UNIVERSIDAD DEL NORTE

Vigilada Mineducación

A characterization of Colombian industries under Schumpeter's patterns of innovation

J. Taborda[1]

[1]Economics Student at Universidad del Norte

Bachelor Thesis Progress. September 30th, 2022

Table of Contents

Table of Contents

Introduction

- **Who drives innovation within an industry?**
- Could it be a small firm.... Or a large corporation?
- We will answer this question using Schumpeterian patterns of innovation –> **Mark I and II**
- How? Characterization –> Cluster analysis
- On the basis of...? EDIT and EAM

Problem Statement

- Characterization exercises "*have been standing the test of time quite well*" (Fontana et al., 2012)
- Useful for policymaking
- But where do these exercises **take place**?
- A gap in the literature for Colombia

Objectives

- **Main objective**: **Characterize** industries in the Colombian manufacturing sector based on the framework of Schumpeter's Mark I and II, aiming to supply a better insight on the differences between industries by finding out what type of firm drives innovation.

Objectives

- **Specific Objectives:**
  - **Utilize information** from the EDIT and EAM surveys to set up a database about firm features in manufacturing industries.
  - Based on EDIT and EAM information, **construct a quantitative analysis** at the firm level that yields a result at the industry level, which in turn gives groundwork to create industry-level comparisons.
  - Employing a cluster algorithm, **group industry-level data** by common patterns and characterize them using Schumpeter's Mark I and II, which will give an insight into the drivers of innovation constrained to industrial structure and dynamics.

Table of Contents

Theoretical Framework

- The concept of innovation (OECD, 2018)
  - What is innovation? (do not confuse with invention)
  - ***New or improved product or process (or a combination thereof) that differs significantly from the unit's previous products or processes and that has been made available to potential users (product) or brought into use by the unit (process)"***
  - What is an innovative activity? (ACTI)

Theoretical Framework

- Taxonomies of innovation
  - By type: Market, process, product, etc... (OECD, 2018)
  - By novelty and impact...
  - Radical (Schumpeter, 1942)
  - Incremental (Kirzner, 1973)

Theoretical Framework

- Schumpeterian patterns of innovation
  - Mark I: Schumpeter early thoughts. **Small firms are the drivers of innovation** (1911)
  - Mark II: Schumpeter late deliberations. **Large firms are the drivers of innovation**, and perfect competition is not only inferior but inefficient.

Theoretical Framework

- **Mark I**: Perfect competition, **radical** innovations
- **Mark II**: Monopoly/Oligopoly, **incremental** innovations
- *According to who*...? Market Structure and innovation:
  - Fontana et al. (2012): Turbulence vs Stability
  - Arrow replacement effect (1962)
  - Baumol proposition (2004)
  - Gilbert (2006) incentives to innovate based on potential profits
  - Shapiro revisit (2012): Unifying principle... **competition**

Literature review

- Previous works on the field: Malerba and Orsenigo (1996), Breschi et al. (2000), Landström & Schön (2010), Castellaci and Zheng (2010), Corrocher et al. (2007).
- The Pavitt's alternative (1984). Which one is better? *spoiler: depends*
  - Pavitt: **Kondratiev waves** (Archibugi, 2001)
  - Schumpeter: **Early/Late stages of an industry** (Malerba, 2005)

Literature review

- Not one size fits all
- Attempts in Colombia with Schumpeter? **Yes, but** not in characterization (Umaña-Aponte et al., 2013; Marroquín, 2010; Arroyo-Mina & Guerrero, 2018; Langebaek-Rueda & Vásquez, 2007).
- Attempts to characterize? **Yes, but** not with Schumpeter (Cerón et al., 2010; Ovallos-Gazabón & Amar-Sepúlveda, 2014).

Table of Contents

The Data

- Cross-section
- Inner join of 2018 **EDIT** ("*Encuesta de Desarrollo e innovación tecnológica*") and **EAM** (*Encuesta Anual Manufacturera*) surveys by DANE (2019:2020)
- EAM is a **census**; EDIT samples EAM **industries** –> Inner join
- Criteria: Employees and profits
- Each firm has a *"Numero de Orden"* (NORDEMP)

The Data

- Number of observations: from almost 8000 to 6405
- I will employ variables that approach Breschi et al. (2000) and Malerba & Orsenigo (1996) dimensions
- That means... **market concentration, stability and technological opportunities**

Table 1: Relevant variables for the study

Dimension	**Concepts**	**Variables (by DANE's code)**
Stability	Amount of radical innovations	I1R4C2N
	Amount of incremental innovations	I1R4C2M
Concentration	Total sales	I3R2C1
	Total spending on innovative activities	I1I1R10C2
	Total Employees	PERTOTAL
	Total Output	PRODBIND
Technological Opportunities	Possession of conventional protection mechanisms valid until 2018 (Patents, IP, Copyright, Trademarks)	VI1R8C2
	Obtention of conventional protection mechanisms between 2017-2018	VI2R8C2
	Usage of non-conventional protection mechanisms (NDA, industrial secrets, high complexity on designs)	VI3R5C2

Source: Own elaboration based on DANE's surveys (2019;2020)

The Measures

- **Concentration (CON)**
- Based on Malerba and Orsenigo (1996) share on ACTI and sales
- Expanded to add supply share and labour share

$$CON = (HH_{ms} * HH_{msa} * HH_{lsd} * HH_{ss})^{1/4} \quad (1)$$

The Measures

- **Stability (STA)**
- The dynamic problem. **EDIT is non comparable**
- Thus, we need another approach. A static approach
- Based on Baumol (2004) proposition

$$STA = Sr - Si \tag{2}$$

The Measures

- **Technological Opportunities (TO)**
- Based on Maleki et al. (2018) approach
- Growth rate of patents
- Non-comparability of EDIT –> Dynamic approach problem, again
- Static approach

$$TO = \frac{PM_{1718} + NCPM_{1718}}{PM} \tag{3}$$

The Methods

- **Important**: I had problems on industries with less than 20 observations. Some industries had 0 spending on ACTI. Hence, I establish 20 as the minimum number of observations in my study. **n = 5986**

The Methods

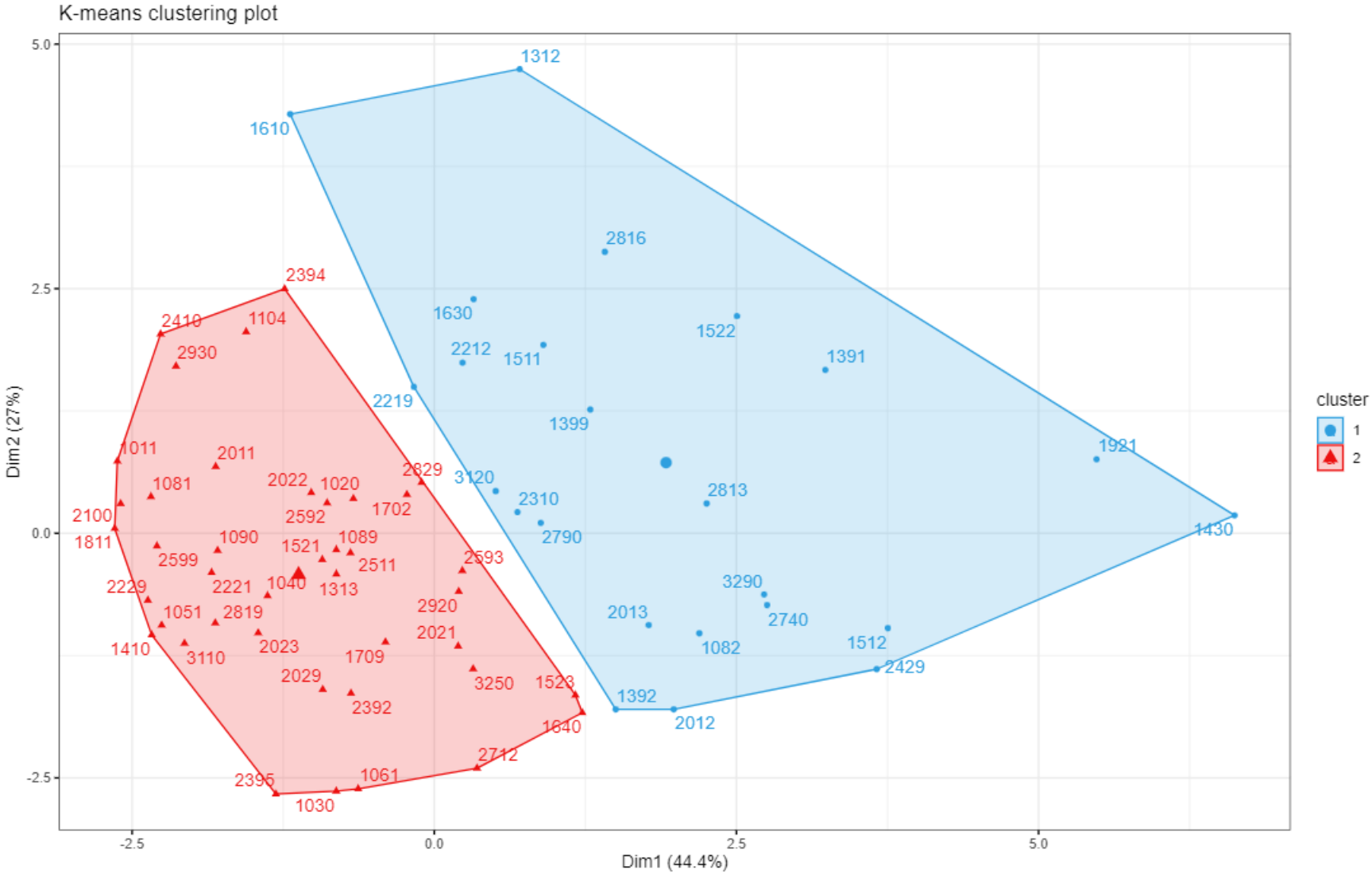
- Method: k-means clustering
  - 2 groups
  - Lloyd algorithm
  - 10 repetitions
- **Where**? in R. *pvclust*, *factoextra* and *stats*

The Methods

- k-means in a nutshell (MacKay, 2003):
  - *Assignation phase*: Each observation gets assigned to the group with the closest mean (by Euclidean distance). Groups have a centroid
  - *Update phase*: Group parameters adjust to match the means of the data points.
  - *Repetition phase*: Assignation and Update phases repeat until they do not change anymore, that is, data points do not change their position in the groups.

Table of Contents

Figure: Preliminary characterization of Colombian Manufacture using a two groups k-means clustering method



Initial impressions

- Dim1 and Dim2
- Two groups: Cluster Group 1 (CG1) and Cluster Group 2 (CG2)
- One is denser than the other
- Also, one has more firms than the other
  - CG1 –> n = 794
  - CG2 –> n = 5192

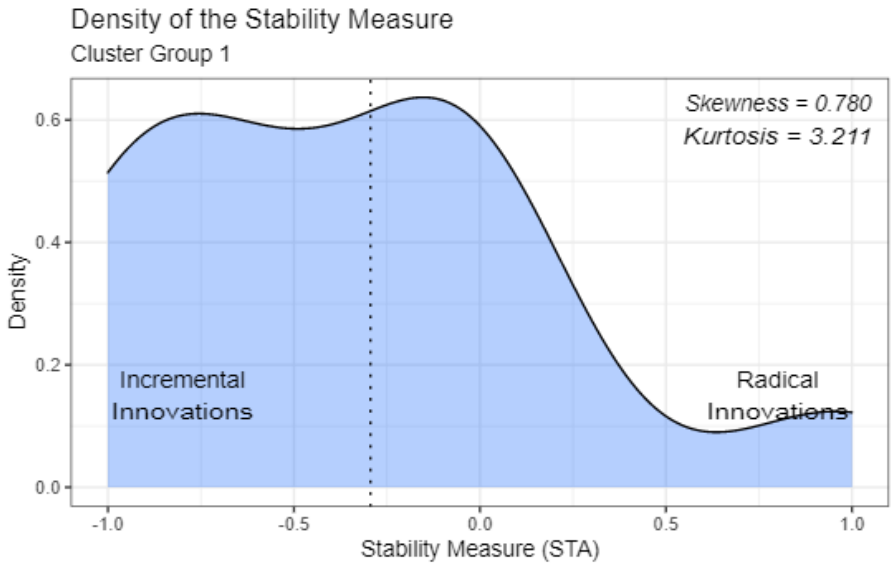
Descriptive statistics

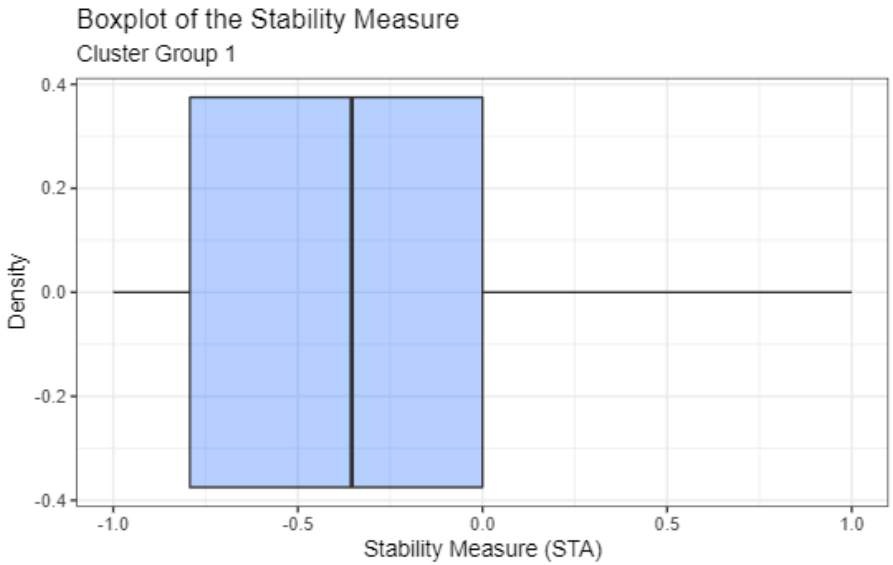
Table 2: Initial descriptive statistics of the two groups k-means clustering

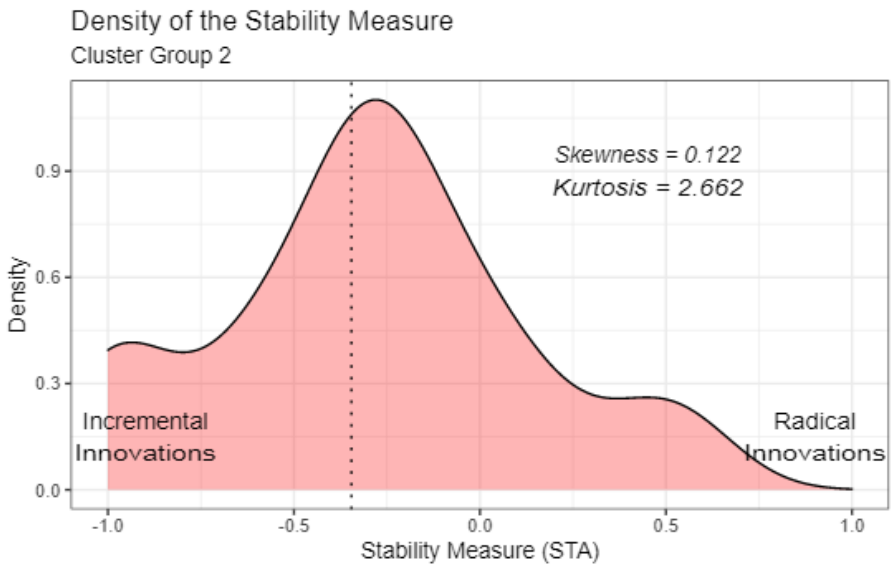
	Cluster Group 2			*Cluster Group 1*		
	TO	*CON*	*STA*	*TO*	*CON*	*STA*
Max	3.500	1023.11	0.586	6.000	2702.96	1.000
Min	0	160.28	-1.000	0.047	1068.11	-1.000
Mean	0.522	572.22	-0.294	0.983	1507.72	-0.345
Std Dev	0.651	262.98	0.414	1.496	416.08	0.550

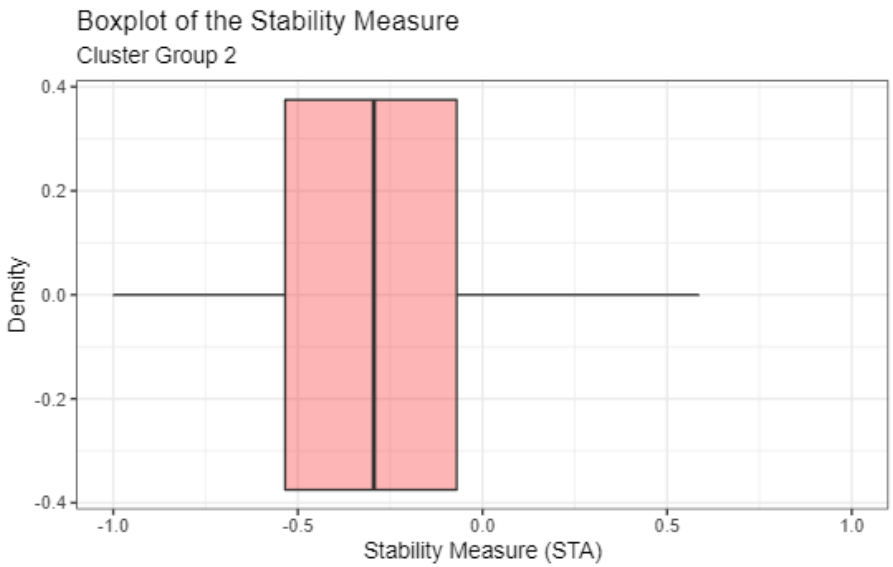
Source: Own elaboration

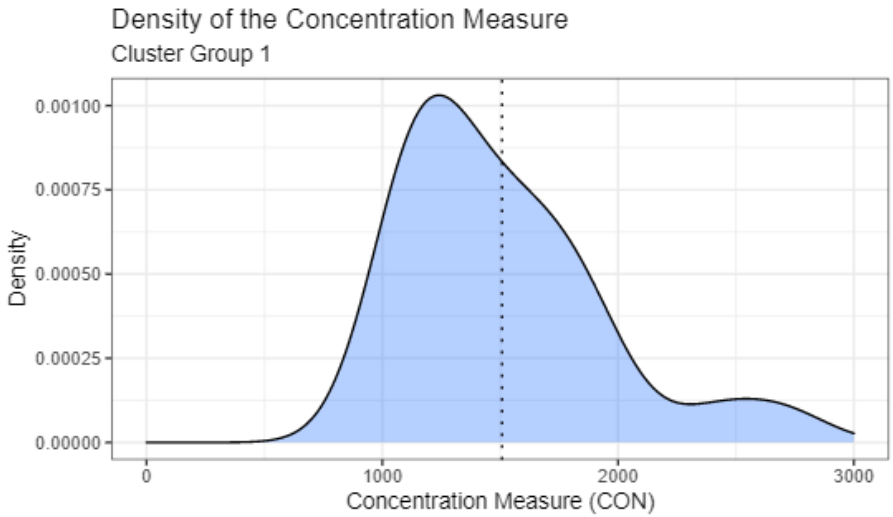
Table of Contents

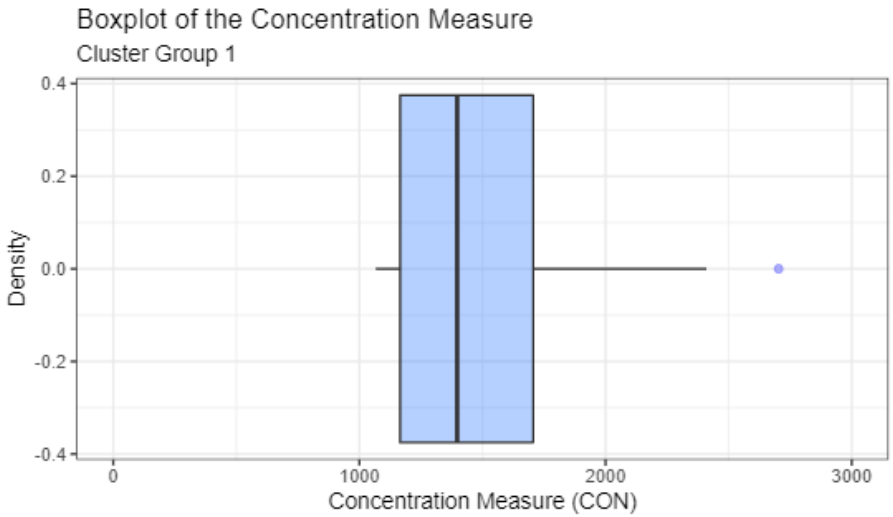

Density of the Stability Measure
Cluster Group 1
Skewness = 0.780
Kurtosis = 3.211
Incremental Innovations
Radical Innovations
Density
Stability Measure (STA)

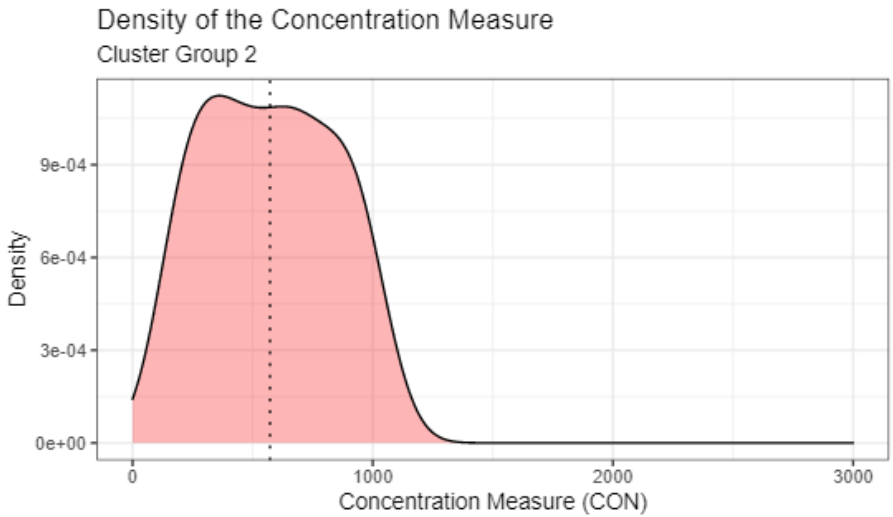

Boxplot of the Stability Measure
Cluster Group 1
Density
Stability Measure (STA)

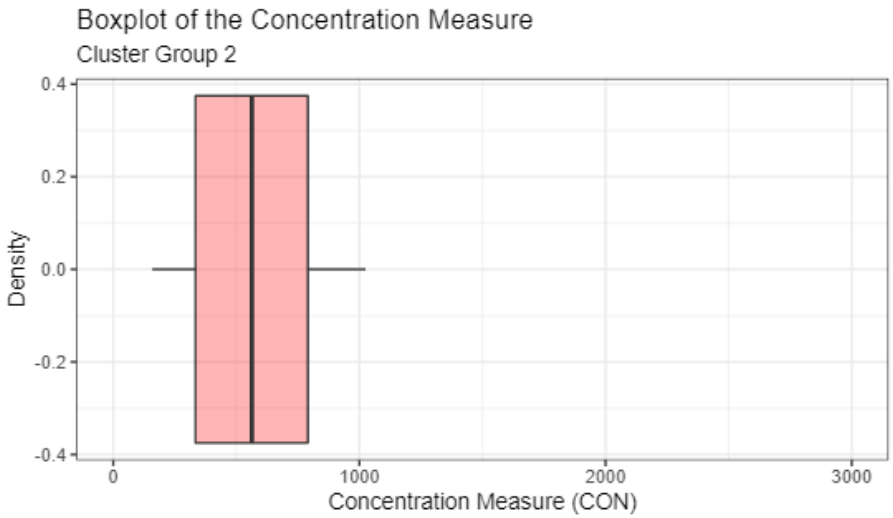

Density of the Stability Measure
Cluster Group 2
Skewness = 0.122
Kurtosis = 2.662
Incremental Innovations
Radical Innovations
Density
Stability Measure (STA)

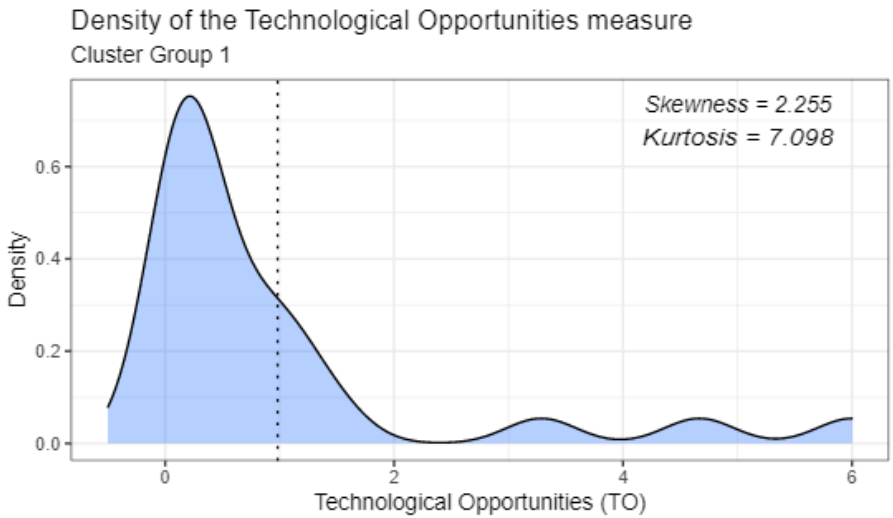

Boxplot of the Stability Measure
Cluster Group 2
Density
Stability Measure (STA)

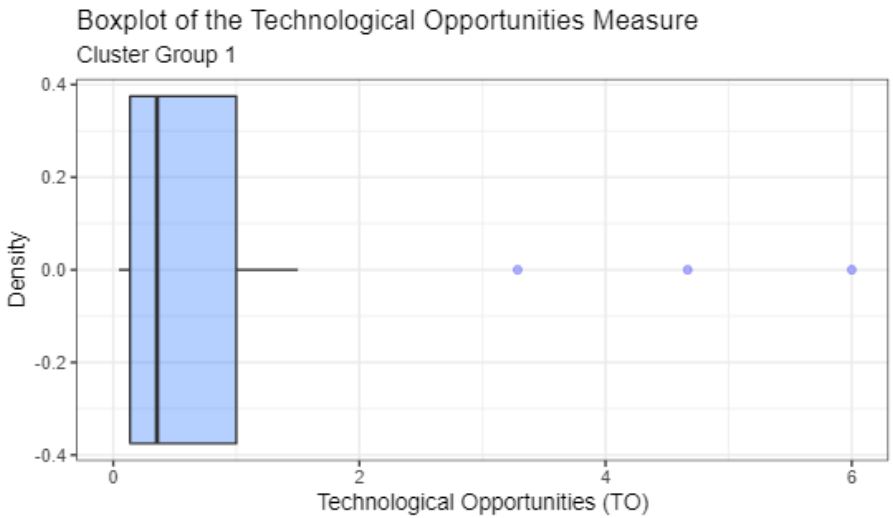

Density of the Concentration Measure
Cluster Group 1
Density
0.00100
0.00075
0.00050
0.00025
0.00000
0
1000
2000
3000
Concentration Measure (CON)

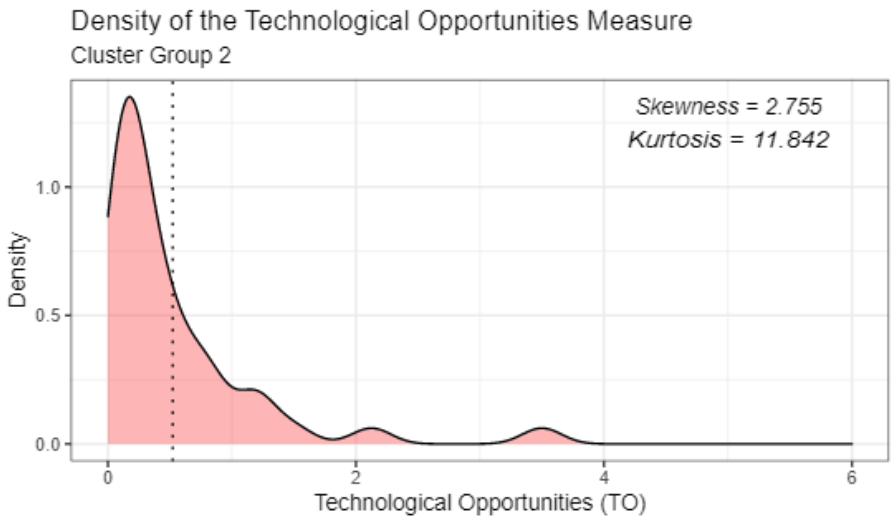

Boxplot of the Concentration Measure
Cluster Group 1
Density
0.4
0.2
0.0
-0.2
-0.4
0
1000
2000
3000
Concentration Measure (CON)


Density of the Concentration Measure
Cluster Group 2
Density
9e-04
6e-04
3e-04
0e+00
0
1000
2000
3000
Concentration Measure (CON)


Boxplot of the Concentration Measure
Cluster Group 2
Density
0.4
0.2
0.0
-0.2
-0.4
0
1000
2000
3000
Concentration Measure (CON)


Density of the Technological Opportunities measure
Cluster Group 1
Skewness = 2.255
Kurtosis = 7.098
Density
Technological Opportunities (TO)


Boxplot of the Technological Opportunities Measure
Cluster Group 1
Density
Technological Opportunities (TO)


Density of the Technological Opportunities Measure
Cluster Group 2
Skewness = 2.755
Kurtosis = 11.842
Density
Technological Opportunities (TO)

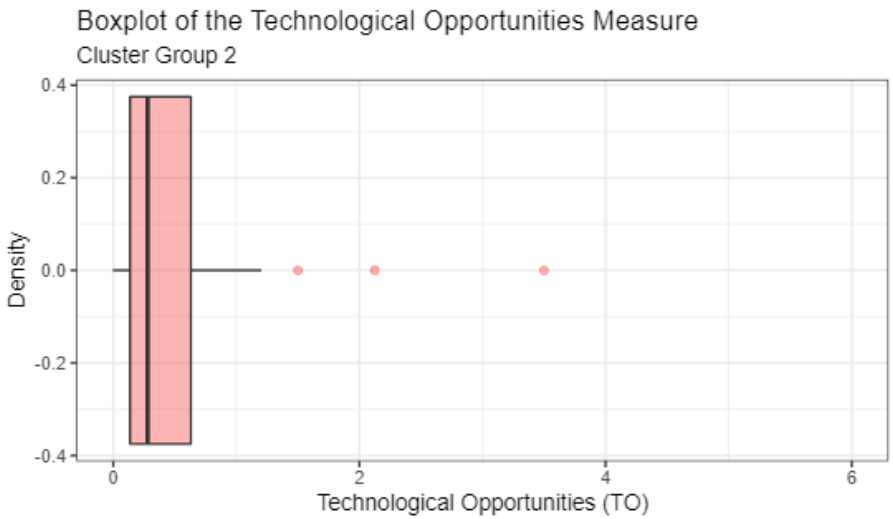

Boxplot of the Technological Opportunities Measure
Cluster Group 2
Density
Technological Opportunities (TO)

Table of Contents

What can we conclude so far

The most important conclusion!!...

- Red cluster (CG2): Mark I industries
- Blue cluster (CG1): Mark II industries

What can we conclude so far

Moreover...

- We have been able to characterize Colombian industries under Schumpeterian patterns of innovation
- We found **Who drives innovation** –> On Red cluster (CG2), small firms drive innovation, but on Blue cluster (CG1)... it is all about large firms
- Concentration is the spearhead of our analysis, but our other two measures provide interesting results too
- Groundwork for policymaking and future studies