Theory & Model Building

Philip D. Waggoner

MACS 30100: Perspectives on Computational Modeling

January 8, 2020

Lecture Outline

1. Building Theories
2. Building Models
3. Putting the Pieces Together: Theories + Models

Why models and theories?

- Implicitly asserting relationships in construction

Why models and theories?

- Implicitly asserting relationships in construction
- Think carefully and intentionally

Why models and theories?

- Implicitly asserting relationships in construction
- Think carefully and intentionally
- Add something useful to the world

A Model vs. A Theory

- **Theory** - a reasoned and precise speculation about the answer to a research question, including a statement about why the proposed solution is correct

A Model vs. A Theory

- **Theory** - a reasoned and precise speculation about the answer to a research question, including a statement about why the proposed solution is correct

- **Model** - a simplification of, and approximation to, some aspect of the world

A Model vs. A Theory

- **Theory** - a reasoned and precise speculation about the answer to a research question, including a statement about why the proposed solution is correct
- **Model** - a simplification of, and approximation to, some aspect of the world
- We will focus first on theories and how to (think about) develop(ing) them, though *this is really hard to do well*
- Transition to link models to theories at the end

Lecture Outline

1. Building Theories
2. Building Models
3. Putting the Pieces Together: Theories + Models

Developing Quality Theories

- A good theory is one that changes the way that we think about some aspect of the world

Developing Quality Theories

- A good theory is one that changes the way that we think about some aspect of the world
- The development of theories comes in where we are curious about how behavior may change when a rule is altered;

Developing Quality Theories

- A good theory is one that changes the way that we think about some aspect of the world
- The development of theories comes in where we are curious about how behavior may change when a rule is altered; or a customer base shifts;

Developing Quality Theories

- A good theory is one that changes the way that we think about some aspect of the world
- The development of theories comes in where we are curious about how behavior may change when a rule is altered; or a customer base shifts; or when an actor behaves unexpectedly;

Developing Quality Theories

- A good theory is one that changes the way that we think about some aspect of the world
- The development of theories comes in where we are curious about how behavior may change when a rule is altered; or a customer base shifts; or when an actor behaves unexpectedly; or when an institution is reformed; and so on

Developing Quality Theories

- A good theory is one that changes the way that we think about some aspect of the world
- The development of theories comes in where we are curious about how behavior may change when a rule is altered; or a customer base shifts; or when an actor behaves unexpectedly; or when an institution is reformed; and so on
- To build a theory, then, we are really just codifying a move from some specific event to a more *general story of behavior, change, action, etc.*

What is a Theory?

What is a Theory?

- Your expectation for the way the world works

What is a Theory?

- Your expectation for the way the world works
- Includes: observation,

What is a Theory?

- Your expectation for the way the world works
- Includes: observation, questions,

What is a Theory?

- Your expectation for the way the world works
- Includes: observation, questions, engagement,

What is a Theory?

- Your expectation for the way the world works
- Includes: observation, questions, engagement, and expectations

What is a Theory?

- Your expectation for the way the world works
- Includes: observation, questions, engagement, and expectations
- Functionally, theories are meant to explain variation in an outcome of interest

What is a Theory?

- Your expectation for the way the world works
- Includes: observation, questions, engagement, and expectations
- Functionally, theories are meant to explain variation in an outcome of interest
- A key in theory building, or story telling too though, is the counterfactual

What is a Theory?

- Your expectation for the way the world works
- Includes: observation, questions, engagement, and expectations
- Functionally, theories are meant to explain variation in an outcome of interest
- A key in theory building, or story telling too though, is the counterfactual
- So how do we create a theory?

Theories are Intimately Related to (Interesting) Questions

- One of the main ways in which theories can be evaluated is in terms of the questions they set out to answer

Theories are Intimately Related to (Interesting) Questions

- One of the main ways in which theories can be evaluated is in terms of the questions they set out to answer
- If the question being addressed by a theory is interesting and important, then that theory has potential

Good Questions Come from Observations

- Paying attention to the world

Good Questions Come from Observations

- Paying attention to the world
- Recognizing patterns in a descriptive sense

Identifying Interesting Variation as a Start

- Begin by identifying interesting variation $\rightsquigarrow$ Measurement of an outcome/response/DV

Identifying Interesting Variation as a Start

- Begin by identifying interesting variation $\rightsquigarrow$ Measurement of an outcome/response/DV
- When we think about measuring a response feature, first identify a dimension: usually either spatial or time

Identifying Interesting Variation as a Start

- Begin by identifying interesting variation $\rightsquigarrow$ Measurement of an outcome/response/DV
- When we think about measuring a response feature, first identify a dimension: usually either spatial or time
    1. The spatial dimension $\rightsquigarrow$ units in comparison to other units

Identifying Interesting Variation as a Start

- Begin by identifying interesting variation $\rightsquigarrow$ Measurement of an outcome/response/DV
- When we think about measuring a response feature, first identify a dimension: usually either spatial or time
  1. The spatial dimension $\rightsquigarrow$ units in comparison to other units
  2. The time dimension $\rightsquigarrow$ points in time *at which* we would like to measure

Dynamic vs. Static Tradeoffs

- Implied tradeoff between dynamic and static

Dynamic vs. Static Tradeoffs

- Implied tradeoff between dynamic and static
- This means that our measures of our DV will most likely be of one of two types:

Dynamic vs. Static Tradeoffs

- Implied tradeoff between dynamic and static
- This means that our measures of our DV will most likely be of one of two types:
    1. **Time-series**: spatial dimension is the same for all cases and the response is measured at multiple points in time

Dynamic vs. Static Tradeoffs

- Implied tradeoff between dynamic and static
- This means that our measures of our DV will most likely be of one of two types:
  1. **Time-series**: spatial dimension is the same for all cases and the response is measured at multiple points in time
  2. **Cross-sectional**: time dimension is the same for all cases and the response is measured for multiple spatial units

How Do I Know If I Have a "Good" Theory?

- Does your theory offer an answer to an interesting research question?

How Do I Know If I Have a "Good" Theory?

- Does your theory offer an answer to an interesting research question?
- Is your theory causal?

How Do I Know If I Have a "Good" Theory?

- Does your theory offer an answer to an interesting research question?
- Is your theory causal?
- Can you test your theory on data that you have not yet observed? (i.e., training and testing)

How Do I Know If I Have a "Good" Theory?

- Does your theory offer an answer to an interesting research question?
- Is your theory causal?
- Can you test your theory on data that you have not yet observed? (i.e., training and testing)
- How general is your theory?

How Do I Know If I Have a "Good" Theory?

- Does your theory offer an answer to an interesting research question?
- Is your theory causal?
- Can you test your theory on data that you have not yet observed? (i.e., training and testing)
- How general is your theory?
- How parsimonious is your theory?

How Do I Know If I Have a "Good" Theory?

- Does your theory offer an answer to an interesting research question?
- Is your theory causal?
- Can you test your theory on data that you have not yet observed? (i.e., training and testing)
- How general is your theory?
- How parsimonious is your theory?
- How new is your theory?

How Do I Know If I Have a "Good" Theory?

- Does your theory offer an answer to an interesting research question?
- Is your theory causal?
- Can you test your theory on data that you have not yet observed? (i.e., training and testing)
- How general is your theory?
- How parsimonious is your theory?
- How new is your theory?
- How non-obvious is your theory?

Lecture Outline

1. Building Theories
2. Building Models
3. Putting the Pieces Together: Theories + Models

Social science model

- Simplified representation of the real world, which often flow from your *original* theory

Social science model

- Simplified representation of the real world, which often flow from your *original* theory
  - Build an initial model

Social science model

- Simplified representation of the real world, which often flow from your *original* theory
  - Build an initial model
  - Evaluate its effectiveness

Social science model

- Simplified representation of the real world, which often flow from your *original* theory
  - Build an initial model
  - Evaluate its effectiveness
  - Revise your model

Social science model

- Simplified representation of the real world, which often flow from your *original* theory
  - Build an initial model
  - Evaluate its effectiveness
  - Revise your model
- Not driven purely by predictive power

7 Steps to Building a Model

1. Get an idea

7 Steps to Building a Model

1. Get an idea
2. Decide to pursue the idea

7 Steps to Building a Model

1. Get an idea
2. Decide to pursue the idea
3. Examine the literature (or don't)

7 Steps to Building a Model

1. Get an idea
2. Decide to pursue the idea
3. Examine the literature (or don't)
4. Build your model

7 Steps to Building a Model

1. Get an idea
2. Decide to pursue the idea
3. Examine the literature (or don't)
4. Build your model
5. Make mistakes

7 Steps to Building a Model

1. Get an idea
2. Decide to pursue the idea
3. Examine the literature (or don't)
4. Build your model
5. Make mistakes
6. (Re)Turn to the literature

7 Steps to Building a Model

1. Get an idea
2. Decide to pursue the idea
3. Examine the literature (or don't)
4. Build your model
5. Make mistakes
6. (Re)Turn to the literature
7. Give a seminar

Step 1: Get an idea

- Ideas are easy to generate - you want a *good* idea (e.g., from a theory)

Step 1: Get an idea

- Ideas are easy to generate - you want a *good* idea (e.g., from a theory)
- Where can you find ideas?
  - Academic journals
  - Newspapers, magazines
  - Talking to people (including non-academics)

Step 1: Get an idea

- Ideas are easy to generate - you want a *good* idea (e.g., from a theory)
- Where can you find ideas?
  - Academic journals
  - Newspapers, magazines
  - Talking to people (including non-academics)
- High-risk vs. high-reward

Step 2: Decide to Pursue the Idea

Step 2: Decide to Pursue the Idea

- Can you phrase it as a question?

Step 2: Decide to Pursue the Idea

- Can you phrase it as a question?
- Can you meaningfully explain it to a non-expert?

Step 2: Decide to Pursue the Idea

- Can you phrase it as a question?
- Can you meaningfully explain it to a non-expert?
- Albert Einstein . . .
  - "[Your model] should be as simple as possible. . . but no simpler."

Step 2: Decide to Pursue the Idea

- Can you phrase it as a question?
- Can you meaningfully explain it to a non-expert?
- Albert Einstein . . .
  - "[Your model] should be as simple as possible. . . but no simpler."
  - "If you can't explain it simply, you don't understand it well enough."

Step 2: Decide to Pursue the Idea

- Can you phrase it as a question?
- Can you meaningfully explain it to a non-expert?
- Albert Einstein . . .
  - "[Your model] should be as simple as possible. . . but no simpler."
  - "If you can't explain it simply, you don't understand it well enough."
- Is it interesting?

Step 3: Examine the literature (or don't)

- Don't look too soon
  - Potentially more work
  - May come up with a different approach than what people have done before you

Step 3: Examine the literature (or don't)

- Don't look too soon
  - Potentially more work
  - May come up with a different approach than what people have done before you

* Eventually you will do a thorough review of the literature

Step 4: Build Your Model

- Identify the outcome of interest (endogenous feature(s))

Step 4: Build Your Model

- Identify the outcome of interest (endogenous feature(s))
- Identify relevant features that will explain, predict, and lead to understanding of the outcome of interest (exogenous features)

Step 4: Build Your Model

- Identify the outcome of interest (endogenous feature(s))
- Identify relevant features that will explain, predict, and lead to understanding of the outcome of interest (exogenous features)
- Work a simple example first

Step 4: Build Your Model

- Identify the outcome of interest (endogenous feature(s))
- Identify relevant features that will explain, predict, and lead to understanding of the outcome of interest (exogenous features)
- Work a simple example first
- Then generalize your model piece-by-piece

Step 4: Build Your Model

- Three rules of thumb for model building

Step 4: Build Your Model

- Three rules of thumb for model building
  1. Think "process"
  2. Develop interesting implications
  3. Look for generality

Step 5: Make Mistakes

- If everything is perfect and works the first time, I wouldn't believe it or the results produced

Step 5: Make Mistakes

- If everything is perfect and works the first time, I wouldn't believe it or the results produced $\rightsquigarrow$ *embrace* mistakes

Step 5: Make Mistakes

- If everything is perfect and works the first time, I wouldn't believe it or the results produced $\rightsquigarrow$ *embrace* mistakes
- Its the same with code

Step 5: Make Mistakes

- If everything is perfect and works the first time, I wouldn't believe it or the results produced $\rightsquigarrow$ *embrace* mistakes
- Its the same with code

$\rightarrow$ "It's easy when you start out programming to get really frustrated and think, "Oh it's me, I'm really stupid," or, "I'm not made out to program." But, that is absolutely not the case. Everyone gets frustrated. I still get frustrated occasionally when writing R code. It's just a natural part of programming."

Step 5: Make Mistakes

- If everything is perfect and works the first time, I wouldn't believe it or the results produced $\rightsquigarrow$ *embrace* mistakes
- Its the same with code

$\rightarrow$ "It's easy when you start out programming to get really frustrated and think, "Oh it's me, I'm really stupid," or, "I'm not made out to program." But, that is absolutely not the case. Everyone gets frustrated. I still get frustrated occasionally when writing R code. It's just a natural part of programming."

- So too with developing theories and deriving models

Step 6: (Re)Turn to the literature

- If you have not already, begin your literature search

Step 6: (Re)Turn to the literature

- If you have not already, begin your literature search
- Search on Google Scholar

Step 6: (Re)Turn to the literature

- If you have not already, begin your literature search
- Search on Google Scholar
- Talk to senior/junior colleagues

Step 6: (Re)Turn to the literature

- If you have not already, begin your literature search
- Search on Google Scholar
- Talk to senior/junior colleagues
- Puts the value of your contribution in context

Step 6: (Re)Turn to the literature

- If you have not already, begin your literature search
- Search on Google Scholar
- Talk to senior/junior colleagues
- Puts the value of your contribution in context
- You may find out someone else already did what you did (it sucks)

Step 7: Give a seminar

- Show your work to other people

Step 7: Give a seminar

- Show your work to other people $\rightsquigarrow$ Giving a seminar/talk is often the easiest way for both parties to communicate

Step 7: Give a seminar

- Show your work to other people $\rightsquigarrow$ Giving a seminar/talk is often the easiest way for both parties to communicate
- You can lose perspective when you are deep in the trenches of your project, e.g., ...
  - You may think something is obvious, when it is not to others
  - You may think something is complicated, when it is really obvious

A Caveat

- Much of our discussion to this point has been (implicitly) on developing original research

A Caveat

- Much of our discussion to this point has been (implicitly) on developing original research
- But importantly (and especially if you are going to industry), modeling is sometimes an exploratory process, rather than pure theory development

A Caveat

- Much of our discussion to this point has been (implicitly) on developing original research
- But importantly (and especially if you are going to industry), modeling is sometimes an exploratory process, rather than pure theory development
- The trick is knowing precisely when to use one or the other, which is most-often dictated by the target audience and contribution (e.g., a pitch to a boss in response to a task vs. an academic, peer-reviewed journal article)

Lecture Outline

1. Building Theories
2. Building Models
3. Putting the Pieces Together: Theories + Models

A Case Study: Rational Choice Theory

A Case Study: Rational Choice Theory

- "Formal theory" or "rational choice" approaches are often used to *develop* causal theories, which translate into models, or simplifications of the world for testing purposes

A Case Study: Rational Choice Theory

- "Formal theory" or "rational choice" approaches are often used to *develop* causal theories, which translate into models, or simplifications of the world for testing purposes
- In this world, we start with basic set of assumptions about behavior and then use game theory and other mathematical tools to build models

A Case Study: Rational Choice Theory

- "Formal theory" or "rational choice" approaches are often used to *develop* causal theories, which translate into models, or simplifications of the world for testing purposes
- In this world, we start with basic set of assumptions about behavior and then use game theory and other mathematical tools to build models
- Assumptions are key:

A Case Study: Rational Choice Theory

- "Formal theory" or "rational choice" approaches are often used to *develop* causal theories, which translate into models, or simplifications of the world for testing purposes
- In this world, we start with basic set of assumptions about behavior and then use game theory and other mathematical tools to build models
- Assumptions are key: We can summarize these assumptions by suggesting *all* individuals are rational utility maximizers...

A Case Study: Rational Choice Theory

- "Formal theory" or "rational choice" approaches are often used to *develop* causal theories, which translate into models, or simplifications of the world for testing purposes
- In this world, we start with basic set of assumptions about behavior and then use game theory and other mathematical tools to build models
- Assumptions are key: We can summarize these assumptions by suggesting *all* individuals are rational utility maximizers... $\rightsquigarrow$?

Rational Utility

- In thinking about an individual's choices of actions, we will talk about things in terms of utility (*intentionally vague*)

Rational Utility

- In thinking about an individual's choices of actions, we will talk about things in terms of utility (*intentionally vague*)
- We can start by defining a key concept that will underlie our theoretical claim: The utility from a particular action is equal to the sum of all benefits minus the sum of all costs from that action

Rational Utility

- In thinking about an individual's choices of actions, we will talk about things in terms of utility (*intentionally vague*)
- We can start by defining a key concept that will underlie our theoretical claim: The utility from a particular action is equal to the sum of all benefits minus the sum of all costs from that action
- If we consider some action Y, we summarize the utility from Y for individual i,

$$U_i(Y) = \sum B_i(Y) - \sum C_i(Y)$$

Rational Utility Maximizers

- **Theoretical claim**: When choosing among a set of possible actions (including the decisions not to act), a rational individual will choose that action that maximizes their utility,

$$\text{given a set of choices } Y = Y_1, Y_2, Y_3, \ldots, Y_n,$$

$$\text{individual } i \text{ will choose } Y_a \text{ such that } U_i(Y_a) > U_i(Y_b) \; \forall \; b \neq a,$$

Modeling Rationality: Expected utility

- What if we don't know exactly what the costs and benefits will be from a particular action?

Modeling Rationality: Expected utility

- What if we don't know exactly what the costs and benefits will be from a particular action? *How do we* **model** *this?*

Modeling Rationality: Expected utility

- What if we don't know exactly what the costs and benefits will be from a particular action? *How do we* **model** *this?*

- Previously, we were assuming complete information, for which we knew exactly what would be the costs (C_i), benefits(B_i), and thus the utility from each possible action ($U_i(Y)$)

Modeling Rationality: Expected utility

- What if we don't know exactly what the costs and benefits will be from a particular action? *How do we* **model** *this?*

- Previously, we were assuming complete information, for which we knew exactly what would be the costs (C_i), benefits(B_i), and thus the utility from each possible action ($U_i(Y)$) **But this seems unrealistic**

Modeling Rationality: Expected utility

- What if we don't know exactly what the costs and benefits will be from a particular action? *How do we* **model** *this?*
- Previously, we were assuming complete information, for which we knew exactly what would be the costs (C_i), benefits(B_i), and thus the utility from each possible action ($U_i(Y)$) **But this seems unrealistic**
- When we relax this assumption, we move our discussion from utility to expected utility (E), i.e., "putting expectations" in front of all utilities

Modeling Rationality: Expected utility

- The term $U_i(Y)$, becomes $E[U_i(Y)]$ under **incomplete information**

Modeling Rationality: Expected utility

- The term $U_i(Y)$, becomes $E[U_i(Y)]$ under **incomplete information**
- So, returning to our rational actor assumption, under incomplete information, for an individual action Y,

Modeling Rationality: Expected utility

- The term $U_i(Y)$, becomes $E[U_i(Y)]$ under **incomplete information**
- So, returning to our rational actor assumption, under incomplete information, for an individual action Y,

$$E[U_i(Y)] = \sum E[B_i(Y)] - \sum E[C_i(Y)],$$

Modeling Rationality: Expected utility

- The term $U_i(Y)$, becomes $E[U_i(Y)]$ under **incomplete information**
- So, returning to our rational actor assumption, under incomplete information, for an individual action Y,

$$E[U_i(Y)] = \sum E[B_i(Y)] - \sum E[C_i(Y)],$$

- and a rational actor will maximize her expected utility where,

given a set of choices $Y = Y_1, Y_2, Y_3, \ldots, Y_n$,

$\rightarrow$ individual i will choose Y_a such that $E[U_i(Y_a)] > E[U_i(Y_b)] \forall b \neq a$.

Practice Building a Model

Practice Building a Model

- Small groups $\rightsquigarrow$ Develop a model

Practice Building a Model

- Small groups $\rightsquigarrow$ Develop a model

- **Possible topics**
  - Murders in Chicago
  - Fake news and elections
  - Lying and political candidates
  - Medical treatments and patient outcomes
  - International war
  - Thriving and struggling economies
  - Successful businesses, etc.

Practice Building a Model

- Small groups $\rightsquigarrow$ Develop a model

- **Possible topics**
  - Murders in Chicago
  - Fake news and elections
  - Lying and political candidates
  - Medical treatments and patient outcomes
  - International war
  - Thriving and struggling economies
  - Successful businesses, etc.

- **Questions to consider**
  - What is the research question?
  - What is the process in question?
  - What are the key factors?
  - What is exogenous, and what is endogenous?
  - How might you get data on these things?