The Anatomy of Simulated Covid-19 Epidemics in

Multiple Interacting Populations and Spaces

Ernie Chang MD PhD, Ken Moselle PhD

Victoria BC

INTRODUCTION

Typically, equation based models fit historical data to predict trends into the future, where the key assumptions are a population uniformly distributed subject to random transmission of the infective agent. These dynamic models do well when populations and times are large, but do not model complex movements of persons in and between different types of social environments very well.

CovidSIMVL is an agent-based model which supports three levels of simulation: the temporal dynamics of viral growth within persons, the interaction between agents that are susceptible and those who are transmitters, and the interaction of populations that share common spaces in time-varying patterns.

The classic paradigm for simulation applies to Covid epidemics as much as to many complex systems: that simple rules governing behaviour at each level, when put together, create dynamic systems whose behaviour is not easily predictable from component rules. Studying real complex systems offer many problems which simulations can avoid by creating and observing many variations of these systems, with the ability to record granular detail at will.

BACKGROUND

We have posted two articles about CovidSIMVL and its dynamics as seen in a single fixed Universe. The first of these addresses the relationships between the population density and agent mingling activity as determinants of the intensity of progression of an epidemic. Here, the population density is expressed as the Hazard Radius of an agent, given that the arena for the model is a fixed 800x600 pixel frame. The Mingle Factor affects the stochastic random walk of an agent – the higher it is, the more likely the agent is to cover the arena, and to encounter other agents.

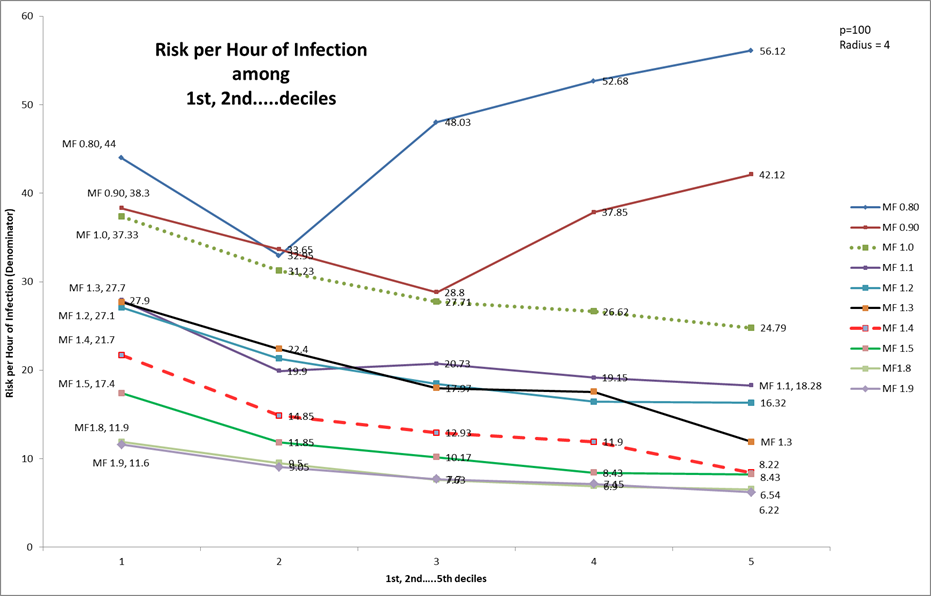
Clearly, the larger the Hazard Radius, the more likely a susceptible agent is to be contacted by a transmitter, and similarly, the larger the Hazard Radius of a transmitter, the more likely it is to contact a susceptible. This is a proxy for population density as well as susceptibility (which would act to change the Hazard Radius of a particular agent).

The Mingle Factor, again, increases the likelihood of contact between agents. It is easy to see that if the Mingle Factors are reduced to zero in the middle of an epidemic, no new infections would occur.

The first article reports on conjoint values of HzR (Hazard Radius) and mF (MingleFactor) that give rise to R0 values at the point either of extinction (no further transmitters) or at 50% of the population infected. From these relationships, one can set simulation parameters for different intensities of Covid epidemics.

In addition, the concept of Risk per Hour (perhaps better termed “theta”) is a measure of the average time (in generations) between infections, at various points in the epidemic. Thus, theta-10 would be the average number of generations for the first 10% of transmissions, and theta-20 would be the average generation count for the first 20%, and so on. Since we use a population of 100, a 10% transmission is the same as 10 agents becoming infected in that period. Theta is, to put it another way, the number of hours between infections within the time period in which 10%, 20% etc of agents are infected.

It was shown that although Mingling Factor changes the base theta-10 values (higher Mingle Factors produce lower theta-10s), the slope of the theta values as the epidemic progresses from theta-10 to theta-50 and theta-60 is a reflection of the intensity of the epidemic. A declining value (downslope) indicates that the epidemic is increasing in intensity, whereas an upward slope reflects an epidemic slowing down and heading to self-extinction. This is illustrated in the following graph, taken from the first article.



The second article looks further into the stochastic morphology of the chains of transmission that are generated in the course of a single source epidemic in a fixed Universe. It is intuitively obvious that rampaging epidemics have many transmitters simultaneously active, and therefore require fewer generations to infect the simulated population, whereas epidemics that run to extinction with partial population penetration must have fewer simultaneously active transmitters, and take more generations to cover the same population infected.

We developed the notion of a Transmission Tree, since the simulation can capture every transmission between agents, and the time of the transmission. The concept was developed of the shape of the Transmission Trees in terms of the depth (longest path), average depth (over all path lengths), the breadth expressed as the number of leaves (which is the same as the number of distinct paths).

To quantify the aggressiveness of the epidemic (the more simultaneity, the shorter the total generations needed, the more agents are infected), we proposed two measures that combined all these factors. The most sensitive, which we adopt, is simply termed Q, and is defined as:

|  |  |
| --- | --- |
| Q = | (Leafs/AvD)/Gen\*100 |
|  |  |



This table shows that as the Mingle Factor and R0 (for 50% population infection) increases, the value of Q increases, in bands, and that these bands reflect the morphology of the Transmission Tree in the metric Leafs/AverageDepth fairly well.

We will apply these metrics to the subject of this article, which is an exploration of the dynamics of Covid-19 epidemics where multiple interacting populations of various structures move between different types of spaces (which we term Universes) that have their own intrinsic mingling factors.

DESCRIPTION OF THE OUT-OF-BOX TRIAL

In the first trial, the simulation is configured as programmed and specified in the .CSV files, without altering any run-time parameters (specifically, Hazard Ratio and MingleFactor). The default Hazard Ratio as an initial setting for every agent is 5, and the MingleFactor for each Universe is set to 10.

We now describe the structure of the 9 Universes, and then the population structures, and then the schedule of movement of the agents.

This set of Trials is created to simulate the areas of a school, and interactions between the students, teachers, their families who may work as staff in a Long Term Care facility, and a bar, and these agents may also be visitors to Long Term Care, and be clients of the bar.

There are 9 Universes in this Trial called U0 to U8. The names assigned to them are:

U0 Classroom 1

U1 Project/Lab Room

U2 Lunchroom

U3 Playground

U4 Classroom 2

U5 Teacher’s Lounge

U6 LTC (Long Term Care)

U7 High Mingle site (Bar, reception, party)

U8 HOME

HOME is a unique Universe, in that transmissions among agents who are in the HOME universe only take place between members of the same family.

There are agents in Long Term Care who do not leave U6, and there are agents in U6 who are not parts of existing families.

Each Universe can be assigned its own intrinsic Mingle Factor at run time, and this acts as a multiplier to the move generated for each agent in the Universe, each such agent having their own Mingle Factor, which may be related to their role in that facility, as set in the population file.

The population of 100 agents is set up in the Population.csv file. As configured, these look like:



The Family Structure

We begin by enumerating the social roles (total 100)

1. Students 3 x 10 in groups A, B, C
2. Teachers 6
3. Grandparents 7
4. LTC residents 28 of which 10 are associated with families
5. LTC staff 14
6. High-mingle site (Bar) staff 10
7. Spouse community work 5

The family structures are complex and we have created 19 families in this community. Before we enumerate them, the general structure of a family is:

1. Heads of family (one or two) with social roles
   1. May be two-person eg Teacher + LTC staff
   2. May be single-person
   3. May be grandparents only (two-person with social role)
2. May have children (0 to 4)
3. May have multi-generational (grandparents not heads of family)
4. May have LTC resident as part of family (grandparents, aunts)

The 19 families have been carefully defined and structured with the use of analog tools (poker chips, pins, paper clips, checkers pieces, anchor bolts, etc). The enumeration of the families follows, in no particular order:

F0 LTC(S), BAR, A, C LTC(S) – staff, BAR – bar staff, A,C – student groups A,C

F1 T, BAR, LTC(R), C T – teacher, LTC(R) – LTC resident

F2 T, LTC(S), LTC(R),G,B,C G – grandparent

F3 LTC(S),BAR,B,C,LTC(R),LTC(R)

F4 G,G,B Grandparents as head of family with one grandchild

F5 T,BAR,A,C

F6 T,LTC(S),A,A,B,C,G,G,LTC(R) biggest family unit – four kids, 2 grandparents, one LTC res

F7 BAR, LTC(S), A,B,B

F8 BAR,LTC(S),A,C

F9 BAR,LTC(S),A,C

F10 BAR,S,G,A,C Bar worker, spouse, grandparent, two children

F11 LTC(S), S Long Term Care staff, spouse – no kids

F12 LTC(S),S

F13 LTC(S),S, LTC(R) this will be among the visitors to LTC

F14 LTC(S), S, LTC(R), LTC(R)

F15 T,LTC(S), G, LTC(R)

F16 BAR,LTC(S),A,B,B three kids, two in same group

F17 BAR, LTC(S), A,B,C

F18 T,B, LTC(R) single parent (teacher), one child, one LTC resident

This microcosm is tightly coupled between the workers who are teachers, LTC staff, and BAR workers (BAR being a place-holder for high-mingle spaces). With so many being in the same family units, this scenario may have a higher rate of crossover transmission between the school and LTC unit not just because of frequenting the high-mingle bar, but also because of the family units, which have most LTC workers with children, and/or spouses who are teachers or work in the HIGH-MINGLE bar.

This information is incorporated in the Population.csv file, which specifies the movements between Universes for each agent, according to the following overall schedule.



This is a single day 24-hour schedule, repeated over and over. However, the capability has been built in to CovidSIMVL to specify different DDHH for movements to take place.



Clearly, this is a deterministic rather than a stochastic schedule. We can easily modify this to be more probabilistic in one of two ways: first, to provide an alternative set of Universes to move to anytime there is a departure for an agent from a Universe; second, to introduce stochastic elements into length of stay, so it might be for example, 4 hours +/- 5%; third, both. For now, most of the non-determinism in CovidSIMVL is in the viral growth dynamics, and the distances moved at each generation.

Whether this pattern of movement is reasonable, too restricted, or too broad is difficult to verify. This is a pattern, and others can be specified, but a true set of movements can be at best a snapshot obtained with great difficulty, without any guarantee of generalizability. Our approach is to try a reasonable configuration, and to be open to any others proposed.

The population.csv file is the technical specification for the movements for each person, and has a structure like this.



These are the .csv lines for the students (agents) 0,1,2 which are the pIDs. The “sno” column is the “STOP number” for the agent, and these must be present, and in order. The ETA is the time of arrival at the universe “@U” and the ETD is the time of departure from that universe “@U” to the new Universe “>U”. Note that when at HOME (U8) the mF is high (8) so family members may contact one another.

The Role is either “R”, or “A”, or “T”. The first, “R”, means that they are “resident” in that universe for the time they are there, and the second “A” means “Attached” as a staff member would be, and “T” would be for “Transient” visitors, shoppers, guests at bars, etc. In general, these are meant to convey different levels of activity. However, each line has its own MingleFactor, which describes the activity level of the agent in that Universe.

For example, a person may be a staff member in LTC, during which the MingleFactor (mF) may be high, say >6. Then in the bar, they may be quiescent, and have an mF of 3. As a visitor to LTC, they may be a volunteer, in which case they might have an mF of 4, and at home they might after 11 have an mF of 1.

Since mF influences the extent of contact with another agent in the same Universe, this fine level of detail can be applied to classes of persons, as well as to specific individuals. CovidSIMVL has the capability of tracking all movements, and transmissions. The movement algorithm generates a proposed move, using a Pareto distribution, then modifies it by the individual mF and the Universe mF for a final position, which is in turn modified if it goes outside the fixed arena.

Note that the FamKey needs to be specified just for the first line of a new agent.

Finally, we come to the Case file, another .csv file. This describes the initial agents that are infective, and the system permits the user to invoke additional case files during a run. The format is simple, but has to include the viral load, and the number of days since the agent was infected, in order to comply with viral dynamics.

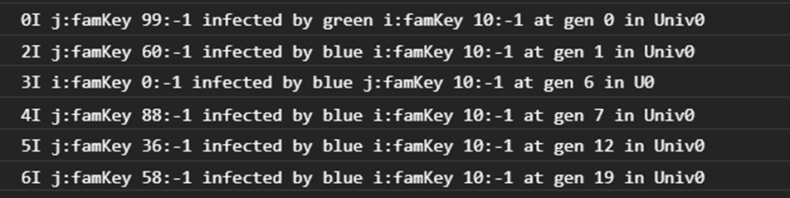


The operations for running CovidSIMVL using the .csv files are found in the /docs directory of the gothub repository github.com/ecsendmail/MulitverseContagion, which contains documentation of how to operate the system. This is an open source, public domain software under the framework of the GNU License.

RUNNING THE TRIAL

The trial “Out-Of-Box” was run by invoking the index.html file in the repository, and then opening the two .csv files which were loaded from the repository into a local directory. The browser used was Chrome, and devtools was invoked by right-click and selecting “inspect” in the window.

“console” was chosen in devtools, so that the transmission details are displayed, which were copied into an Excel spreadsheet at the end of the run. It looks like this:



As the mF was 10 for each Universe out of the box, the simulation was expected to go to completion, and to end after relatively few generations. The model as programmed does not remove symptomatic agents from general circulation, nor place them into hospital or transition them to death, or to quarantine. Thus, this simulation is set for mild cases with no mitigation.

At the end of the run, the console.log was analyzed as discussed below, using the tools. Screen shots were captured, as shown here.



There’s a lot going on here. First, the top left boxes in gray show that the run stopped on Day 41 Hr 17, with 9 survivors (Green), and 91 Oranges (Inert after 10 days symptomatic).

The rightmost charts show the cumulative status on each day for the entire population of 100 agents. The top is the typical SEIR graph, with two Y-axes….the right showing Yellows (incubating), and Reds (symptomatic), while the left Y-axis shows Greens (susceptibles), Blues (pre-symptomatic infectious) and Orange (inert – next phase of Red).

The top right SEIR chart shows the symptomatic cases in red peaking at Day 20, and the susceptibles (green) reaching 50% at about Day 12. A closer examination of the red curve reveals 2-3day plateaus around days 21, 26, 30 and 33, lasting for 2-3 days each. These small increases in transmissions might be explained by within family transmission after a family member gets infected. HOME is a Universe that everyone except LTC residents go to, and to model family members contact we increased the mF for persons at HOME in their schedules. Thus, in the console.log, we might expect to see a cluster of transmissions in U8 at around the times indicated. Indeed, in the Universe detail windows (the 3x3 grid of charts and tallies) we can see for U8 (right bottom) that there are four peaks at the times in question.

Despite the initial seed of 5 students as transmitters (pID 10 to 14) the detail charts show that subsequent transmission among the student and teacher population is not high. The way to interpret the individual Universe charts is to keep in mind that populations are moving in and out of these spaces periodically.

Take the top left window for Classroom 1. We know that at most 10 students and one teacher occupies it, and that there are 3 such groups. So there would be 11 susceptibles (green) counted when they are present, while the group with the 5 transmitting students would show up as a column of 5 REDs. Indeed we see that the Greens start at just above 10, and decline relatively slowly, while the REDs (infected) peak around day 22.

The Project/Lab room (U1) see the Greens (susceptibles) maintain their numbers into day 14. The decline for the Lunchroom (U3) is interesting. It shows the Greens level at 20 to about Day 10, and similarly the Reds are at the same height, but because a different axis is used, their values are at 5.

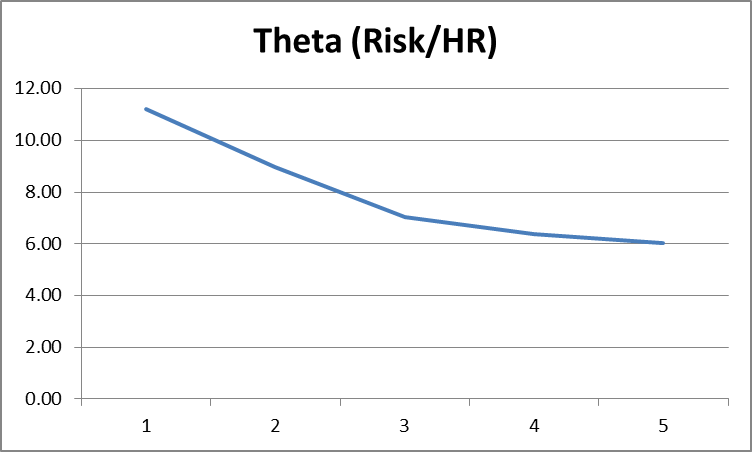
On the other hand, the LTC (U6) chart shows that the susceptibles (Greens) fall sharply so that by Day 10 they have gone from an initial daily low of 32 (and high of 38) with staff and visitors, to half that number of 20 Greens. Clearly this is the hot spot for this complexity of Universes, at these model parameters.

The Bar (U7) is interesting in that the visits are of course from susceptibles (Green) initially, and they are maintained at the same level till about day 6, when the composition of visitors shift to have fewer Greens and more infected.

These general observations can be reflected in more detail in the metrics that have been proposed above for characterizing Covid-19 epidemics.

Out-Of-Box Trial and Theta Values

Theta is the Risk Per Hour of infection, which is computed by the count of generations for the first 10%, 20% etc. of transmissions. This can be done for the Metaverse of all the Universes seen as a whole population.



The first two columns are derived from the console.log trace. From this we get the table of Theta values and the graph with the downslope showing the acceleration with time of the epidemic, but slowing.

Next, we look at the school Universes from the console.log (sorted by U).

Theta-n is the number of generations to reach n% of infections, and for a population of 100, n can be considered the number of agents to be infected.

For small numbers, such as U0, the number of generations to reach its one and only infection is 396. That is its theta.

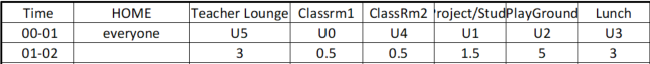
For U1, it takes 62 generations to reach the first infection, and 135 to reach the second infection. Thus, the theta-1 value is 62, theta-2 is 67 and theta-3 is 68.

Thus, for larger numbers, we can use theta-n%, and for smaller numbers, we would use theta-k.

For U3, theta-1 is 180, theta-2 is 186, theta-3 is 180, and theta-4 is 171.T

For U4, theta-1 is 178, and theta-2 is 173.

The simulation shows that in spite of the larger numbers in Playground (U2) and Lunchroom (U3) with 2 cohorts each of 30 sharing the space, the theta-k values are worse for Project/Lab (U1). A possible explanation is that the Playground and Lunchrooms are used for one hour each cohort, whereas they spend more time in the other rooms. The worse theta-k for the Project/Lab (U1) can be further predicted by the mF used when students arrive in these rooms:

The Project/Lab is assigned a MingleFactor three times higher than the two Classrooms. Thus, we see here the complexity of detailed simulated spaces: duration of exposure, activity, and density all matter.

Turn now to the LTC, which leads in the speed of progression of transmissions.

Now, should we use theta-n% since the numbers are larger and we want to know the time between infections within the space.

There are 55 infections.

10% is 5.5, round to 6 gen 148

20% is 11 gen 182

30% is 16.5, round to 17 gen 214

40% is 22 gen 253

Theta-10% is 14.8

Theta-20% is 9.1

Theta-30% is 7.1

Theta-40% is 4.4

Recall that the lower the number, the faster the speed of transmission. Thus, an infection every 4.4 generations is much worse than 173.

In this simulation, the Bar is relatively harmless. Only four transmissions took place within U7.

Theta-1 is 164

Theta-2 is 126

Theta-3 is 94

Theta-4 is 105

All rather safe, compared to the theta of <10 for LTC.

The situation for HOME transmissions between family members is interesting. The 10 families involved have four repeats – F07 with three, and F2, F8 and F17 with two each. Referring back to the Family Structure, these are not the largest families, which are F06 with 9, and F2 and F3 with 6 members each.

A member of F06 (pID 25) did get infected in generation 47, but this did not lead anywhere. To understand this, we turn to the Transmission Tree(s) for this trial.

Transmission Tree and Intensity of the Epidemic

The initial conditions were set for five transmitter agents, so we end up with five distinct transmission trees, each rooted at one of the initial infected agents, pID 10-14. They look like this:

Recall that each row of the Transmission Tree (in its path-representation) is a path of the tree, which represents a chain of transmission from an index case to the leaf of the tree, from which no further transmissions take place in this simulation.

The previous article on Transmission Trees presented the metric Q that takes the shape of the tree, as well as the time involved, and the number of distinct agents infected, to be an estimate of the aggressiveness of the epidemic, with higher numbers being a faster expanding epidemic.

We work through the example of the largest tree, rooted at pID=12, and leave the others to the reader.



The values of Q for the other trees are:

Root 10 T/E 17.06 Q 19.61

Root 11 T/E 15.94 Q 18.22

Root 12 T/E 16.10 Q 28.12

Root 13 T/E 4.21 Q 0.67

Root 14 T/E 8.08 Q 2.69

Note the asymmetry in the size of tree that follow from the initial seed of transmitters, and the heterogeneous property of the path lengths, which vary from a high of 6 in Tree 12 to lows of 1 in both Tree 10 and Tree 11.

Indeed, it appears not just natural, but inevitable, that stochastic approaches to the modelling of contagion-based epidemics should end with trees that are not uniform in shape, where some nodes, near the root, give rise to many leafs, while others wither. For example, in tree 12, node 12 gives rise to all 35 new infections, but node 75, an immediate descendant of node 12, is the ancestor of 14 chains and 28 new infections.

From this simulation (and many others), it would appear that a stochastic approach predicts an inherent asymmetry in path generation, and that so-called super-spreader persons may be artifacts of the natural probabilistic progression of a contact-based epidemic.

This simulation trial also presents the inherent asymmetry in spaces with different mingling characteristics visited by varying mixes, and the notion that some time-based occupations of these spaces would give rise to more transmissions than others is also borne out as a natural consequence of stochastic behaviours.

For example, in the span between generations 396 to 399 (3 hrs) there were 4 infections in U6; between 248 to 255 (7hrs) there were 5 infections in U6; between 301 to 311 (10 hrs) there were 6 infections.

EPIDEMIC DRIVERS

In a complex interacting set of spaces (Universes), some naturally have more conditions conducive to rapid spread than others (LTC vs Classroom2), which can be identified by their Q metrics. For the less contagious Universes, they must be supplied with more transmitters as earlier ones go past their viral best-by dates. These are Universes that lead in transmissions and the creator of new ones that feed into the others. We can identify these by their velocity graphs, as follows.

These charts show, for selected Universes, the relationship between time and the percentage increase of new cases per day over the previous day.

The clustering makes it clear that U6 rises first, followed closely by U8, then the school rooms, with the Teacher Lounge and Bar the latest in terms of days since the start of the trial.

This kind of information is useful, as it shows the places that are the drivers where new transmitters are created, and the places that lag behind and transmit new cases as a result of the activity in the lead spaces.

Here, it is no surprise that the LTC and HOMEs are the first places where transmissions take place. In the case of LTC, the residents are densely situated with the staff and visitors, and have long exposure times to one another.

In the case of the HOMEs, everyone (except LTC residents) go home, and the mF level for agents are set high, so family members get the opportunity to contact and transmit to one another.

Perhaps it is less expected that the bar as a high mingle Universe should be the laggard, peaking at almost 40 days. A closer examination shows that U7 has an average of 6.6 agents present and a maximum of 13. Since the initial see of transmitters are students, and they don’t frequent the bar, the active transmitters would have to come from teachers or LTC staff. So even if the bar has a higher mF than the classrooms, the transmitters and the prolonged opportunity to mix gives the schools a higher infection velocity than the bar.



Indeed, if we look at the first 25 transmissions, we see that 14 of them occurred in U6 (LTC), 5 at HOME (U8), 5 at school (two in U1, and one each in U2, U3 and U4) while only one transmission happened in the Bar, by pID=31 (a Teacher).

The first school transmission was in U1, in generation 62 which is on the 3rd day.

Of the first 10 infections, all but one were by the initial seeds (pID 10 to 14), and only one was in a school space. The others were either at HOME or in LTC, as visitors.

This shows the power of population density (LTC) and mF (HOME).

COMPARING THE OUT-OF-BOX TRIAL TO OTHERS

In what follows, we will briefly look at the metrics for some other simulations using the same population structure, but changing the mF for selected Universes, to see whether we can change the Universes that are more aggressive as sources of new transmissions, and thus are ht spots that fuel the overall contagion.