in supply chain risk
  - Outsourcing and offshoring
    - Supply chain is geographically more diverse
  - Lean manufacturing
    - Just-in-time (JIT) manufacturing and low inventory levels

Intel Sales are down

Giant blames Thai flood for $1B drop in sales goals. Toyota, Honda, Goodyear, Canon, Nikon, Sony… have cut production and lowered financial forecasts because of the flooding in Thailand.

The Wall Street Journal, 2011

Everywhere You Look, the Global Supply Chain Is a Mess

Supply chain woes mounted world-wide for makers of everything from cars and clothing to home siding and medical needle containers, as the extreme Texas weather and port backlogs compounded problems for manufacturers already beset by pandemic disruptions.

The Wall Street Journal, 2021

Managing Critical Supply Chain Risk: Policy Challenge

- Firms do not want to share supply chain data
- Impact of disruptions propagate across sectors
- Impact on social welfare, jobs, national security
- Critical technology supply chains have unknown dependence on other countries
- No existing policy compliance framework to impose supply chain resilience

Supply Chain Stress Test?

- Similar to the financial industry, the government could create a "critical supply chain stress test" framework to:
  1. Help firms identify their vulnerabilities
  2. Collect information from firms to assess macro-level risk, without having to access firm's internal data
  3. Incentivize firms to improve supply chain resilience
  4. Challenges:
     - Verifying compliance to resilience targets while treating supply chain as a black box
     - Designing meaningful scenarios to incent changes in system design

Risk Awareness And Resilience

- Fundamental in many engineering domains
  - much to learn from for supply chains
- Reliability in Power systems
  - N-1 reliability: ability to survive the loss of a generator or of a transmission
  - e.g., Security-Constrained Unit Commitment
- Risk in Financial Engineering
  - Conditional value at risk
  - transferable to many engineering domains and supply chains!
  - tractable computationally

Time to Recover and Time to Survive

Time-to-Recover (TTR): The time for a node in the supply chain to return to full functionality after a disruption.

Time-to-Survive (TTS): The maximum duration for a node in the supply chain to match supply with demand

$$TTR(j) < TTS(j) \text{ for all nodes } (j)$$

↓

Robust Supply Chain

Simchi-Levi and Zhang (2020/2021)

Policy Level Task

Identify risks in critical supply chains:

- Semiconductor manufacturing and advanced packaging;
- Large capacity batteries, like those for electric vehicles;
- Critical minerals and materials;
- And pharmaceuticals and active pharmaceutical ingredients.

White House (2021): https://www.whitehouse.gov/cea/blog/2021/06/17/why-the-pandemic-has-disrupted-supply-chains/

Supply Chain Stress Test Framework

1. Policymakers identify a list of supply chain disruption scenarios regularly (e.g., each quarter or every year), based on current evaluation of climate, geopolitical, and financial risks
   - Each scenario contains: the disruption event and its duration (TTR)
2. Firms in a critical supply chain receive the list, and evaluate their performance (e.g., using an optimization model)
3. Firms report the impact of these scenarios to policymakers, e.g., TTR and TTS on its output:
   - Solvency
   - Performance impact for its products (shortage, delay, etc)
   - Potential adjustment made in its business process / production process to achieve the performance
4. Repeat the process if there are cycles in the entire supply chain
5. If not risk compliant, firms adjust supply chain (data analytics + optimization)

Carnegie Mellon University HeinzCollege
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Helping Firms Respond To Resilience Requirements

- How to survive various type of shocks?
  - e.g., weather, political instabilities, …
  - by upgrading your supply chains and networks
- Complex stochastic optimization problem

$$
\begin{aligned}
\min \quad & c^T w \\
\text{s.t.} \quad & w \geq w^s, & \forall s \in \mathcal{S}, \\
& w^s \in \mathfrak{Q}(s), & \forall s \in \mathcal{S}, \\
& w \in \{0,1\}^m.
\end{aligned}
$$

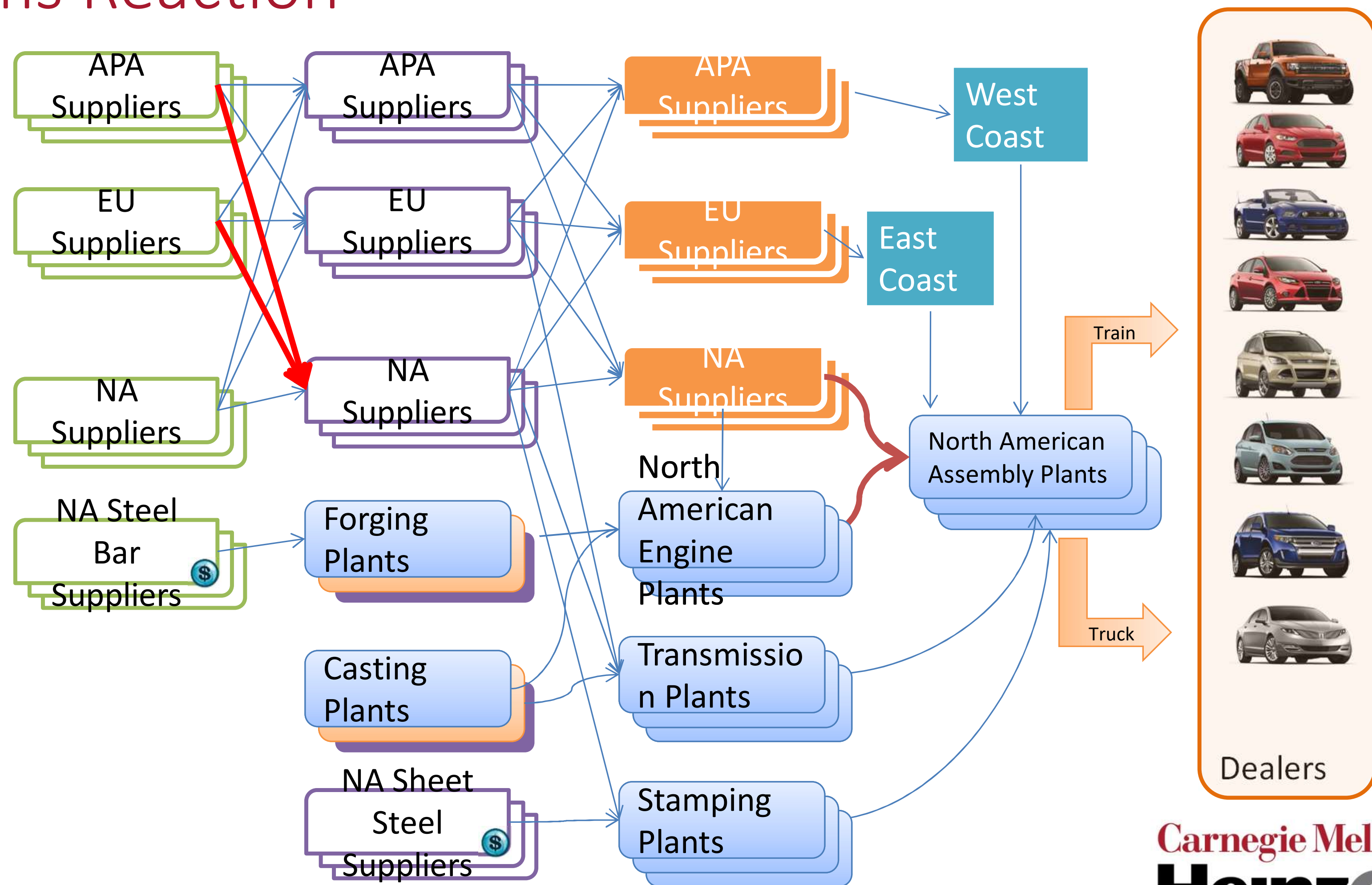
minimizing costs

choosing resiliency upgrades

surviving scenarios

  - tractable using bespoke branch and price algorithms

Firms Reaction



Questions for Research to assist Policymakers?

- How to design the scenarios?
- How often to update the scenarios?
- What risk metrics should be used to set targets? CVAR, TTR/TTS?
- Should the policymaker industry-level TTS/TTR targets while providing some guarantees on supply (e.g., via trade agreements, national reserve) for critical ingredients?

- Zero Knowledge Proof – how to verify that supply chain is robust without detailed knowledge of the supply chain?

Opportunities for Collaboration

- Between AI, Data Science and the OR communities at three layers
  - Methods
    - ML, high dimensional statistics, causal inference and optimization
  - Data
    - Creating benchmark data sets (like the ARPA-E Power grid for societal systems like critical supply chains will be hugely beneficial
  - Policy
    - Opportunity to work on problems that have direct policy impact

[redacted]