Theorizing in social science

Formalization and mathematics in social science

Giuseppe Vittucci Marzetti[1]

Doctoral Program in Analysis of Social and Economic Processes (ASEP)
Department of Sociology and Social Science
University of Milano-Bicocca

November 30, 2017

[1]Department of Sociology and Social Research, University of Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126, Milan, E-mail: giuseppe.vittucci@unimib.it

Explanation and formalization

- Social science aims at explaining social phenomena.
- What is an "explanation" of a phenomenon? (Abbott (2004) builds on Morris' classification of relations in symbolic systems)
    - *Pragmatic view*: an account that enables us to intervene
        - It works best for phenomena with "a narrow neck of necessary causality" (2004, p. 9);
    - *Semantic view* (*reduction*):
        - an account that enables us to stop looking for further accounts;
        - a translation of a phenomenon into a realm we think we understand intuitively.
    - *Syntactic view*: a demonstration that the phenomenon fits the condition of some general "covering law" (Hempel, 1942).
- Each sense defines an *explanatory program*.
- According to Abbott (2004):
    - *formalization* is an "abstract version of the syntactic explanatory program" (p. 27);
    - formal modeling and simulation "attempt to improve syntactic explanation by making it more abstract" (p. 34).

Modeling

- According to Frigg & Hartmann (2012), models perform two fundamentally different "representational functions": they represent
  - a selected part of the world (the "target system");
  - a theory in the sense of interpreting its laws and axioms.

  These notions are not mutually exclusive as scientific models can be representations in both senses at the same time.
- A (formal) mathematical model is a particular representation of a phenomenon – i.e. a relatively stable and general feature of the world that are interesting from a scientific point of view – that allows one to analyze the phenomenon itself by means of math.

The role of math in models

"Mathematical translation is itself a substantive contribution to the theory. [...] Mathematics has become the dominant language of the natural sciences not because it is quantitative – a common delusion – but primarily because it permits clear and rigorous reasoning about phenomena too complex to be handled in words".

(Simon, 1957, p. 89)

"Mathematics is a language, and all that mathematical sociology means to me is that sociological ideas are expressed in mathematical terms, and that we try and take advantage of using the mathematics. I don't see mathematics as being that special, at least in my own work".

(Patrick Doreian, quoted by Edling, 2002, p. 200)

Krugman (1997) on the aim and scope of formal model-building in social sciences

The objective of the most basic physics is a complete description of what happens. ... But most things we want to analyze, even in physical science, cannot be dealt with at that level of completeness. The only exact model of the global weather system is that system itself. Any smaller-scale model of that system is therefore to some degree a falsification: it leaves out many aspects of reality.

And how do you know that the model is good? It will never be right in the way that quantum electrodynamics is right. At a certain point you may be good enough at predicting that your results can be put to repeated practical use ...; in that case predictive success can be measured ..., and the improvement of models becomes a quantifiable matter.

In the early stages of a complex science, however, the criterion for a good model is more subjective: it is a good model if it succeeds in explaining or rationalizing some of what you see in the world in a way that you might not have expected. [...]

The important point is that any kind of model of a complex system – a physical model, a computer simulation, or a pencil-and-paper mathematical representation – amounts to pretty much the same kind of procedure. You make a set of clearly untrue simplifications to get the system down to something you can handle; those simplifications are dictated partly by guesses about what is important, partly by the modeling techniques available. And the end result, if the model is a good one, is an improved insight into why the vastly more complex real system behaves the way it does.

(Krugman, 1997)

Once upon a time there was a field called development economics...

So what is it that makes some ideas acceptable, while others are not? The answer which is obvious to anyone immersed in economic research yet mysterious to outsiders is that to be taken seriously an idea has to be something you can model.
A properly modeled idea is, in modern economics, the moral equivalent of a properly surveyed region for eighteenth-century mapmakers.

(Krugman, 1997)

- In the 1940s and 1950s the ideas of development economics were regarded as revolutionary and important.
- Development economics attracted creative minds and was marked by a great deal of intellectual excitement (e.g. Hirschman, 1958; Lewis, 1954; Rosenstein-Rodan, 1943).
- Between 1960-1980, this "high development theory was buried".
- According to Krugman (1997), this was due to the inability of the creators to express their ideas in a way suitable for the modeling techniques available at the time.

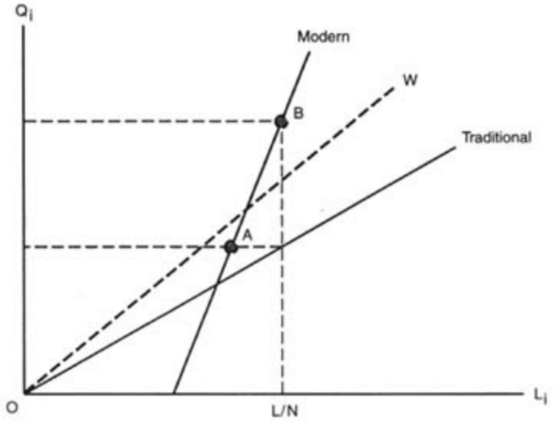
The Big Push

- Rosenstein-Rodan's (1943) paper on the idea of "Big Push" is a short non-mathematical paper.
- The paper has inspired many different interpretations and other papers on "low-level equilibrium traps".
- The original idea was formalized only in the late 1980s (Murphy, Shleifer, and Vishny, 1989).

Murphy et al.'s (1989) Big Push model

- Symmetric Cobb-Douglas utility function over N goods $\Rightarrow$ Same constant fraction of income spent in each good ($1/N$).
- Single factor of production (labor) in fixed total supply (L).
- Closed economy producing N goods: each good i can be produced
    - in a "traditional" sector:
        - constant returns to scale:
          $$Q_i = L_i$$
        - perfect competition $\Rightarrow$ *marginal cost pricing*: price equals wage in the traditional sector (set to 1).
    - in a "modern" sector:
        - increasing returns to scale:
          $$L_i = F + cQ_i \quad \Rightarrow \quad Q_i = \frac{L_i - F}{c}$$
          where $c < 1$.
        - each producer is a monopolist facing a unit-elastic demand for her good and competition from the traditional sector $\Rightarrow$ *limit pricing*: price set a shade below 1.
- Labor must be paid a premium to move from traditional to modern: w (> 1) is the ratio of modern sector wage over traditional one.

Murphy et al.'s (1989) Big Push model



Source: Krugman, 1997.

Murphy et al.'s (1989) Big Push model

- If all labor is allocated to:
  - the traditional sector:
    - L/N workers are employed in the production of each good.
    - each good has an output $Q_i = L/N$.
  - the modern sector:
    - L/N workers are employed in the production of each good.
    - each good has an output $Q_i = \frac{L/N - F}{c}$
- $\frac{L/N - F}{c} > L/N$ as long as:
  - the marginal cost advantage of modern production is sufficiently large (c is low), or
  - fixed costs (F) are not too large.
- But even if the economy can produce more in the modern sector, it must be profitable for each firm to do so and this depends on:
  - the wage premium (w);
  - the decisions of all the other firms.

Murphy et al.'s (1989) Big Push model

- If an individual firm starts modern production when all other goods are produced using traditional techniques:
  - the firm will charge (almost) the same price as that on other goods and hence sell roughly the same amount (when N is large income effects are negligible) (point A);
  - given the wage premium (w), it is profitable for the firm to move from the traditional to the modern sector iff: $wL_i < Q_i$;
  - in the case shown in the figure, since at point A $\frac{Q_i}{L_i}$ (the slope of a ray connecting the point with the origin) is smaller than w, the move is not profitable.
- On the contrary, if all modern firms start simultaneously leading to production and employment at point B, for each individual firm the move is profitable.

Krugman (1997) on the Big Push model

- As argued by Krugman (1997), "it is easy to critique the plausibility of the assumptions. Yet the model can serve as a useful jumping-off point for thinking about development models."
- The Big Push model is a simple model very useful to understand the potential role of pecuniary external economies for development:
  - no *technological external economies* in the model,
  - it is the interaction between:
    - economies of scale at the plant level
    - elastic supply of factors of production

    that gives rise to de facto external economies.
- The model is simple but very smart: it uses simplifying assumptions and the bag of tricks developed by industrial organization theorists in the 1970s (e.g. limit-pricing monopolists to deal with imperfect competition) to simplify the matter and better emphasize the hidden relations and insights.

Krugman (1997) on the fall of high development theory

Yet neither declining external demand for development economists nor their practical failures fully explain the sputtering out of the field. Purely intellectual problems were also extremely important. In particular, during the years when high development theory flourished, the leading development economists failed to turn their intuitive insights into clear-cut models that could serve as the core of an enduring discipline.

From the point of view of a modern economist, the most striking feature of the works of high development theory is their adherence to a discursive, nonmathematical style. Economics has, of course, become vastly more mathematical over time. Nonetheless, development economics was archaic in style even for its own time. [...]

Some development theorists responded by getting as close to a formal model as they could, ... but others at least professed to see a less formal, less disciplined approach as a virtue. It is in this light that one needs to see Hirschman and Myrdal. [...]

What marked Myrdal and Hirschman was not so much the novelty of their ideas but their stylistic and methodological stance. ... Myrdal and Hirschman ... eventually in effect took stands on principle against any effort to formalize their ideas.

One imagines that this was initially very liberating for them and their followers. Yet in the end it was a vain stance. Economic theory is essentially a collection of models. Broad insights that are not expressed in model form may temporarily attract attention and even win converts, but they do not endure unless codified in a reproducible and teachable form. You may not like this tendency; certainly economists tend to be too quick to dismiss what has not been formalized (although I believe that the focus on models is basically right). Like it or not, however, the influence of ideas that have not been embalmed in models soon decays.

(Krugman, 1997)

General heuristics: search and arguments heuristics

- "The aim of heuristic is to study the methods and rules of discovery and invention" (Polya, 1957, p. 112).
- Abbott (2004) discusses two kinds of *general heuristics*, i.e. heuristics not specifically aimed at any particular phase or aspect of the research process:
  - *search heuristics*: "ways of getting new ideas from elsewhere" (p 113)
    - making an analogy;
    - borrowing a method.
  - *argument heuristics*: "ways of turning old and familiar arguments into new and creative ones" (p. 120)
    - problematizing the obvious;
    - making a reversal;
    - making an assumption;
    - reconceptualizing.
- The role of formalization and mathematical models is to:
  - ensure internal consistency in the process;
  - provide a rigorous analysis of the implications and causal chains.

Search heuristics in Kirman's (1993) model of rational herding

- Kirman (1993) makes an analogy between ants searching for food and boundedly rational agents in markets.
- He can borrow from experiments and models of entomologists to account for the observed patterns of herding behavior in market participants:
- In an apparently symmetric situation (e.g. one food source and two symmetric bridges leading to it), ants behave collectively in an asymmetric way (Deneubourg *et al.*, 1987; Pasteels *et al.*, 1987):
  - the ants exploit one bridge more intensively than the other;
  - from time to time they switch their attention to the bridge that they had previously neglected.
- The entomologists suggest that the asymmetric aggregate behavior arises from the interaction between identical individuals, each of whom acts in a simple way.

Recruitment in ants

- When an ant finds food, it typically "recruits" other ants to that food, by returning to the nest and:
  - physically stimulating, either by contact or by chemical secretion, another ant to follow it to the food source (tandem recruiting);
  - recruiting a group, or laying a chemical trail (by leaving pheromones) which attracts other ants.
- To ensure survival, it is always useful to have some ants searching, even when there is a currently productive source: a percentage of ants searching, sacrificing short-term gains for longer run survival.

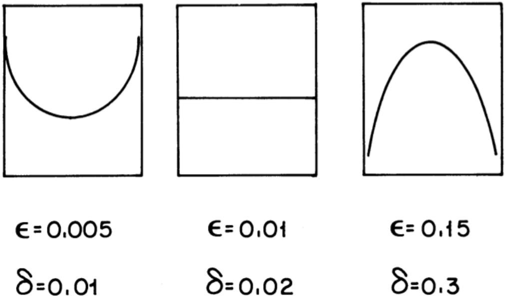
Recruitment in ants modeled as a urn model

- N ants and two sources of food (black and white).
- The state of the system is defined as the number k of ants feeding at the black source: $k \in (0, 1, \ldots, N)$.
- The system evolves as follows:
  - In each period, two ants meet at random, like drawing two balls from an urn containing black and white balls.
  - The first is converted to the second's color with probability $(1-\delta)$;
  - There is also a small probability ϵ that the first will change his own color independently before meeting anyone (to prevent the process getting "stuck" at $k=0$ or $k=N$).
- The dynamic evolution of the process is then given by:

$$k_{t+1} = \begin{cases} k_t + 1 & \text{with probability} \quad p_1 = \left(1 - \frac{k}{N}\right)\left(\epsilon + (1-\delta)\frac{k}{N-1}\right) \\ k_t - 1 & \text{with probability} \quad p_2 = \frac{k}{N}\left(\epsilon + (1-\delta)\frac{N-k}{N-1}\right) \\ k_t & \text{with probability} \quad p_3 = 1 - p_1 - p_2 \end{cases}$$

- This process defines a finite Markov chain related to several standard *urn models*.

Possible equilibrium distributions of the model



Source: Kirman, 1993.

Figure: Equilibrium distributions for the model over state spaces with three different values of ϵ and δ and $N = 100$

From entomology to economics

- Kirman (1993) recaps the principal features of the ant process and see how these correspond to particular aspects of economic models:
  - The process involves agents making different choices (made on the basis of different tastes, different expectations, etc.);
  - Agents "recruit" other agents to their particular choice. The recruitment may take three forms:
    - An agent may persuade another of the superiority of his choice (e.g. for better information);
    - The fact that the first agent makes a particular choice may lead, through externality, to the second agent concurring in that choice;
    - There may be a general externality of the Keynesian "beauty queen" type or resulting from technological spillovers.
  - The process is intrinsically dynamic: it can stay in one "regime" for a considerable period before switching to another.
- The "ant model" (a urn model) has close connections with:
  - Becker's (1991) restaurant example;
  - Foster and Young's (1990) notion of stochastic stability;
  - Arthur (1989) and David (1985) models on path dependence in technological progress;
  - Herding and contagion in financial markets (e.g. Day and Huang,

Argument heuristics in Spence's (1973, 1974) signaling model of education in labor market

- Spence (1973,1974) formally analyzes the possible role of "signaling" of education in labor markets.
- Signaling theory has been applied in several fields: evolutionary biology (Zahavi's handicap principle), finance, industrial organization, etc.
- Spence starts purposely from the strong (and hardly true) assumption that formal education has *no effect* on people's skills and competences.
- We consider the simplest version of the model discussed in Bolton & Dewatripont, 2005, § 3.1, pp. 100-106

Spence's model: assumptions

- Two productivity levels of workers:
  - high (H): r_H;
  - low (L): r_L ($< r_H$).
- Firms pay a wage (w) equal to expected productivity;
- Workers can study e years having a cost:

$$c(e) = \theta_i e \quad i \in \{L, H\}$$

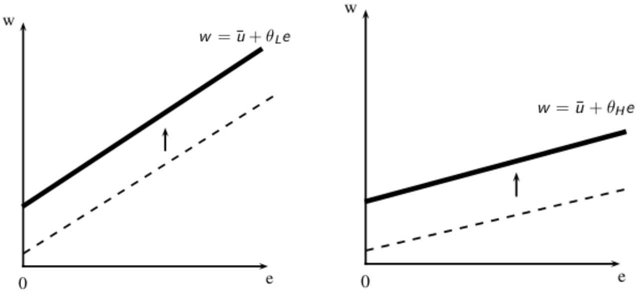
- Formal education:
  - has no effect on productivity;
  - has a lower marginal cost for productive workers:

$$\theta_H < \theta_L$$

Utility functions and indifference curves

- Worker's utility function:

$$u_i(w, e) = w - c(e) = w - \theta_i e \quad i \in \{L, H\}$$



(a) Low-productivity workers (b) High-productivity workers

Figure: Indifference curves

Signaling game

- Perfect information (the hiring firm can observe r_i):
  - nobody invests in formal education ($e_H = e_L = 0$);
  - firms pay each worker according to her productivity ($w_H = r_H > r_L = w_L$);
- Asymmetric information: *signaling game*
  - Two-stage game:
    1. each worker chooses her investment of education (e);
    2. firm pay a wage equal to the expected level of productivity given the observed level of education:

$$w(e) = p(\theta_H|e)\, r_H + (1 - p(\theta_H|e))\, r_L$$

  - An equilibrium of this game can be defined as:
    - for each worker the investment in education e maximizes her own utility given firms' expectation $p(\theta_H|e)$;
    - firms' expectation is *ex post* confirmed.

Equilibria in the signaling game

The game has got mainly (although not exclusively) two kinds of equilibria:

- *Separating equilibria*, where each type of worker (H and L) chooses a different level of education ($e_H \neq e_L$), so that:
  - $p(\theta_H|e_H) = 1$ e $p(\theta_H|e_L) = 0$;
  - $w_H = r_H > r_L = w_L$.
- *Pooling equilibria*, where all workers (H and L) choose the same level of education ($e_H = e_L$), so that:
  - $p(\theta_H|e_H) = p(\theta_H|e_L) = p(\theta_H)$;
  - $w_H = w_L = p(\theta_H)r_H + (1 - p(\theta_H))\, r_L$.

Separating equilibria

- Low-productive workers do not acquire education if the minimum level needed to be wrongly identified as a high-productive worker $\hat{e}$ is such that:

$$r_H - \theta_L \hat{e} \leq r_L \qquad \Rightarrow \qquad \hat{e} \geq \frac{r_H - r_L}{\theta_L} \tag{1}$$

- High-productivity workers acquire education as a "signal" to be identified as long as:

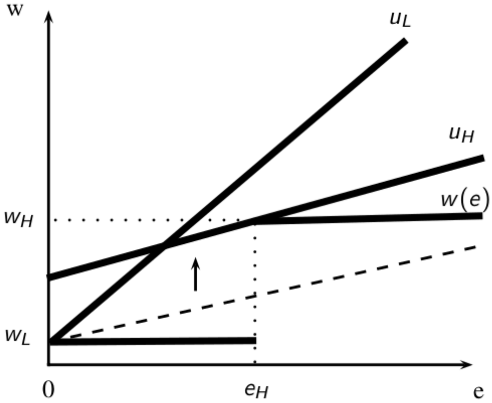
$$r_H - \theta_H \hat{e} \geq r_L \qquad \Rightarrow \qquad \hat{e} \leq \frac{r_H - r_L}{\theta_H} \tag{2}$$

- If conditions 1 and 2 are satisfied and firms believe that:

$$p(\theta_H|e) = \begin{cases} 1 & \text{if } \; e \geq \hat{e} \\ 0 & \text{otherwise} \end{cases}$$

then $(e_L, e_H) = (0, \hat{e})$ is a separating equilibrium.

Separating equilibrium



Pooling equilibria

- Low-productivity workers acquire education if:

$$w - \theta_L \hat{e} \geq r_L \qquad \Rightarrow \qquad \hat{e} \leq \frac{w - r_L}{\theta_L} \tag{1}$$

- If condition (1) is satisfied and firms believe that:

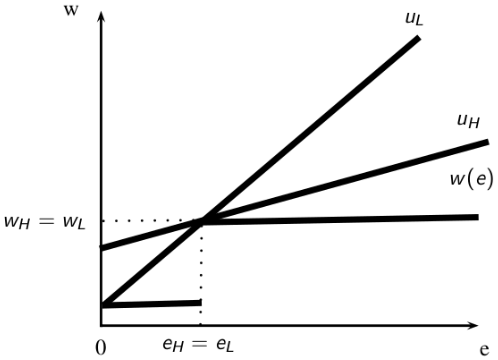
$$p(\theta_H | e) = \begin{cases} p(\theta_H) & \text{if } \; e \geq \hat{e} \\ 0 & \text{otherwise} \end{cases}$$

then $(e_L, e_H) = (\hat{e}, \hat{e})$ is a pooling equilibrium with:

$$w_L = w_H = p(\theta_H) r_H + (1 - p(\theta_H))\, r_L$$

$$0 \leq \hat{e} \leq p(\theta_H) \frac{r_H - r_L}{\theta_L}$$

Pooling equilibrium



Argument heuristics in Schelling's (1969, 1971, 1978) segregation model

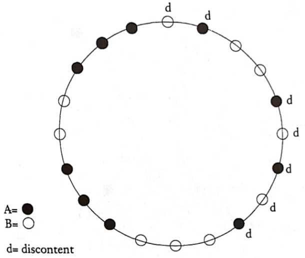
- The Schelling's (1969, 1971, 1978) model of segregation effectively shows how very simple formal models can be extremely powerful in "problematizing the obvious":
  - Racism is a straightforward and apparently convincing explanation of the emergence of racially divided neighborhoods in cities.
  - By using a very simple model, Schelling actually shows that this is not necessary case: even very small preferences among otherwise civic individuals can lead to segregation.
- The original model is an *agent-based model*, i.e. a class of models that study, via simulation, global patterns that emerge from the interaction between autonomous "agents" acting according to simple rules, that might also involve learning or randomness.
- In Schelling's (1998) words, micro-motives lead to macro-behavior.

Young's (1998) variant of Schelling's segregation model

- Schelling's original model took place on a checkerboard with two different types of agents. He studied the equilibrium outcomes of the game by means of simulation.
- Young (1998) considers a variant of Schelling model that takes place on a circle and studies the equilibrium outcomes of the game using the notion of stochastic stability (Foster and Young, 1990).
- Loosely speaking, given an evolutionary process subjected to small but persistent random shocks, a *stochastically stable state* of the process is the state that occurs more frequently than the others over the long run.

Young's (1998) variant of Schelling's segregation model

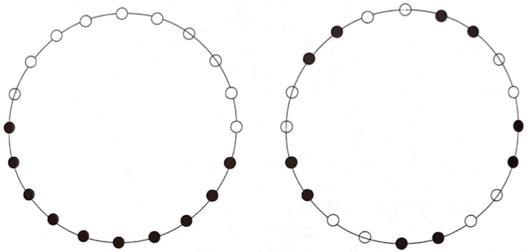
- Two types of agents – As and Bs – choose where to locate on a circle (there are at least two agents of each type).
- Their utility ("happiness" in Schelling's words) for a given location depends on the composition of their neighborhoods.
- Let us suppose that an individual is:
  - discontent if her two immediate neighbors are unlike herself;
  - otherwise she is content.



Source: Young, 1998.

Young's (1998) variant of Schelling's segregation model

- An *equilibrium state* is one in which there is no pair of agents such that at least one is currently discontent and both would be content after they trade locations.



Source: Young, 1998.

Figure: Two equilibrium configurations: one "segregated" and one "integrated"

Young's (1998) variant of Schelling's segregation model

- Initial random distribution of the two types of agents on the circle.
- In each period, one pair of agents is selected uniformly at random and they are allowed to trade their respective locations at a cost.
- The probability that the two agents trade depends on their gains from trade (positive gains only if the agents are of opposite types and at least one of them was discontent before the trade and content afterwards).
    - There exist real numbers $0 < a < b < c$ such that the probability of a trade with no positive gains is:
        - ϵ^a if one partner's increase in contentment is offset by the other decrease in contentment (the net losses involve only moving costs);
        - ϵ^b if both partners were content before and one is discontent after;
        - ϵ^c if both partners were content before and both are discontent after;
    - A trade with positive gains occurs with probability: $1 - \epsilon^a - \epsilon^b - \epsilon^c$.
- The resulting process P^ϵ is an *irreducible Markov process* as long as $\epsilon > 0$.

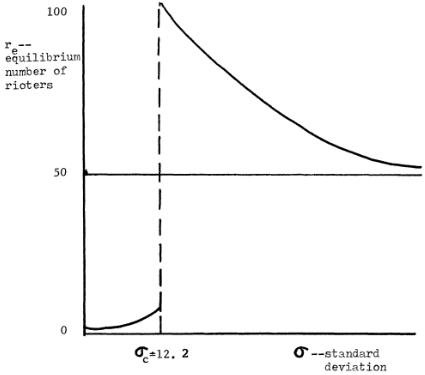
Absorbing states and stochastically stable states

- An *absorbing state* of P^{ϵ} is an equilibrium configuration of the process P^0, where nobody is discontent
  NB: with $\epsilon = 0$ the Markov process is reducible and therefore nonergodic.
- The *stochastically stable state* is the absorbing state that has the associated highest probability in the *stationary distribution* of the Markov process P^{ϵ} ($\epsilon > 0$).
- Young (1999, p. 62-65) shows that the segregated states (configuration where the As form a continuous group and the Bs therefore form the complementary group) are stochastically stable.

Argument heuristics on Granovetter's (1978) threshold models of collective behavior

- Granovetter's (1978) threshold models of collective behavior effectively show that the representative-agent approach for the analysis of group behavior can be strongly misleading.
- The example used by Granovetter to illustrate his model is whether or not a given population of individuals would take part in a riot.
- Assumptions:
  - Agents make a binary decision (whether or not to join the riot);
  - Individual decision influenced by the others (the more the others join in the riot, the less "costly" it is to do the same).
  - Each agent has a different threshold.
- Imagine two towns, each with a population of 100:
  - In the first town, individual thresholds are uniformly distributed between 0 and 99 $\Rightarrow$ the agent with a threshold of 0 starts the riot and produces a domino effect that leads to widespread unrest.
  - In the second town, the thresholds of the population are the same, but the agent with threshold 1 is missing $\Rightarrow$ no domino effect and the instigator is a "lone vandal".

Heterogeneity and riots



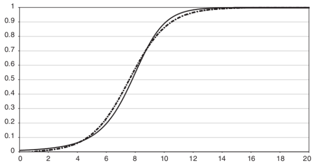
Source: Granovetter, 1978.

Figure: Equilibrium number of rioters plotted against standard deviation of normal distributions of thresholds (mean = 25, N = 100)

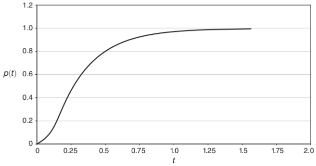
Diffusion and adoption in heterogeneous populations with homogeneous mixing

- Young (2009) distinguishes three broad classes of diffusion and adoption mechanisms:
  - *contagion*: agents adopt when they come in contact with other adopters ⇒ Epidemic models.
  - *social influence*: people adopt when enough other people in the group have adopted ⇒ Threshold models.
  - *social learning*: people adopt once they see enough empirical evidence to convince them that the innovation is worth adopting ⇒ Models of Bayesian learning.
- By maintaining the assumption of *homogeneous mixing*, Young (2009) formally derives the patterns of diffusion/adoption in the three "mechanisms" when agents are heterogeneous.

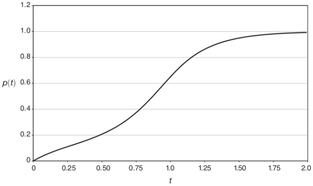
Diffusion and adoption in heterogeneous populations with homogeneous mixing



(a) Contagion



(b) Social influence (threshold)



(c) Social (Bayesian) learning

The role of network structure in diffusion and adoption

- By relaxing the homogeneity assumption one can analyze the role of network structure on the speed and the degree of diffusion/adoption (Watts, 2002):
    - In *small-world networks* the threshold is significantly reduced (Zanette, 2001);
    - In *scale-free networks* the critical point approaches to 0 (Pastor-Sartorras & Vespignani, 2001, 2002).
- Network-based approaches to model innovation diffusion:
    - Percolation models in:
        - regular lattices (Ising models) (e.g. Allen, 1982; Mort, 1991; David & Foray, 1994; Solomon *et al.*, 2000; Goldenberg *et al.*, 2000; Hohnisch *et al.*, 2008).
        - complex networks (e.g. Silverberg & Verspagen, 2005).
    - Agent-based models of technology diffusion in complex networks (e.g. Delre, Jager & Janssen, 2007; Lee, Lee & Lee, 2006; Uchida & Shirayama, 2008; Pegoretti, Rentocchini & Vittucci Marzetti, 2012).

Arthur's (1994) "El Farol Bar" model

- Arthur (1994) starts from the results of modern psychology that as humans, in complicated or ill-defined situations, we do not use deductive logic, but "we look for patterns and simplify ... the problem by using these to construct temporary internal models or hypotheses or schemata to work with" (p. 406).
- How to model such "inductive reasoning"?
- "El Farol Bar" problem:
  - N $(= 100)$ people decide independently each week whether to go to a bar in Santa Fe that offers Irish music on Thursday nights;
  - The evening is enjoyable if less than 60 of the possible 100 are present, therefore an agent:
    - goes if she expects smaller than 60 to show up;
    - stays home otherwise.
  - Choices are unaffected by previous visits.
  - No collusion or prior communication among the agents, and the only information available is the numbers who came in past weeks.

Arthur's (1994) "El Farol Bar" model

- From the agents' viewpoint, the problem is ill-defined and they are propelled into inductive reasoning:
  - no deductively rational solution, no "correct" expectational model;
  - any commonalty of expectations gets broken up:
    - if all believe few will go, all will go, invalidating the belief.
    - if all believe most will go, nobody will go, invalidating the belief.
- Each agent:
  - possesses and keeps track of an individualized subset of k focal predictors, i.e. rules/hypothesis in the form of functions that map the past d weeks' attendance figures into next week's (e.g. predict next week's number to be. the same as two weeks ago; etc.).
  - decides to go or stay according to the currently most accurate predictor in his set (her active predictor);
  - learns the new attendance figure and updates the accuracies of her predictors.

Arthur's (1994) "El Farol Bar" model

- To analyze how attendance behave dynamically over time, Arthur (1994) uses computer simulation.
- Results show that:
  - cycle-detector predictors are quickly "arbitraged" away, so there are no persistent cycles;
  - mean attendance converges always to 60.
- In Arthur's words, "the predictors self-organize into an equilibrium pattern or ecology".
- The bar problem has been generalized along several directions (see, for instance, Challet, 2004, Minority games).

Costs and risks of mathematical models in social science

- Blind spots:

  The strategic omissions involved in building a model almost always involve throwing away some real information. ... And yet once you have a model, it is essentially impossible to avoid seeing the world in terms of that model, which means focusing on the forces and effects your model can represent and ignoring or giving short shrift to those it cannot. The result is that the very act of modeling has the effect of destroying knowledge as well as creating it. A successful model enhances our vision, but it also creates blind spots, at least at first.

  The cycle of knowledge lost before it can be regained seems to be an inevitable part of formal model-building.

  (Krugman, 1997)

- Research agenda dictated by "mathematical resistance":

  ...the theory of international trade followed the perceived line of least mathematical resistance.

  (Krugman, 1990, p.4)

On common objections of using mathematical models in social science

When it comes to physical science, few people have problems with the idea that to study complex systems it is necessary to build simplified models. When we turn to social science, however, the whole issue of modeling begins to raise people's hackles. Suddenly the idea of representing the relevant system through a set of simplifications that are dictated at least in part by the available techniques becomes highly objectionable. Everyone accepts that it was reasonable for meteorologists to represent the earth, at least for a first pass, with a flat dish, because that was what was practical. But what do you think about the decision of most economists between 1820 and 1970 to represent the economy as a set of perfectly competitive markets, because a model of perfect competition was what they knew how to build? It's essentially the same thing, but it raises howls of indignation.

Why is our attitude so different when we come to social science?

(Krugman, 1997)

On the objections of using formal models in social science

Many of those who reject the idea of economic models are ill-informed or even ... intellectually dishonest. Still, there are highly intelligent and objective thinkers who are repelled by simplistic models for a much better reason: they are very aware that the act of building a model involves loss as well as gain.
Model-building, especially in its early stages, involves the evolution of ignorance as well as knowledge; someone with powerful intuition and a sense of the complexities of reality, may well feel that from his point of view more is lost than is gained.
The problem is that there is no alternative to models. We all think in simplified models, all the time. The sophisticated thing to do is not to pretend to stop, but to be self-conscious to be aware that your models are maps rather than reality.
There are many intelligent writers ... who are able to convince themselves and ... other people ... that they have found a way to transcend the narrowing effect of model-building. Invariably they are fooling themselves. If you look at the writing of anyone who claims to be able to write about social issues without stooping to restrictive modeling, you will find that his insights are based essentially on the use of metaphor. And metaphor is, of course, a kind of heuristic modeling technique. [...]
For the most part ... thinkers who imagine that they have broadened their vision by abandoning the effort to make simple models have done no such thing. All that they have really done is to use high-flown rhetoric to disguise, not least from themselves, their lack of clear understanding.
So modeling, which may seem simplistic, is in practice often a discipline that helps you avoid being even more simplistic. But there is more: a formal model, which may seem like a ridiculously stylized sketch of reality, will often suggest things that you would never think of otherwise.

(Krugman, 1997)

What the Big Push tale can teach us

By the late 1950s, ... high development theory was in a difficult position. Mainstream economics was moving in the direction of increasingly formal and careful modeling. ... But it was difficult to model high development theory more formally, because of the problem of dealing with market structure.

The response of some of the most brilliant high development theorists, above all Albert Hirschman, was simply to opt out of the mainstream. They would build a new development school on suggestive metaphors, institutional realism, interdisciplinary reasoning, and a relaxed attitude toward internal consistency. The result was some wonderful writing, some inspiring insights, and (in my view) an intellectual dead end.

[...] And yet in the end it turned out that mainstream economics eventually did find a place for high development theory. ... And it was not simply a matter of rediscovery: the restatement of high development theory, in terms of such models as the Murphy et al. version of the Big Push, is not only clearer but in some ways deeper than the original statement.

(Krugman, 1997)

What the Big Push tale can teach us

The Murphy et al. (1989) Big Push model:

1. *shows that it is possible to tell high development-style stories in the form of a rigorous model.*
2. *shows that the essential logic ... emerges even in a highly simplified setting.*
   It is common for those who haven't tried the exercise of making a model to assert that underdevelopment traps must necessarily result from some complicated set of factors, irrationality or shortsightedness on the part of investors, cultural barriers to change, inadequate capital markets, problems of information and learning, and so on. Perhaps these factors play a role, perhaps they don't: what we now know is that a low-level trap can arise ... without [them].
3. *unlike a purely verbal exposition, reveals the sensitivity of the conclusions to the assumptions.*
   In particular, verbal expositions of the Big Push story make it seem like something that must be true. In models we see that it is something that might be true. A model makes one want to go out and start measuring, to see whether it looks at all likely in practice, whereas a merely rhetorical presentation gives one a false feeling of security in one's understanding.

What the Big Push tale can teach us

4. *tells us something about what attitude is required to deal with complex issues in economics.*
Little models like the Murphy et al. Big Push may seem childishly simple, but I can report from observation that until they published their formalization of Rosenstein-Rodan, its conclusions were not obvious to many people, even those who have specialized in development. Economists tended to regard the Big Push story as essentially nonsensical: if modern technology is better, then rational firms would simply adopt it! (They missed the interaction between economies of scale and market size.) Non-economists tended to think that Big Push stories necessarily involved some rich interdisciplinary stew of effects, missing the simple core. In other words, economists were locked in their traditional models, non-economists were lost in the fog that results when you have no explicit models at all.

How did Murphy et al. break through this wall of confusion? Not by trying to capture the richness of reality, either with a highly complex model or with the kind of lovely metaphors that seem to evade the need for a model. They did it instead by daring to be silly ... to get at an essential point.

(Krugman, 1997)

All that said...

...yes, there's a lot of excessive and/or misused math in economics; plus the habit of thinking only in terms of what you can model creates blind spots ... So yes, let's critique the excessive math, and fight the tendency to equate hard math with quality. But in the course of various projects, I've seen quite a lot of what economics without math and models looks like – and it's not good.

(Krugman, 2013, More economath, NYT)

I am a strong believer in the importance of models, which are to our minds what spear-throwers were to stone age arms: they greatly extend the power and range of our insight. In particular, I have no sympathy for those people who criticize the unrealistic simplifications of model-builders, and imagine that they achieve greater sophistication by avoiding stating their assumptions clearly. The point is to realize that economic models are metaphors, not truth. By all means express your thoughts in models, as pretty as possible (more on that below). But always remember that you may have gotten the metaphor wrong, and that someone else with a different metaphor may be seeing something that you are missing.

(Krugman, 1992, How I work)

Everything should be made as simple as possible, but not simpler

Fortunately, there is a strategy that does double duty: it both helps you keep control of your own insights, and makes those insights accessible to others. The strategy is: always try to express your ideas in the simplest possible model. The act of stripping down to this minimalist model will force you to get to the essence of what you are trying to say (and will also make obvious to you those situations in which you actually have nothing to say). And this minimalist model will then be easy to explain to other economists as well.

(Krugman, 1992, How I work)

It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience.

(Einstein, A., 1934, "On the Method of Theoretical Physics" The Herbert Spencer Lecture)