Policy Simulation Using Evaluation

May 6, 2021

Canadian Evaluation Society 2021

Gregory Mason
University of Manitoba

TRADITIONAL TERRITORIES ACKNOWLEDGEMENT

The University of Manitoba campuses are located on original lands of Anishinaabeg, Cree, Oji-Cree, Dakota, and Dene peoples, and on the homeland of the Métis Nation.

We respect the Treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in a spirit of reconciliation and collaboration.

umanitoba.ca

Plan for the workshop

Hour 1

- Part A – Define/position simulation as part of evaluation
- Part B – Features of a policy that support simulation
- Part C – Deterministic simulation (Forecasting)

Hour 2

- Part D – Scenario design (Rate impact on energy poverty)

Hour 3

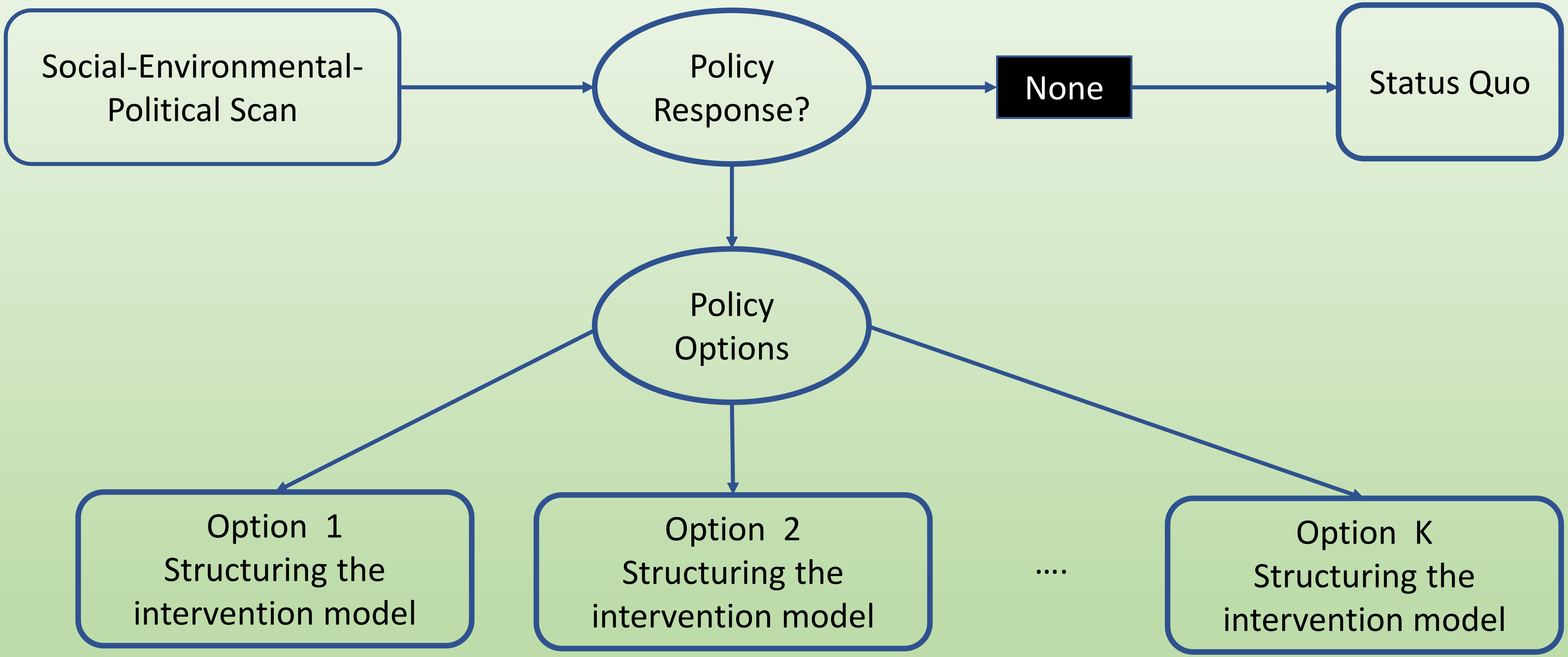
- Part E – Monte Carlo Simulation (Vaccination Benefits)

Note: all images/videos in this presentation have been acquired under licence, creative commons, or lie in the public domain (YouTube) for the "live" presentation.

Part A

Define/position simulation as part of evaluation

The policy development cycle



Social-Environmental-Political Scan

Policy Response?

None

Status Quo

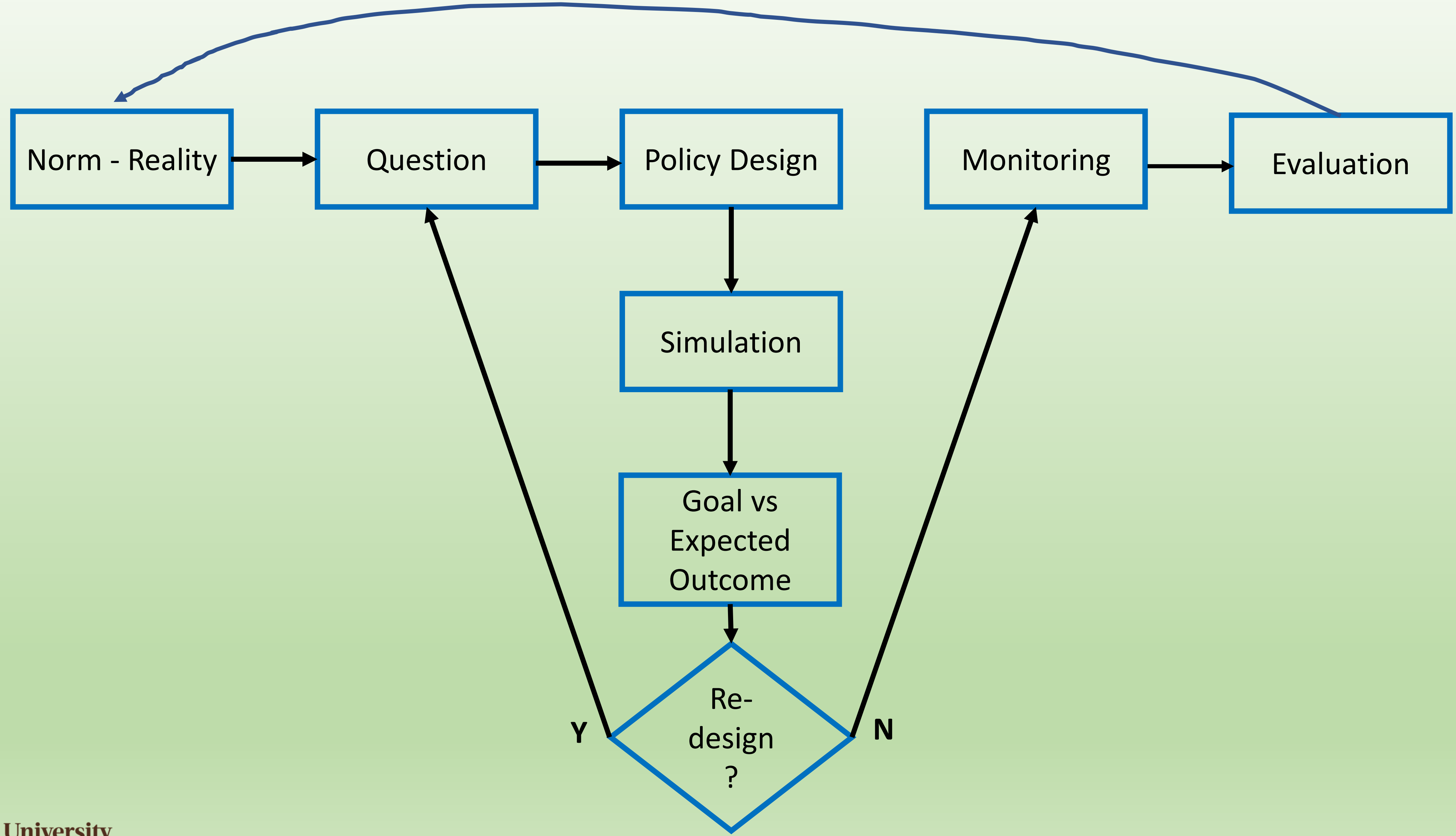
Policy Options

Option 1
Structuring the intervention model

Option 2
Structuring the intervention model

….

Option K
Structuring the intervention model


Norm - Reality
Question
Policy Design
Monitoring
Evaluation
Simulation
Goal vs Expected Outcome
Re-design ?
Y
N

Simulation is embedded in all policy evaluation

- A universal child care program will raise GDP by 1.5% and pay for itself.
- Social distancing and masking will reduce COVID hospitalizations by ?%.
- A minimum wage/universal basic income will reduce poverty.
- Raising electricity rates will increase energy poverty.
- Vaccinations have a benefit-cost ratio that exceeds 1.

Simulation resembles theatre

- Rehearsal may involve
  - A single line, part, scene,
  - Or the whole play, culminating in the dress rehearsal
- But simulation is much more:
  - Different actors rehearing the same lines
  - Rehearing the play with different cast
  - Randomly assigning roles to different cast members for the entire play
  - Randomly assigning cast members to different roles for each scene/act/play.

"That was just a simulation. Nothing can prepare you for the kind of monkey bars you'll find in an actual war zone."

CartoonStock.com

Simulation as *prophecy*

"a statement that says what is going to happen in the future, especially one that is based on what you believe about a particular matter rather than existing facts.

https://dictionary.cambridge.org/dictionary/english/prophecy

- Biblical prophecy are conditional forecasts - Statements of outcome <u>conditional</u> on causal factors

> "You've got to change your evil ways, baby
> Before I stop lovin' you
> You've got to change, baby
> And every word that I say is true"
> *Santana – Evil Ways*

- Simulations present one or more potential futures arising from systematically altering parameters and/or extending variables in a ***causal*** model.

Part B
Features of a policy that support simulation

Simulations rest on causal models that are numerate

> “Causal models are mathematical models representing causal relationships within an individual system or population. They facilitate inferences about causal relationships from statistical data.”
>
> Stanford Encyclopedia of Philosophy

- In evaluation we often use causal descriptions/explanations that are verbal (Contribution Analysis), tabular, or graphical.
- Simulation requires must perforce a formal structure that may be
  - Algebraic
  - Logical
  - Decision sequential
- This gives the technique great power
- But, rests on assumptions.

Policy features the support simulation

Policies that *may* be simulated

- Tax/rate changes
- Statistical models (regressions and other multivariate models)
- Accounting structures
- Financial models
- Decision models

Policies that cannot easily be simulated

- Aspirational statements and moral suasion
- Organizational changes
- Guidance on changed practice
- Rules and norms drawn from qualitative data
- Non-quantifiable statements on causal impact
- Case Studies

The distinction has potential for ambiguity

Hard to simulate program policies: example

Evaluation of Transportation Development Centre's rail research and development

- Conducts research on behalf of the industry to promote safety, efficiency, advance knowledge, and expand training
- Like all R&D programs, success may require years to realize and outcomes may be difficult to attribute to program activities.
- Key informant data and case studies formed the core of the program.

- Transport Canada should strengthen its participation in the RRAB through establishing formal linkages with international counterparts, in order to ensure maximum complementarity and leveraging of rail R&D.

- Transport Canada, in consultation with the RRAB, should develop and implement a targeted, outcomes-based rail R&D plan, identifying specific information/technology needs, how they will be addressed, timelines, and the specific role for Transport Canada in each R&D project/program.

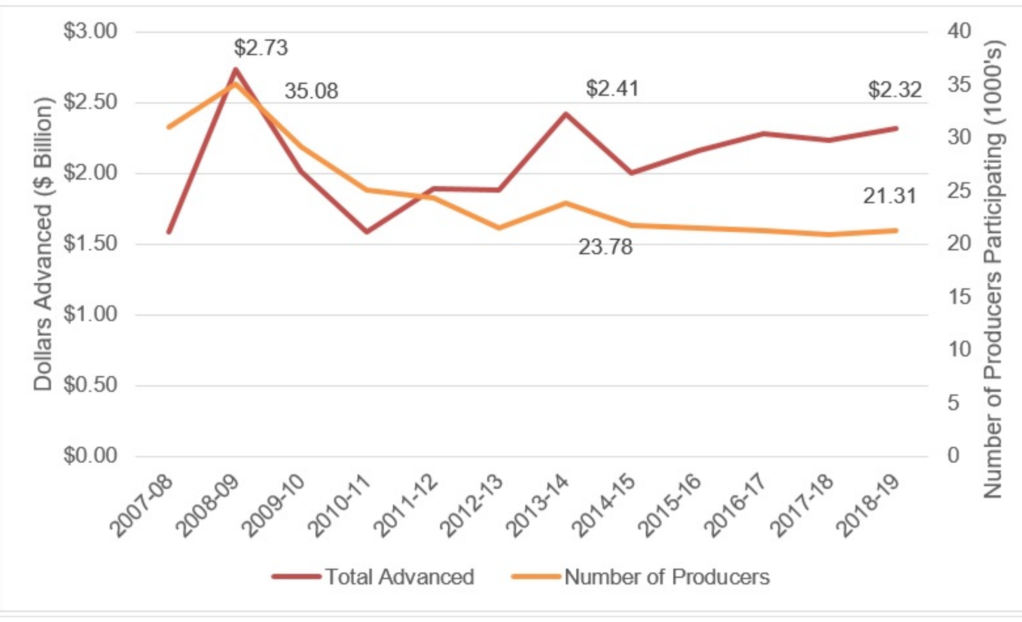
- Transport Canada should develop and implement a knowledge management strategy for its rail R&D.

Little opportunity for causal models

Program policies that could support simulation

<u>Evaluation of the programs under the *Agricultural Marketing Programs Act* (2014 to 2019)</u>

- The APP is administered by third party administrators and financing is guaranteed by the federal government to enable administrators borrow money from lenders to deliver advances to producers.
- Methodology – case studies, interviews, program activity measures, surveys, linked producer data to Statistics Canada.
- The availability of a micro dataset of participants and non-participants offered the opportunity to create a causal model of program uptake



- The micro data offers the opportunity to statistically estimate a causal model to explain program participation based on – interest rates, farm size, sector.
- Such causal models support simulation

Tools in causal modelling: Variables, logic and probability -1

Variables

- **Functions** that can take on a range of values
  - Interest rate
  - Infection rate
  - Population of (participants, non-participants, …)
  - Income
  - …
- **Parameters** modify variables and relationships among variables
- **Continuous/discrete**
  - Income (measured to the nearest dollar on an annual, monthly, weekly, hourly, minutely?
  - D = 1 if male and 0 if female (indicator or dummy variable)
- **Independent/Dependent**
  - Cause (with the important subset being the policy variables)
  - Effect or outcome

Tools in causal modelling: Variables, logic and probability - 2

Probability 1: Simple Gambling Models

Gambling has inspired he development of statistics and serves as a simple framework for understanding the probabilistic basis for policy simulation.

The simple coin toss uses the excel function **=RAND()**, and it is simple to model a

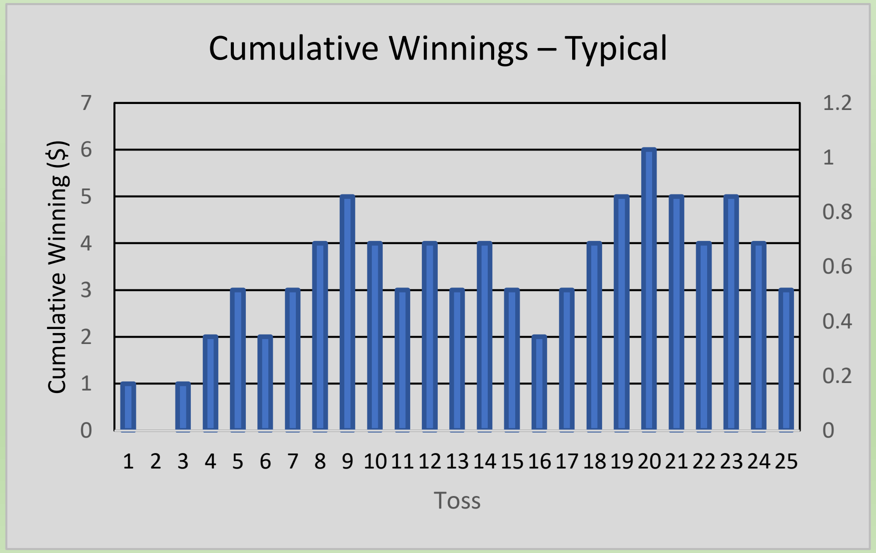
- **fair** (equal chances of head or tail (50-50 chance)
- **unfair** coin.

See Coin Toss.xlsx

Gambler's ruin

Rubbing or blowing on dice is assumed to place "Lady Luck" on one's side. Imagine this simple gambling game involving a coin-toss (fair coin).

"For every head, win a dollar, for every tail lose a dollar. Calculate net winnings at each toss"



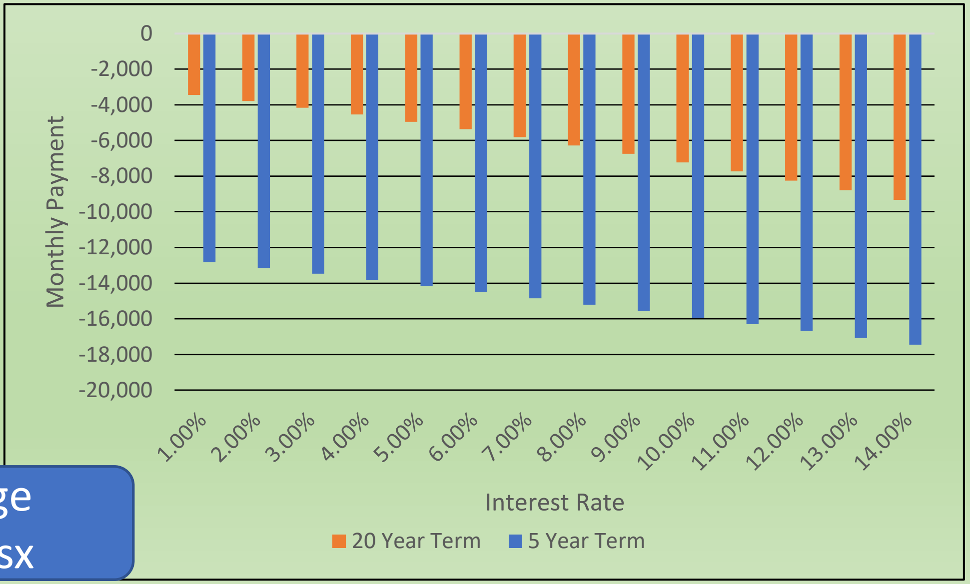
Tools in causal modelling: Variables, logic and probability - 2

Probability 2

- **Simulation models** are either
  - Deterministic
  - Probabilistic
- **Deterministic simulations** typically vary one variable/parameter
  - Mortgage applicants typically must ensure a “stress” test to determine whether the monthly payments will be “affordable” as the interest rate increases.
  - Many forecasts are deterministic, where the future replicates the past.
- **Probability models** exploit the rules of probability and usually must operate within a programming language structure such as Excel.

Deterministic Simulation: Mortgage Stress Test
With housing prices on the rise and the prospect of inflation with rising interest rates, prospective homeowners will face increased scrutiny on their capacity to service debt.

Excel offers a simple method for exploring the implications of varying the impact of interest and term (number of years to pay-down the loan)



See Mortgage Calculator.xlsx

Part C - Forecasting

Possibly the most common form of simulation is the forecast.

- Pre-pandemic federal budgets featured an annual forecast for key economic variables (growth in GDP, growth in personal income, growth in corporate income…)
- AKA … Simulation using regression
- These projections served to anchor revenue predictions.
- Taken with spending, this creates future projections of deficits/surpluses (what are those?)
- These simulations fall into three classes
  - **Simple projections** (where the future replicates the past)
  - **Recursive models** (where the projected deficit leads to increased spending to service the bonds covering the debt, which leads to revised spending, revised deficits, revised requirements to service the debt… and so.
  - **Recursive models with probabilistic parameters**, combine recursion and probability generated parameters; and seem to exist primarily to drive PhD students in economics completely insane.

University of Manitoba

Forecasting – Two traditions

Regression on times series data

- For most this is the common application
- Comprises one or more dependent variables as a function of one or more independent variables.

$$Y_t = f(X_t)$$

$$Y_t = a_0 + a_1 X_{1t} + a_2 X_{2t} \ldots + e_t$$

Time Series Analysis

- Used with high frequency data (monthly, daily, hourly….)
- Just the information in the dependent variable becomes the source for the forecasting model (Autoregression)

$$Y_t = a_1 + a_1 Y_{t-1} + a_2 Y_{t-2} \ldots.. + e_t$$

Forecasts as simulations

- Extending the regression into the future uses information based on the past
- Regression models rely on plausible projections of the independent variables
- Time series models also use information from the past, but rely solely on the parameters of the model. No independent variables require projection

Basic Forecasts

- The simplest simulation that assumes the future will track linearly based on the value of a_1 and time (t)

$$Y_t = a_0 + a_1 t$$

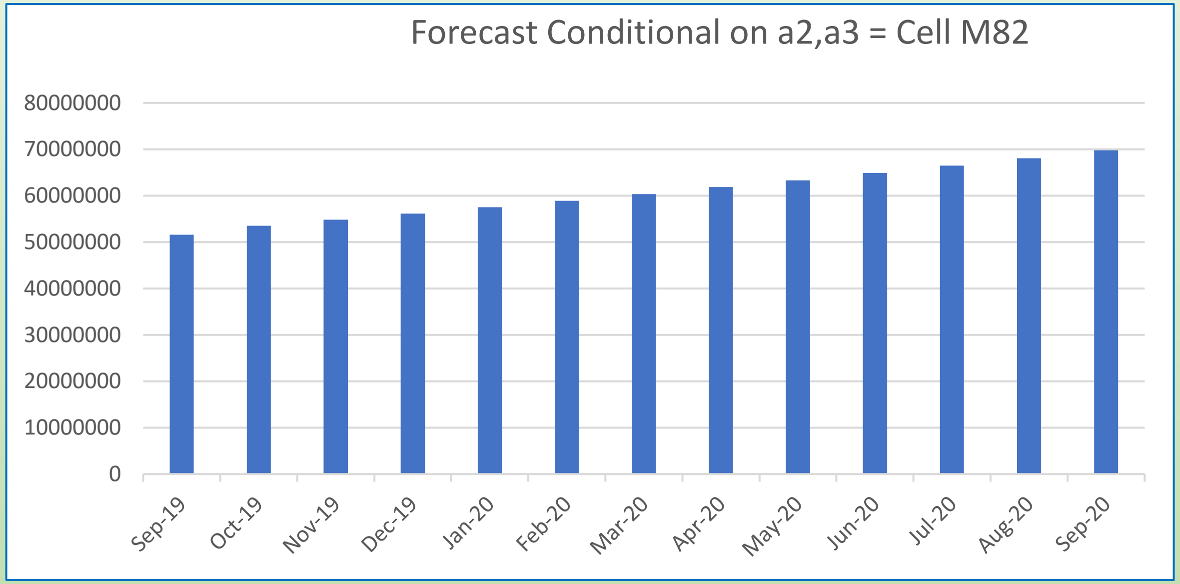
- This is a very naïve model, since time is not a variable … it is a crude proxy for social/economic/environmental change
- Two more useful models are

$$Y_t = a_0 + a_1 X1_t + a_2 X2_t + u_t$$

$$Y_t = a_0 + a_1 X1_t + a_2 Yt - 1 + a_3 D1t + a_5 D2t \ldots + a_{14} D11t + u_t$$

Retail sales forecasts

- This example (Forecast Models.xlsx) uses unemployment and wages to create model to create a model to predict retail sales.
- Such models require sufficient data, something that can challenge evaluations for many programs.
- Annual data usually presents insufficient information to create effective forecasts



See Forecasts.xlsx

Part D – Scenario design (Rate impact on energy poverty)

Analysis of the impact of rate changes on energy poverty

In Order 73/15, the Public Utilities Board (PUB) directed Manitoba Hydro to initiate a collaborative process to enhance bill affordability programming offered by Manitoba Hydro to its low-income customers.

- Energy poverty refers to circumstances in which a household is, or would be, required to make sacrifices or trade-offs that would be considered unacceptable by most Manitobans in order to procure sufficient energy from Manitoba Hydro.

- The basic measure is the ***simple ratio of income approach*** (SRIA) also referred as the *energy burden* .

- A household is energy poor if more than 6% to 10% of their income on energy (gas plus electricity)

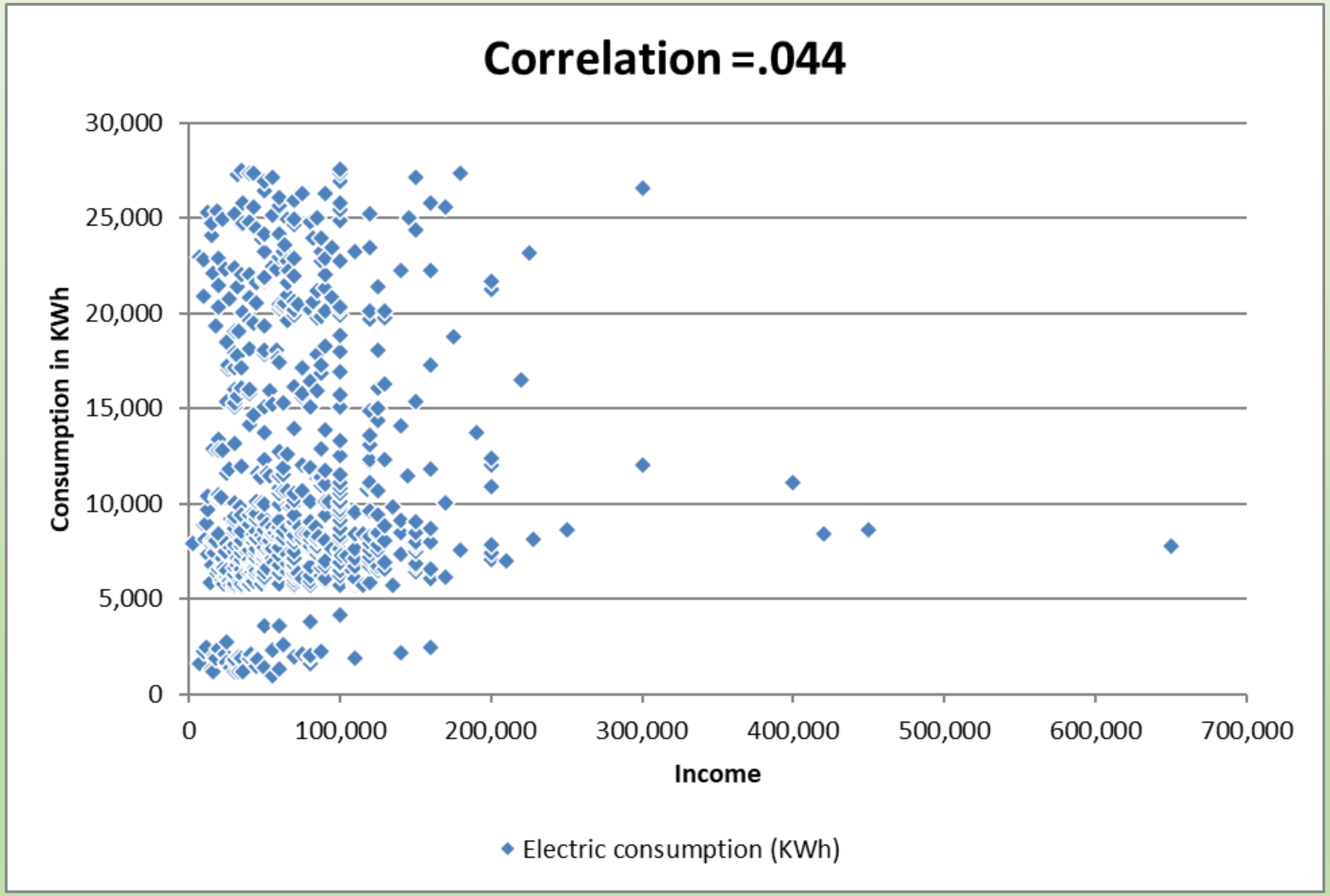
- Manitoba Hydro's residential use survey (2014) found that 14.3% of households had energy burdens exceeding 6% and 4.2% had burdens of exceeding 10%

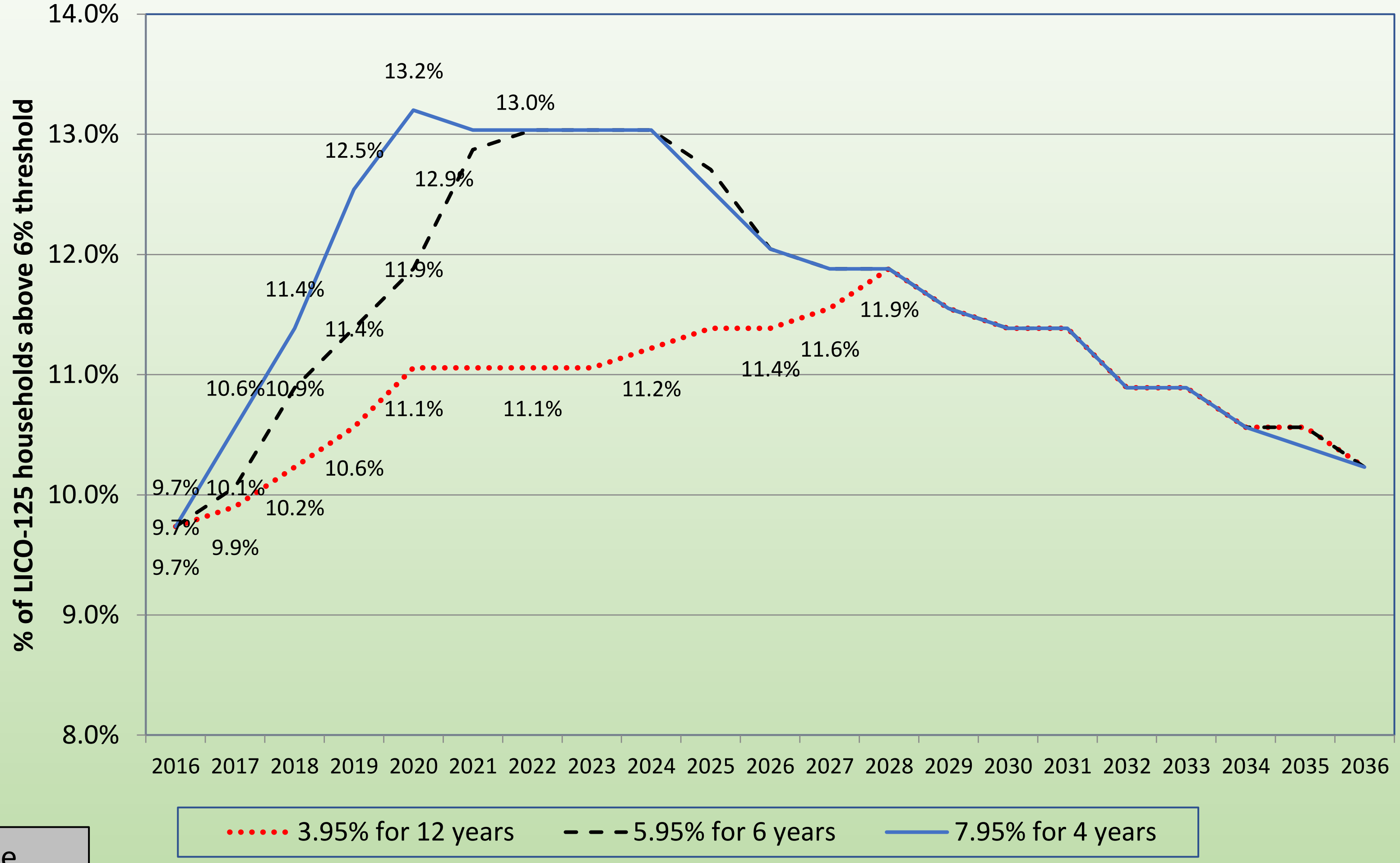
The simulation examined the impact of propose rate increases

Key assumptions

- The survey represents the population characteristics
- Consumers make no adjustments to energy consumption due rate increases
- No cost reduction in collateral costs due to rate changes:
  - The costs of collecting arrears do no rise (rate increase) or fall (special measures to mitigate rate increases)
- No interaction with other initiatives (demand side management such as subsidies to go solar).

Income and electricity consumption are not correlated

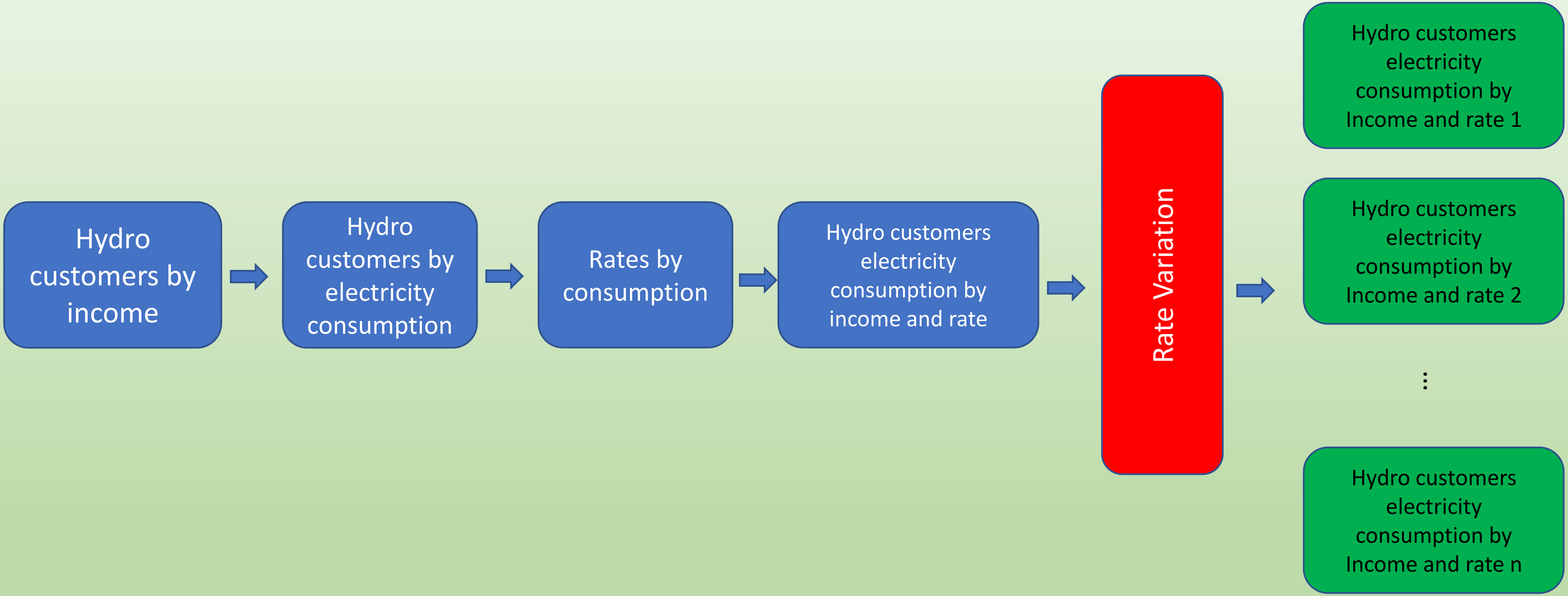




LICO-125 measures the poor + near poor (the numbers of customers whose income is less than 25% above the poverty line)

3.95% nominal electricity rate increases for 12 years
5.95% nominal electricity rate increases for 6 years
7.95% nominal electricity rate increases for 4 years

Modelling Proposed rate increases



Hydro customers by income

Hydro customers by electricity consumption

Rates by consumption

Hydro customers electricity consumption by income and rate

Rate Variation

Hydro customers electricity consumption by Income and rate 1

Hydro customers electricity consumption by Income and rate 2

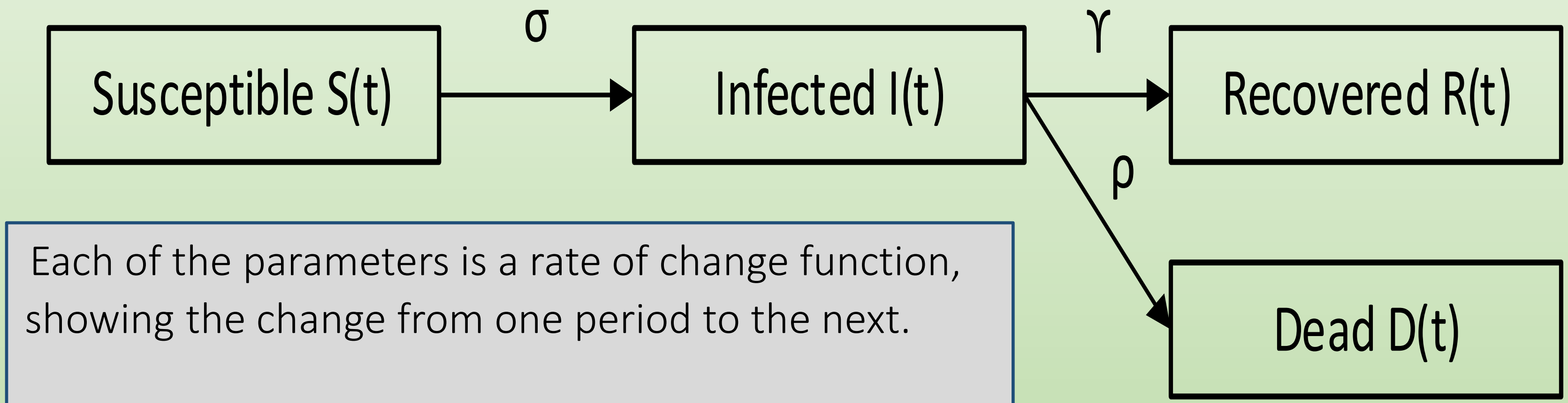
⋮

Hydro customers electricity consumption by Income and rate n

Part E – Monte Carlo Simulation (Disease progression and Vaccination Benefits)

Epidemiology – Flatten the curve

- We have all become familiar with “flattening the curve”, the evocative term to mark the progress of any controlled viral/bacterial outbreak
- The modelling for the advance-retreat of COVID rest on a classic SIR(D) model.



Each of the parameters is a rate of change function, showing the change from one period to the next.

- σ is the transmission rate (the rate at which those who are susceptible become infected.
- ϒ is the recovery rate
- ρ is the death rate.

Identities and initial conditions

- Simulation models have identifies to “glue” everything together.

$$S(t) + I(t) + R(t) = N$$

- N is the total population.
- The initial conditions are
  - S(0) … number of susceptible people at t=0
  - I(0) … number of infected people at t=0
  - R(0) … number of recovered at t=0

A gentle introduction to differential equations

- The essence of the SIR model are three equations that describe change over time.
- We can imagine discrete change over a week, such as the change in COVID cases in (name the jurisdiction).

$$Change = \frac{C(t+7) - Ct}{t+7-t} = \frac{\Delta Ct}{\Delta t}$$

- Now if the rate is constant, say 300 cases per day, we can write

$$\frac{\Delta Ct}{\Delta t} = 300t$$

- If we imagine shrinking Δt to an instant, we write $\frac{dC}{dt} = 300t$, a simple differential equation.

SIR System

$$\frac{dS}{dt} = -\sigma S * I \qquad ...1$$

$$\frac{dI}{dt} = \sigma S * I - \gamma I \qquad ...2$$

$$\frac{dR}{dt} = \gamma I \qquad ...3$$

See SIR.xlsx

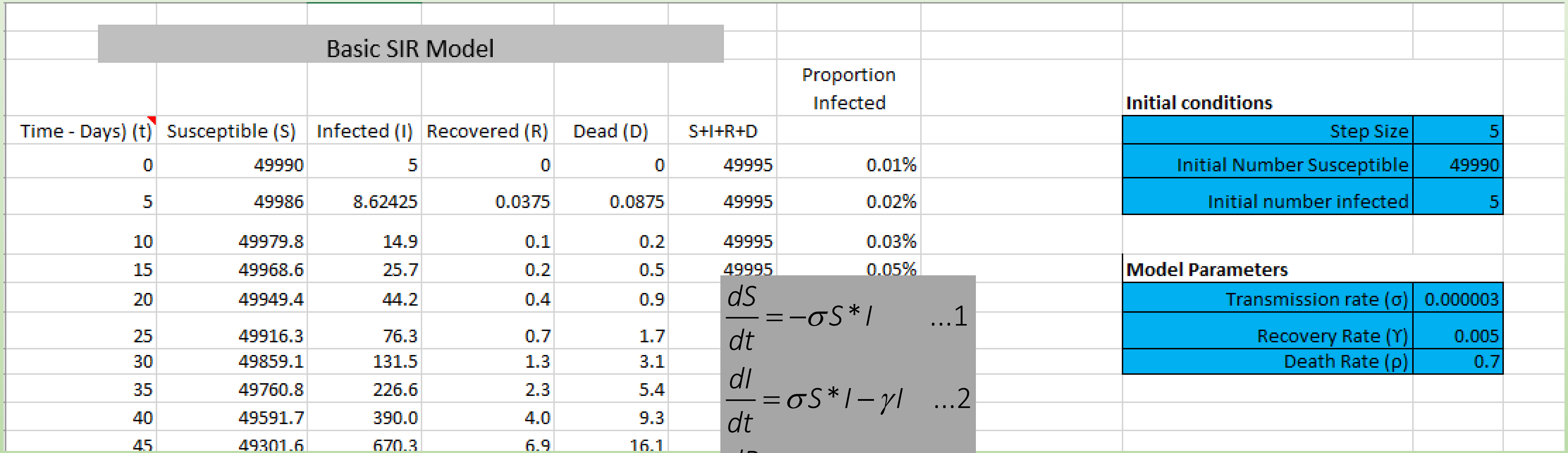
- The change in susceptibles (dS/dt) is negative since it depends on how the infected and susceptibles interact. The term S*I is the simplest description of this interaction and the simple
- The change in the numbers infected is (dI/dt), and this is growth in the numbers who become infected less those that recover.
- The change in recovered dR/dt is the proportion of infected who recover.

See SIR.MP4

These equations must be solved at the same time and each is a differential equation. Euler's ("oilers") method is a basic numerical calculation process that avoids the need to solve the solve the differential equation explicitly.

The spreadsheet (SIR.xlsx)

- It s possible to program the SIR model in Excel. The core part of the model appears below



Basic SIR Model

Time - Days) (t)	Susceptible (S)	Infected (I)	Recovered (R)	Dead (D)	S+I+R+D	Proportion Infected
0	49990	5	0	0	49995	0.01%
5	49986	8.62425	0.0375	0.0875	49995	0.02%
10	49979.8	14.9	0.1	0.2	49995	0.03%
15	49968.6	25.7	0.2	0.5	49995	0.05%
20	49949.4	44.2	0.4	0.9		
25	49916.3	76.3	0.7	1.7		
30	49859.1	131.5	1.3	3.1		
35	49760.8	226.6	2.3	5.4		
40	49591.7	390.0	4.0	9.3		
45	49301.6	670.3	6.9	16.1		

Initial conditions

Step Size	5
Initial Number Susceptible	49990
Initial number infected	5

Model Parameters

Transmission rate (σ)	0.000003
Recovery Rate (ϒ)	0.005
Death Rate (ρ)	0.7

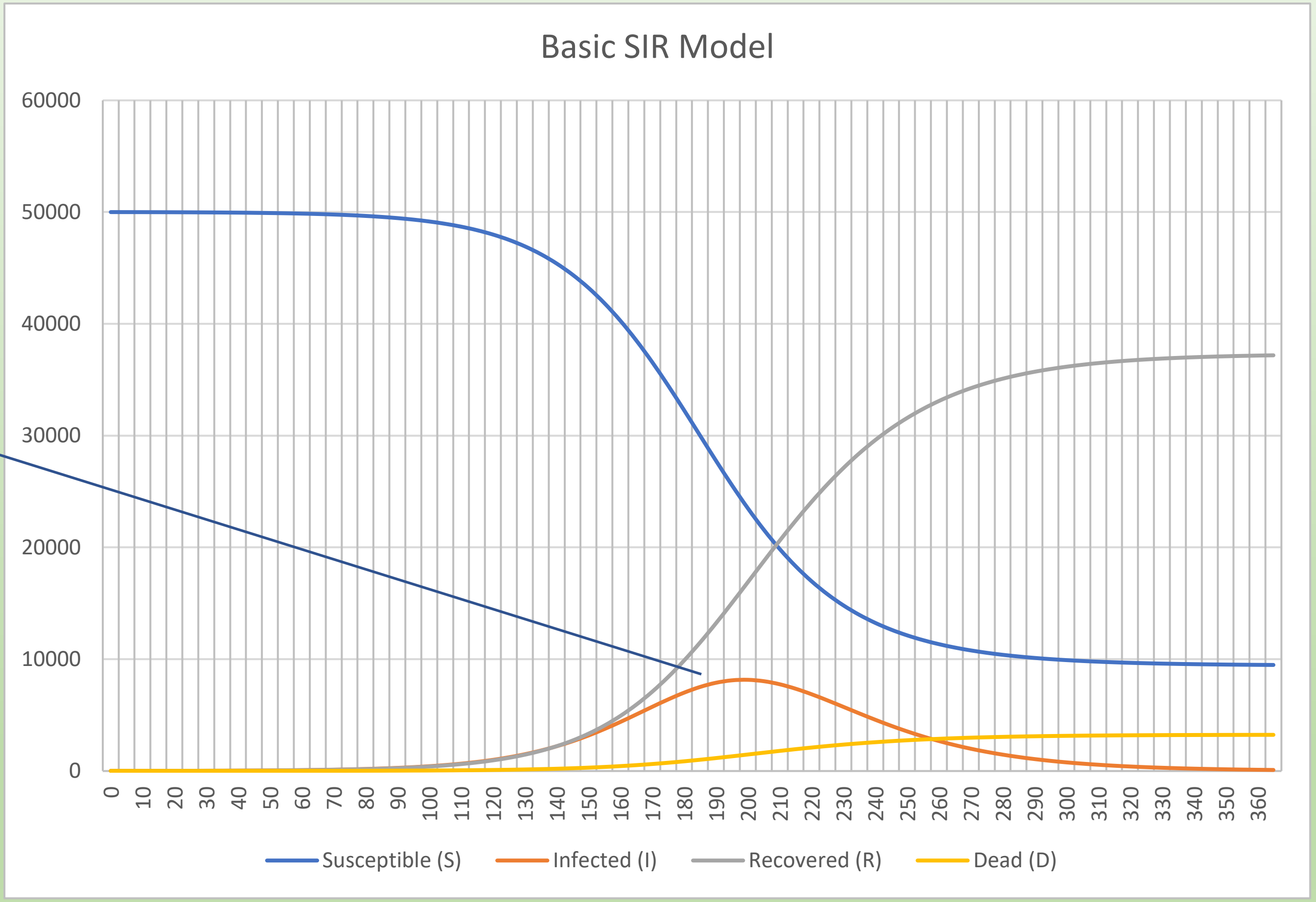
$$\frac{dS}{dt} = -\sigma S * I \qquad ...1$$

$$\frac{dI}{dt} = \sigma S * I - \gamma I \qquad ...2$$

$$\frac{dR}{dt} = \gamma I \qquad ...3$$

This example is adapted from, Verschuuren, G.M, *Excel Simulations*, Union Town Ohio, Holy Macro! Books, 2014

SIR Model Output

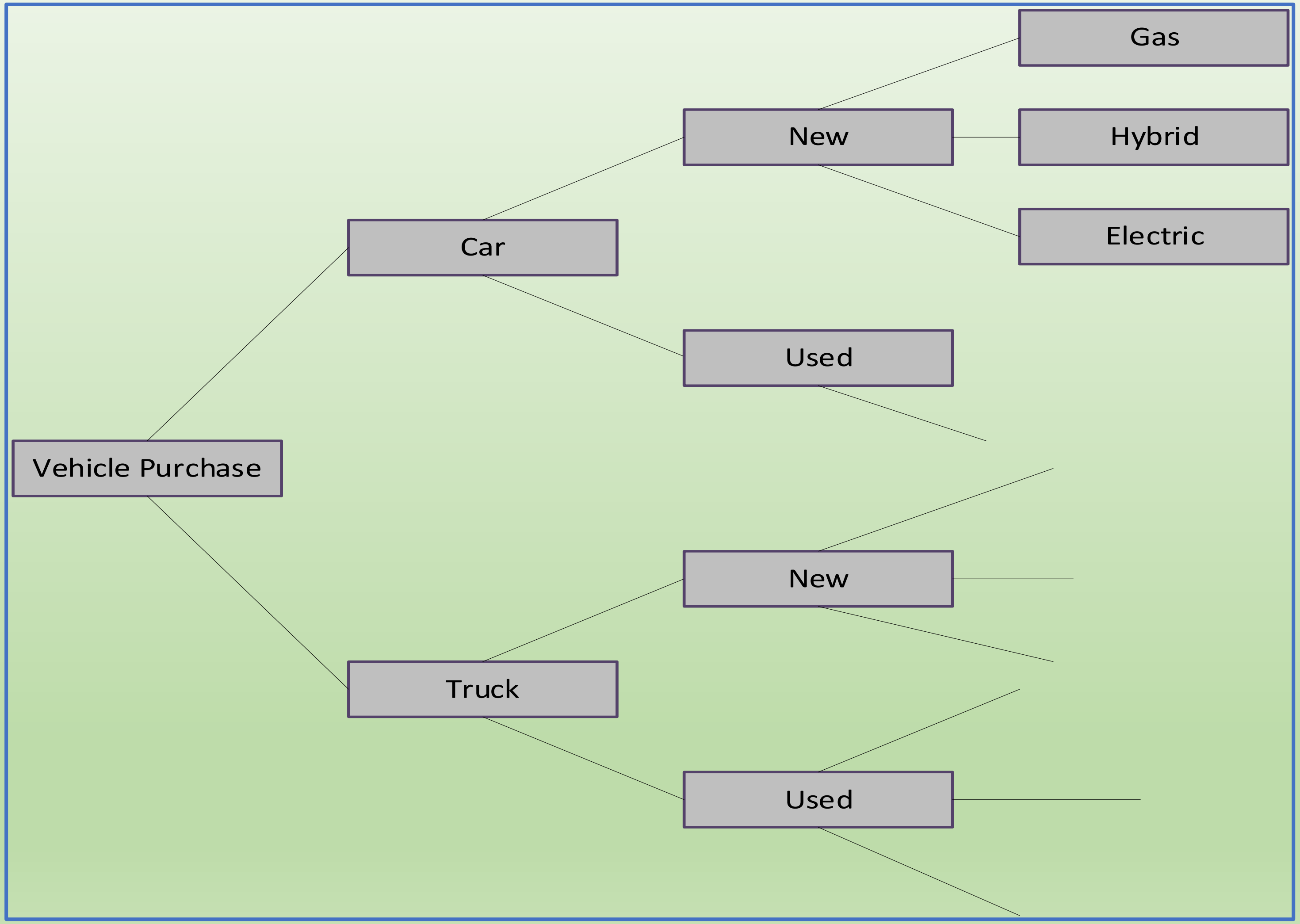


The curve to flattened

©Gregory Mason 2021

Vaccine Model

- Many policy simulations involve binary decisions, disease and vaccination being especially apt.
  - I get sick, or not
  - I get vaccinated, or not
  - I recover, or not.
- The basic mechanism is a decision tree.
- Here is the decision tree for purchasing a vehicle



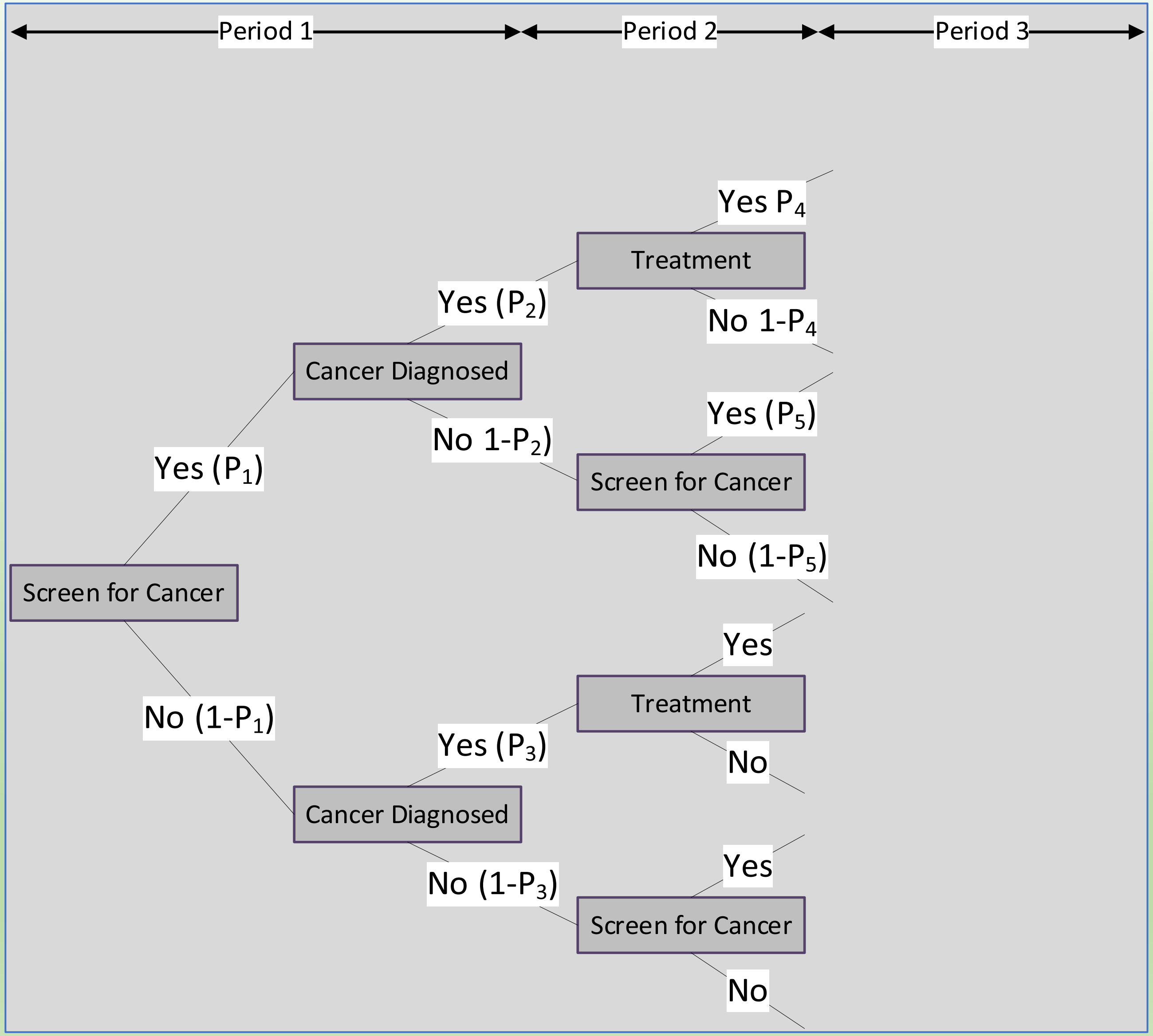
Cancer Decision Tree

This decision tree represents the process of cancer screening.

The is a three period model, and the convention is to set a period as one year.

The decision at each node is binary, with a probability of P and 1-P. (As a reminder a fair coin has P = .5 for heads and 1-P=.5 for tails.) This means that health simulations often uses the =RAND and =RANDBETWEEN functions.

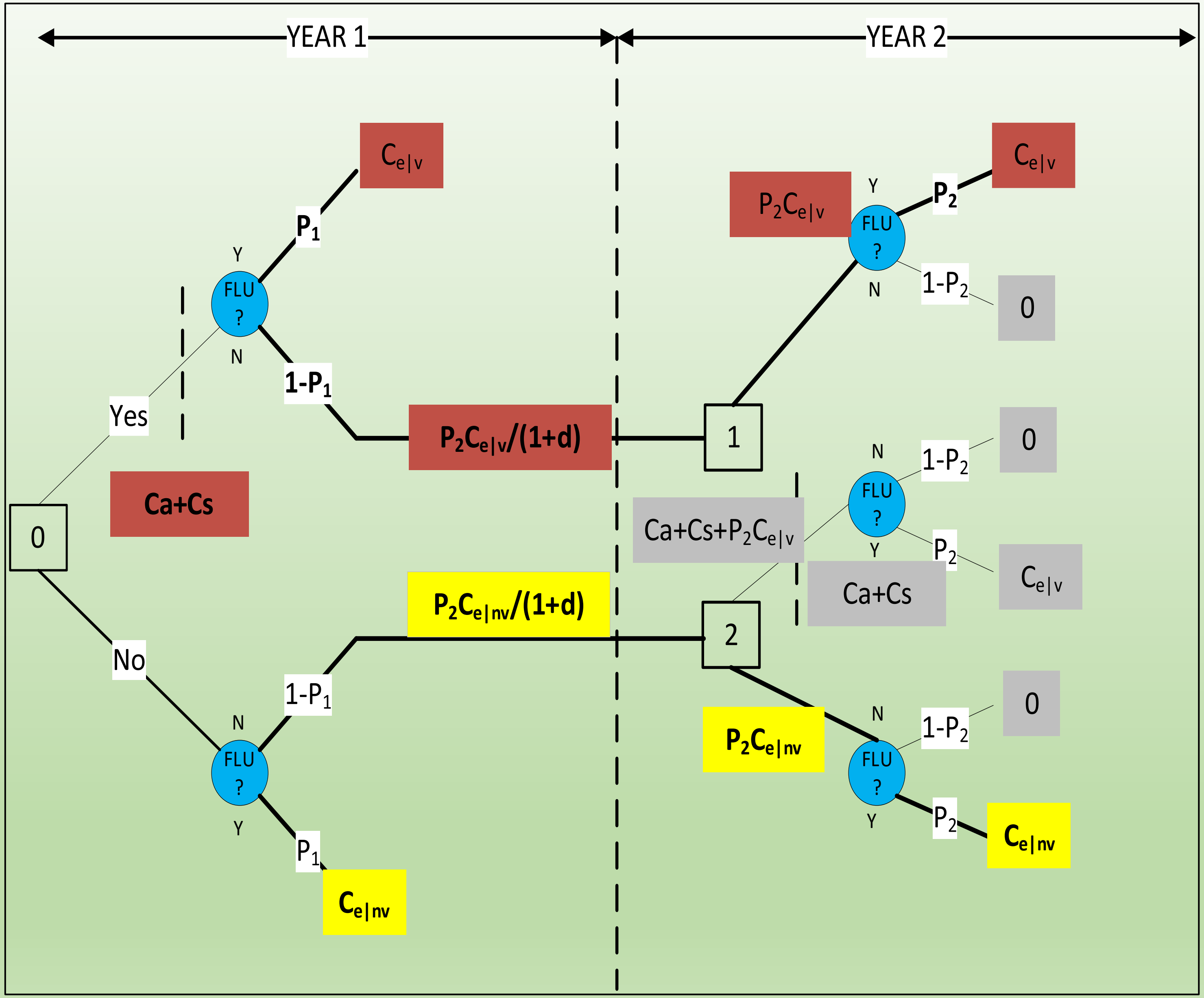
In this example, each period uses the same probabilities for each period.



Flu Vaccine Decision Tree

1. This is a model of the net benefit of government provision of a flu vaccine program.
2. It is not the decision tree for an individual user.
3. The vaccine offers immunity for two years.
4. The model examines the cost of a flu season with and without the vaccine program.
5. The flu occurs in Year 1 and Year 2 with the same probability (P1 and P2) respectively; there the probability of the flu not occurring is 1-P1 and 1-P2 respectively.

See Vaccine Model Explained.docx



Parameter Table for Vaccine Model

Model Parameters				
Parameter		Value	Range	Explanation
Population	N	380000		Total population
Fraction high risk	r	0.06	.04 - .08	1/2 those over 65
Low-risk vaccination rate	vl	0.05	.03 - .07	Proportion of low-risk persons vaccinated
High-risk vaccination rate	vh	0.6	.4 - .8	Proportion of high-risk persons vaccinated
Adverse reaction rate	a	0.03	.01 - .05	Fraction of those vaccinated who have an adverse reaction
Low-risk mortality rate	ml	0.00005	.000025 - .000075	Mortality of low-risk persons who die after contracting flu
High-risk mortality rate	mh	0.001	.0005 - .002	Mortality of high-risk persons who die after contracting flu
Herd immunity effect	H	1	.5 - 1.0	Fraction of those vaccinated who contribute to herd immunity
Vaccine effectiveness rate	e	0.75	.65 - .85	
Hours lost	t	24	18 - 30	Average hours lost due to flu
Infection rate	i	0.25	.20 - .30	Proportion of unvaccinated who contract flu
Year 1 Epidemic Prob	P1	0.4		
Year 2 Epidemic Prob	P2	0.2		
Vaccine dose cost per dose	q	$9		
Overhead cost	o	120000		
Opportunity cost of Time (w) hourly	w	12		Average wage
Value of Life	L	3000000		
Discount Rate	d	0.05		
Number of High-Risk Vaccinations	Vhi	13680	13680	= Vh*r*N
Number of Low-Risk Vaccinations	Vlo	17860	17860	Number of low-risk persons vaccinated = vl*(1-r)*N
Faction Vaccinated	v	0.083	0.083	Fraction of the total population vaccinated = r*vh+vl*(1-r)

Model Equations

Eqn #	Variable	Equation (see Model Parameters)
1	Ca	o+(Vhi+Vlo)*q
2	Cs	a*(Vhi+Vlo)*(w*t+mh*L)
3	Ce|nv	i*(r*N*(w*t+mh*L)+(1-r)*N*(wt+ml*L))
4	Ce|v	(i-H*v*e)((r*N-e*Vh)(w*t+mh*L)+((1-r)*N-e*Vl)*(w*t+ml*L))
5	ECv	Ca+Cs+P1*Ce|v+(1-P1)*P2*Ce|v/(1+d)
6	ECnv	P1*Ce|nv+(1-P1)*P2*Ce|nv/(1+d)
7	E(NB)	ECe|v-ECe|nv

Explanation of equations

1. The administrative cost of the program is the overhead costs plus the total cost of doses (number vaccinated times the cost per dose).
2. This the cost of adverse reaction which is the adverse reaction rate times the total numbers vaccinated (Vh+Vl) times the sum of opportunity cost of lost time plus the value of lost lives (w*t+mh*L)
3. The costs of the flu with no vaccine program is infection rate (i) times the number of high risk individuals (r*N) times the sum of the opportunity cost of hours lost plus the value of lost lives(w*t+mh*L), plus the numbers not infected (1-r)*N times the sum of the opportunity cost of hours lost plus the value of lost lives(w*t+mh*L).
4. The costs of the flu with a vaccine program is the net infection rate or (i) less the herd immunity times the fraction vaccinated times the vaccine effectiveness times 1) the sum of the net numbers at high risk (r*N - e*Vh) times the sum of the opportunity cost of hours lost plus the value of lost lives(w*t+mh*L), plus 2) the numbers not at high risk (N-e*Vl) times the opportunity cost of hours lost plus the value of lost lives(w*t+mh*L.)
5. The expected cost of flu with the vaccine program is the sum of admin costs and cost of adverse reactions, plus the expected cost of the vaccine in Year 1 (P1*Ce|v) plus the discounted expected cost if the program in Year 2 (1-P1)*P2*Ce|v/(1+d).
6. The expected cost if the flu without the program is the expected cost of the flu in Year 1 (P1*Ce|nv) plus the discounted expected cost of the flu without the program in Year 2 (1-P1)*p2*Ce|nv/(1+d).
7. The expected net benefits is the difference between 5 and 6.