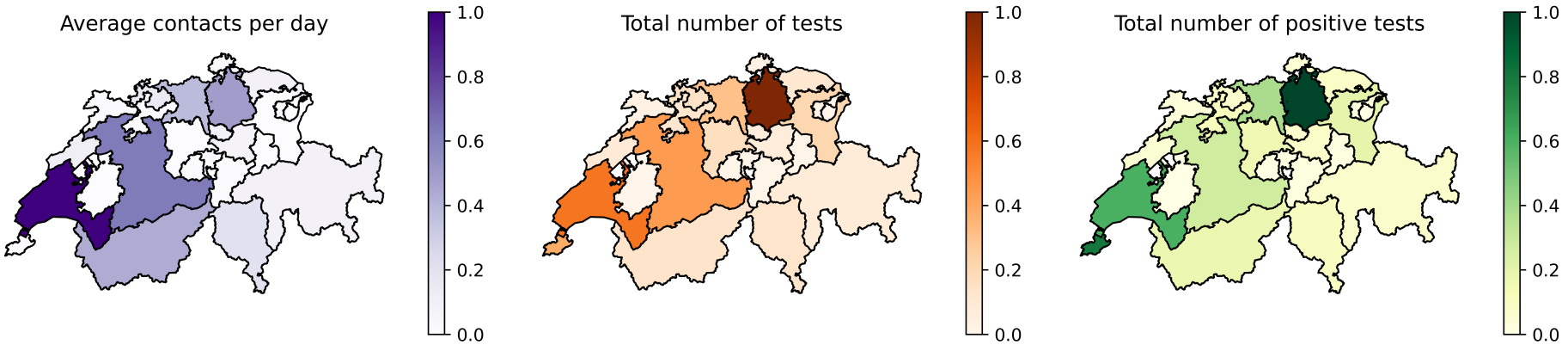
Activity-Based Modeling for capturing Dynamics of Mobility, Awareness, and Individual Choices in Disease Spread

Internal Seminar

Transport and Mobility Laboratory

Cloe Cortes Balcells and Michel Bierlaire

cloe.cortesbalcells@epfl.ch



- What epidemiological **data** is **available**?
- What does it mean then?
- How can we **use** this information?
- How do we link it with **travel-behavior**?
- How **travel-behavior** will **change**?

So many questions...

Research Gaps

- Traditional models often overlook the **heterogeneity of individual actions** and their impact on the spread of disease (Hackl and Dubernet 2019; Eubank et al. 2004; Perez and Dragicevic 2009).
- Existing research does not fully integrate **socioeconomic and health factors** influencing mobility and epidemiological outcomes.
- There is a **lack of emphasis on individual choices**, particularly regarding testing and the subsequent behavioral adjustments (Cui, Ni, and Shen 2021; Brotherhood et al. 2020).
- Need for models that can **dynamically capture individual decisions** and their effects on the pandemic's trajectory (Tuomisto et al. 2020).

Outline of this talk

1. Importance of **individual-based modeling** in understanding spreading.
2. Discussion on the **integration of latent states**
3. Tracking **individual movements and health states**, and the role of **awareness**.
4. Exploration of the **potential impacts of policy decisions**.

Overall framework

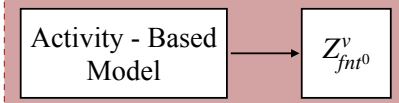
Pedex

$f =$ facility
$n =$ individual
$t =$ timestep

Apex

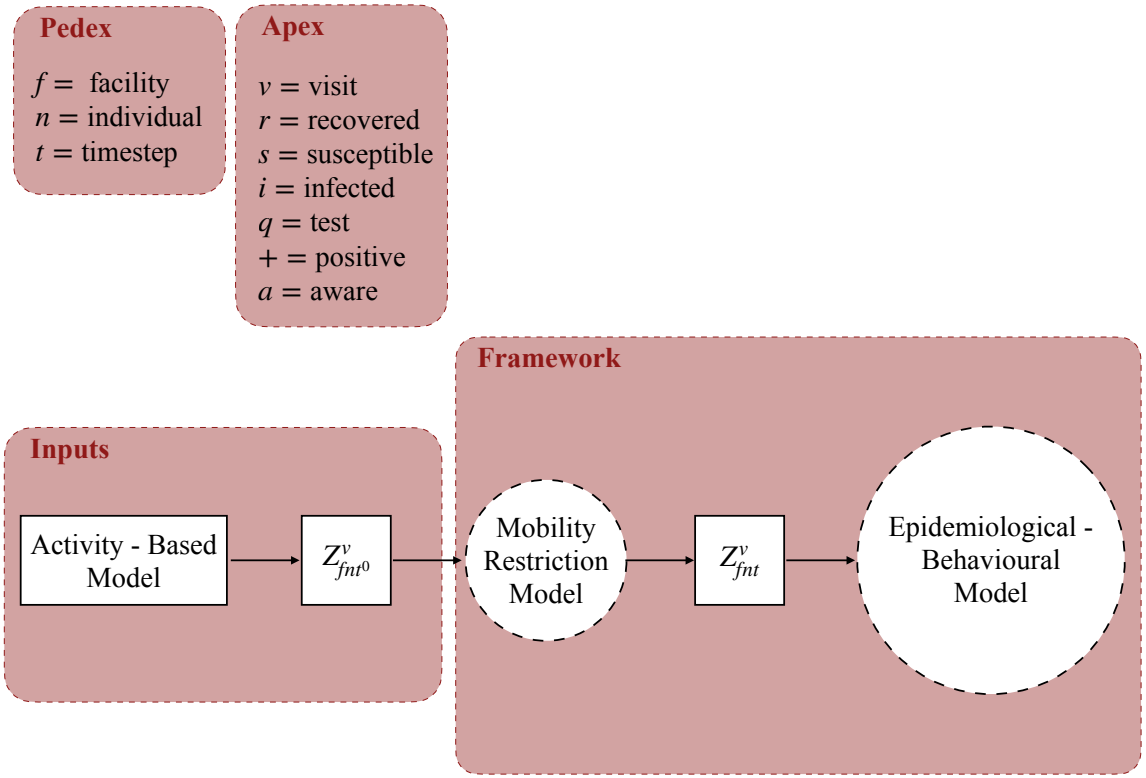
$v =$ visit
$r =$ recovered
$s =$ susceptible
$i =$ infected
$q =$ test
$+ =$ positive
$a =$ aware

Inputs



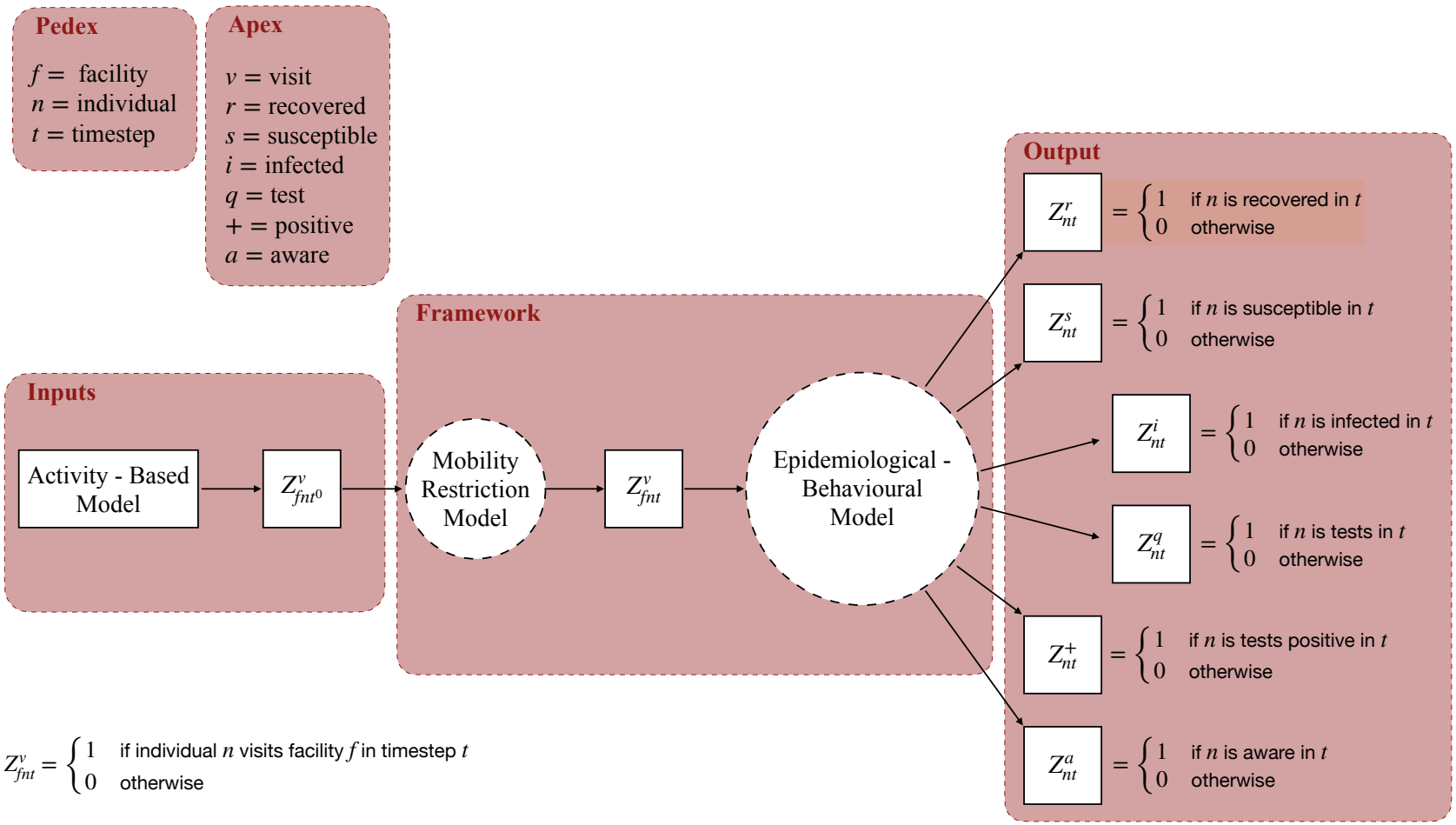
$$Z^v_{fnt} = \begin{cases} 1 & \text{if individual } n \text{ visits facility } f \text{ in timestep } t \\ 0 & \text{otherwise} \end{cases}$$

Overall framework



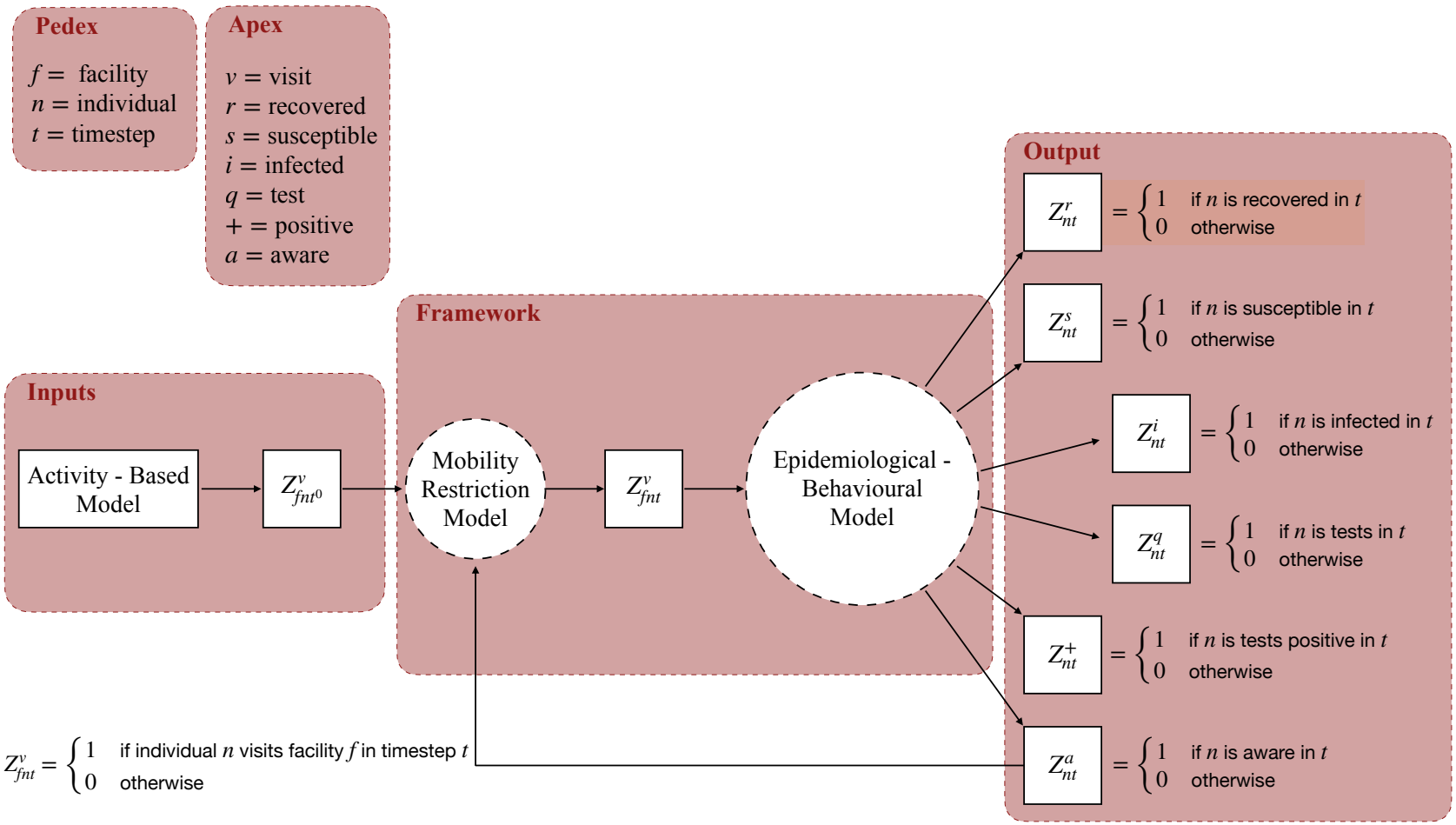
$$Z^v_{fnt} = \begin{cases} 1 & \text{if individual } n \text{ visits facility } f \text{ in timestep } t \\ 0 & \text{otherwise} \end{cases}$$

Overall framework

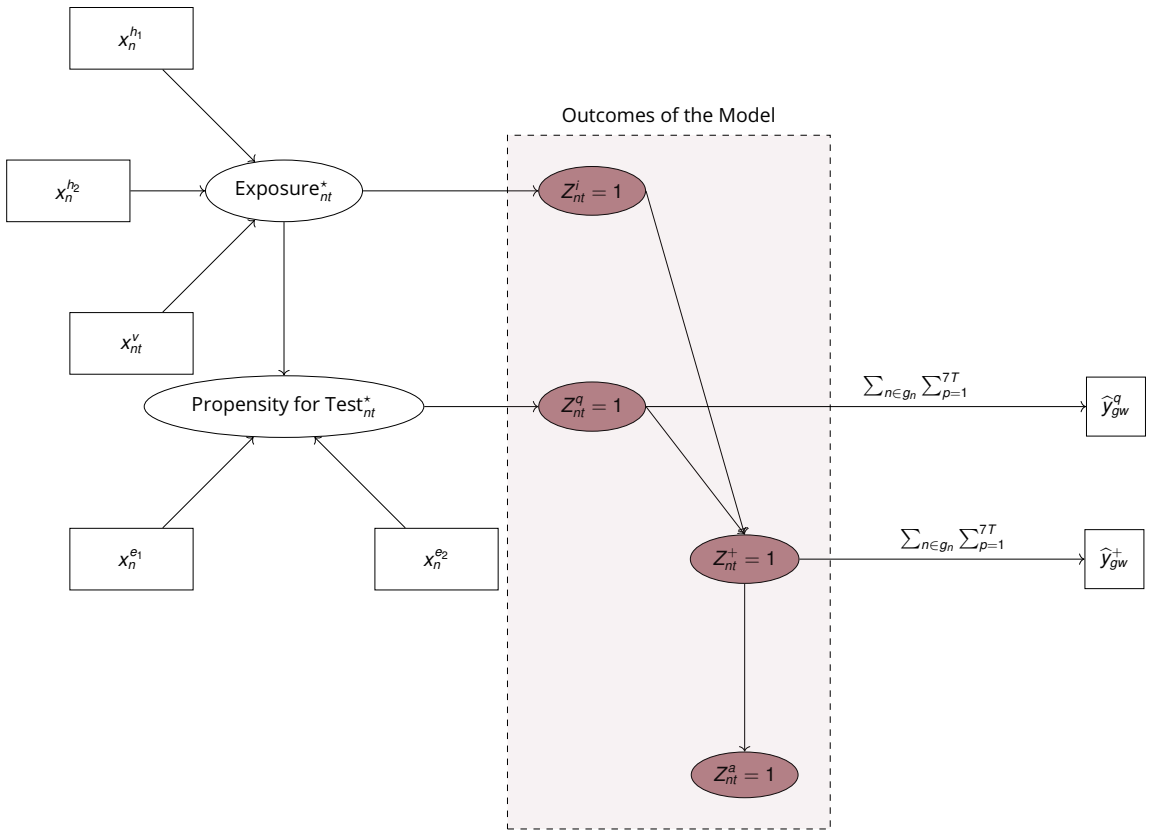


$$Z^v_{fnt} = \begin{cases} 1 & \text{if individual } n \text{ visits facility } f \text{ in timestep } t \\ 0 & \text{otherwise} \end{cases}$$

Overall framework



Epidemiological-Behavioural Model



Epidemiological-Behavioural Model

Exposure State Modeling

$$E^*_{nt} = \beta^0_{E^\star} + \sum_{k=1}^{K_{E^\star}-1} \beta^h_k x^h_{nk} + \beta^v x^v_{nt} + \varepsilon_{E^\star}$$

- x^v_{nt}: the proportion of infected individuals that individual n encounters in a facility f at timestep t, where

$$x^v_{nt} = \sum_f \sum_{m \neq n} Z^v_{fnt} Z^v_{fmt} Z^i_{mt}.$$

- x^h_{nk}: $K^\star_E$ health characteristics of indiviudal n.
- $\beta^0_{E^\star}$, β^h_k, and $\varepsilon_{E^\star}$, parameters to be calibrated.

Epidemiological-Behavioural Model

- x^e_{nk}: $K^\star_Q$ socioeconomic characteristics of indiviudal n.
- $\beta^0_{Q^\star}$, $\beta^{E^\star}_k$, $\eta_{E^\star}$ and $\varepsilon_{Q^\star}$, parameters to be calibrated.

Propensity to Test Modeling

$$Q^\star_{nt} = \beta^0_{Q^\star} + \sum_{k=1}^{K_{Q^\star}-1} \beta^{E^\star}_k x^e_{nk} + \eta_{E^\star} E^\star_{nt} + \varepsilon_{Q^\star}$$

Health State Transition

A **hidden Markov chain** is employed to model the **transitions** between **health states**. The **state transition matrix** $\mathbb{B}$ is:

$$\begin{bmatrix} P(Z^s_{n(t+1)} = 1 | Z^s_{nt} = 1) & P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1) & 0 \\ 0 & P(Z^i_{n(t+1)} = 1 | Z^i_{nt} = 1) & P(Z^r_{n(t+1)} = 1 | Z^i_{nt} = 1) \\ P(Z^s_{n(t+1)} = 1 | Z^r_{nt} = 1) & 0 & P(Z^r_{n(t+1)} = 1 | Z^r_{nt} = 1) \end{bmatrix}$$

Health State Transition

Is the individual **becoming infected or remaining susceptible**?

$$P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1) = \frac{1}{1 + e^{-\mu E^\star_{nt}}} \tag{1}$$

$$P(Z^s_{n(t+1)} = 1 | Z^s_{nt} = 1) = 1 - P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1) \tag{2}$$

$$Z^i_{n(t+1)} = \begin{cases} 1, & \text{if } \pi < P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1) \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

π: random uniform value

Health State Transition

Is the individual **becoming recovered or remaining infected**?

$$P(Z^r_{n(t+1)} = 1 | Z^i_{nt} = 1) = \Phi_{\gamma_n}(t - t^i_n)$$

Φ_{γ_n}: cumulative distribution function of the log-normal distribution of γ_n (recovery rate (Kerr et al. 2020)).
$t - t^i_n$: number of timesteps that satisfy $Z^i_n = 1$.
For each individual n at time t, we decide if they will be recovered by time t given that they were infected at time t^i_n, by computing:

$$Z^r_{n(t+1)} = \begin{cases} 1, & \text{if } \pi < P(Z^r_{n(t+1)} = 1 | Z^i_{nt} = 1) \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

I am not sure of this, basically now what i do is

- simulating recovery times for a number of agents based on a lognormal distribution
- i return an array of recovery times for each agent, expressed in the number of periods
- then we decrease an infection timer for all currently infected agents, identifying which agents have reached the end of their infection period (based on the timer reaching zero), and updating their state to ”recovered”

Health State Transition

Finally, we assume that an individual who recovers becomes immediately susceptible until the end of the simulation.

$$P(Z^r_{n(t+1)} = 1 | Z^r_{nt} = 1) = 0,$$

$$P(Z^s_{n(t+1)} = 1 | Z^r_{nt} = 1) = 1.$$

Testing process

What about the testing process?

We model the probability of an individual to test through a logit:

$$P(Z^q_{n(t+1)} = 1) = \frac{1}{1 + e^{-\mu Q^\star_{nt}}}, \tag{5}$$

and each individual n at time t, decides to get tested by generating π, and following:

$$Z^q_{n(t+1)} = \begin{cases} 1, \text{if } \pi < P(Z^q_{nt}) \\ 0, \text{otherwise.} \end{cases} \tag{6}$$

Testing process

The outcome of the test is computed by:

$$P(Z^+_{nt}=1|Z^i_{nt}=1)=P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^i_{nt}=1)P(Z^q_{nt}=1 \text{ and } Z^i_{nt}=1)=$$
$$P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^i_{nt}=1)P(Z^i_{nt}=1|Z^q_{nt}=1)P(Z^q_{nt}=1)=$$
$$P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^i_{nt}=1)P(Z^i_{nt}=1)P(Z^q_{nt}=1),$$

$$P(Z^+_{nt}=1|Z^s_{nt}=1)=P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^s_{nt}=1)P(Z^q_{nt}=1 \text{ and } Z^s_{nt}=1)=$$
$$P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^s_{nt}=1)P(Z^s_{nt}=1|Z^q_{nt}=1)P(Z^q_{nt}=1)=$$
$$P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^s_{nt}=1)P(Z^s_{nt}=1)P(Z^q_{nt}=1),$$

$$P(Z^+_{nt}=1|Z^r_{nt}=1)=P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^r_{nt}=1)P(Z^q_{nt}=1 \text{ and } Z^r_{nt}=1)=$$
$$P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^r_{nt}=1)P(Z^r_{nt}=1|Z^q_{nt}=1)P(Z^q_{nt}=1)=$$
$$P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^r_{nt}=1)P(Z^r_{nt}=1)P(Z^q_{nt}=1).$$

$P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^i_{nt}=1)$, $P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^s_{nt}=1)$, $P(Z^+_{nt}=1|Z^q_{nt}=1 \text{ and } Z^r_{nt}=1)$ are taken from Ai et al. 2020.

Mobility Restriction Model

What about awareness?

Assumption: only people that are infected and test positive will quarantine.

$$Z^a_{nt} = Z^i_{nt} Z^+_{nt} \tag{7}$$

Now, we apply the mobility restriction model, and the new dynamics can be written as:

$$Z^v_{fn(t+1)} = \begin{cases} Z^v_{fnt} & \text{if individual's } n \text{ outcome is 0, and} \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Note on the simplification. It is like this because we do not have symptom data so we assume that infected agents if they test positive and have symptoms will stay home. If you do not have symptoms and test positive you might go out.

Measurement equations

$$N_g^q = \sum_{n \in g} \sum_{t=1}^{7T} Z_{nt}^q \quad (9)$$

$$N^q = \sum_{t=1}^{T} Z_{nt}^q \quad (10)$$

$$N_{hjk}^q = \sum_{n \in h,j,k} \sum_{t=1}^{T} Z_{nt}^q \quad (11)$$

$$\overline{P_{gw}^+} = \frac{\sum_{n \in g} \sum_{t=1}^{7T} P(Z_{nt}^+ = 1)}{7TN_g} \quad (12)$$

$$\overline{P_{\ell}^+} = \frac{\sum_{t=1}^{T} P(Z_{nt}^+ = 1)}{TN} \quad (13)$$

$$\overline{P_{hjk\ell}^+} = \frac{\sum_{n \in h,j,k} \sum_{t=1}^{T} P(Z_{nt}^+ = 1)}{TN_{hjk}} \quad (14)$$

Problem of not having panel data, since the choice of testing strongly depends on the previous choice of testing...

Data & Likelihood functions

$\widehat{y}^{+}_{gw}$: positive tests per age group g and week w

$\widehat{y}^{+}_{\ell}$: positive tests per day ℓ

$\widehat{y}^{+}_{hjk\ell}$: positive tests per age h, gender j ,and municipality k ,per day ℓ

$$\log L_1 = \sum_g \sum_w \left[\log \binom{N^q_g}{\widehat{y}^{+}_{gw}} + \widehat{y}^{+}_{gw} \log(\overline{P^{+}_{gw}}) + (N^q_g - \widehat{y}^{+}_{gw}) \log(1 - \overline{P^{+}_{gw}}) \right] \tag{15}$$

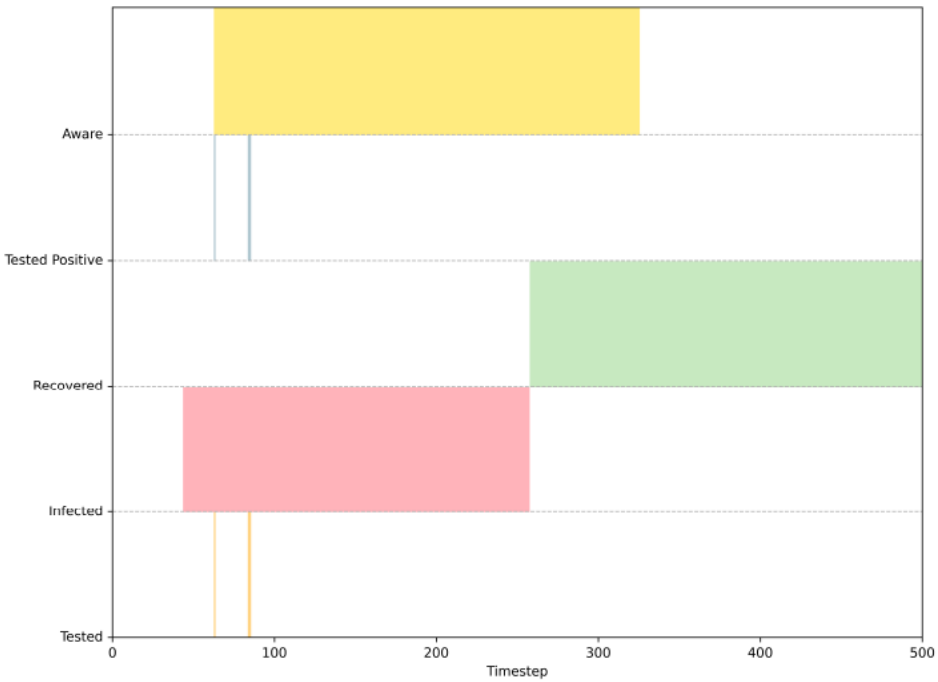
$$\log L_2 = \sum_\ell \left[\log \binom{N^q}{\widehat{y}^{+}_{\ell}} + \widehat{y}^{+}_{\ell} \log(\overline{P^{+}_{\ell}}) + (N^q - \widehat{y}^{+}_{\ell}) \log(1 - \overline{P^{+}_{\ell}}) \right] \tag{16}$$

$$\log L_3 = \sum_h \sum_j \sum_k \sum_\ell \left[\log \binom{N^q_{hjk}}{\widehat{y}^{+}_{hjk\ell}} + \widehat{y}^{+}_{hjk\ell} \log(\overline{P^{+}_{hjk\ell}}) + \right. \tag{17}$$

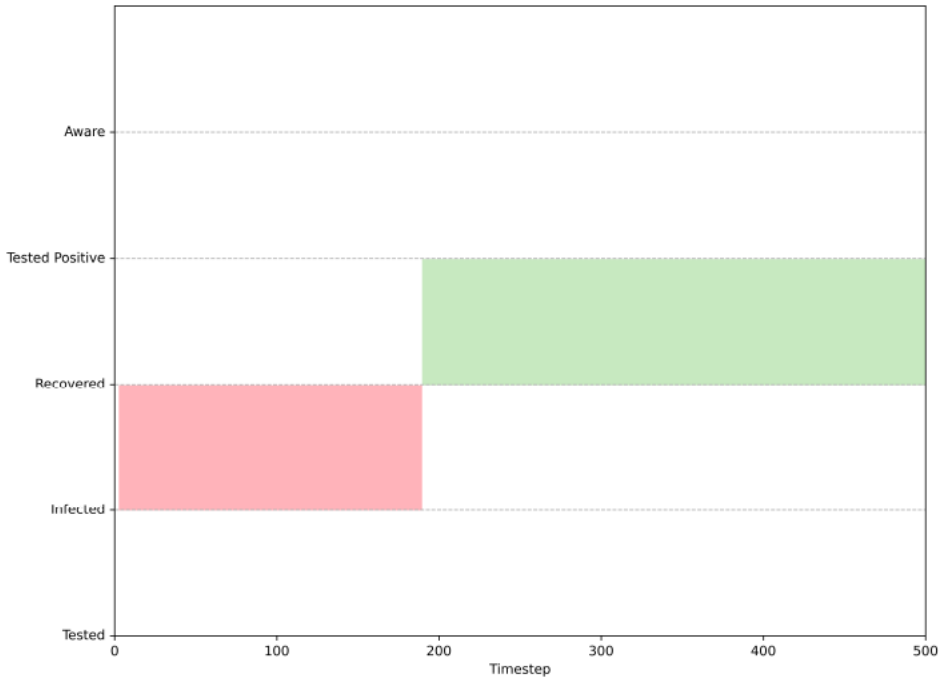
$$\left. (N^q_{hjk} - \widehat{y}^{+}_{hjk\ell}) \log(1 - \overline{P^{+}_{hjk\ell}}) \right] \tag{18}$$

Results across individuals: insights on behavior

The 'ideal' behavior:

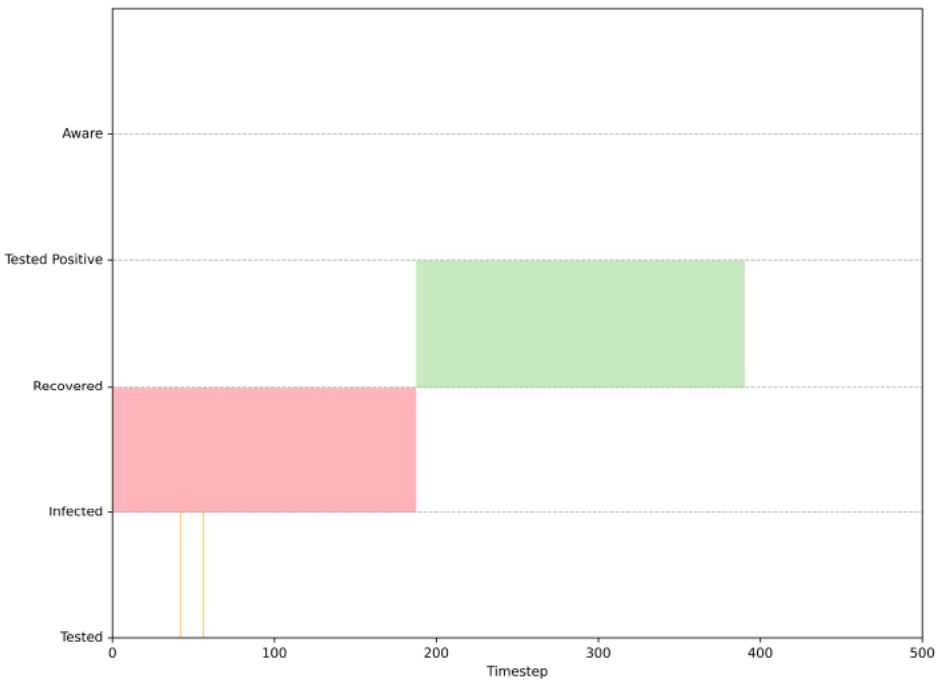


The 'unaware' behavior:

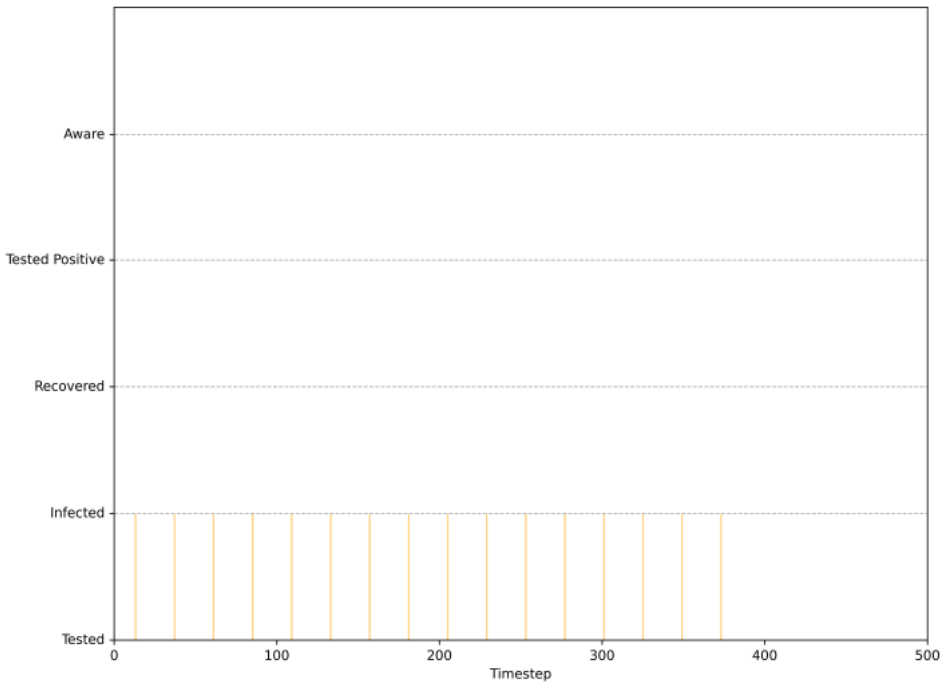


Results across individuals: insights on behavior

The 'unlucky' behavior:



The 'obsessive' behavior:



Conclusion and Future Work

Conclusions:

1. A computationally efficient tool (6m 11s for 800k individuals and 90 days with a timestep of 30m).
2. Lack of disaggregated data always makes it hard to calibrate the models.
3. Bridging epidemiology, transportation, and discrete choice communities for a interdisciplinary model that can better explain how and why a spreading occurs.

Future work:

1. Include health characteristics and calibrate the model.
2. Run this model together with the policy optimization framework from Cortes Balcells 2021.
3. With more data we could study any behavior phenomena related to the spreading of a disease, like for instance people choosing to travel by plane often.

Questions and comments

Negative Binomial

Assuming r is the number of failures until the experiment is stopped and p is the probability of success in each trial (which seems to correspond to avgprobhjkl in your function), the PMF of the negative binomial is:

$$P(Y = k) = \binom{k + r - 1}{k} \cdot p^k \cdot (1 - p)^r$$

where $\binom{k + r - 1}{k}$ is the binomial coefficient, **k** is the number of successes (which seems to correspond to $\widehat{y}^+$, and r is the number of failures until the experiment is stopped. The loglikelihood of this PMF can be calculated by taking the logarithm of the above expression.

Literature review I


Ai, Tao et al. (Aug. 2020). "Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases". en. In: *Radiology* 296.2, E32–E40. ISSN: 0033-8419, 1527-1315. DOI: 10.1148/radiol.2020200642. URL: http://pubs.rsna.org/doi/10.1148/radiol.2020200642 (visited on 03/17/2023).


Brotherhood, Luiz et al. (2020). *An Economic Model of the Covid-19 Epidemic: The Importance of Testing and Age-Specific Policies*. en. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.3618840. URL: https://papers.ssrn.com/abstract=3618840 (visited on 11/27/2023).


Cortes Balcells, Cloe (Sept. 2021). "Activity-based modeling and simulation of epidemics". en. In: URL: http://www.strc.ch/2021.php.


Cui, Yapeng, Shunjiang Ni, and Shifei Shen (Jan. 2021). "A network-based model to explore the role of testing in the epidemiological control of the COVID-19 pandemic". en. In: *BMC Infectious Diseases* 21.1, p. 58. ISSN: 1471-2334. DOI: 10.1186/s12879-020-05750-9. URL: https://doi.org/10.1186/s12879-020-05750-9 (visited on 11/27/2023).

Literature review II


Eubank, Stephen et al. (May 2004). "Modelling disease outbreaks in realistic urban social networks". en. In: *Nature* 429.6988. Number: 6988 Publisher: Nature Publishing Group, pp. 180–184. ISSN: 1476-4687. DOI: 10.1038/nature02541. URL: https://www.nature.com/articles/nature02541 (visited on 09/24/2022).


Hackl, Jürgen and Thibaut Dubernet (Apr. 2019). "Epidemic Spreading in Urban Areas Using Agent-Based Transportation Models". en. In: *Future Internet* 11.4. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute, p. 92. ISSN: 1999-5903. DOI: 10.3390/fi11040092. URL: https://www.mdpi.com/1999-5903/11/4/92 (visited on 11/24/2023).


Kerr, Cliff C. et al. (May 2020). *Covasim: an agent-based model of COVID-19 dynamics and interventions*. en. Pages: 2020.05.10.20097469. DOI: 10.1101/2020.05.10.20097469. URL: https://www.medrxiv.org/content/10.1101/2020.05.10.20097469v1 (visited on 03/10/2023).

Literature review III


Perez, Liliana and Suzana Dragicevic (Aug. 2009). "An agent-based approach for modeling dynamics of contagious disease spread". In: *International Journal of Health Geographics* 8.1, p. 50. ISSN: 1476-072X. DOI: 10.1186/1476-072X-8-50. URL: https://doi.org/10.1186/1476-072X-8-50 (visited on 11/27/2023).


Tuomisto, Jouni T. et al. (Apr. 2020). *An agent-based epidemic model REINA for COVID-19 to identify destructive policies*. en. preprint. Infectious Diseases (except HIV/AIDS). DOI: 10.1101/2020.04.09.20047498. URL: http://medrxiv.org/lookup/doi/10.1101/2020.04.09.20047498 (visited on 09/24/2022).