

DISCUSSION

Patents to Products:
Innovation, Product Creation, and Firm Growth

Antonin Bergeaud
Banque de France and CEP

CEPR MG annual meeting - Queen Mary University - Nov 2019

Takeaways

- ▶ Very large number of papers use (very rich) patent data and patent data matched with firm → imperfect/incomplete measure of innovation.
- ▶ Complete patents data with information on consumer goods at an extreme detailed level
- ▶ Products are matched to patents and to firms → study the strategic use of patents by type of firms.
- ▶ Very exciting project with countless potential applications

Key findings

- ▶ Never patenting firms introduce more than 50% of new consumer goods products (and generate close to 50% of CPG sales)
 - ▶ Suggest patents imperfectly captured these innovations
 - ▶ Yet conditional on patenting, patents are associated with new goods
- ▶ Large firms introduction rate is lower (15% for largest vs 20% for smallest)
 - ▶ But they patent more
 - ▶ New products per patents are lower for larger firms
 - ▶ Suggest big firm use patents to deter product entry by other firms

Comment 1: Representativeness of Nielsen data

- ▶ Total CPG sales in data: \$220b against roughly \$630b.
 - ▶ Mostly food and drugs
 - ▶ What about online retail?
- ▶ Potential differences with aggregate?
 - ▶ in terms of composition wrt CES (Jaravel, 2018)
 - ▶ in terms of turnover rate wrt CPI (Klenow and Malin, 2011)
- ▶ Patenting activity in CPG:
 - ▶ CPG has a lot of patents (10% patenting firms in sample vs less than 1% in all sectors, and 5% in manuf)
 - ▶ But not a lot of new products compared to rest of CPI

Comment 2: Matching (1/2)

- ▶ Need for an aggregation of product based on similarity in underlying technologies.
 - ▶ Classification problem identified by Griliches (1990) → hard to classify without an pre-existing classification.
 - ▶ Circumvented using text from Wikipedia and k-mean algorithms → usually very sensitive to the number of clusters
 - ▶ New automated patent landscaping literature might help (Abood and Feltenberger, 2018; Bergeaud and Verluise, 2019)
 - ▶ How is the new classification nested into the existing ones?

Comment 2: Matching (2/2)

- ▶ Patents then matched with product modules via text similarity
 - ▶ Relying on a fully unsupervised method → validation is crucial (especially for Match 1)
 - ▶ Limitations of tf-idf

Word embedding

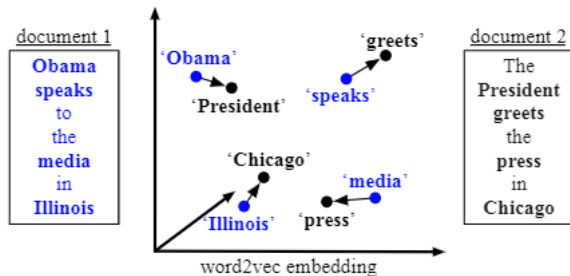


Figure: Word2vec embedding, taken from Krusner et al. (2015)

Comment 2: Matching (2/2)

- ▶ Patents then matched with product modules via text similarity
 - ▶ Relying on a fully unsupervised method → validation is crucial (especially for Match 1)
 - ▶ Limitations of tf-idf
 - ▶ New easy to use techniques based on word embedding → word mover's distance (Krusner et al., 2015) and many others since
- ▶ Here again, supervised learning methods could help
 - ▶ Virtual Patent Marking webpages
 - ▶ Physically marked products
 - ▶ IPRoduct project at EPFL
 - ▶ Patent citations?
- ▶ Problem of false positive?

Comment 3: What is a new product?

- ▶ Very nice “newness” index based on characteristics of products and module specific weights (i.e. distance to average related products).
 - ▶ How about new combination of existing characteristics?
 - ▶ Small marketing changes (color, size...) still increase the novelty
 - ▶ Why is newness so low in electronics?

Comment 4: Timing

- ▶ Extensive margin: number of new products increases simultaneously after first patent application.
- ▶ Same at the intensive margin
 - ▶ Timing suggests that firm anticipate
 - ▶ Causality? Size effect?
 - ▶ Use rejected patents
 - ▶ Or patent marking for case studies
- ▶ Evidence that some type of patents (technologies) are closer to the final product than other (i.e. going beyond the product/process split)?

Other comments

- ▶ Licensing of patents?
- ▶ Evaluate the breadth of a patent from claims to measure defensive patenting activities
- ▶ Ideas are harder to get. Can a patent corresponds to more products today than in the past?

Recap

- ▶ Great idea to match product and patent data which is a momentous task.
- ▶ Tons of interesting findings for the CPG, namely that most product innovation comes from non patenting firms.
- ▶ So many possibilities for further studies using patent specific information (citations), product details and firm level data.