Innovation process: an analysis of European CIS data

Authors: Océane Vernerey and Jimmy Lopez

Preliminary results

Université de Bourgogne (UB)
Laboratoire d’Economie de Dijon (LEDi)
(ANR-19-CE26-0008-01)

Mannheim 16-17 March 2023
Table of contents

1. Introduction
   - Introduction and motivation
   - Related literature

2. Specification of our Model
   - Intuition
   - Model and estimation

3. Data
   - Sample and variables
   - Descriptives Statistics

4. Empirical results

5. Conclusion and extensions
→ **Innovation** : plays a crucial role as an economic engine.
→ **Innovation process** : studying the link between input and output innovations

A lot of studies use the **Community Innovation Survey (CIS)**:
- **Main source** of statistics on business innovation in Europe:
  - to study the determinants of innovation,
  - the complementarity between the types of innovation,
  - the environmental benefit...

- **Currently available on a broad time scale**:
  - Ranging from 1998 to 2018.
  - Firms in each member State of the European Union.
Our contribution consists of a version of the CDM model (Crépon, Duguet, Mairesse 1998).

Many studies use the CDM model: but this work concerns a single country, or a comparison between two countries or between two waves of surveys...

→ We go further by benefiting from 8 waves and 9 countries.
  
  - to compare between countries and over time.
  - We have countries that have not been much studied in the literature.
Related literature

Crépon, Duguet, Mairesse (1998) : CIS0 (France)
- Use the number of patents or sales of innovative products.
- Estimation : Asymptotique Least Squares.
- Positive effect of R&D on innovation and of innovation on productivity.

Mairesse Mohnen Kremp (2005) : CIS3 (France)
- Use sales of innovative products (continuous measure).
- Estimation : two generalized Tobit.
- Positive effect of R&D on innovation and of innovation on productivity.

Parisi et al (2005) : Italian manufacturing firm
- Use product and process innovation (discret measure)
- Estimation : Bivariate Probit
- Positive effect of R&D on product innovation (not on process innovation).
- Positive effect of Process innovation on productivity.
Related literature

- Use product and process innovation (discrete measure).
- Bivariate Probit.
- Positive effect of R&D on product and process innovation.
- Positive effect of product innovation on productivity but not process innovation.

Hajjem, Garrouste, Ayadi (2015) : CIS6 (France)
- Use product, process, marketing and organisational innovation (discret).
- Estimation : Multivariate Probit.
- Positive effect of R&D on all of types of innovation.
- Positive effect of marketing and organisational innovation on productivity.
- Negative effect of product innovation on productivity.

→ From this overview, we see that there is degree of heterogeneity in the findings according to specification.
Intuition of the CDM model

CDM = Crépon, Duguet, Mairesse (1998)

⇒ original empirical approach in order to clarify the stages of the innovation process.

STEP 1: Firms invest in R&D.
STEP 2: R&D expenditures generate (or not) innovations
STEP 3: Innovative firms improve their productivity.

Research return is therefore the result of two processes:

- a research-innovation link,
- an innovation-productivity link.
Model

\[
\begin{align*}
(1) \quad R_{d_i} &= \sum_k x_{1i}^k b_1^k + u_{1i} \\
(2) \quad \ln(R_i) &= \sum_k x_{2i}^k b_2^k + u_{2i} \\
(3) \quad I_{d_i} &= R_{d_i} \beta_3 + \sum_k x_{3i}^k b_3^k + u_{3i} \\
(4) \quad \ln(I_i) &= \ln(R_i) \gamma_4 + \sum_k x_{4i}^k b_4^k + u_{4i} \\
(5) \quad \ln(Y_i) &= \ln(I_i) \gamma_5 + \sum_k x_{5i}^k b_5^k + u_{5i}
\end{align*}
\]

- \( R_{d_i} \) and \( \ln(R_i) \) are respectively probability and intensity of R&D.
- \( I_{d_i} \) and \( \ln(I_i) \) are respectively probability and intensity of innovation.
- \( \ln(Y_i) \) is the performance of the firm \( i \).
- \( x_{.i} \) five vectors of \( k \) covariates.
- \( b_., \beta_., \) and \( \gamma_., \) the parameters, and \( u_{.i} \) five error terms.

→ Estimation by maximum likelihood:
- two Type II tobit models \([eq(1-2) and eq(3-4)]\),
- and one linear regression for \([eq(5)]\).
Sample and variables

Use the **European Community Innovation Survey** (CIS) by Eurostat:
- Harmonized survey that covers firms in each EU Member State in several waves (≈ three years).

**Our sample** (500 000 firms):
- 8 waves from 1998 to 2016, and 9 countries: (Bulgaria, Czech Republic, Estonia, Spain, Hungary, Lithuania, Portugal, Romania, Slovakia)

**Our dependant variable**:
- \( \ln(R_i) \rightarrow \) R&D intensity: the logarithm of the ratio of continuous R&D expenditure to the turnover.
- \( \ln(I_i) \rightarrow \) Innovation intensity: the logarithm of the share of sale of new product in turnover.
- \( \ln(Y_i) \rightarrow \) Firm’s performance: approximated by the growth of turnover over the period.

**Covariates**: size of firm, share of market, group, cooperation and fixed effects...
Selection extend:

- CIS3 contains the largest share of firms without innovation activity (71.22%).
- Other CIS waves quite similar (contain on average 60% of firms without innovation activity).

Innovation sample:

- Innovative firms are more likely to be part of a group (39.5% versus 21.3%).
- Innovative firms are more engaged in R&D (52% versus 5%).
- Innovative firms cooperate more (38.1% versus 4%).
- Innovative firms sell more on the international market (67.3% versus 40.1%).
  - Less difference for the local and national market.
## Main results

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) R&amp;D (proba)</th>
<th>(2) R&amp;D (log)</th>
<th>(3) Innov (proba)</th>
<th>(4) Innov (log)</th>
<th>(5) Perform (log)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SAMPLE</strong></td>
<td>Full</td>
<td>R&amp;D&gt;0</td>
<td>Full</td>
<td>Innov&gt;0</td>
<td>Full</td>
</tr>
<tr>
<td>[50;249] employees (dummy)</td>
<td>0.005</td>
<td>-0.19*</td>
<td>0.135***</td>
<td>-0.071***</td>
<td>0.019***</td>
</tr>
<tr>
<td></td>
<td>[0.0074]</td>
<td>[0.0102]</td>
<td>[0.0056]</td>
<td>[0.0157]</td>
<td>[0.0029]</td>
</tr>
<tr>
<td>&gt; 250 employees (dummy)</td>
<td>0.115***</td>
<td>0.170***</td>
<td>0.227***</td>
<td>-0.099***</td>
<td>0.021***</td>
</tr>
<tr>
<td></td>
<td>[0.0115]</td>
<td>[0.0160]</td>
<td>[0.0086]</td>
<td>[0.0226]</td>
<td>[0.0041]</td>
</tr>
<tr>
<td>Group (dummy)</td>
<td>0.117***</td>
<td>0.113***</td>
<td>0.088***</td>
<td>0.010</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>[0.0067]</td>
<td>[0.0085]</td>
<td>[0.0060]</td>
<td>[0.0131]</td>
<td>[0.0029]</td>
</tr>
<tr>
<td>Market share (log)</td>
<td>0.129***</td>
<td>-0.317***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.0026]</td>
<td>[0.0039]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National market (dummy)</td>
<td>0.455***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.0111]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>International market (dummy)</td>
<td>0.927***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.0109]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooperation (dummy)</td>
<td></td>
<td></td>
<td></td>
<td>1.103***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.0071]</td>
<td></td>
</tr>
<tr>
<td>R&amp;D probability (dummy)</td>
<td></td>
<td></td>
<td></td>
<td>1.529***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.0175]</td>
<td></td>
</tr>
<tr>
<td>R&amp;D intensity (log)</td>
<td></td>
<td></td>
<td></td>
<td>0.226***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.0210]</td>
<td></td>
</tr>
<tr>
<td>Innovation intensity (log)</td>
<td></td>
<td></td>
<td></td>
<td>0.024***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.0014]</td>
<td></td>
</tr>
</tbody>
</table>

Observations: 504,660

Wald chi²(443) = 205484.42 | Prob > chi² = 0.0000

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1
### Results by countries and by waves

<table>
<thead>
<tr>
<th>COUNTRY</th>
<th>BG</th>
<th>CZ</th>
<th>EE</th>
<th>ES</th>
<th>HU</th>
<th>LT</th>
<th>PT</th>
<th>SK*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R&amp;D intensity (log)</strong>&lt;br&gt;(on innovation intensity (log))</td>
<td>-0.105</td>
<td>0.512***</td>
<td>0.417***</td>
<td>0.152***</td>
<td>0.347</td>
<td>0.913***</td>
<td>0.437***</td>
<td>0.418***</td>
</tr>
<tr>
<td></td>
<td>[0.1293]</td>
<td>[0.06821]</td>
<td>[0.2109]</td>
<td>[0.0229]</td>
<td>[0.2478]</td>
<td>[0.1495]</td>
<td>[0.0080]</td>
<td>[0.1905]</td>
</tr>
<tr>
<td><strong>Innovation intensity (log)</strong>&lt;br&gt;(on performance (log))</td>
<td>0.026***</td>
<td>0.028***</td>
<td>0.054***</td>
<td>0.012***</td>
<td>0.037***</td>
<td>0.023***</td>
<td>0.032***</td>
<td>0.028***</td>
</tr>
<tr>
<td></td>
<td>[0.0043]</td>
<td>[0.0038]</td>
<td>[0.0078]</td>
<td>[0.0021]</td>
<td>[0.0054]</td>
<td>[0.0086]</td>
<td>[0.0035]</td>
<td>[0.0095]</td>
</tr>
<tr>
<td>Observations</td>
<td>98 518</td>
<td>41 278</td>
<td>12 441</td>
<td>197 244</td>
<td>31 486</td>
<td>14 110</td>
<td>42 947</td>
<td>14 945</td>
</tr>
</tbody>
</table>

With waves and sectors fixed effects  
Standard errors in brackets  
*** p<0.01, ** p<0.05, * p<0.1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R&amp;D intensity (log)</strong>&lt;br&gt;(on innovation intensity (log))</td>
<td>0.138</td>
<td>0.236***</td>
<td>0.284***</td>
<td>0.275***</td>
<td>0.192***</td>
<td>0.145***</td>
<td>0.203***</td>
<td>0.241***</td>
</tr>
<tr>
<td></td>
<td>[0.1173]</td>
<td>[0.0699]</td>
<td>[0.0643]</td>
<td>[0.0558]</td>
<td>[0.0492]</td>
<td>[0.0552]</td>
<td>[0.0549]</td>
<td>[0.0546]</td>
</tr>
<tr>
<td><strong>Innovation intensity (log)</strong>&lt;br&gt;(on performance (log))</td>
<td>0.036***</td>
<td>0.039***</td>
<td>0.034***</td>
<td>0.019***</td>
<td>0.020***</td>
<td>0.022***</td>
<td>0.011**</td>
<td>0.021***</td>
</tr>
<tr>
<td></td>
<td>[0.0088]</td>
<td>[0.0048]</td>
<td>[0.0041]</td>
<td>[0.0036]</td>
<td>[0.0038]</td>
<td>[0.0040]</td>
<td>[0.0045]</td>
<td>[0.0034]</td>
</tr>
<tr>
<td>Observations</td>
<td>19 521</td>
<td>43 611</td>
<td>61 267</td>
<td>68 453</td>
<td>63 795</td>
<td>59 881</td>
<td>57 660</td>
<td>63 836</td>
</tr>
</tbody>
</table>

With country and sector fixed effects  
Standard errors in brackets  
*** p<0.01, ** p<0.05, * p<0.1
Conclusion:

- We find a **positive effect** of R&D intensity on innovation intensity.
- We find a **positive effect** of innovation intensity on performance.
- We find a certain **heterogeneity** between countries and over time.

Extensions:

- Add control variables: for example, financial support.
- Testing alternative estimation methods and sensitivity analysis.
- Add other types of innovation and check the complementarity.
- Integrate the innovation process with the impact of anti-competitive regulations on the labor market.
Thank you for your attention!