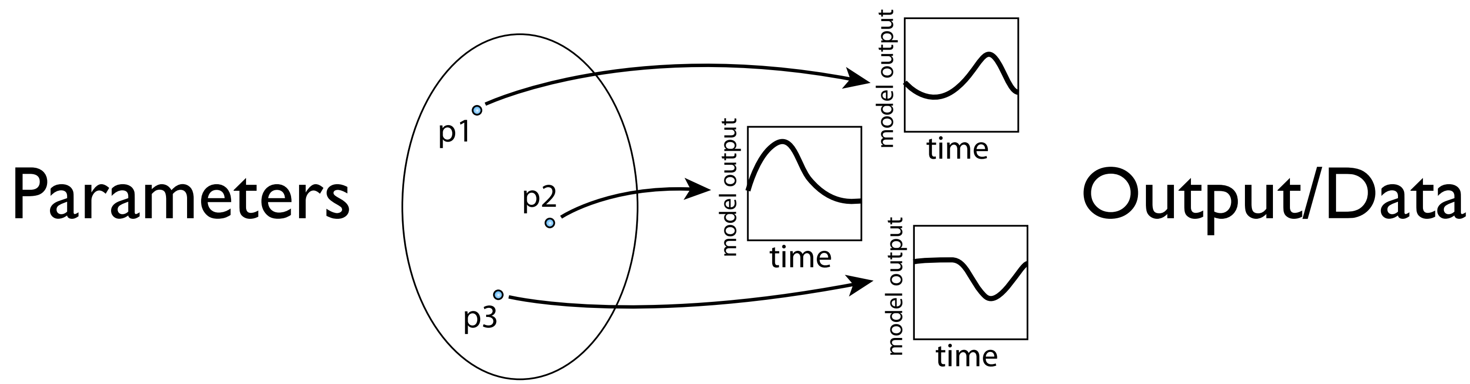
Introduction to Structural & Practical Identifiability

Marisa Eisenberg
University of Michigan, Ann Arbor

Identifiability

- Identifiability—Is it possible to uniquely determine the parameters from the data?

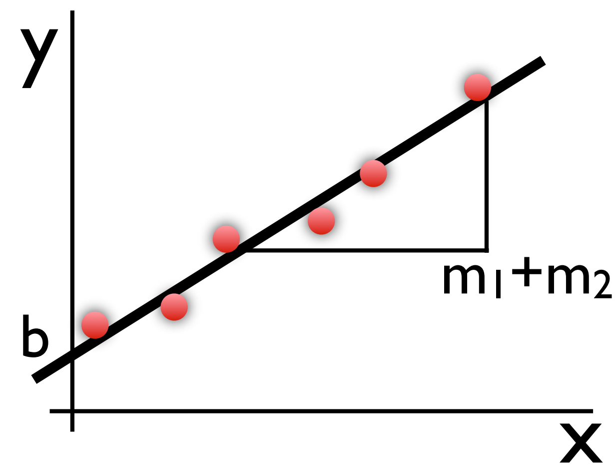


- Important problem in parameter estimation
- Many different approaches - statistics, applied math, engineering/systems theory

Ollivier 1990, Ljung & Glad 1994, Evans & Chappell 2000, Audoly et al 2003, Hengl et al. 2007, Chis et al 2011

Identifiability

- Practical vs. Structural
  - Broad, sometimes overlapping categories
  - Noisy vs. perfect data
- Example: $y = (m_1 + m_2)x + b$
- Unidentifiability - can cause serious problems when estimating parameters
- Identifiable combinations

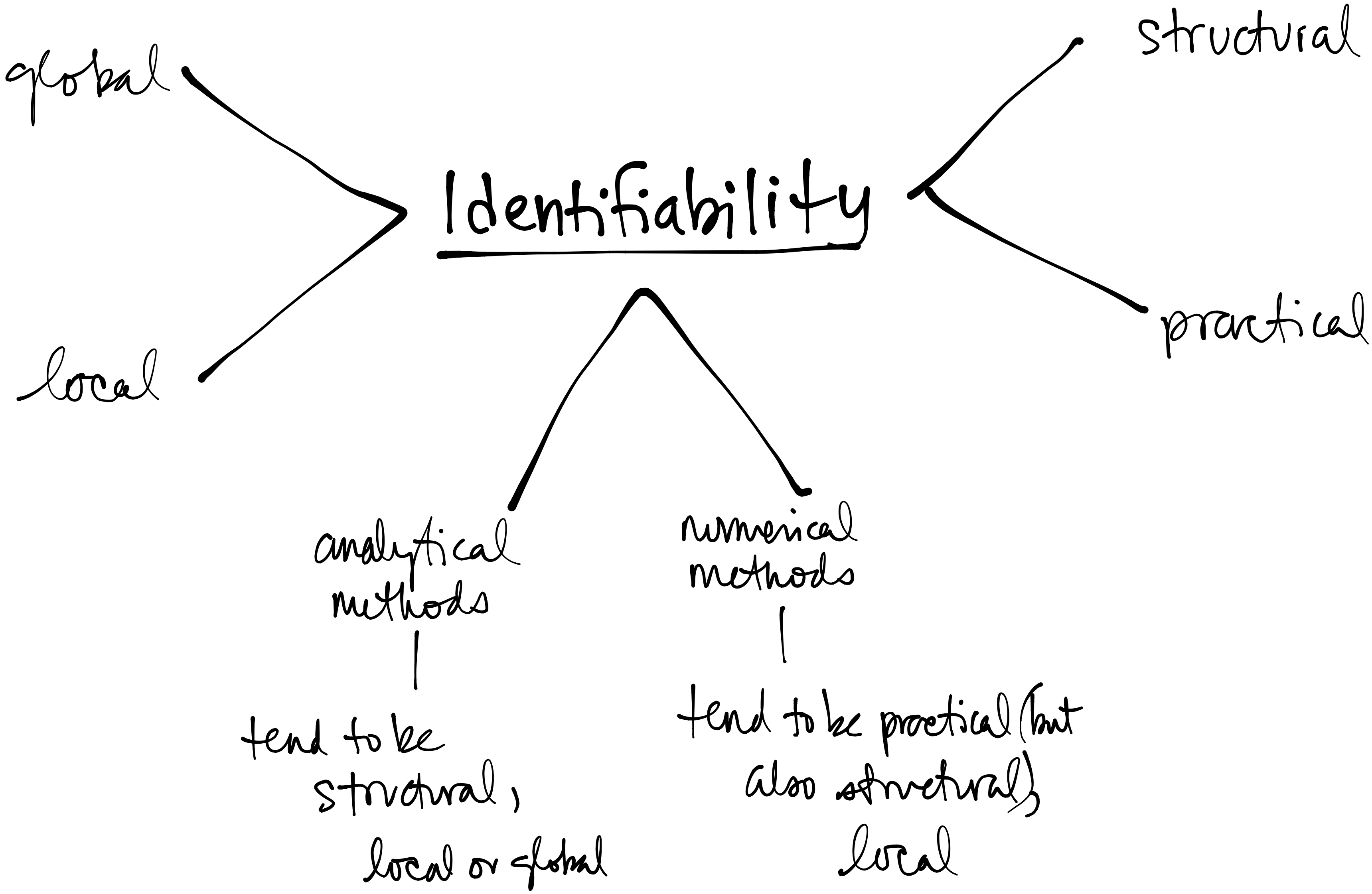


Structural Identifiability

- Assumes best case scenario - data is known perfectly at all times
- Unrealistic!
- But, necessary condition for practical identifiability with real, noisy data

Structural Identifiability

- Reveals identifiable combinations and how to restructure the model so that it is identifiable
- Can give a priori information, help direct experiment design
- Global vs. local methods


Identifiability
global
local
structural
practical
analytical methods
tend to be structural, local or global
numerical methods
tend to be practical (but also structural), local

Key Concepts

- Identifiability vs. unidentifiability
  - Practical vs. structural
  - Can be in between, e.g. quasi-identifiable
  - Locally identifiable
- Identifiable Combinations
- Reparameterization

Reparameterization

- Identifiable combinations - parameter combinations that can be estimated
- Once you know those, why reparameterize?
- Estimation issues - reparameterization provides a model that is input-output equivalent to the original but identifiable
- Often the reparameterized model has 'sensible' biological meaning (e.g. nondimensionalized, etc.)

note about scaling

Methods we'll talk about today

- Differential Algebra Approach - structural identifiability, global, analytical method
- Fisher information matrix - structural or practical, local, analytical or numerical method
- Likelihood Profiling - structural or practical, local, numerical method

Simple Methods

- Simulated data approach
- If you have a small system, you can even plot the likelihood surface (typically can't though—more on this with profile likelihoods)

Analytical Methods for Structural Identifiability

Methods for Structural Identifiability

- **Laplace transform** - linear models only
- **Taylor series approach** - more broad application, but only local info & may not terminate
- **Similarity transform approach** - difficult to make algorithmic, can be difficult to assess conditions for applying theorem
- **Differential algebra approach** - rational function ODE models, global info

Bellman 1970, Cobelli & DiStefano 1980, Evans & Chappell 2000, Ollivier 1990, Ljung & Glad 1994, Audoly et al 2003

Methods for Structural Identifiability

- **Laplace transform** - linear models only
- **Taylor series approach** - more broad application, but only local info & may not terminate
- **Similarity transform approach** - difficult to make algorithmic, can be difficult to assess conditions for applying theorem
- **Differential algebra approach** - rational function ODE models, global info

Bellman 1970, Cobelli & DiStefano 1980, Evans & Chappell 2000, Ollivier 1990, Ljung & Glad 1994, Audoly 2003

Structural Identifiability Analysis

- Basic idea: use substitution & differentiation to eliminate all variables except for the observed output (y)
- Clear (divide by) the coefficient for highest derivative term(s)
- This is called the **input-output equation(s)**
- Contains all structural identifiability info for the model

Structural Identifiability Analysis

- Use the coefficients to solve for identifiability of the model
- If unidentifiable, determine identifiable combinations
- Find identifiable reparameterization of the model?
- Easier to see with an example—

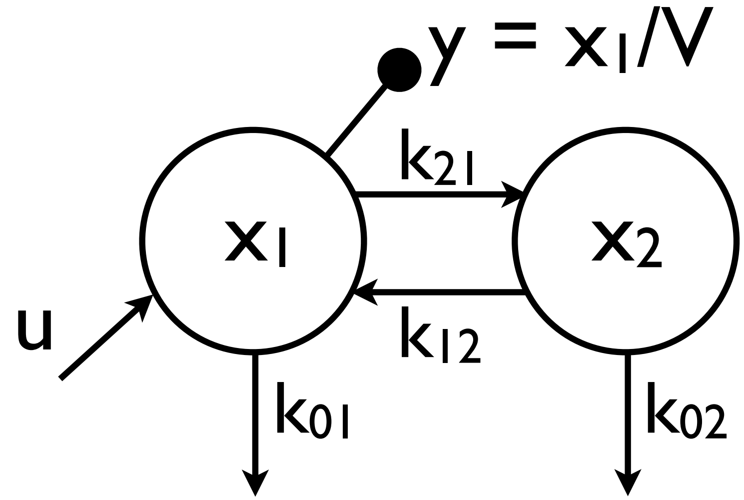
2-Compartment Example

- Linear 2-Comp Model

$$\dot{x}_1 = u + k_{12}x_2 - \left(k_{01} + k_{21}\right)x_1$$

$$\dot{x}_2 = k_{21}x_1 - \left(k_{02} + k_{12}\right)x_2$$

$$y = x_1 / V$$

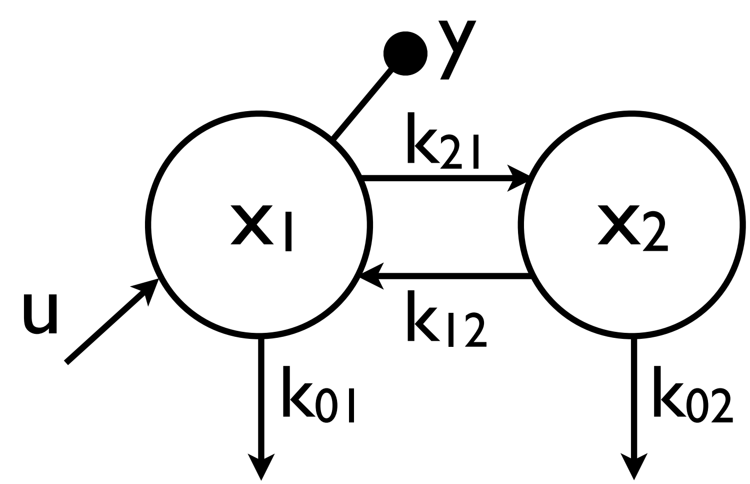


- state variables (x)
- measurements (y)
- known input (u) (e.g. IV injection)

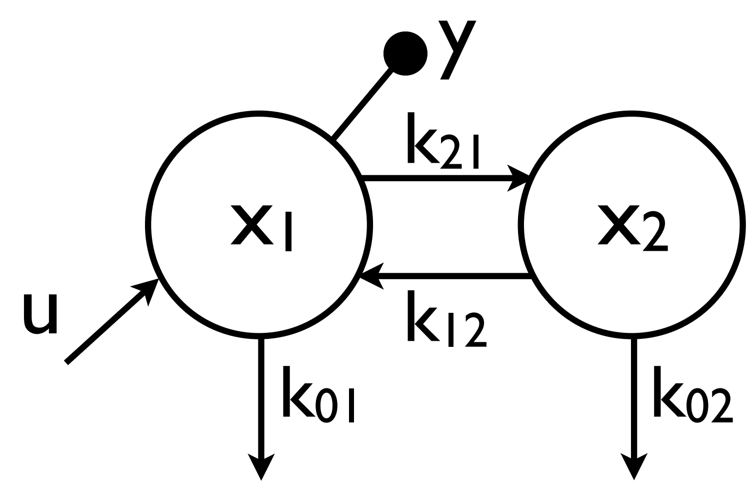
2-Compartment Example

$$\dot{y}V = u + k_{12}x_2 - \left(k_{01} + k_{21}\right)yV$$

$$\dot{x}_2 = k_{21}x_1 - \left(k_{02} + k_{12}\right)x_2$$

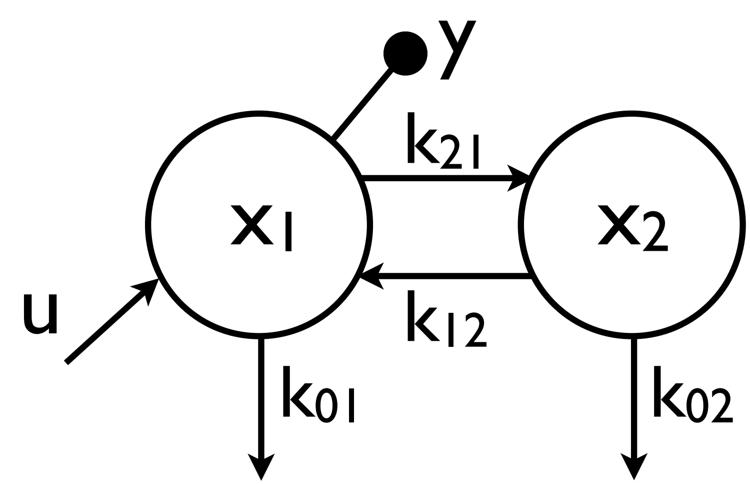


2-Compartment Example



$$\ddot{y} + \left(k_{01} + k_{21} + k_{12} + k_{02}\right)\dot{y} -$$
$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right)y - u\left(k_{12} + k_{02}\right)/V - \dot{u}/V = 0$$

2-Compartment Example



$$1/V = a_1$$

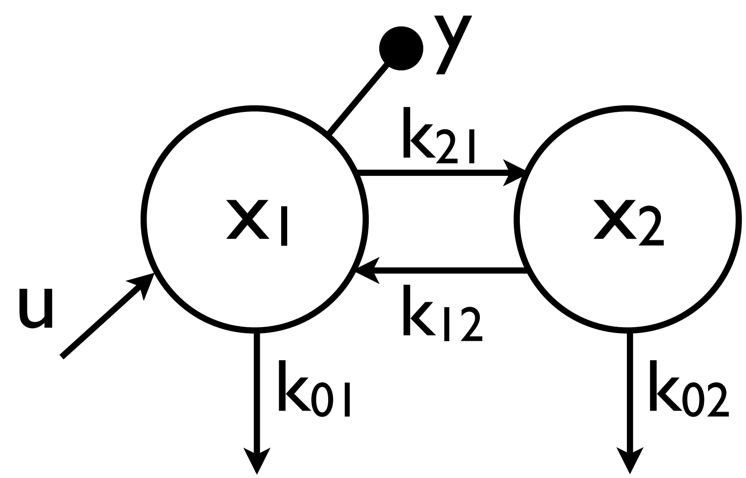
$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$

2-Compartment Example

$$1/V = a_1 \Rightarrow V = 1/a_1$$



$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$

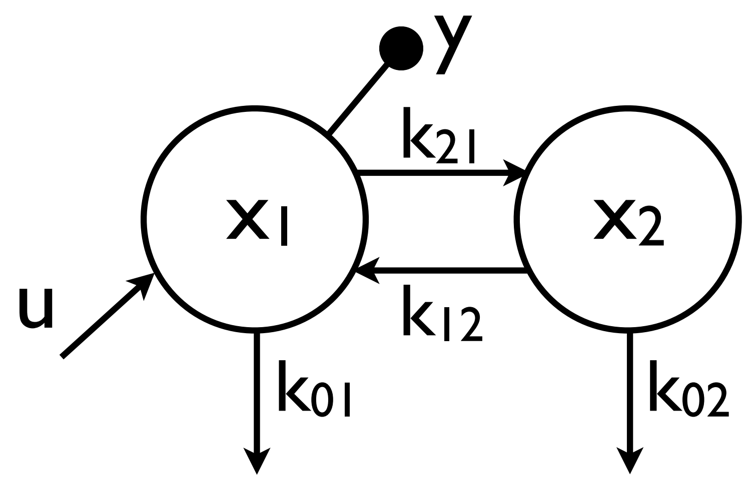
2-Compartment Example

$$1/V = a_1 \Rightarrow V = 1/a_1$$

$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$



Unidentifiable

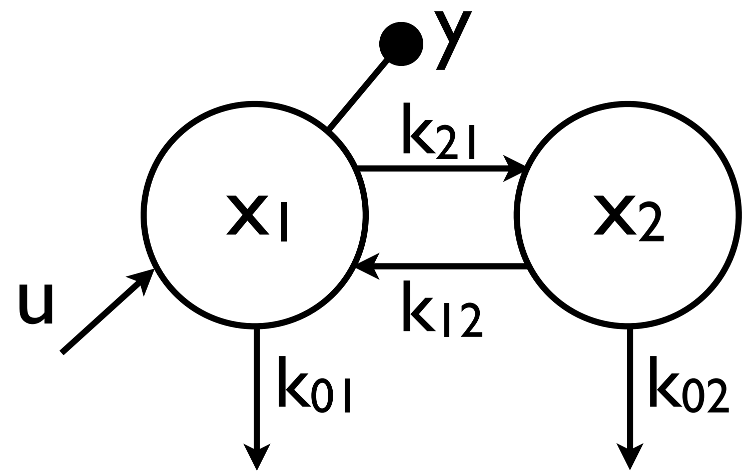
2-Compartment Example

$$1/V = a_1 \Rightarrow V = 1/a_1$$

$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$



Unidentifiable

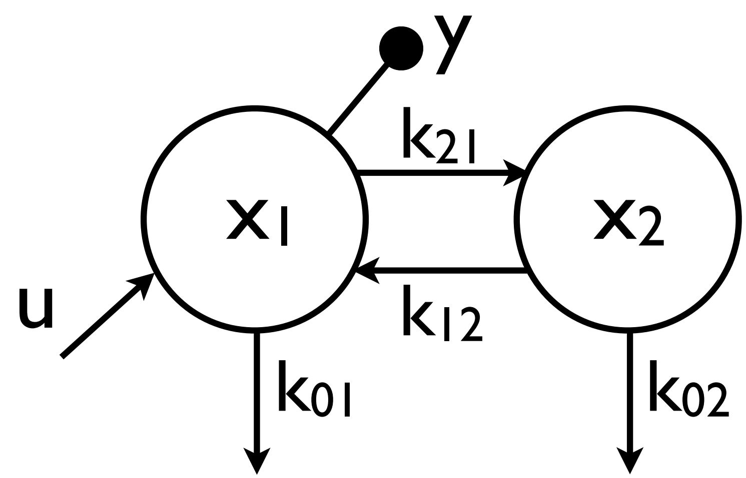
2-Compartment Example

$$1/V = a_1 \Rightarrow V = 1/a_1$$

$$\left(k_{12} + k_{02}\right)/V = a_2$$

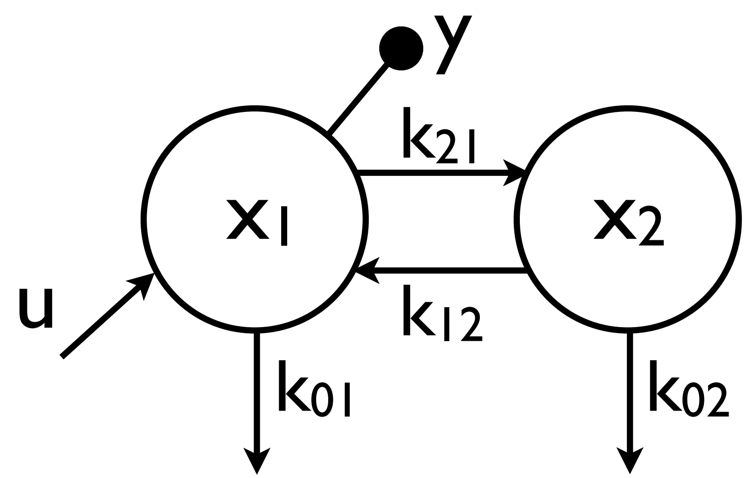
$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$



Unidentifiable

2-Compartment Example



Unidentifiable

$$1/V = a_1 \Rightarrow V = 1/a_1$$

$$\left(k_{12} + k_{02}\right)/V = a_2$$

$$\left(k_{01} + k_{21} + k_{12} + k_{02}\right) = a_3$$

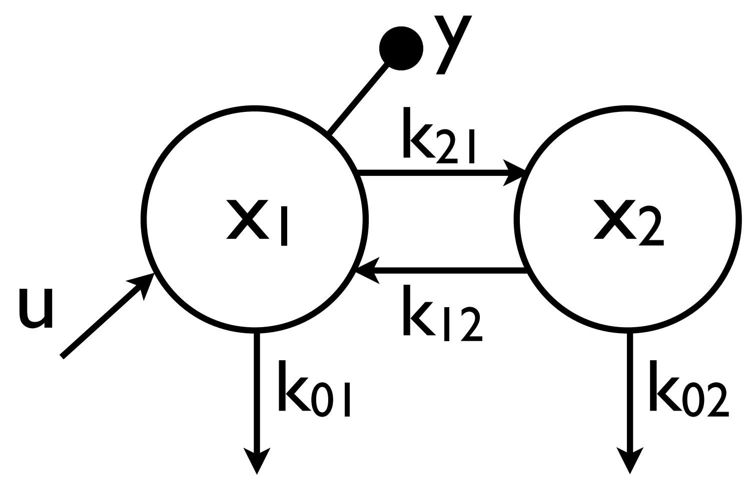
$$\left(k_{12}k_{21} - \left(k_{02} + k_{12}\right)\left(k_{01} + k_{21}\right)\right) = a_4$$

2-Compartment Example

$$\dot{x}_1 = u + \underline{k_{12}}x_2 - \left(\underline{k_{01} + k_{21}}\right)x_1$$

$$\dot{x}_2 = \underline{k_{21}}x_1 - \left(\underline{k_{02} + k_{12}}\right)x_2$$

$$y = x_1 / \underline{V}$$



Let $\underline{x}_2 = k_{12}x_2$

$$\dot{x}_1 = u + \underline{x}_2 - \left(\underline{k_{01} + k_{21}}\right)x_1$$

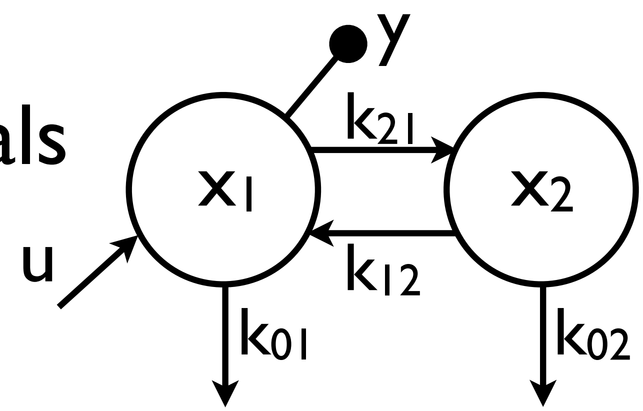
$$\dot{\underline{x}}_2 = \underline{k_{12}k_{21}}x_1 - \left(\underline{k_{02} + k_{12}}\right)\underline{x}_2$$

$$y = x_1 / \underline{V}$$

Or add information about one of the parameters

Differential Algebra Approach

- View model & measurement equations as differential polynomials



- Reduce the equations using grobner bases, characteristic sets, etc. to eliminate unmeasured variables (x)
- Yields **input-output equation**(s) only in terms of known variables (y, u)
- Use coefficients to test model identifiability

Ollivier 1990, Ljung & Glad 1994, Audoly et al 2003, etc.

Differential Algebra Approach

- From the coefficients, can often determine:
  - Simpler forms for identifiable combinations
  - Identifiable reparameterizations for model
- But not always easy by eye—more on this in the next talk!

Numerical Methods for Identifiability Analysis

Numerical Approaches to Identifiability

- Analytical approaches can be slow, sometimes have limited applicability
- Wide range of numerical approaches
  - Sensitivities/Fisher Information Matrix
  - Profile Likelihood
  - Many others (e.g. Bayesian approaches, etc.)

Numerical Approaches to Identifiability

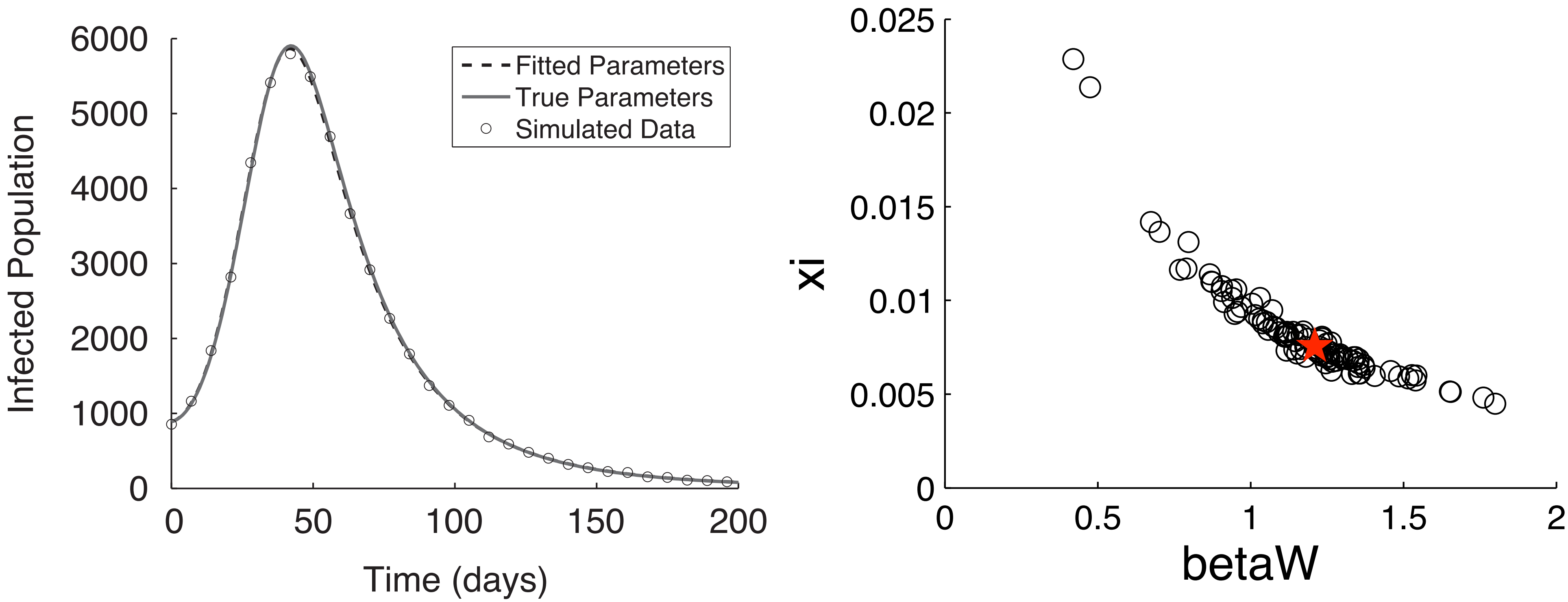
- Most can do both structural & practical identifiability
- Wide range of applicable models, often relatively fast
- Typically only local
- Less attention to the problem of identifiable combinations

Simple Simulation Approach

- Simulate data using a single set of ‘true’ parameter values
  - Without noise for structural identifiability
  - With noise for practical identifiability (in this case generate multiple realizations of the data)

Simple Simulation Approach

- Fit your simulated data from multiple starting points and see where your estimates land
- If they all return to the 'true' parameters, likely identifiable, if they do not, examine the relationships between the parameters
- Note—unidentifiability when estimating with 'perfect', noise-free simulated data is most likely structural



6000
5000
4000
3000
2000
1000
0
Infected Population
0
50
100
150
200
Time (days)
Fitted Parameters
True Parameters
Simulated Data
0.025
0.02
0.015
0.01
0.005
0
xi
0
0.5
1
1.5
2
betaW

Parameter Sensitivities

- Design matrix/output sensitivity matrix
- Closely related to identifiability
- Insensitive parameters
- Dependencies between columns

$$X = \begin{pmatrix} \frac{\partial y(t_1)}{\partial p_1} & \cdots & \frac{\partial y(t_1)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y(t_m)}{\partial p_1} & \cdots & \frac{\partial y(t_m)}{\partial p_n} \end{pmatrix}$$

Fisher Information Matrix

- FIM - N_P x N_P matrix
- Useful in testing practical & structural ID - represents amount of information that the output **y** contains about **p**
- Cramer-Rao Bound: $FIM^{-1} \leq Var(\mathbf{p})$
- Rank(FIM) = number of identifiable parameters/combinations
- Identifiable Combinations

Fisher Information Matrix

- Special case when errors are normally distributed

$$F = X^T W X$$

W = weighting matrix

$$X = \begin{pmatrix} \frac{\partial y(t_1)}{\partial p_1} & \cdots & \frac{\partial y(t_1)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y(t_m)}{\partial p_1} & \cdots & \frac{\partial y(t_m)}{\partial p_n} \end{pmatrix}$$

Design Matrix

Fisher Information Matrix

- For looking at structural ID, often just use

$$F = X^T X$$

$$X = \begin{pmatrix} \frac{\partial y(t_1)}{\partial p_1} & \cdots & \frac{\partial y(t_1)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y(t_m)}{\partial p_1} & \cdots & \frac{\partial y(t_m)}{\partial p_n} \end{pmatrix}$$

Design Matrix

Identifiability & the FIM

- Covariance matrix/confidence interval estimates from Cramer Rao bound
  - e.g. large confidence interval —>probably unID
  - Often can detect structural unID as ‘near-infinite’ (gigantic) variances in Cov ~ FIM^{-1}

Identifiability & the FIM

- **Rank of the FIM** is number of identifiable combinations/parameters - can do a lot by testing sub-FIMs and versions of the FIM
- Identifiable combinations - can often see what parameters are related, but don't know form
  - Interaction of combinations

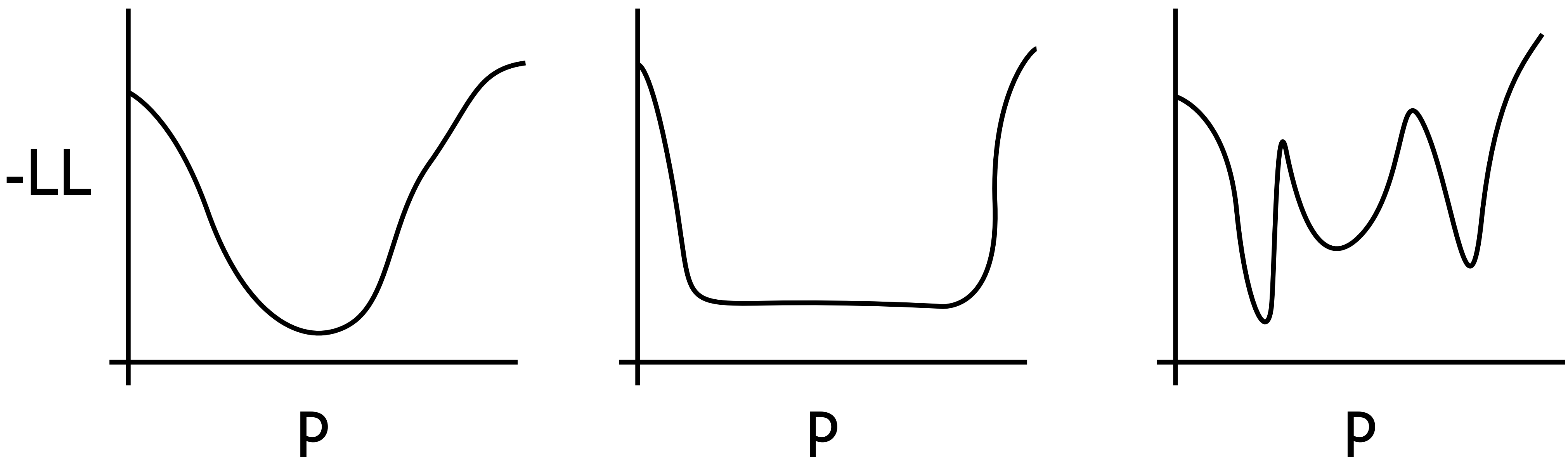
Profile Likelihood

- Basic Idea: 'profile' one parameter at a time, by fixing it to a range of values & fitting the rest of the parameters
- Gives best fit at each point
- Evaluate curvature of likelihood to determine confidence bounds on parameter (and to evaluate parameter uncertainty)

Profile Likelihood

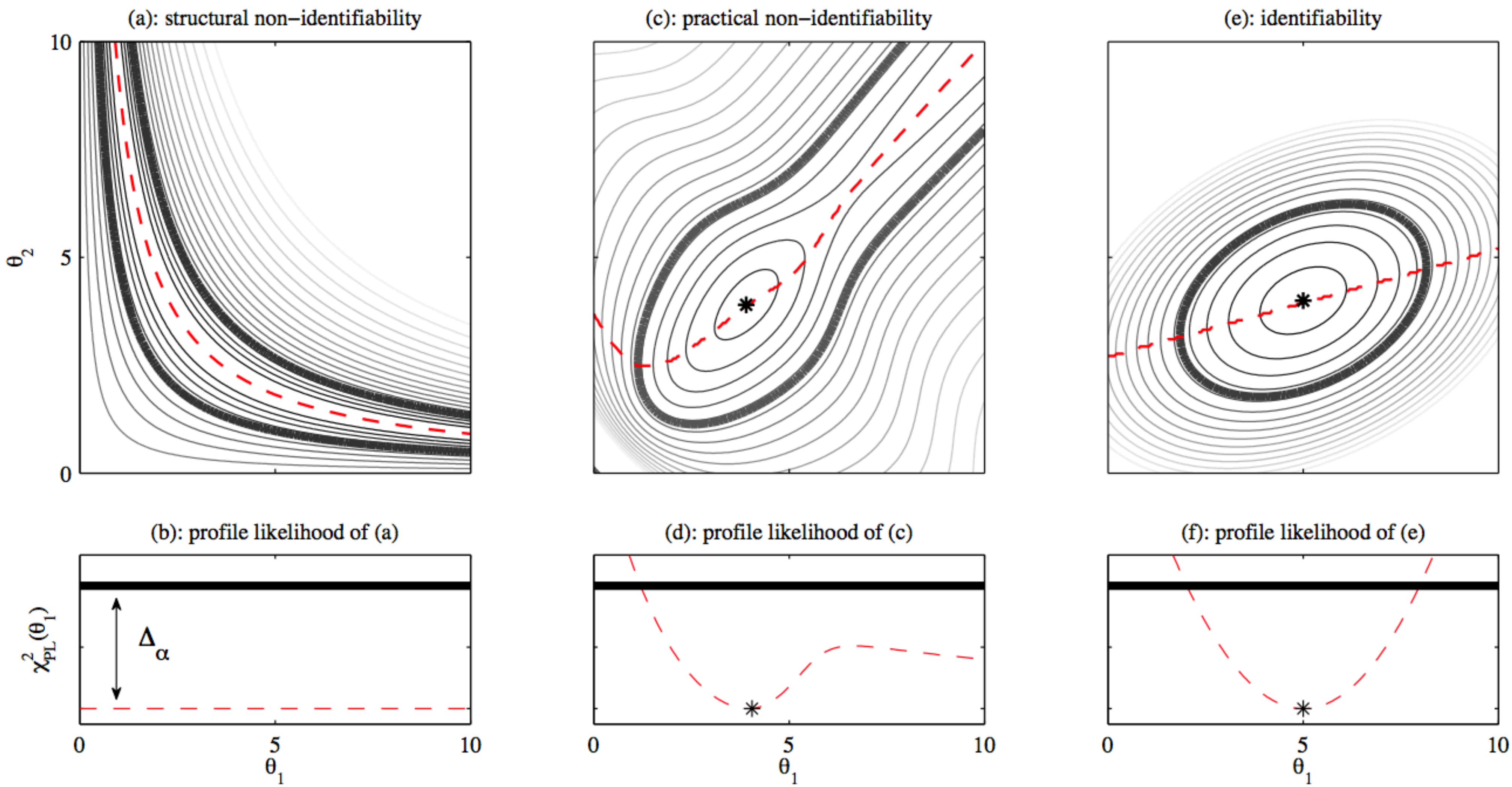
- Choose a range of values for parameter p_i
- For each value, fix p_i to that value, and fit the rest of the parameters
- Report the best likelihood/RSS/cost function value for that p_i value
- Plot the best likelihood values for each value of p_i—this is the profile likelihood

Profile Likelihood



Profile Likelihood & ID

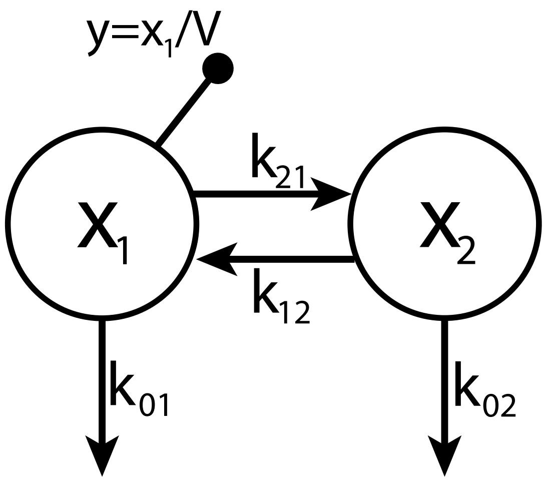
- Can generate confidence bounds based on the curvature of the profile likelihood
- Flat or nearly flat regions indicate identifiability issues
- Can generate simulated ‘perfect’ data to test structural identifiability

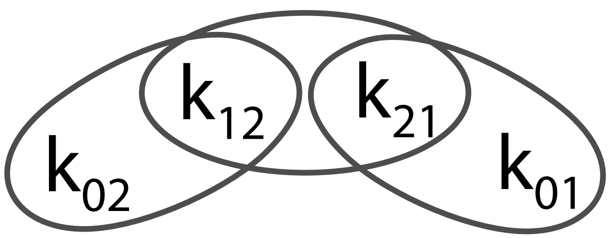


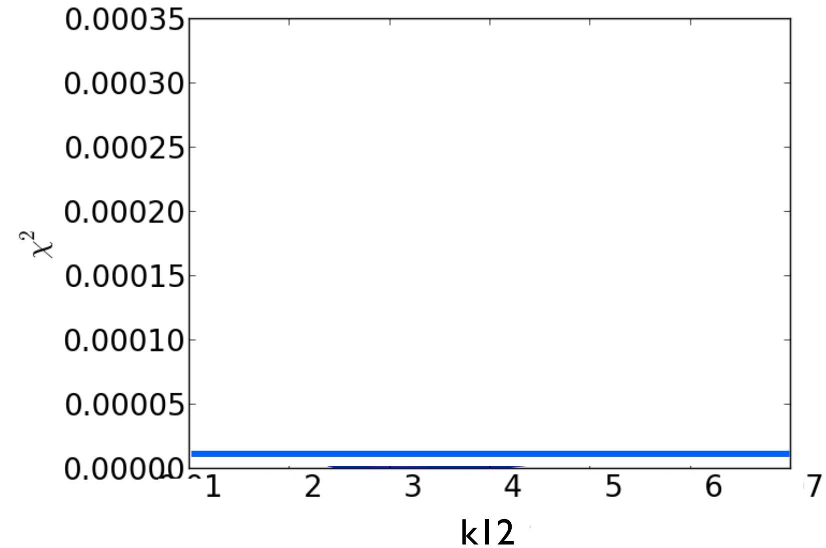
Raue et al. 2010

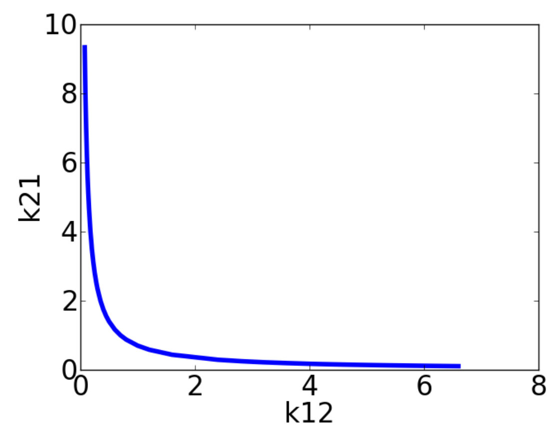
Profile Likelihood

- Can also help reveal the form of identifiable combinations
- Look at relationships between parameters when profiling
- However, can be problematic when too many degrees of freedom


y=x1/V
x1
x2
k21
k12
k01
k02


k02
k12
k21
k01


0.00035
0.00030
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000
χ2
1
2
3
4
5
6
7
k12

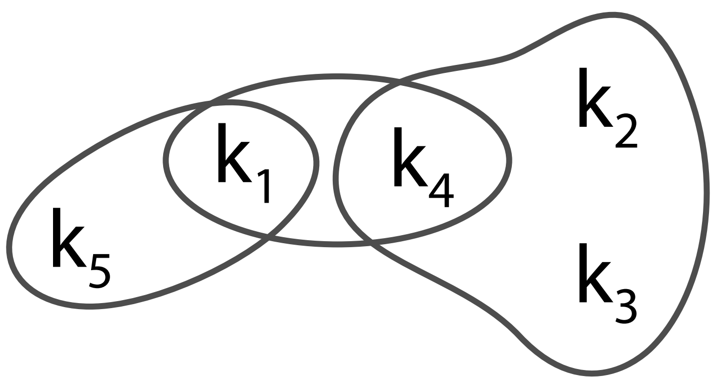

10
8
6
4
2
0
k21
0
2
4
6
8
k12

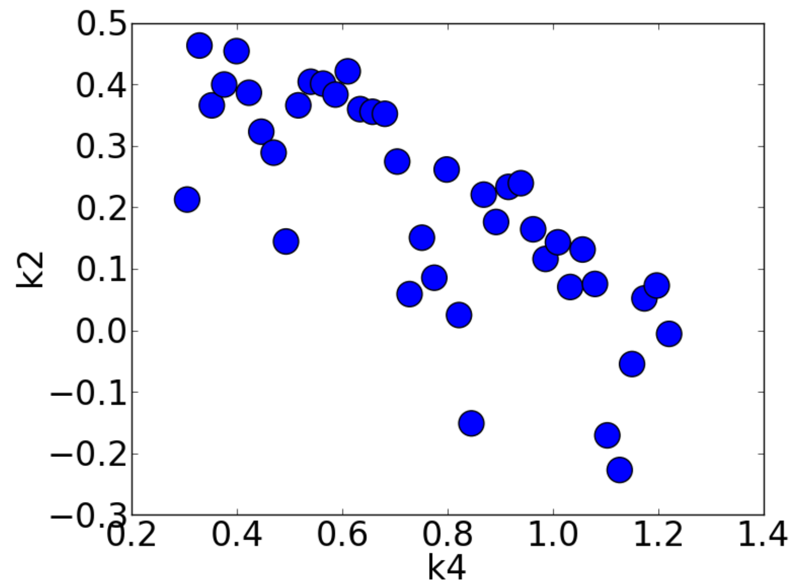
Likelihood Profiling Example

$$\dot{x}_1 = k_1 x_2 - (k_2 + k_3 + k_4) x_1$$

$$\dot{x}_2 = k_4 x_1 - (k_5 + k_1) x_2$$

$$y = x_1 / V$$



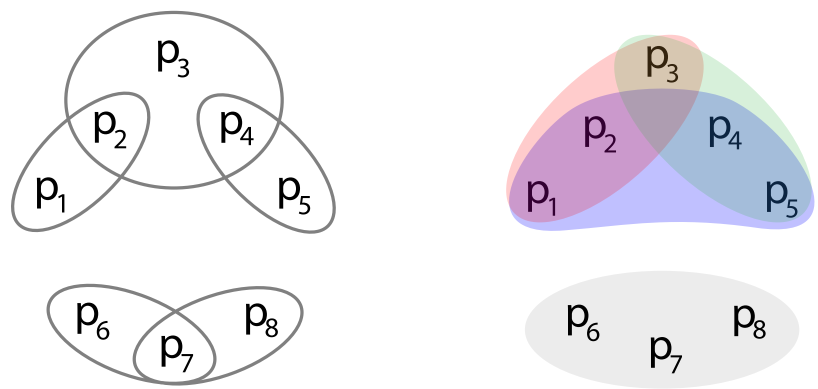


FIM Subset Approach

http://arxiv.org/abs/1307.2298

FIM Subset Approach

- Basic idea - evaluate the rank of the FIM for subsets of parameters to elucidate the structure of the identifiable combinations



- Can then combine this with profile likelihood approach by Raue et al. to determine the form of the combinations

FIM Subset Approach

- Use the FIM rank to select subsets of parameters which are *nearly full rank* (i.e. which become full rank if any single parameter is fixed)
- Use these subsets when likelihood profiling to determine all parameter relationships
- Polynomial interpolation to recover identifiable combinations

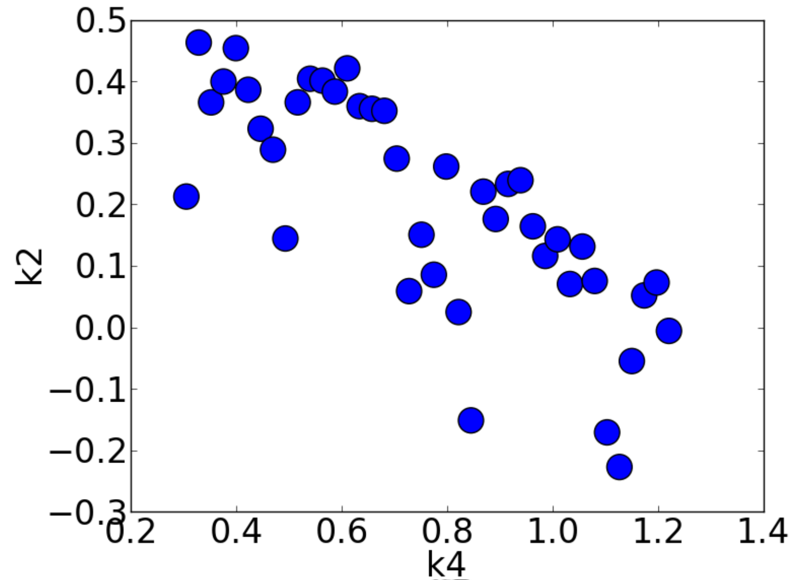
Eisenberg & Hayashi 2014, *in review*

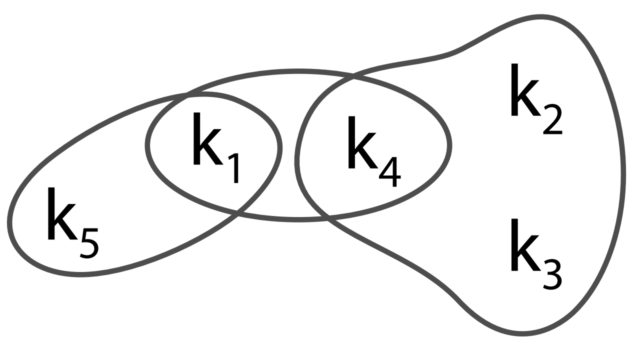
Example Model

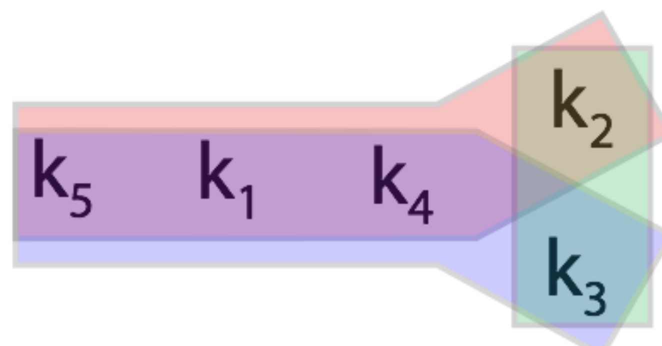
$$\dot{x}_1 = k_1 x_2 - (k_2 + k_3 + k_4) x_1$$

$$\dot{x}_2 = k_4 x_1 - (k_5 + k_1) x_2$$

$$y = x_1 / V$$



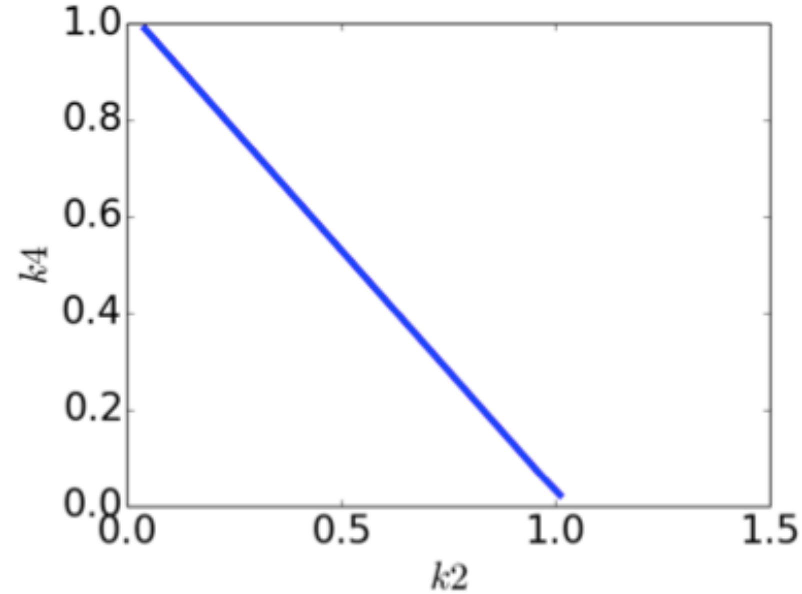


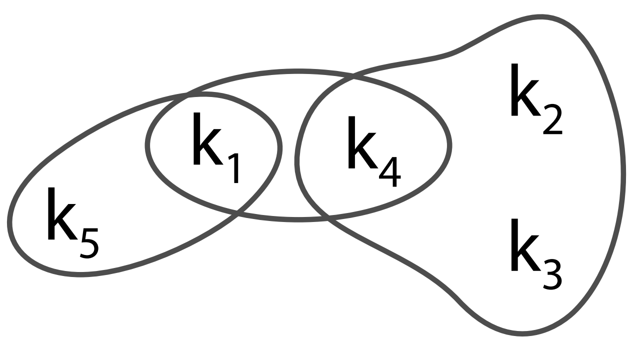


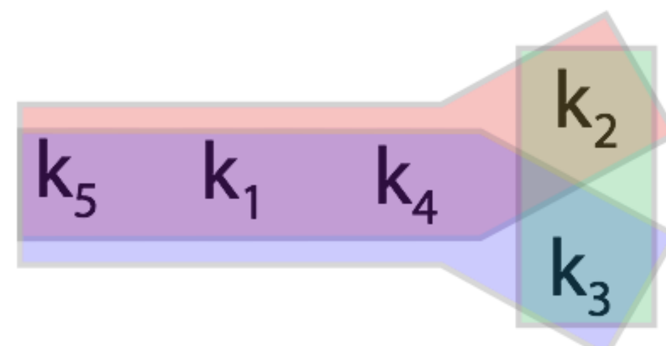
Eisenberg & Hayashi 2014, *in review*

Example Model

$$\dot{x}_1 = k_1 x_2 - (k_2 + k_3 + k_4) x_1$$
$$\dot{x}_2 = k_4 x_1 - (k_5 + k_1) x_2$$
$$y = x_1 / V$$

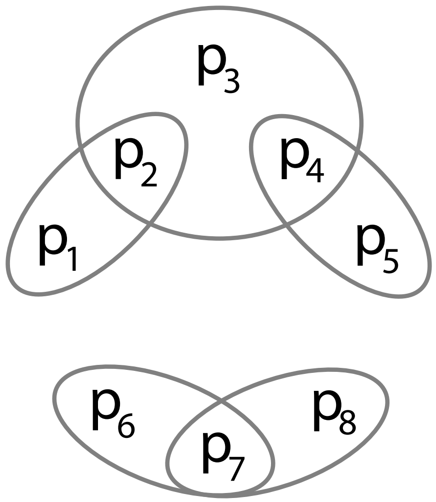


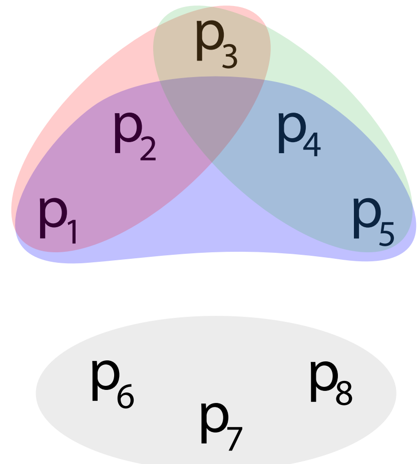


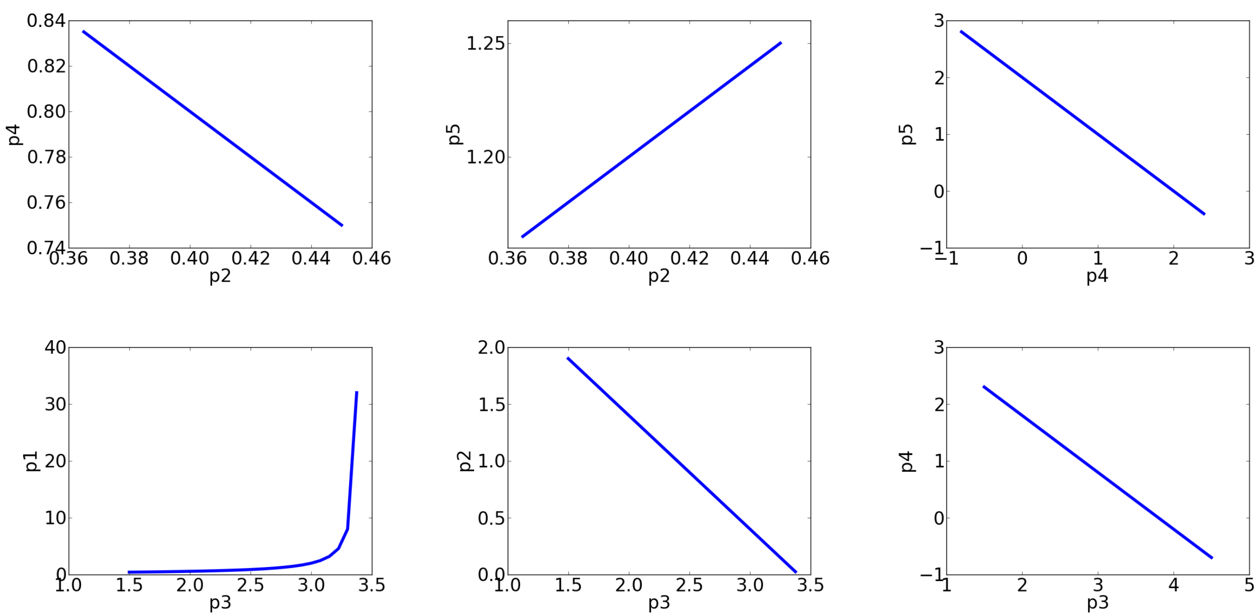


Eisenberg & Hayashi 2014, *in review*

Example Model







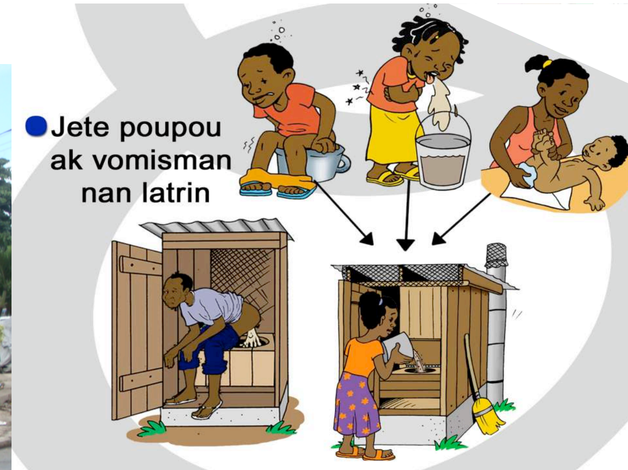
$p_1 p_2$

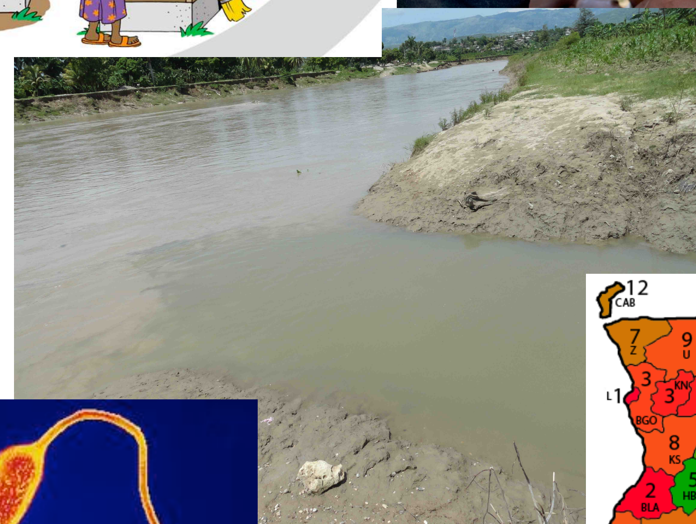
$p_2 + p_3 + p_4$

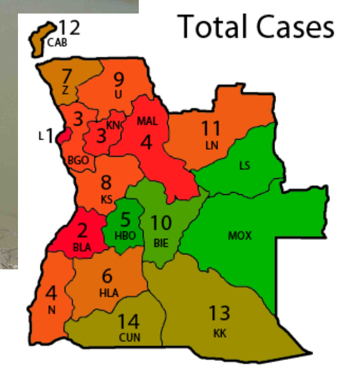
$p_4 + p_5$

$p_6 + p_7$ $p_7 + p_8$

Eisenberg & Hayashi 2014, *in review*

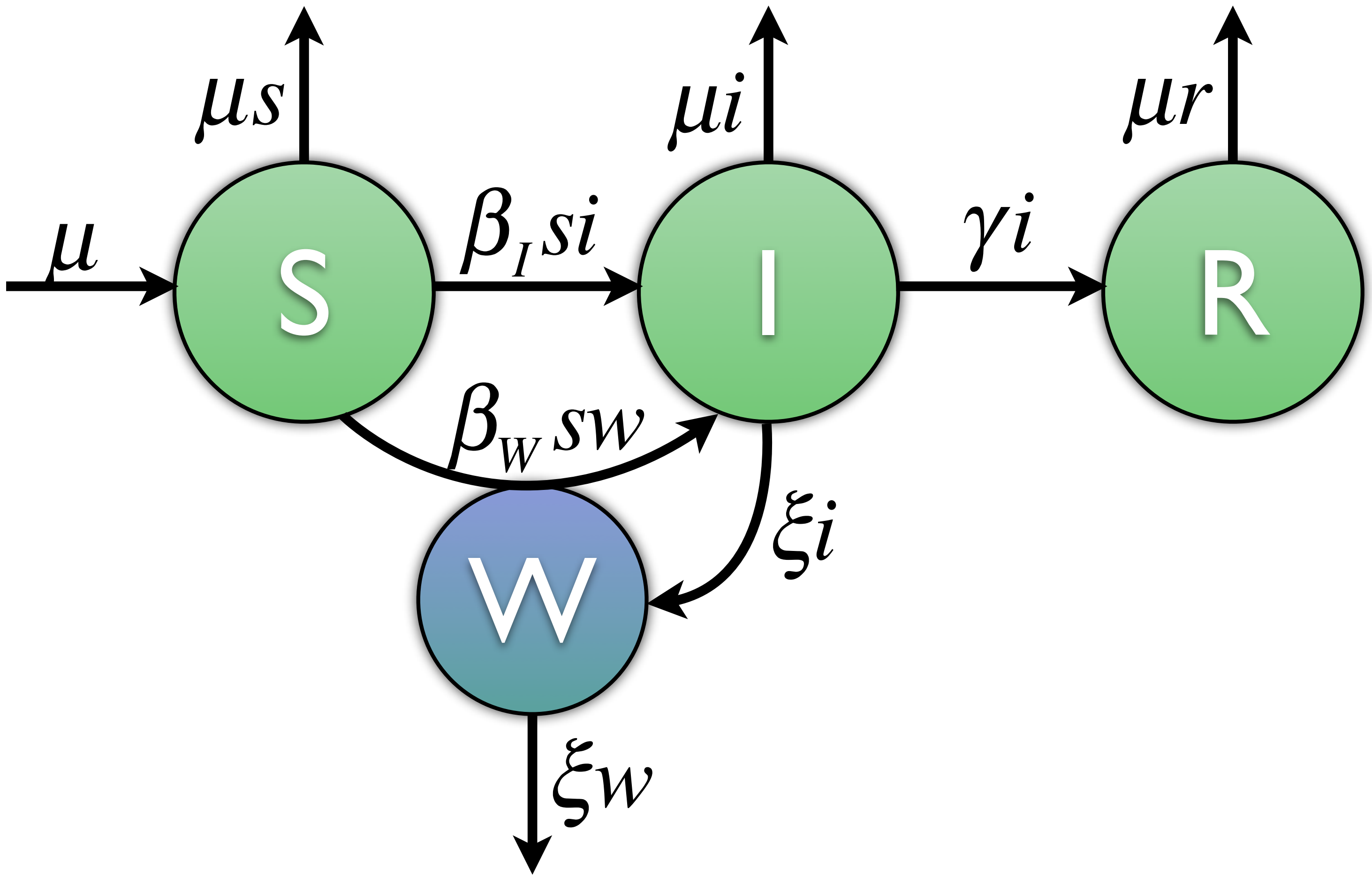






Example: Modeling Cholera

Cholera: SIWR Model



W = pathogen concentration in water reservoir

Tien & Earn, *Bull. Math. Biol.* 2010

SIWR Model Equations

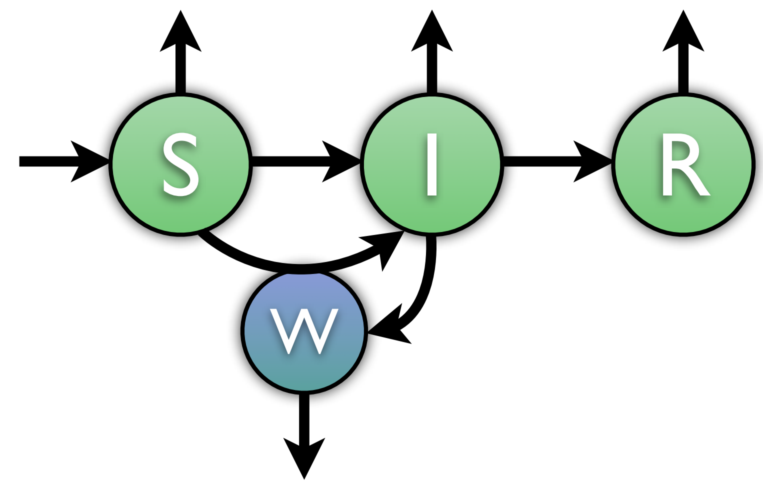
$$\frac{ds}{dt} = \mu - \beta_W ws - \beta_I si - \mu s$$

$$\frac{di}{dt} = \beta_W ws + \beta_I si - \gamma i - \mu i$$

$$\frac{dw}{dt} = \xi(i - w)$$

$$\frac{dr}{dt} = \gamma i - \mu r$$

$$y = ki$$

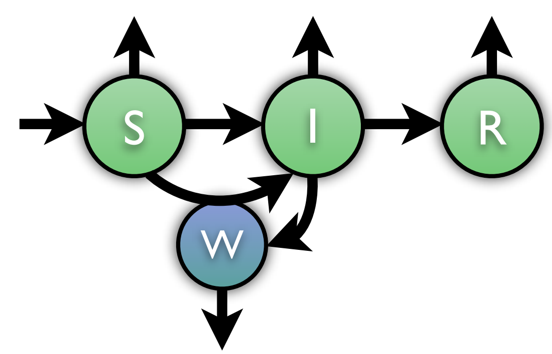


Same process,
but the polynomials aren't as nice...

Tien & Earn, *Bull. Math. Biol.* 2010

SIWR Identifiability Results

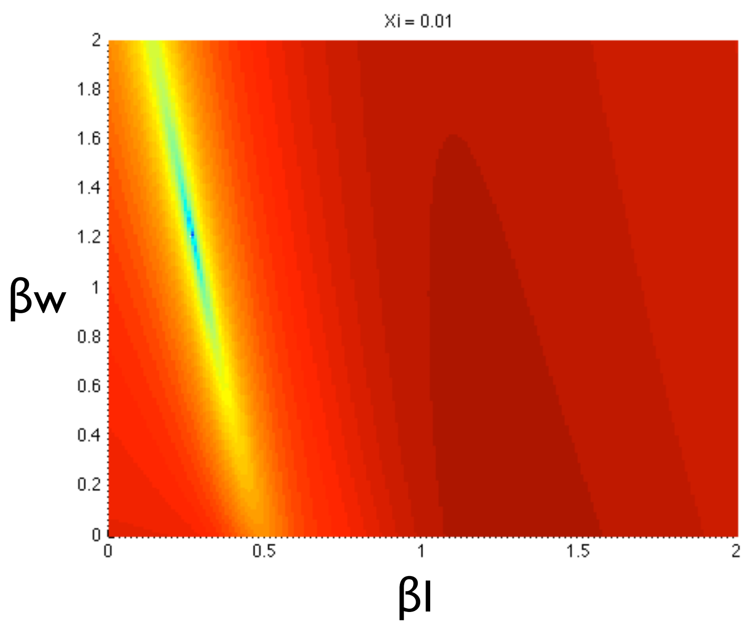
- The scaled SIWR model is uniquely (globally) structurally identifiable for measurements of a portion of the infected population, $y = ki$.
- However, identifiability can be lost in the limit as pathogen lifetime decreases ($\xi \to \infty$)



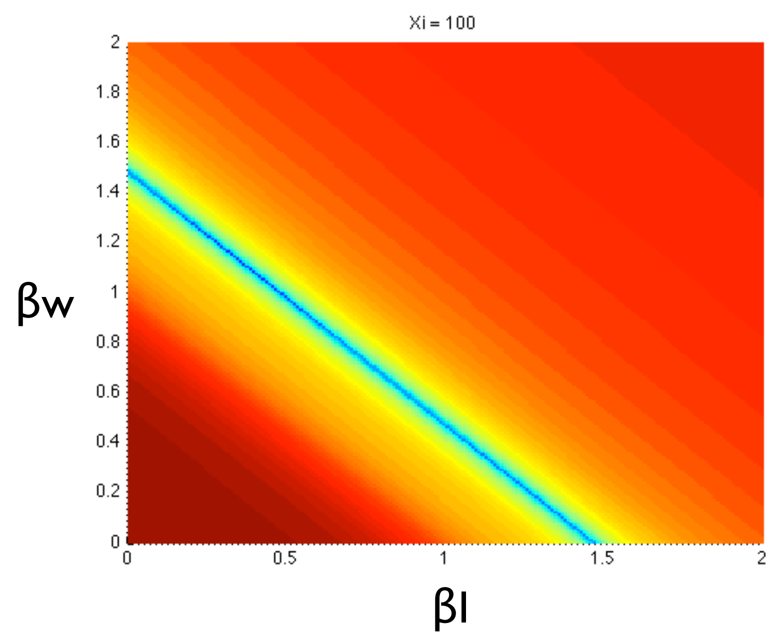
Eisenberg, Robertson, Tien 2013 (*JTB*)

Identifiability as $\xi \to \infty$

$\xi = 0.01$



$\xi = 100$

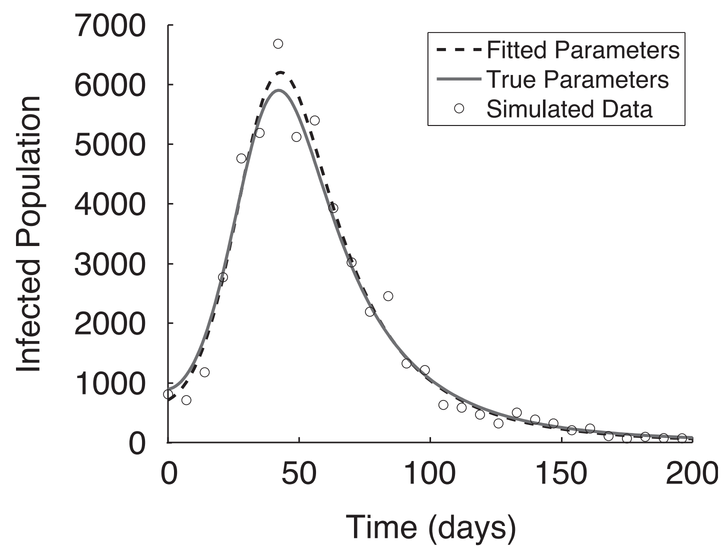


$\beta_W + \beta_I$ forms an identifiable combination

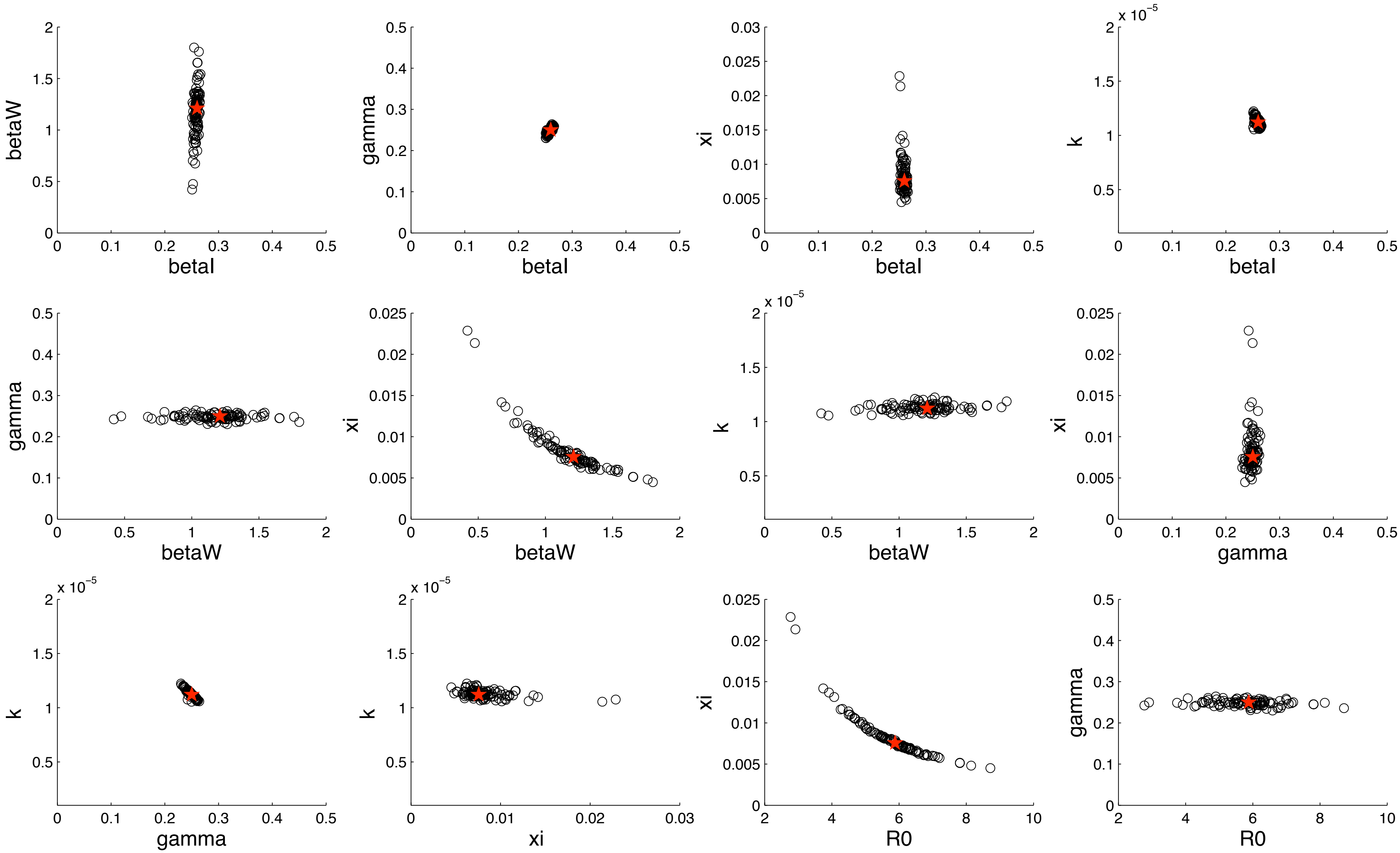
Eisenberg, Robertson, Tien 2013 (JTB)

Practical Identifiability for the SIWR model

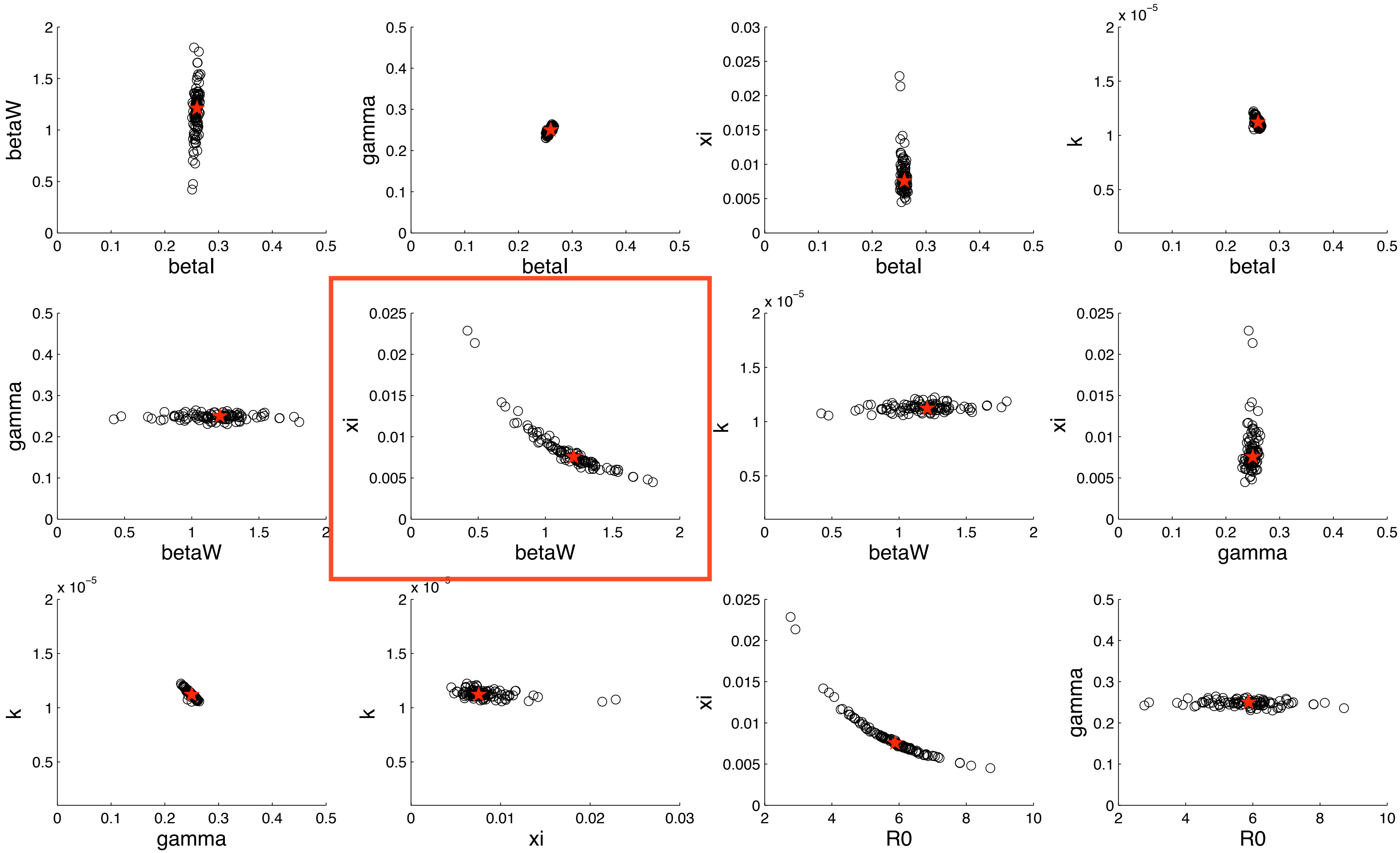
- Simple simulated data approach - simulate noisy data and see if you can estimate parameters
- Simulate data + noise
  - Poisson, negative binomial, normal
- Repeated runs—how well do estimates match true values?



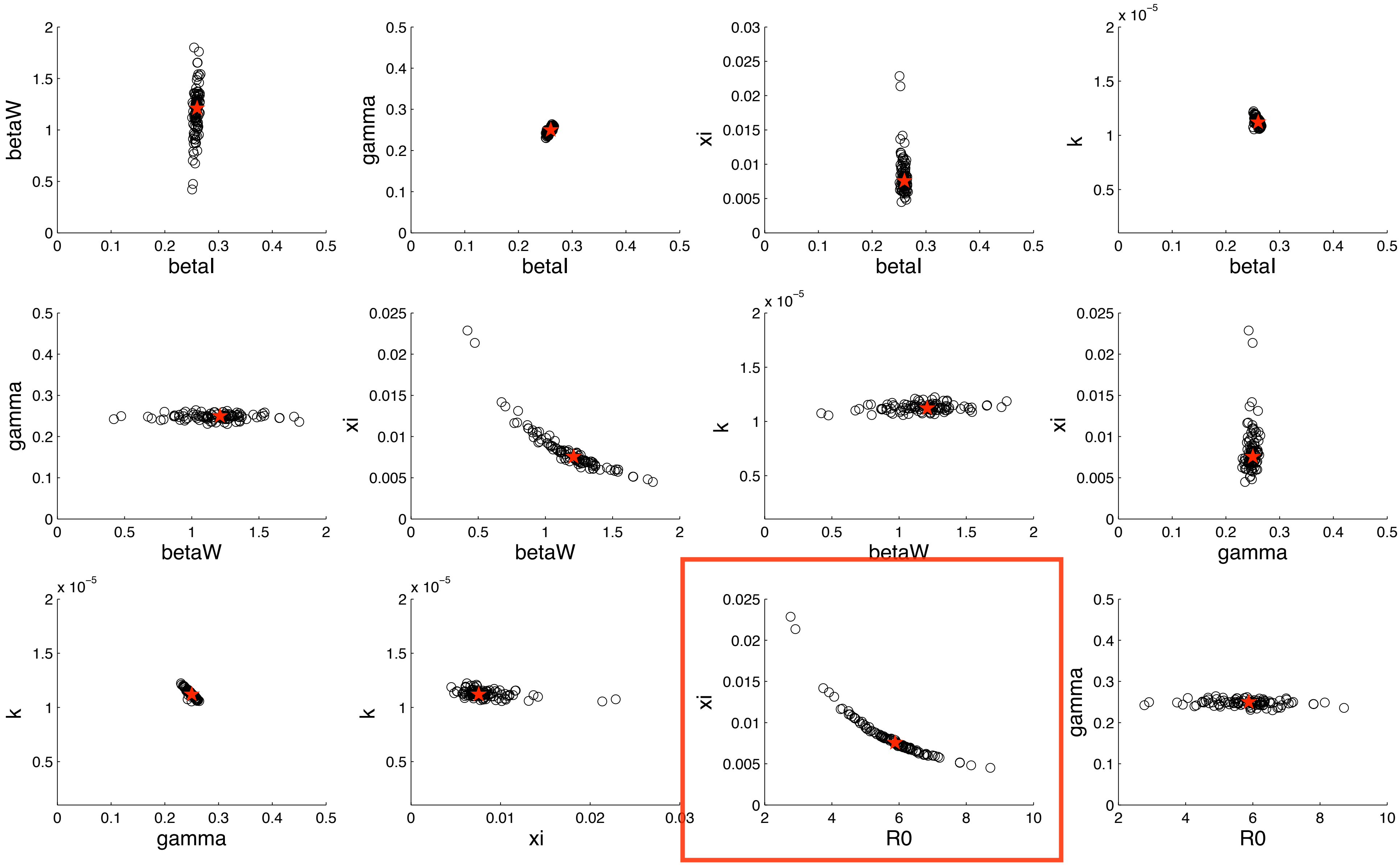
Poisson Noise



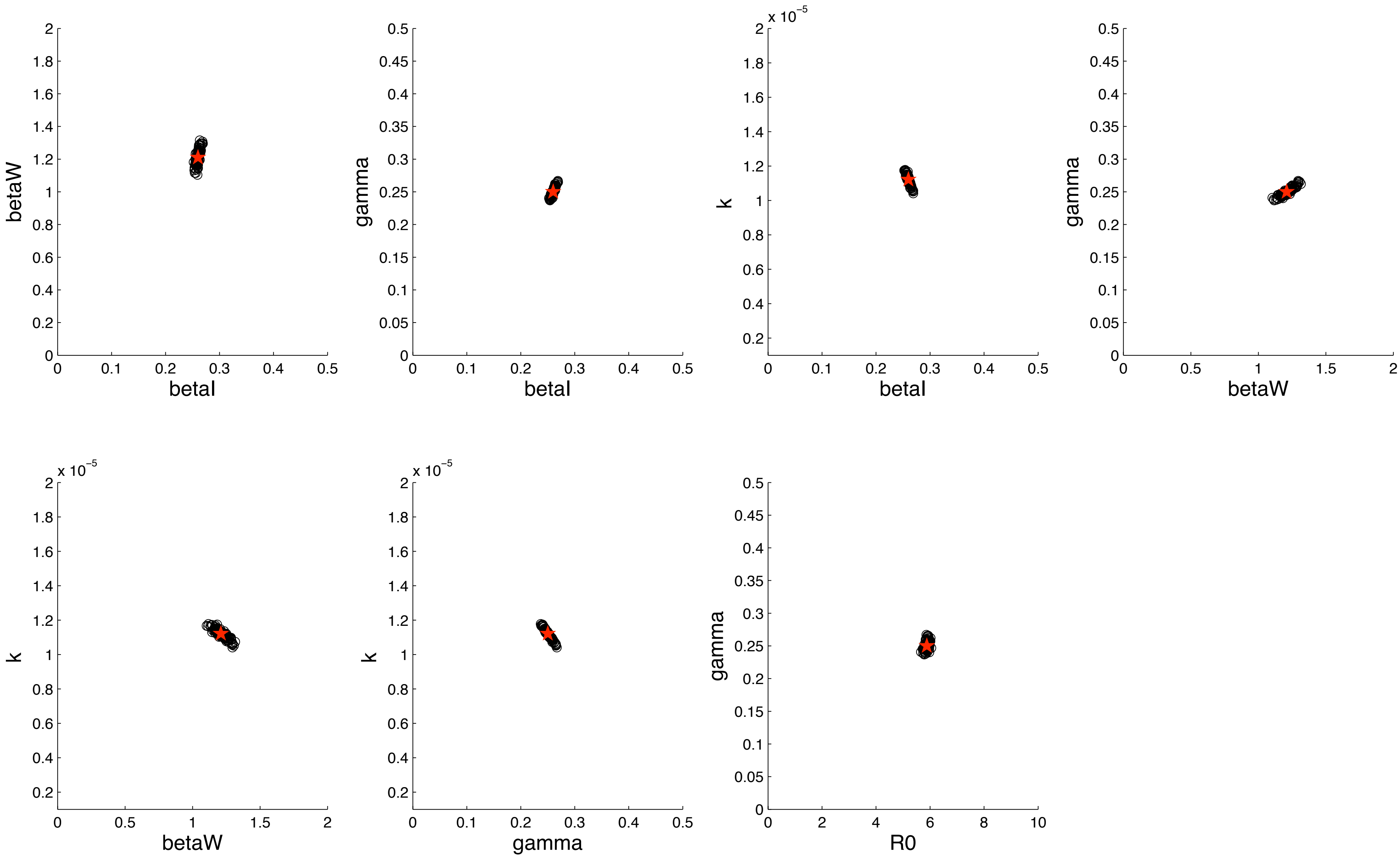
Poisson Noise



Poisson Noise

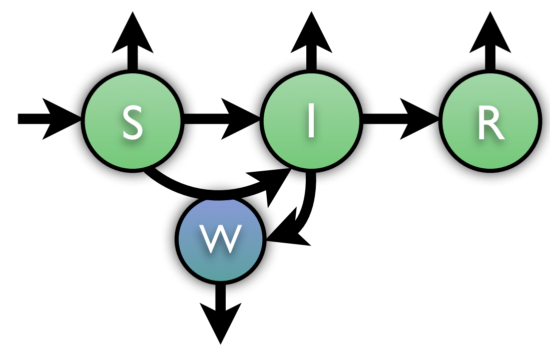


Water Info Improves Estimates



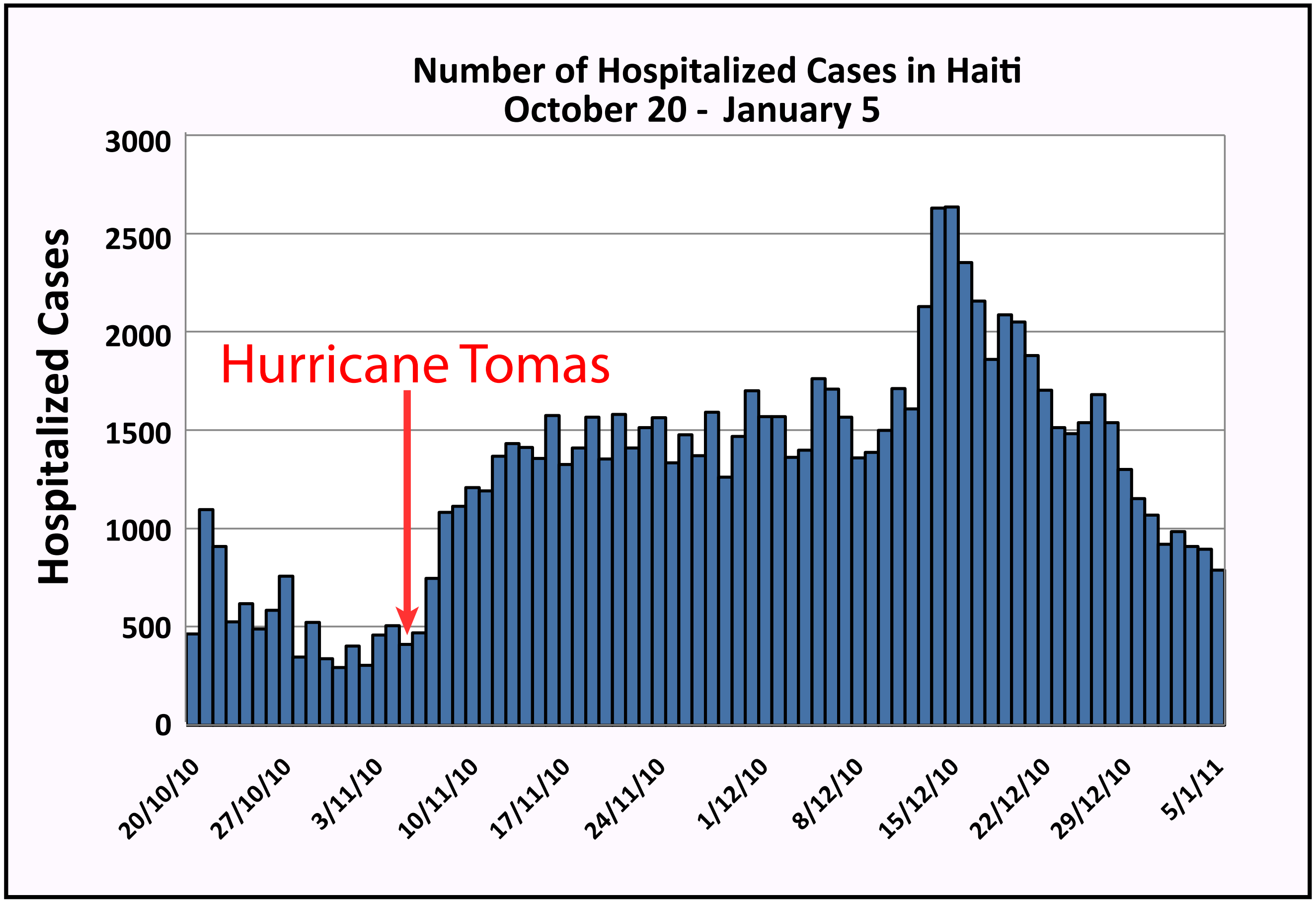
SIWR Identifiability Results

- SIWR model is uniquely structurally identifiable
- However, identifiability can be lost in the limit as pathogen lifetime decreases ($\xi \to \infty$)
- With noise, can also lose practical identifiability
- Adding water information can improve identifiability & make it possible to estimate relative contribution of transmission pathways



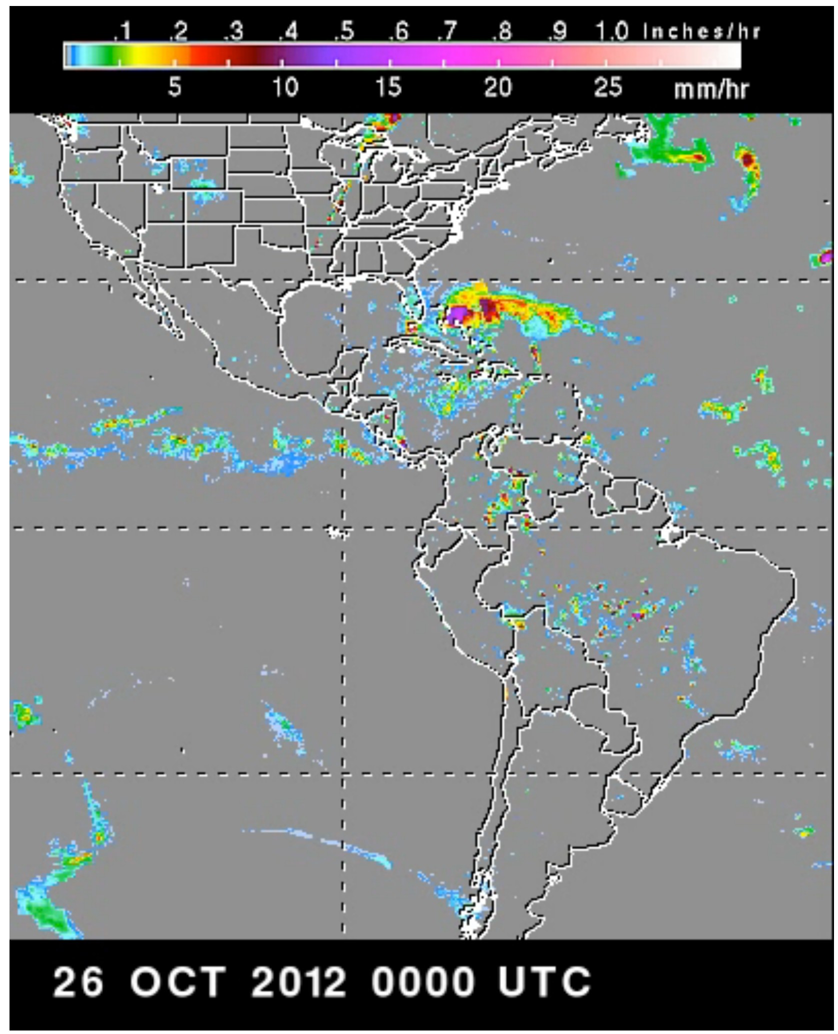
Eisenberg, Robertson, Tien 2013 (*In press*)

Cholera & the environment

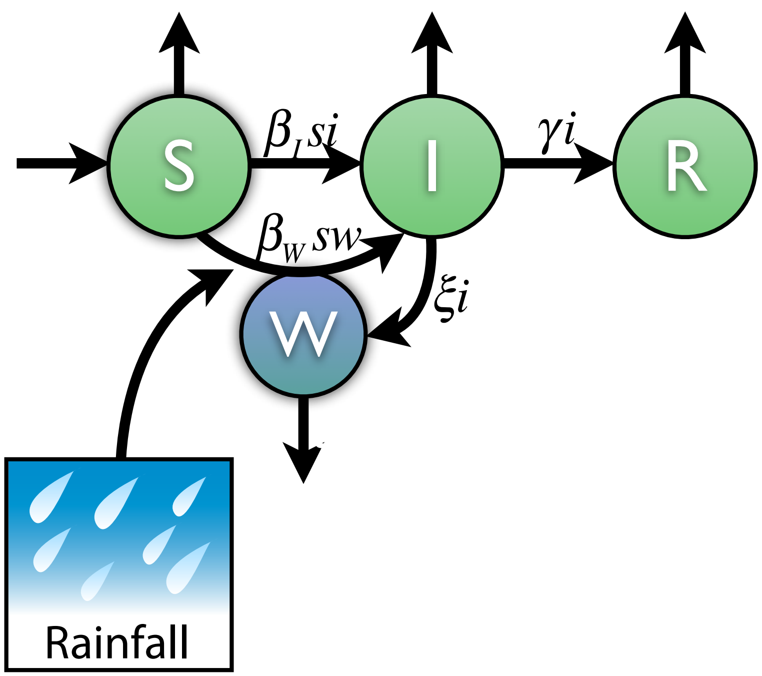


Rainfall Data

- NASA TRMM Data - satellite precipitation data (resolution 0.25° × 0.25°) averaged over each area
- USGS Rain Gauges in the Morne Gentilehomme and Foret de Pins regions



SIWR Model & Rainfall



$$\frac{ds}{dt} = \mu - \beta_W f_{rain}(t) ws - \beta_I si - \mu s$$

$$\frac{di}{dt} = \beta_W f_{rain}(t) ws + \beta_I si - \gamma i - \mu i$$

$$\frac{dw}{dt} = \xi (i - w)$$

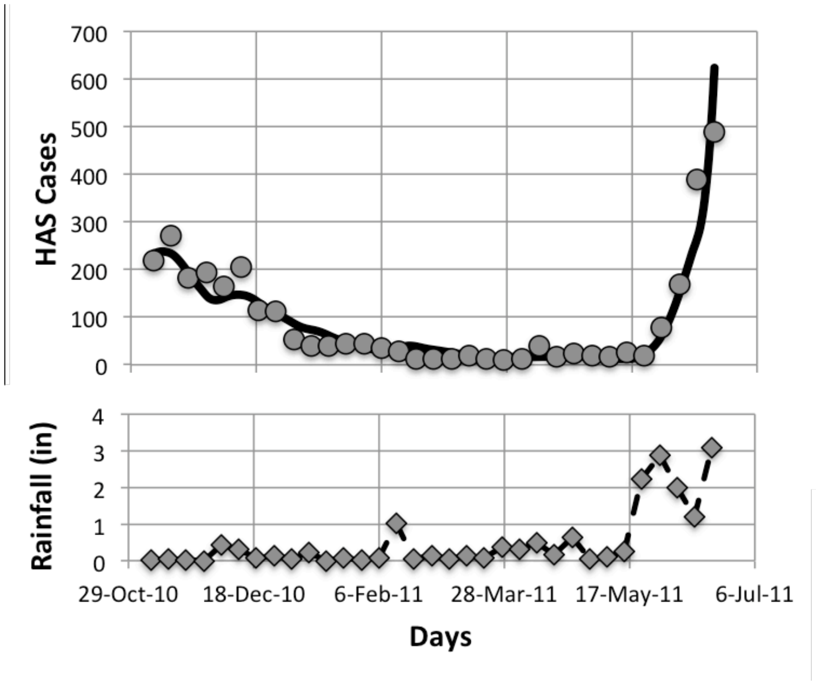
$$\frac{dr}{dt} = \gamma i - \mu r$$

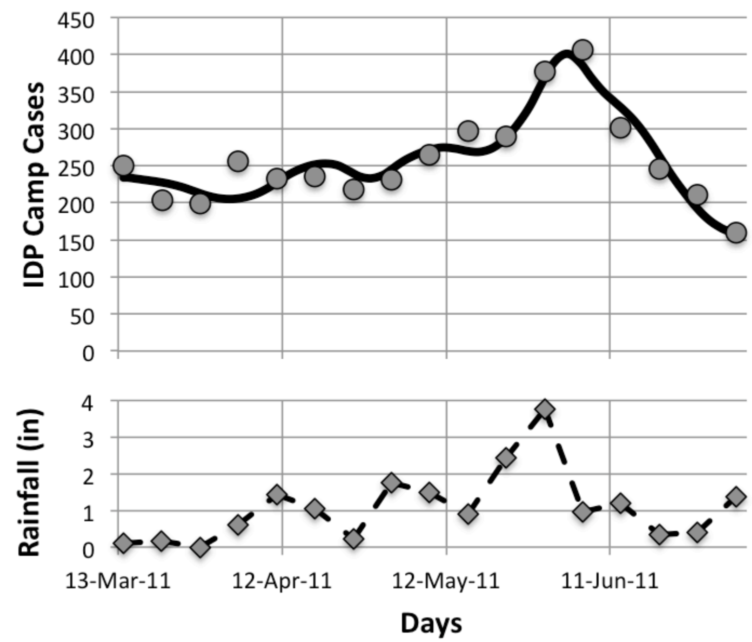
$$y = ki$$

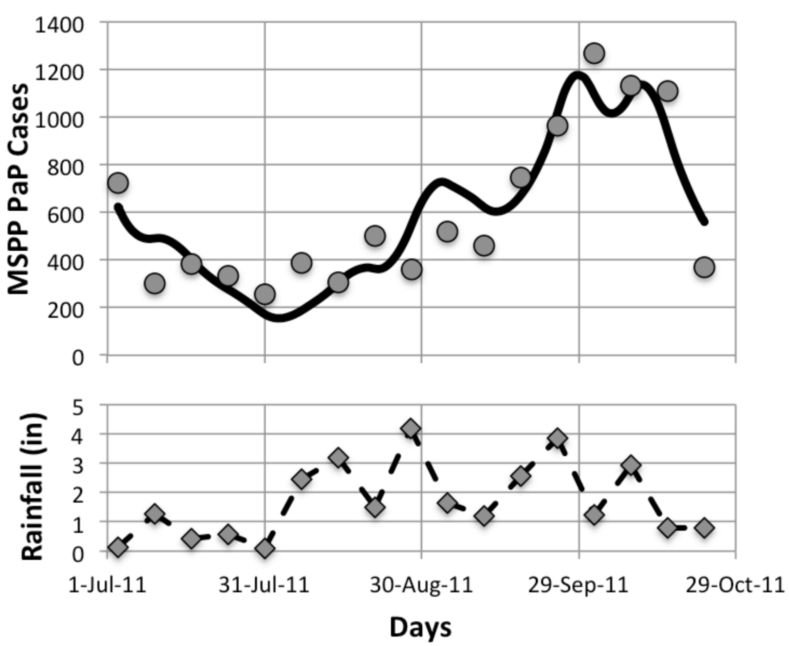
Eisenberg, Kujbida, Tuite, Fisman Tien, *Epidemics* 2013

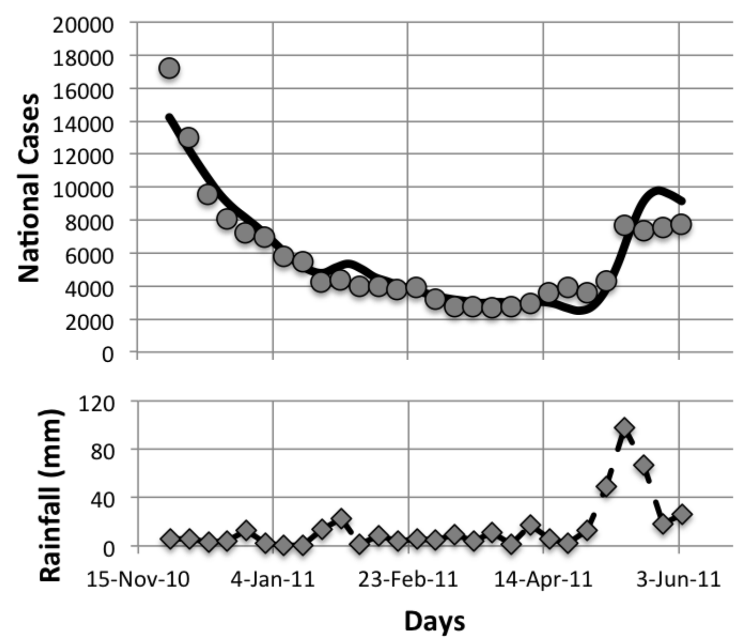
Rainfall Forcing & Identifiability

- Adding satellite rainfall data corrects the structural identfiability problem when $\xi \rightarrow \infty$
- Allows β_W and β_I to be estimated separately
- Can also improve practical identifiability

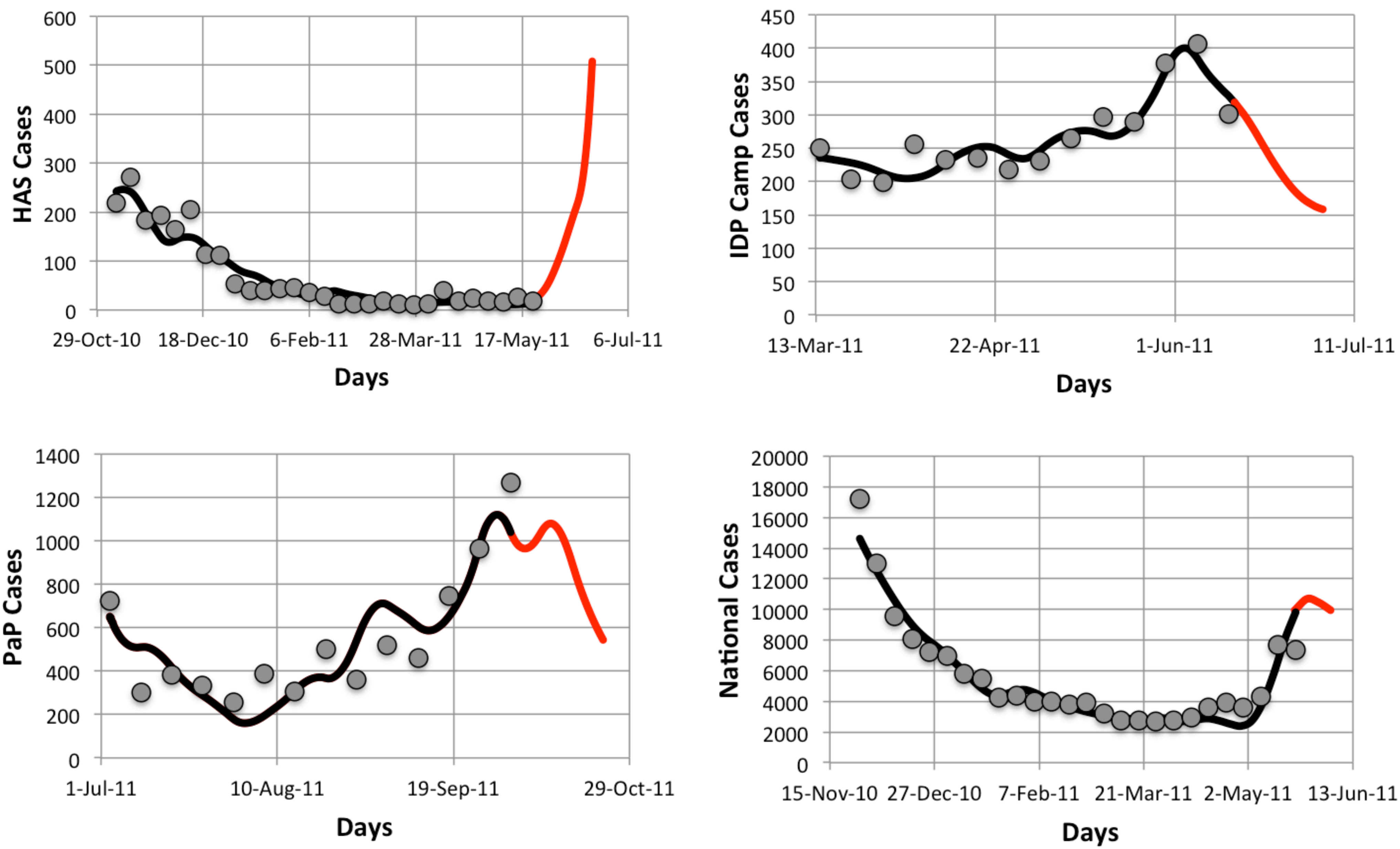




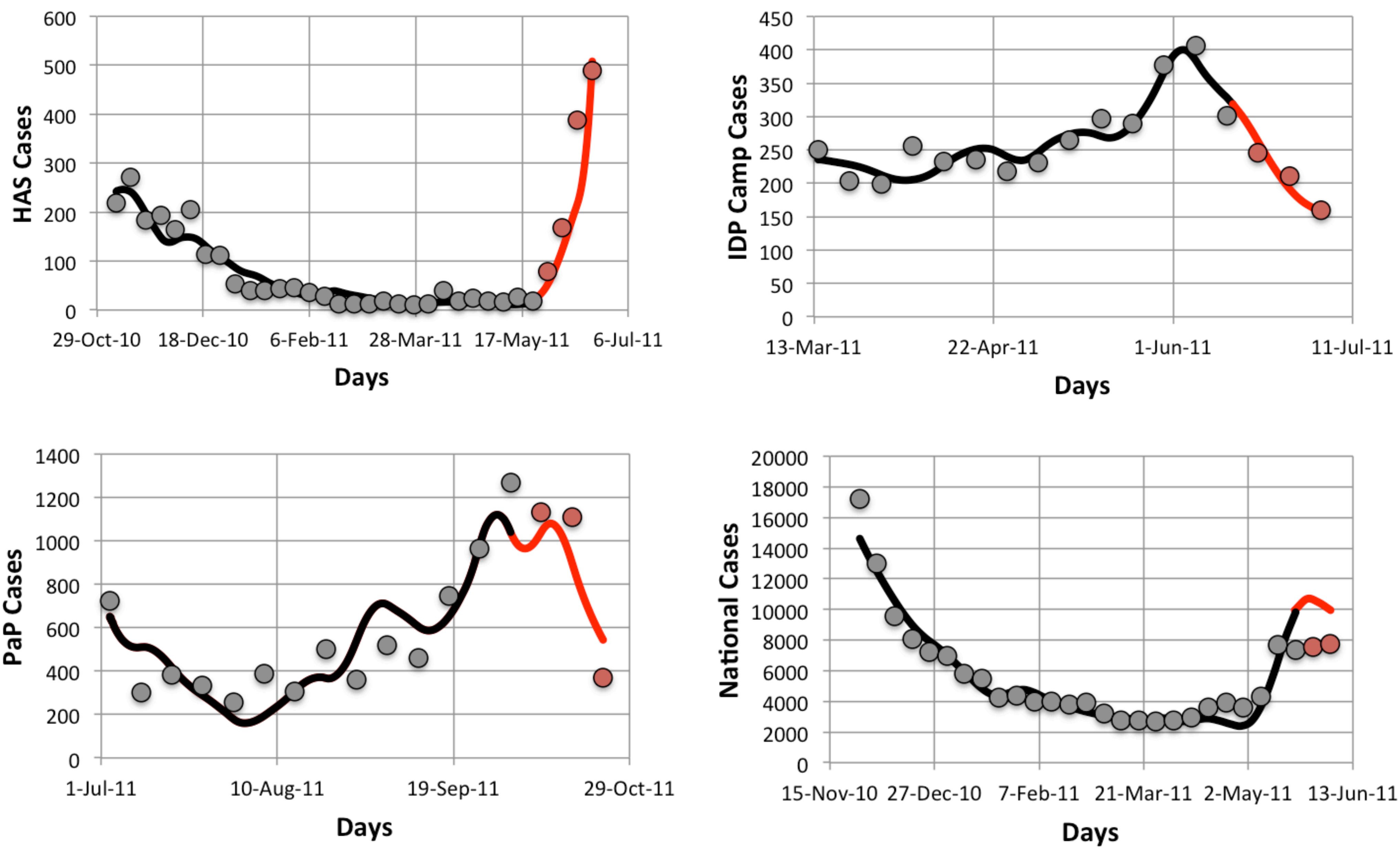




Eisenberg, Kujbida, Tuite, Fisman Tien 2013

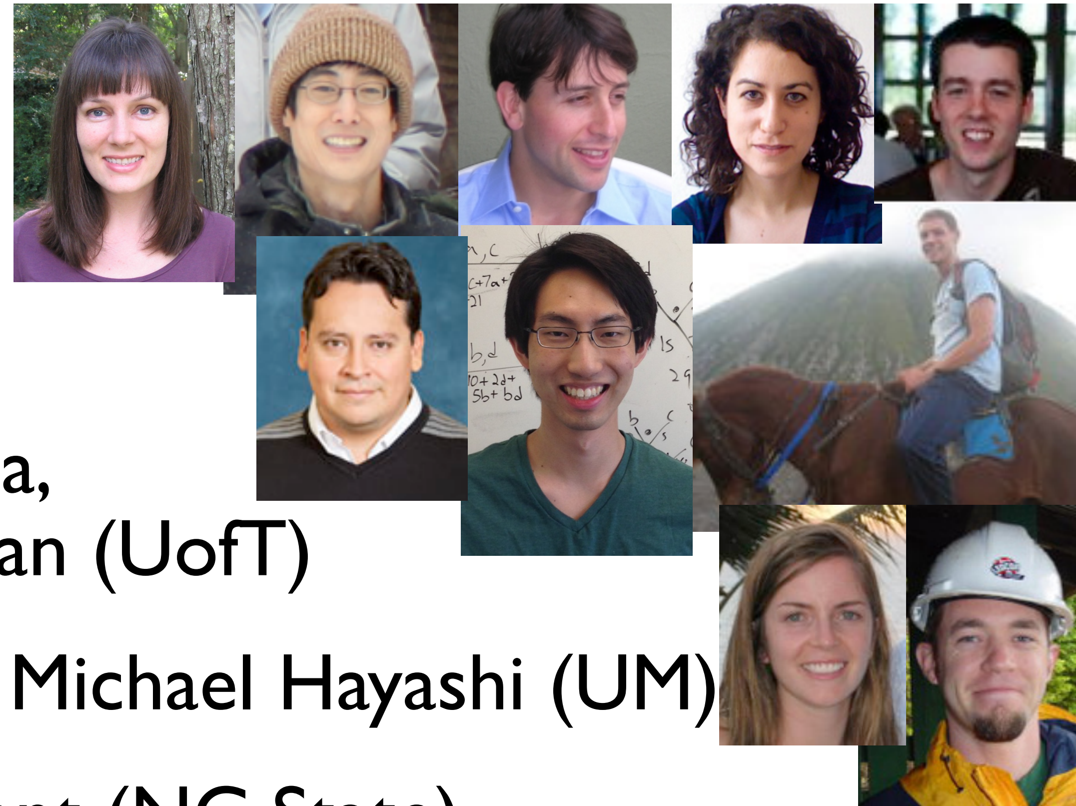


Eisenberg, Kujbida, Tuite, Fisman Tien, *Epidemics*, 2013



Eisenberg, Kujbida, Tuite, Fisman Tien, *Epidemics*, 2013

Thank you!

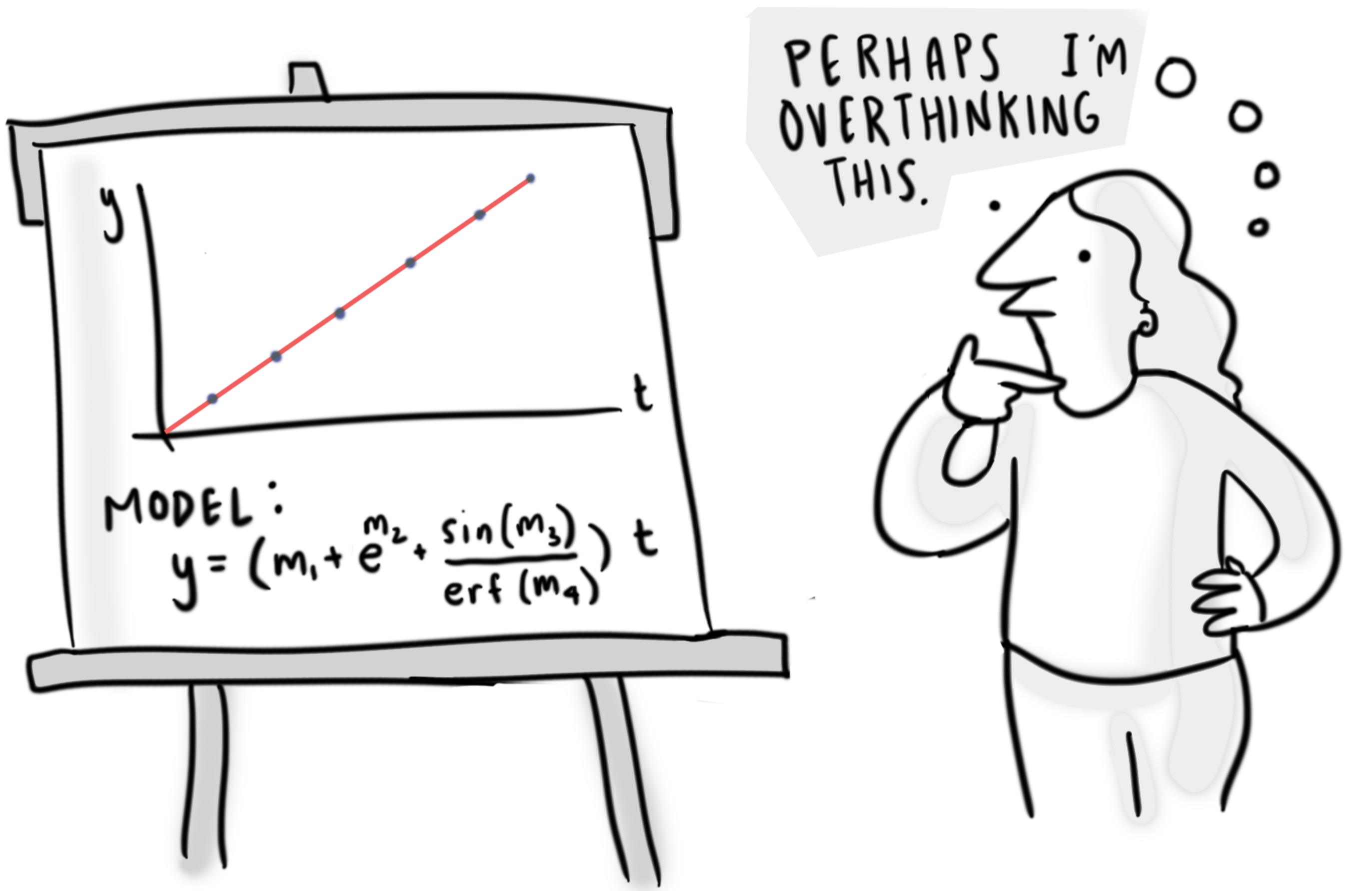


- Suzanne Robertson (VCU),
  Joe Tien (OSU), Greg Kujbida,
  Ashleigh Tuite, & David Fisman (UofT)
- Josh Havumaki, Rafael Meza, Michael Hayashi (UM)
- Nikki Meshkat & Seth Sullivant (NC State)
- Hôpital Albert Schweitzer, UN WASH Cluster,
  PU-AMI, Solidarités, Chris Phares & CDC
- Mathematical Biosciences Institute
- National Science Foundation (EEID Grant 1115881)

mbi
Mathematical Biosciences Institute
at The Ohio State University

University of Michigan

National Science Foundation
WHERE DISCOVERIES BEGIN



comic by Olivia Walch (UM):
http://imogenquest.net