Mobility and Individual Choices in the Spread of Infectious Diseases: Enhancing Activity-Based Models with Awareness and Testing Dynamics

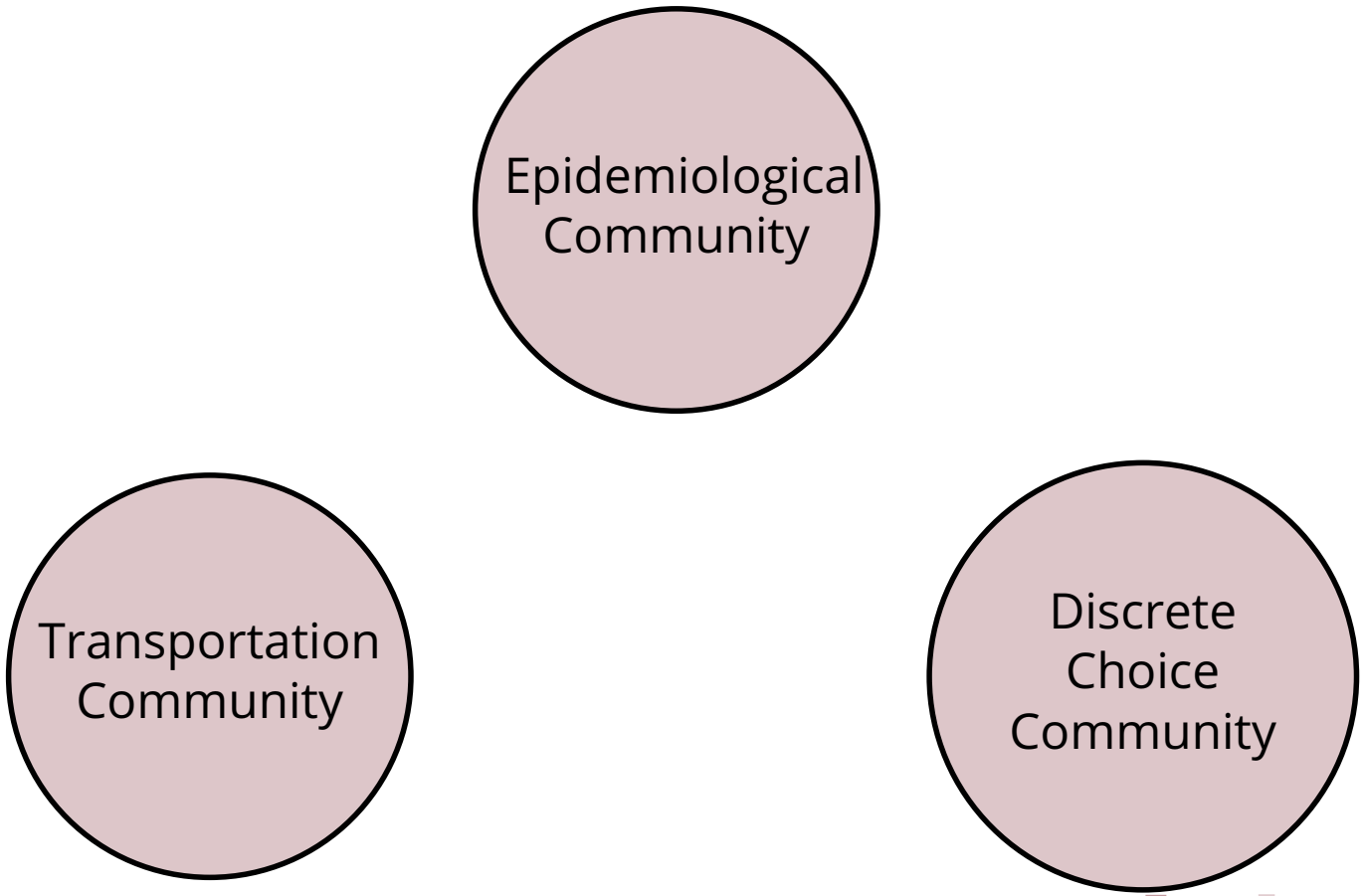
ICMC 2024

Transport and Mobility Laboratory

Cloe Cortes Balcells and Michel Bierlaire

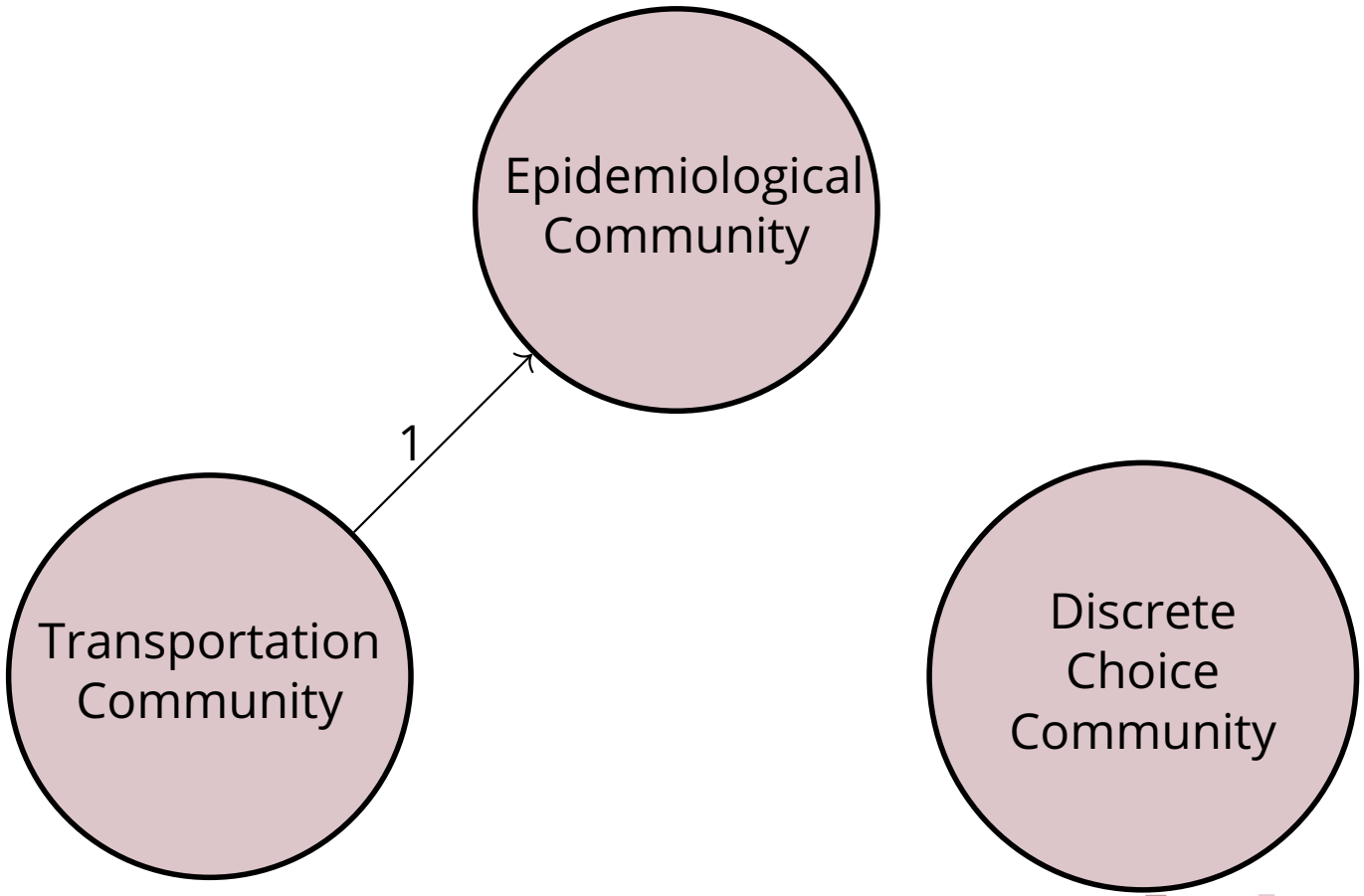
cloe.cortesbalcells@epfl.ch

How do we link these communities?



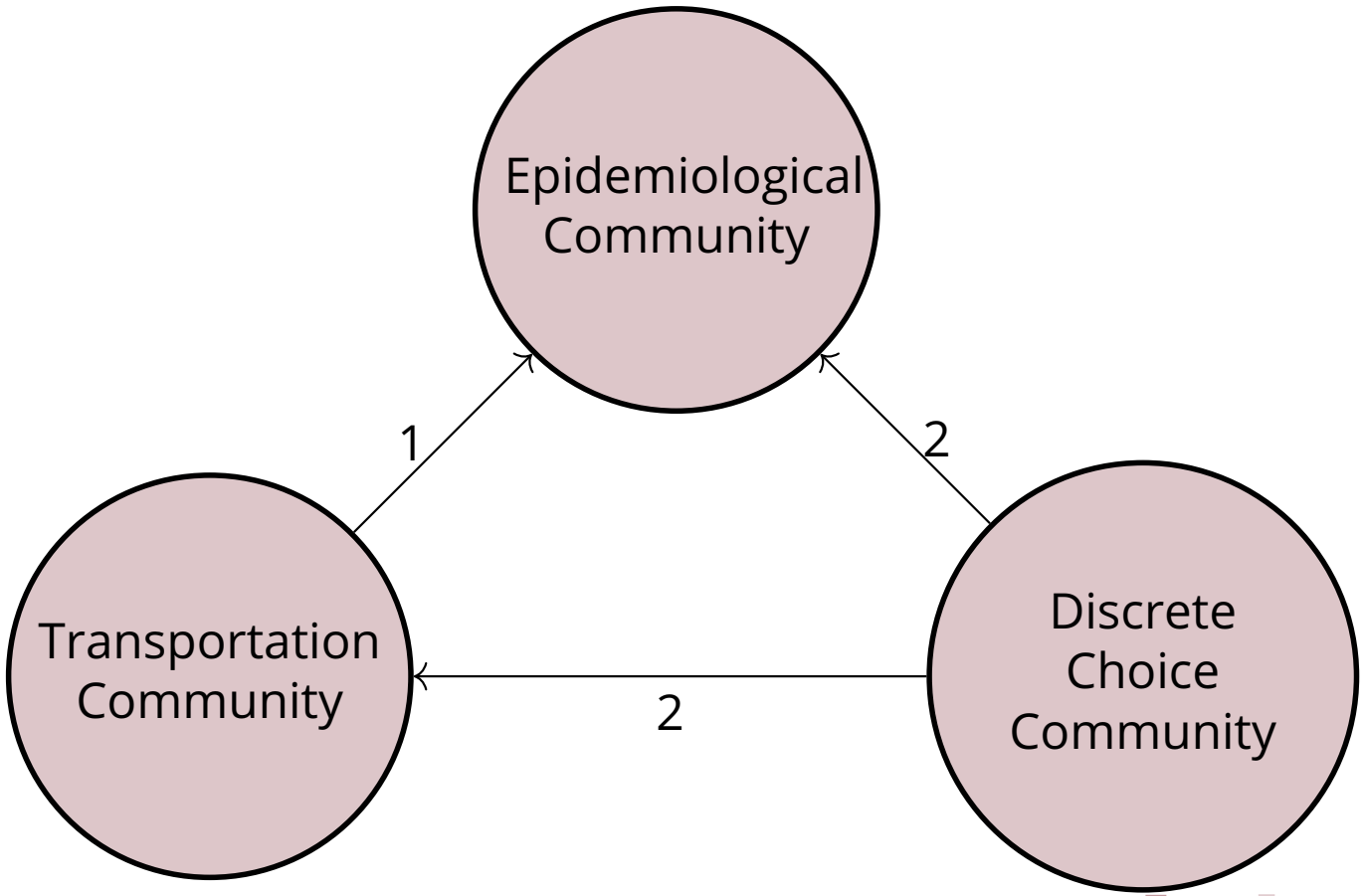
How do we link these communities?

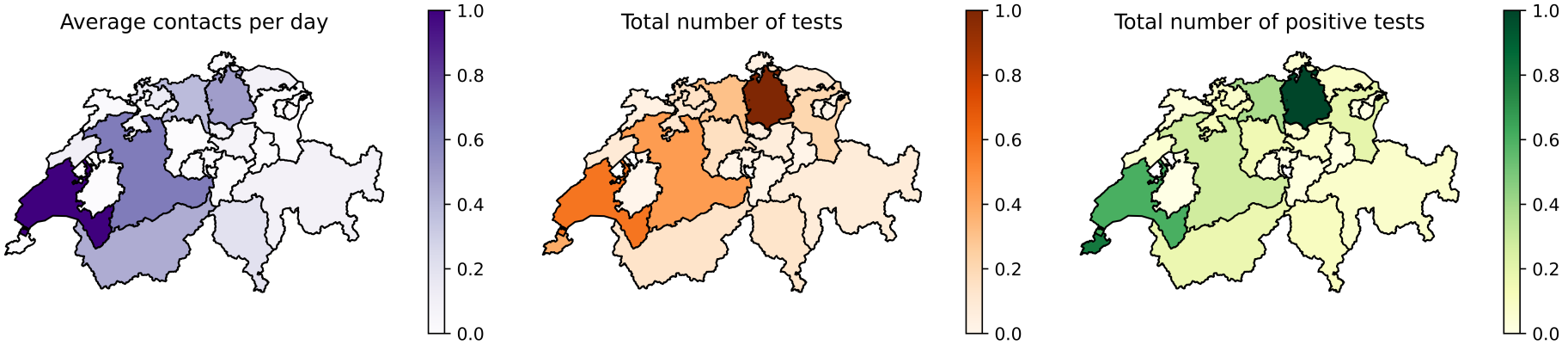
1. Activity-travel behavior impacts the spread of a disease.



How do we link these communities?

1. Activity-travel behavior impacts the spread of a disease.
2. Testing choices changes individual's behavior and therefore the spread of a disease.





The Thinker's Corner:

- What epidemiological **data** is **available**?
- How can we **leverage** this information?
- How do we link epidemiological data with **travel-behavior**?
- How will **travel-behavior change**?

Research Gaps

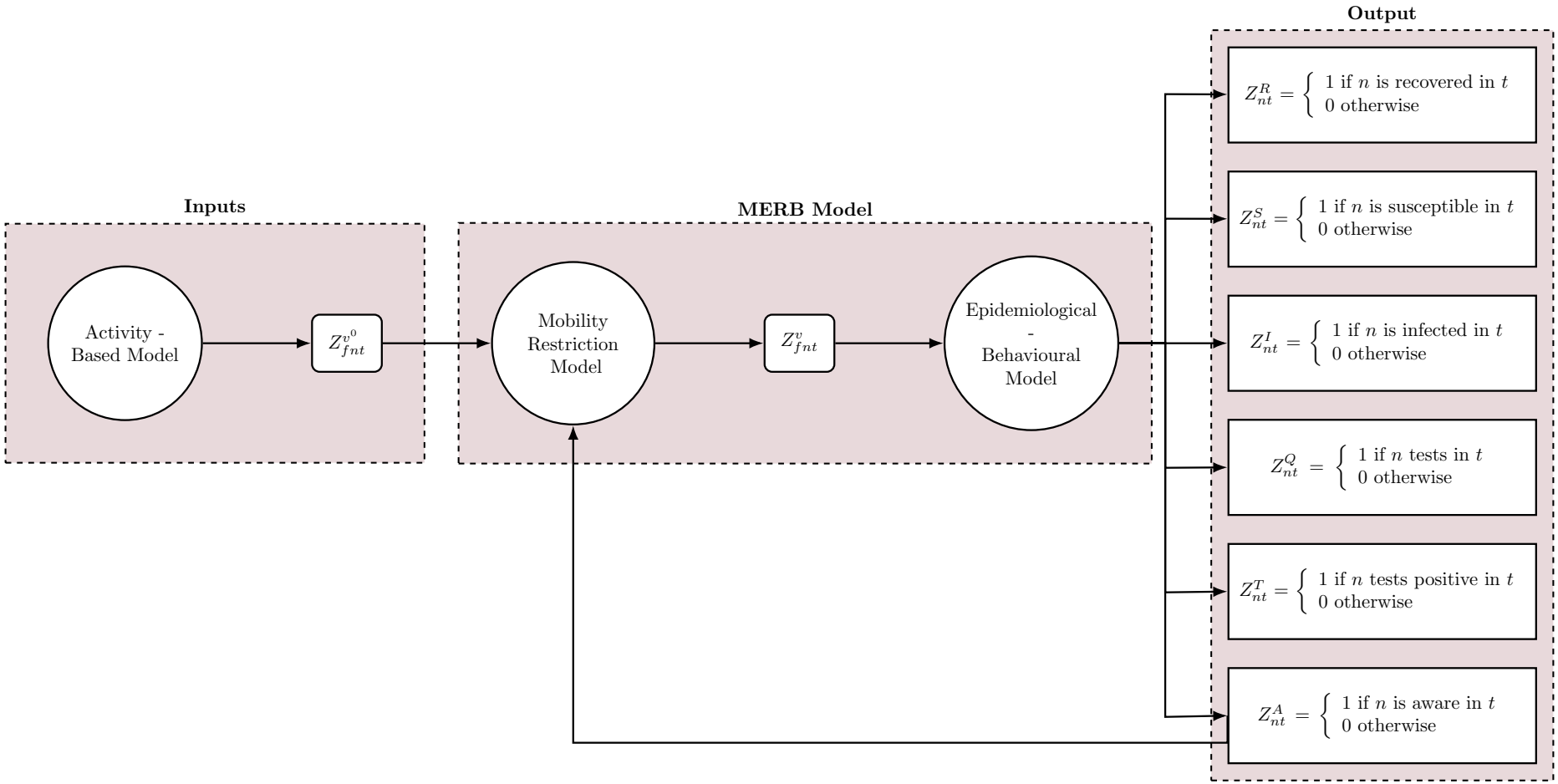
- Traditional models often overlook the **heterogeneity of individual actions** and their impact on the spread of disease (Hackl and Dubernet 2019; Eubank et al. 2004; Perez and Dragicevic 2009).
- Existing research does not fully integrate **socioeconomic and health factors** influencing mobility and 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).

Main Problem: Lack of disaggregated data available.

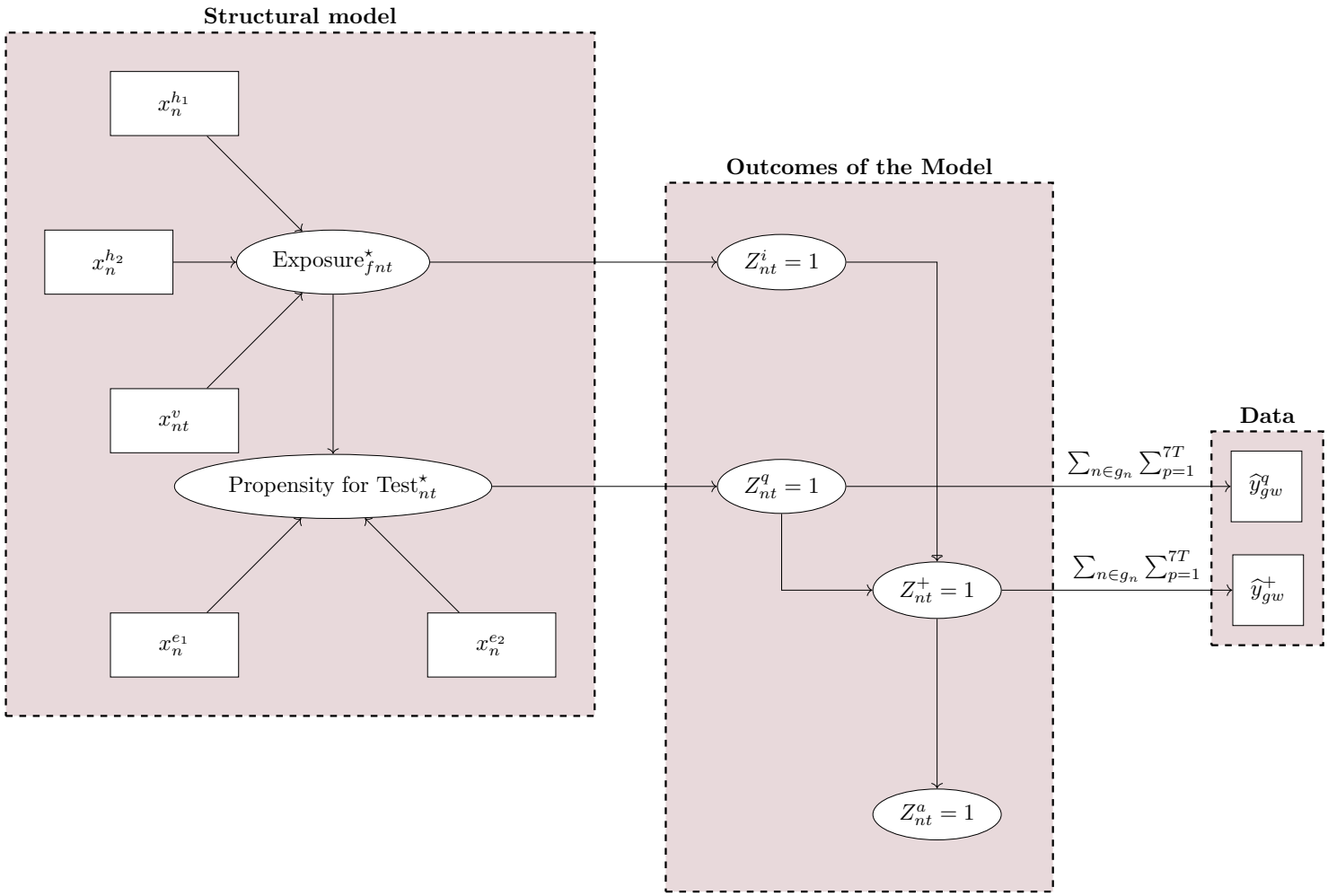
Outline of this talk

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



Epidemiological-Behavioural Model



Epidemiological-Behavioural Model: Structural equations

Exposure State

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

Propensity to Test

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

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

Measurement equations

Since available **data** tends to be **aggregated**,

we link the tests performed for each individual and timestep, with the observed number of tests by defining:

$$P^q_{gw} = \sum_{n \in g} \sum_{t=1}^{7T} P(Z^q_{nt} = 1), \tag{1}$$

$$P^+_{hjk\ell} = \sum_{i \in h,j,k\ell} \sum_{t=1}^{T} P(Z^+_{nt} = 1). \tag{2}$$

Beware: The lack of panel data presents a challenge because the modeling of testing choices is influenced by preceding ones.

Epidemiological-Behavioural Model: Measurement equations

Data & Likelihood functions

We use the **negative binomial distribution** for modeling the count data.

$\mathcal{L}_1(\theta)$ - testing

$$= \sum_{i \in g} \sum_{w \in W} \Big(\log \Gamma(\widehat{y}^q_{gw} + r_1) - \log \Gamma(r_1)$$

$$- \log \Gamma(\widehat{y}^q_{gw} + 1) + r_1 \cdot \log \left(\frac{r_1}{r_1 + P^q_{gw}} \right)$$

$$+ \widehat{y}^q_{gw} \cdot \log \frac{P^q_{gw}}{r_1 + P^q_{gw}} \Big),$$

$\mathcal{L}_2(\theta)$ - testing positive

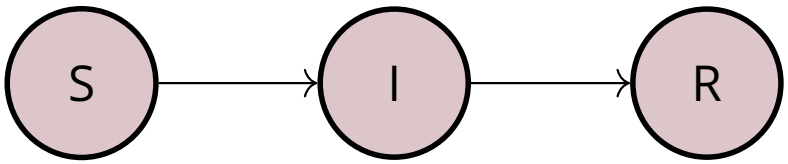
$$= \sum_{i \in h,j,k\ell} \sum_{\ell \in L} \Big(\log \Gamma(\widehat{y}^+_{hjk\ell} + r_2) - \log \Gamma(r_2)$$

$$- \log \Gamma(\widehat{y}^+_{hjk\ell} + 1) + r_2 \cdot \log \left(\frac{r_2}{r_2 + P^+_{hjk\ell}} \right)$$

$$+ \widehat{y}^+_{hjk\ell} \cdot \log \frac{P^+_{hjk\ell}}{r_2 + P^+_{hjk\ell}} \Big).$$

where r_1 and r_2 are the parameters of the negative binomial, $\widehat{y}^+_{gw}$ are the positive tests per age group, g, and week w, and $\widehat{y}^+_{hjk\ell}$ are the positive tests per age, h, gender, j, and municipality k, per day ℓ.

Dynamics of the MERB model: Health State Transition



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

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

Dynamics of the MERB model: Health State Transition

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

$$P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1) = \frac{1}{1 + e^{-\mu E^\star_{nt}}} \tag{3}$$

$$P(Z^s_{n(t+1)} = 1 | Z^s_{nt} = 1) = 1 - P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1) \tag{4}$$

$$Z^i_{n(t+1)} = \begin{cases} 1, & \text{if } \pi < P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1) \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

π: random uniform value

Dynamics of the MERB model: 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 infected 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:

$$Z^r_{n(t+1)} = \begin{cases} 1, & \text{if } \pi < P(Z^r_{n(t+1)} = 1 | Z^i_{nt} = 1) \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

Dynamics of the MERB model: Health State Transition

Finally, we assume that an individual who recovers remains recovered until the end of the simulation.

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

Dynamics of the MERB model: Testing process

What about the testing process?

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

$$P(Z^q_{n(t+1)} = 1) = \frac{1}{1 + e^{-\mu Q^\star_{nt}}}, \tag{7}$$

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

$$Z^q_{n(t+1)} = \begin{cases} 1, \text{if } \pi < P(Z^q_{nt}) \\ 0, \text{otherwise.} \end{cases} \tag{8}$$

Dynamics of the MERB model: Mobility Restriction Model - What about awareness?

Assumption:

Only people that are infected and test positive will quarantine.

$$Z^a_{nt} = Z^i_{nt} Z^+_{nt} \tag{9}$$

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

$$Z^v_{fn(t+1)} = \begin{cases} Z^v_{fnt} & \text{if individual's } n \text{ outcome is 0, and} \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

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.

Results: Case Study - Population of Vaud

Study Focus

- Our study examines the population of **Vaud in Switzerland** (**823'456 individuals**).
- A **synthetic population** is generated to simulate **socio-economic characteristics** and **daily schedules** for each individual (Horl and Balac 2021).

Data requirements

- **Open-source data** including **number of tested individuals** ($\widehat{y}^{q}_{gw}$) and **positive tests** ($\widehat{y}^{+}_{gw}$) individuals (see CloudPlatform 2021).
- Data from **Federal Office of Public Health** including the **tested positive** individuals ($\widehat{y}^{+}_{hjk\ell}$) with their **age, gender and municipality** information (see Riou et al. 2021).

Results: Case Study - Population of Vaud

Data and Model Development

Due to the unavailability of disaggregated data:

- The exposure level of an individual (E^*_{nt}) is calculated as:

$$E^*_{nt} = \beta^0_{E^\star} + \beta^h_1 x^h_{n1} + \beta^v x^v_{nt}, \tag{11}$$

- The propensity to test of an individual ($Q^\star_{nt}$) is calculated as:

$$Q^\star_{nt} = \beta^0_{Q^\star} + \beta^e_1 x^e_{n1} + \eta_{E^\star} E^\star_{nt}, \tag{12}$$

where:

- x^h_{n1} represents the age of the individual, x^e_{n1} represents the employment of the individual.
- β^h_1, β^v,β^e_1,$\beta^0_{Q^\star}$ and $\beta^0_{E^\star}$ are parameters to be estimated.
- $\eta_{E^\star} = 1$.

Results: Case Study - Population of Vaud

How do we deal with endogeneity?

$$E^*_{nt} = \beta^0_{E^\star} + \beta^h_1 x^h_{n1} + \beta^v \circled{x^v_{nt}},$$

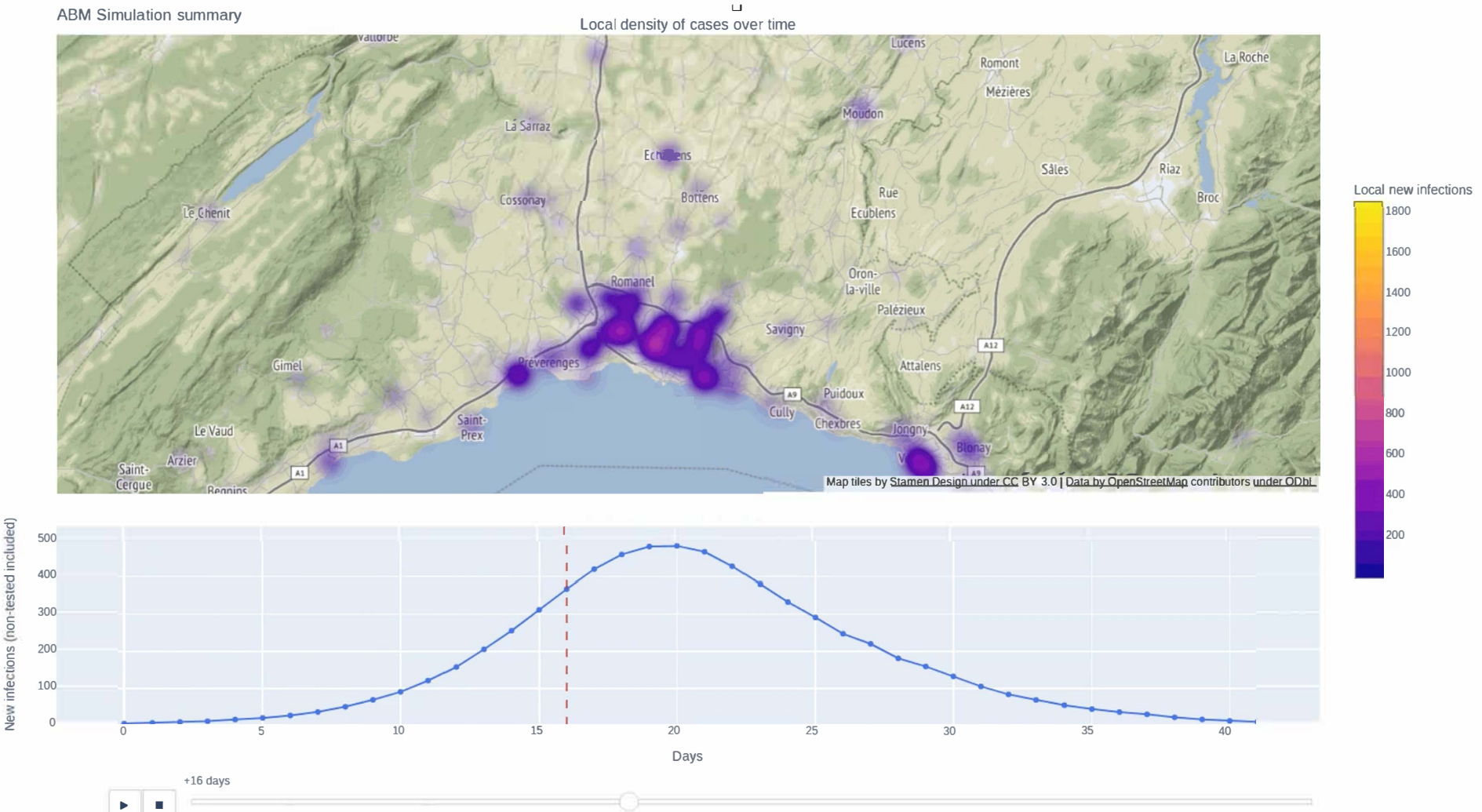
x^v_{nt} is endogenous in time

Algorithm 1 Model Calibration

Require: Simulation model, Optimization solver, MSE threshold
Ensure: Calibrated parameter values

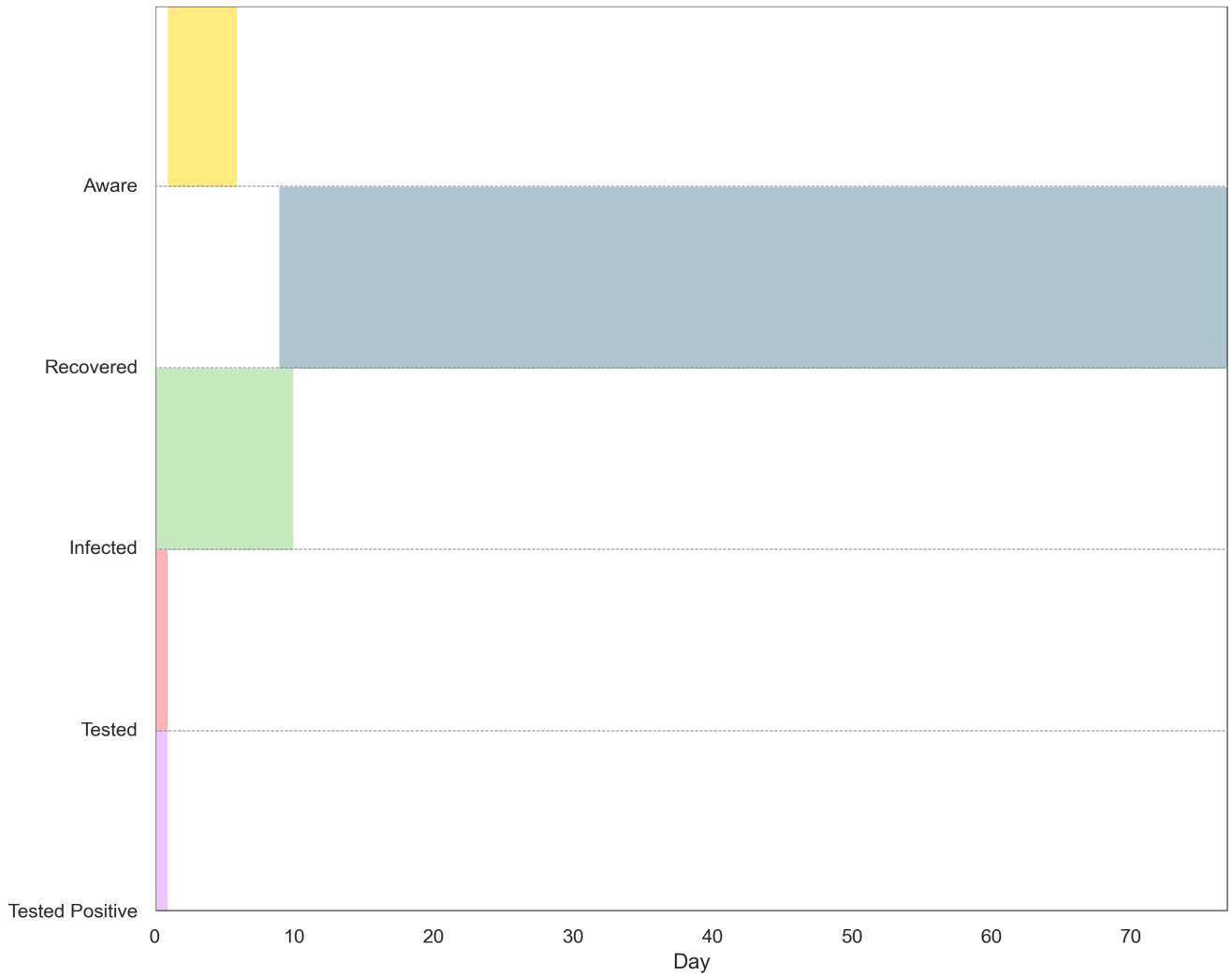
1: Initialize MSE $= \infty$
2: Run simulation to obtain x^v_{nt} for each n and t
3: Use obtained x^v_{nt} as an exogenous variable
4: Compute Equations for (8) and (9) using obtained x^v_{nt}
5: **while** MSE > 0.05 **do**
6: Solve optimization problem (31), with constraints (32)–(38) to find θ^* with objective function and constraints
7: Run simulation with parameters from θ^* to obtain simulated $x^{v'}_{nt}$
8: Calculate MSE between simulated $x^{v'}_{nt}$ and obtained x^v_{nt}
9: **if** MSE ≤ 0.05 **then**
10: Output calibrated parameter values θ^*
11: **else**
12: Repeat steps 4-8
13: **end if**
14: **end while**

Results across geographical area



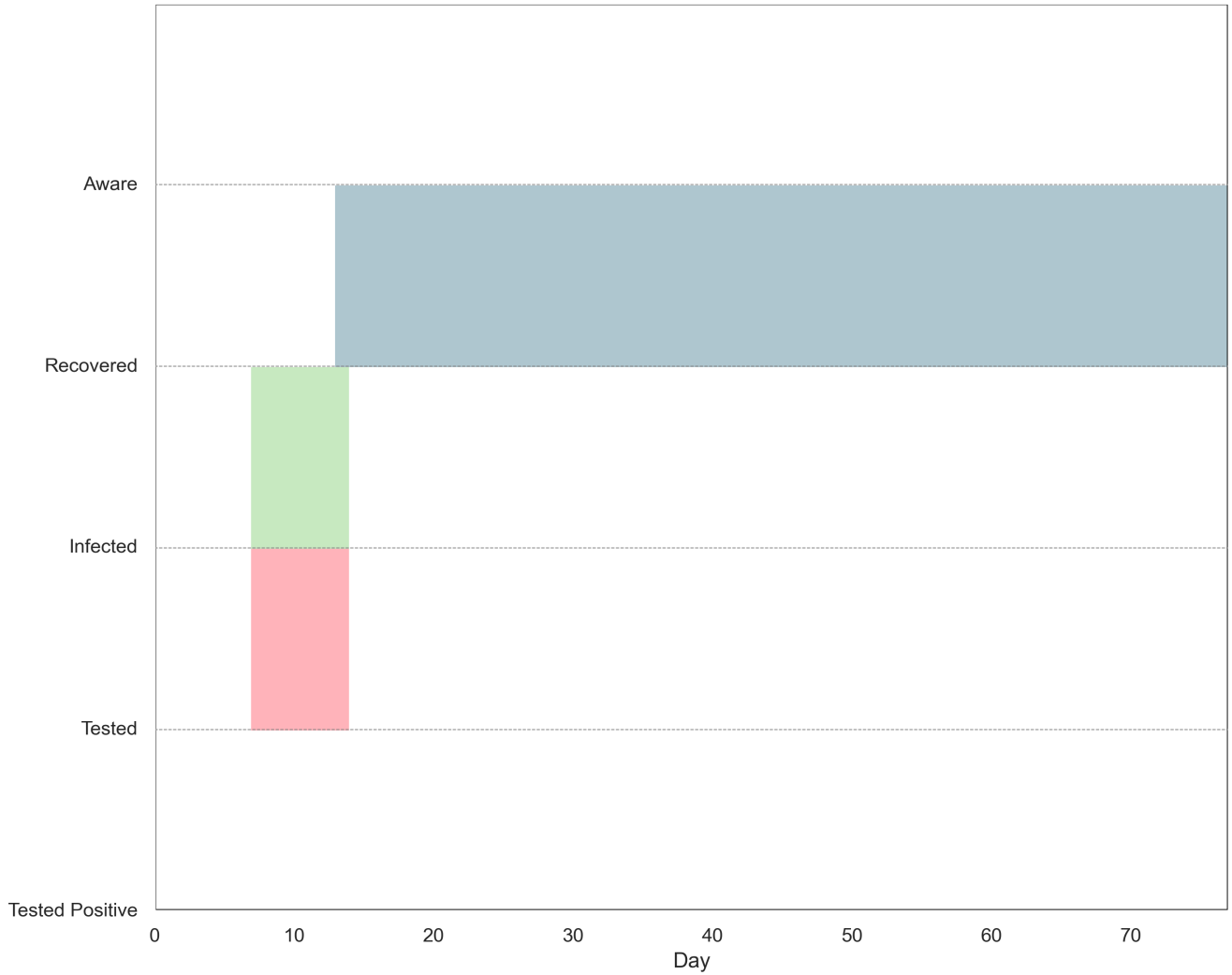
Results across individuals: insights on behavior

The 'ideal' behavior:



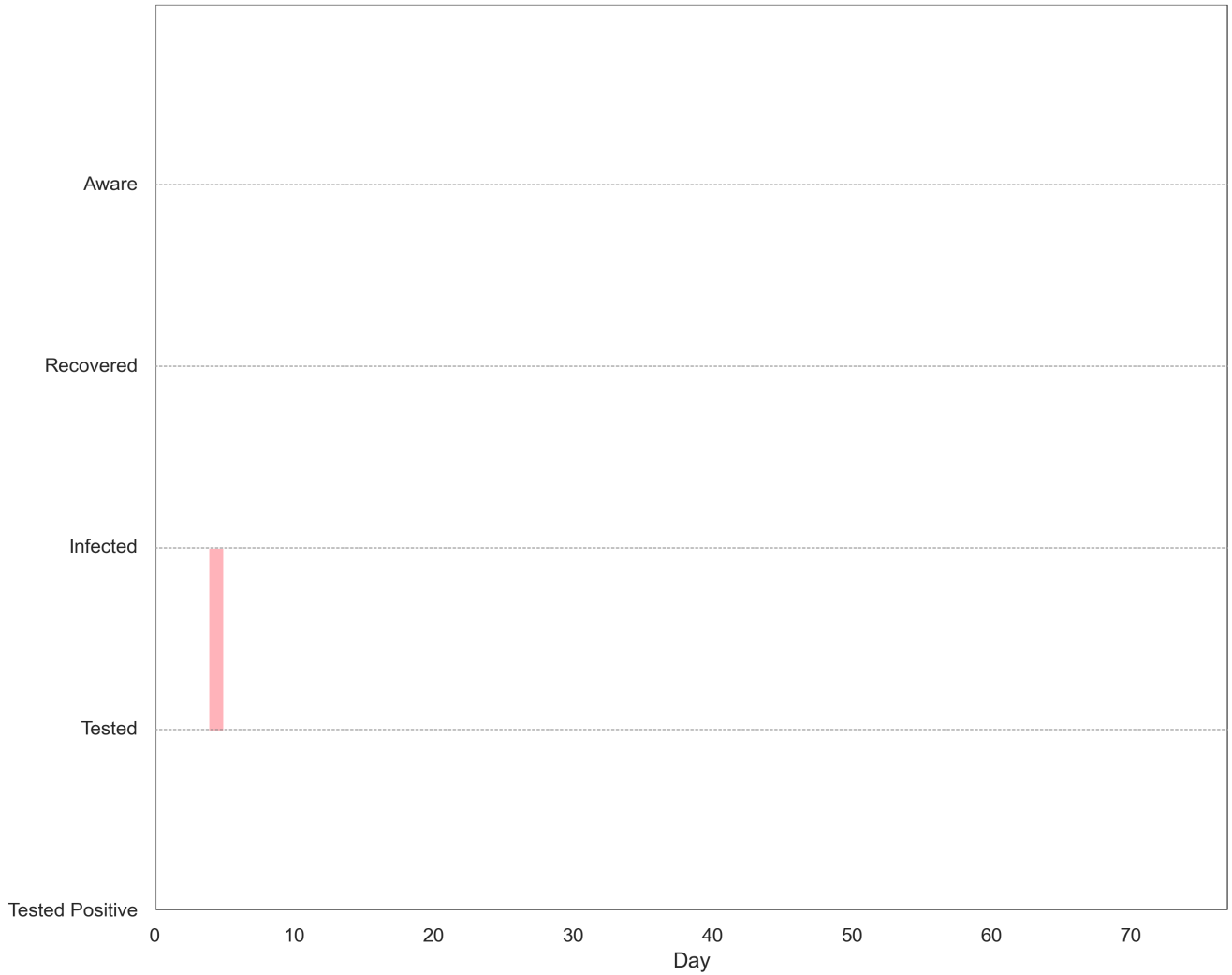
Results across individuals: insights on behavior

The 'unaware' behavior:

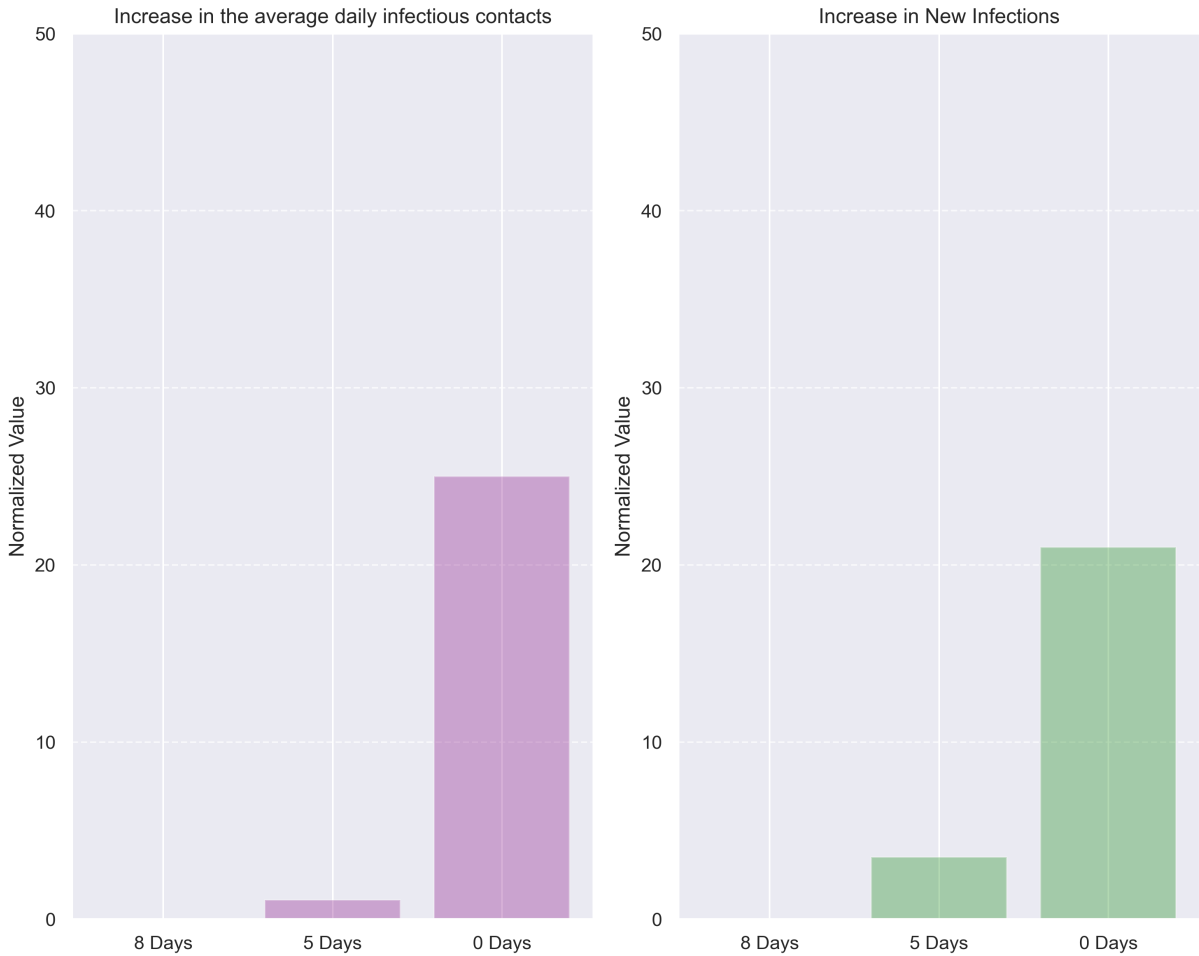


Results across individuals: insights on behavior

The 'cautious' behavior:



Results by applying different confinements



Conclusion and Future Work

Conclusions:

1. A computationally efficient tool (4m 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. Run the model for different cantons and see how the testing behavior impacts the spreading.
4. 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.

Thank you for your attention

Testing process

The outcome of the test is computed by:

$$P(Z_{nt}^{+}=1|Z_{nt}^{i}=1)=P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{i}=1)P(Z_{nt}^{q}=1 \text{ and } Z_{nt}^{i}=1)=$$
$$P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{i}=1)P(Z_{nt}^{i}=1|Z_{nt}^{q}=1)P(Z_{nt}^{q}=1)=$$
$$P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{i}=1)P(Z_{nt}^{i}=1)P(Z_{nt}^{q}=1),$$

$$P(Z_{nt}^{+}=1|Z_{nt}^{s}=1)=P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{s}=1)P(Z_{nt}^{q}=1 \text{ and } Z_{nt}^{s}=1)=$$
$$P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{s}=1)P(Z_{nt}^{s}=1|Z_{nt}^{q}=1)P(Z_{nt}^{q}=1)=$$
$$P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{s}=1)P(Z_{nt}^{s}=1)P(Z_{nt}^{q}=1),$$

$$P(Z_{nt}^{+}=1|Z_{nt}^{r}=1)=P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{r}=1)P(Z_{nt}^{q}=1 \text{ and } Z_{nt}^{r}=1)=$$
$$P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{r}=1)P(Z_{nt}^{r}=1|Z_{nt}^{q}=1)P(Z_{nt}^{q}=1)=$$
$$P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{r}=1)P(Z_{nt}^{r}=1)P(Z_{nt}^{q}=1).$$

$P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{i}=1)$, $P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{s}=1)$, $P(Z_{nt}^{+}=1|Z_{nt}^{q}=1 \text{ and } Z_{nt}^{r}=1)$ are taken from Ai et al. 2020.