Title

Identification Strategies for Models of Innovation,R&D, and Productivity

Author

Sanchez, Juana

Publication Date

2015-08-10

Identification Strategies for Models of Innovation, R&D, and Productivity

Juana Sanchez[1]
University of California Los Angeles
jsanchez@stat.ucla.edu

Presented at the Joint Statistical Meetings of the American Statistical Association
Seattle, August 11, 10:00 AM, 2015

Session 176

[1] The research in this paper was conducted while the author was Special Sworn Status researcher of the U.S. Census Bureau at the Center for Economic Studies. Research results and conclusions expressed are those of the author and do not necessarily reflect the views of the Census Bureau. This paper has been screened to insure that no confidential data are revealed.

Summary

1. Introduction
   - R&D and Innovation
   - Economics of innovation
   - Innovation Policy
2. Challenges
   - Many causal assumptions, no counterfactuals
   - Lack of innovation and innovation sale data
   - Innovation data gives mixed messages
3. R&D as a mediating agent
   - Multiple equation approaches
   - Model the causal mechanisms with graphical models
4. Conclusions

R&D

Facts

- Innovative capacity of a country measured almost exclusively by firms' R&D outlays (somewhat over 2.8% of GDP in the U.S.A. in 2010)
- R&D outlays are expenditures on innovation activities which, accumulated, are believed to create the stock of knowledge.
- BRDIS and other surveys reveal that
  - there are firms without R&D outlays that innovate.
  - there are companies with R&D outlays that do not innovate.

R&D (Innovative activities) do not always result in innovation

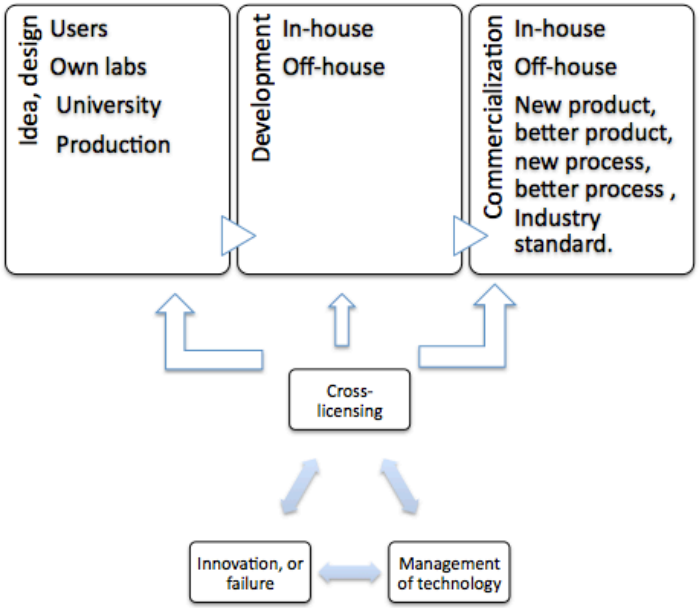


Figure 1 : Innovative activities

Definition

An innovation is the **implementation** of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method in business practices, workplace organization or external relations.

Theoretical Framework

Knowledge production function New knowledge (K) depends on current and past investment in new knowledge (e.g. current and past R&D expenditures) and knowledge flows from outside the firm.

Output production function The augmented Cobb-Douglas production function.

$$Y = AL^{\alpha}C^{\beta}[K]^{\lambda}[K^{0}]^{\phi}e^{\mu}$$

Y= total production (usually total output is used instead of output derived from innovation only).

Different operationalizations, different conclusions: fifty years of empirical research.

Economics of innovation: focus on determinants of R&D

Examples

- Appropriability of returns
- Market demand conditions
- Firm characteristics: size, market power, diversification
- Market structure
- Management of technology (management literature)
- Financial constraints on the decision of companies to invest in R&D and to patent.

R&D tax credits in the U.S.

- Science and innovation policy under the umbrella of the America COMPETES Act: permanent R&D tax credit.
- R&D tax credit: market-based tool aimed at reducing the marginal cost of R&D activities
- They allow firms to decide which R&D projects to fund.
- Critics:
  - Innate ability of a company provides returns to R&D, not R&D tax credits (management literature)
  - Firms' elasticities varying would justify a more targeted innovation policy.
  - Credits favor the process of catching up of firms lagging behind the technological frontier rather than pushing the country's frontier further
  - Perhaps an improvement in open innovation mechanisms or other policies would be more effective.

Challenges in evaluating determinants and effects of R&D

Methodological challenges

- Third variables
- Is total sales or output the appropriate outcome or effect to look at?
- Estimating the counterfactual: what would have happened without the policy or factor?
- Self-selection

Different data collection methods

- Researchers must conduct their own data collection to obtain innovation data. Small studies.
- The Business Research and Development and Innovation Survey (since 2009) asks companies
  - whether they actually innovated or not.
  - whether they actually obtained sales from their innovations and how much (self-reported)
  - whether the innovation are new to the market or new to the firm.

Innovation sales in the U.S. do not follow R&D

Table 1 : Percent of sales due to “new-to-market (ntm)“ and “new-to-company“ (ntc) innovations. Weighted. Source: BRDIS 2009-2011. Preliminary results.

R&D Status	Sector	N	Variable	Average % sales	sd
Not active	Service	4700	% sales due to ntm	13	29
			% sales due to ntc	14	27
Not active	Manufacturing	4600	% sales due to ntm	10	23
			% sales due to ntc	12	23
Active	Service	6500	% sales due to ntm	24	39
			% sales due to ntc	16	29
Active	Manufacturing (Have largest R&D outlays)	11100	% sales due to ntm	13	24
			% sales due to ntc	12	22

SUR models. Correlated errors in multiple equations

Dependent variables : log Research, log Development, log innovation sales, log productivity.

Independent variables : lagged factors mentioned earlier, lagged values of the dependent variables to avoid reverse causality and endogeneity that could arise due to heterogeneity of omitted variables

Fixed effects : 11 industry fixed effects and year.

SUR models

Table 2 : SUR3. Four Seemingly Unrelated Regressions. The model fits also 11 industry groups but those are not reported here.

Sample size used for computations=2200 (rounded) (those observations with complete data). System weighted R-square=0.67.

Preliminary results. (In parenthesis: standardized coefficients. Source: BRDIS 2008-2011

Lagged Independent Variable	Research	Development	Innovation sales	Productivity.
lagsize1		0.42(0.52)*		0.57(1.26)*
lagsize1sq		−0.03(−0.48)*		−0.04(−1.28)*
lagikht		0.07(0.02)*	0.18(0.07)*	
lagcapital		−0.14(−0.24)*		−0.05(−0.17)
lagcapital2		0.01(0.21)*		0.005(0.26)*
lagdemand	0.10(0.03)*			
lagtechop1	0.16(0.04)*			
lagmapprop	0.46(0.02)			
laglrrdint	0.76(0.79)*		0.06(0.07)*	
lagldrdint		0.8(0.90)*	0.14(0.17)*	
lagsales	−0.13(−0.07)	−0.16(−0.09)*	−0.21(−0.15)*	0.59(0.64)*
group			0.21(0.08)*	
laginovps			0.25(0.05)*	
lagmarkpower				0.34(0.96)*
lagmarkpower2				0.03(1)*

Structural Equation Models

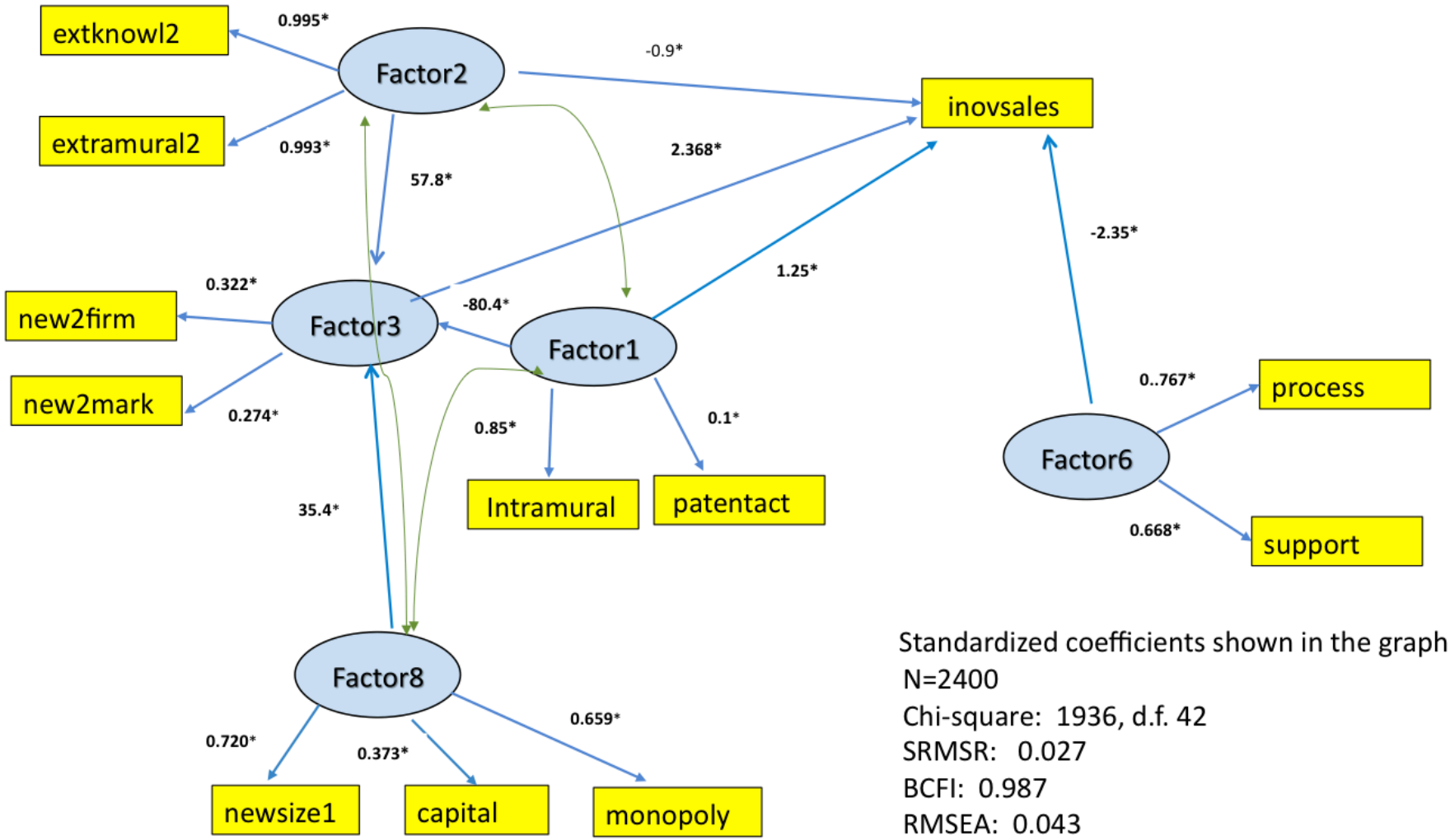


Figure 2 : Structural Equation Model. Preliminary results.Data source: BRDIS

Direct and Indirect Effects in SEM

Table 3 : Structural equation modeling (N=24000, rounded). Preliminary results. Data source:BRDIS 2008-2011. Standardized direct and indirect effects.

Variables	Total	Direct	Indirect
factor1 on inovsales	−189.228	1.253	−190.481
factor2 on inovsales	135.952	−0.900	136.852
factor8 on inovsales	83.837	0	83.837

Advantages and disadvantages of SEM

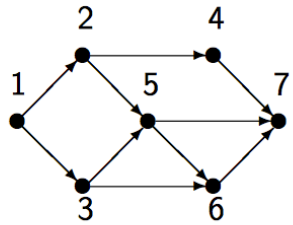
Estimating covariance matrix in SEM: can do available-case analysis (advantage)

Restrictions in SEM: by excluding some variables in the models we state that their effect is 0 on the other variables. Explicit causal assumptions. (advantage)

Functional form: must specify a functional form (disadvantage).

Directed acyclic Graphs. Think deeper about the mechanisms.

- Causal interpretations in terms of conditional probabilities.
- Distinguishes between observed and caused.
- No functional form required.



$$p(x \,||\, x_5^*) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_1)p(x_4 \mid x_2) \times p(x_6 \mid x_3, x_5^*)p(x_7 \mid x_4, x_5^*, x_6)$$

whereas

$$p(x \mid x_5^*) \propto p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_1)p(x_4 \mid x_2) \times p(x_5^* \mid x_2, x_3)p(x_6 \mid x_3, x_5^*)p(x_7 \mid x_4, x_5^*, x_6)$$

Figure 3 : Directed Acyclic Graph (source: Lauritzen, 2001)

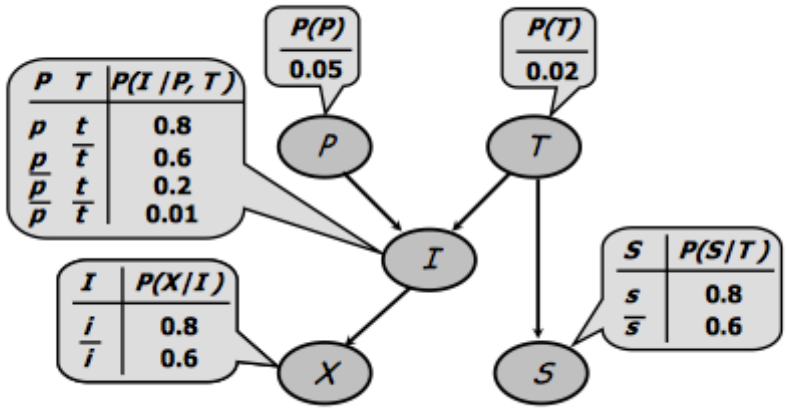


Figure 4 : Directed Acyclic Graph (source: Koller et al. 2009)

Main components

- Markovian causal theory
- $$P(x_1, x_2, x_3, ..., x_k) = \prod_{i=1}^{k} P(x_i \mid Pa_i)$$
- Given its parent Pa_i, each variable is conditionally independent of all its predecessors (Markovian independency).

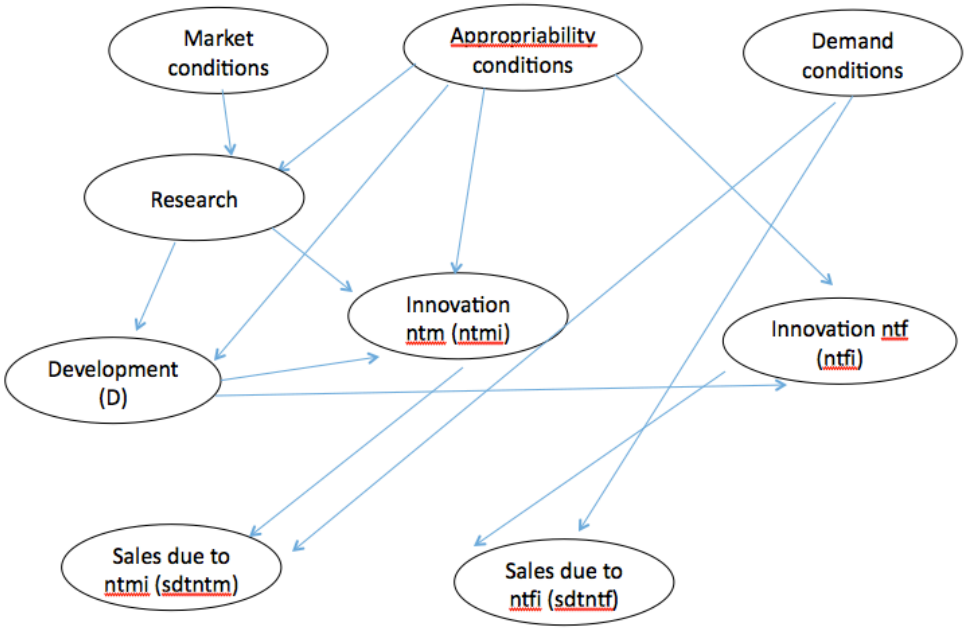


Figure 5 : Directed Acyclic Graph

- Multiple equation models that have as outcome innovation per se and innovation sales are more promising to study the innovativeness of U.S. companies than models that proxy innovation with R&D and R&D returns with total productivity.
- Multiple equation models (SEM and graphical models) allow the separation of direct and indirect (via R&D) of variables on innovation sales.
- Future work that accounts for the effect in missing data and improvements in the collection of innovation data will help clarify the complexity of the associations between R&D, innovation and innovation sales.