## I. Introduction

The aim of this research paper is to use patent data to study the patenting behavior of firms during the dot-com bubble. There are considerable anecdotal evidence that a lot of firms during that period of time are concentrating their research efforts on a selected few technology fields that were "hot" at the time. This research paper aims to provide concrete evidence to the phenomenon discussed above by examining all the patents applied in the US during the dot-com bubble periods and see if firms indeed adopted different patenting strategies during the bubble compared with other time periods. A more detailed introduction of the topic can be found in the Rmd file.

II. Detailed discussion of each graph

(a)
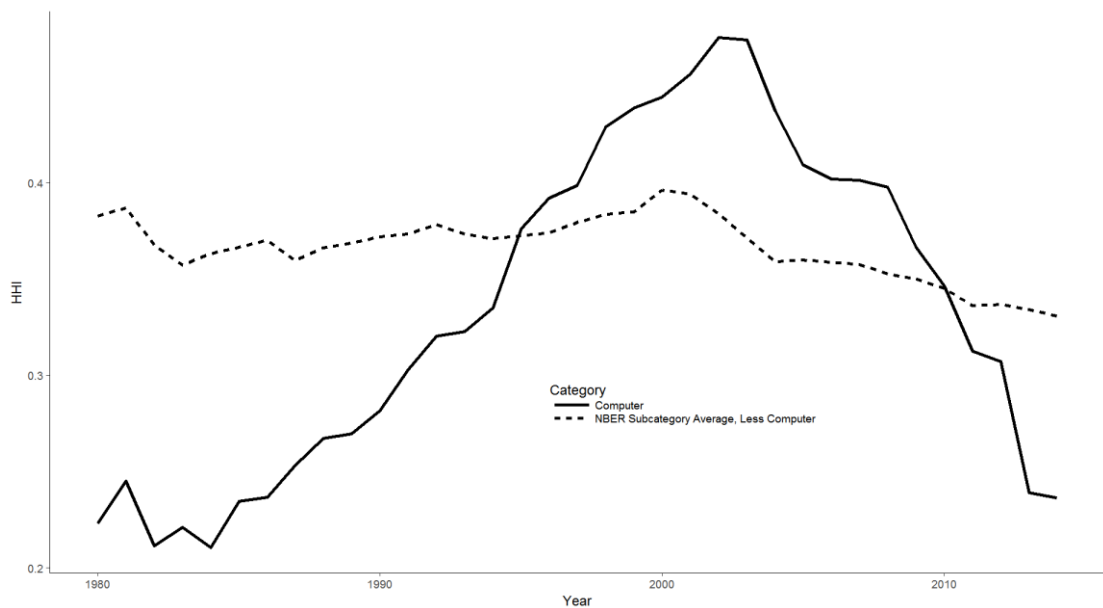
This graph shows the result that Computer patents during the dot-com bubble display extremely high level of concentration in IPC subclasses.

This graph is quite truthful. The x-axis is year and the y-axis is the HHI of every year. The dash line shows the HHI of the subcategory average less computer, and the other one is the computer & communications subcategories. Both year and HHI are objective indicators so it is definitely truthful. Moreover, it is really functional. As we mentioned earlier, our aim is to show Computer patents during the dot-com bubble display an extremely higher level of concentration than other IPC subclasses. The pattern of the lines exactly shows this. As we can see, at the peak of the computer category, the HHI of which is significantly higher than the index shown by the dash line. This is quite persuasive and works really well in terms of telling the audience what we want them to learn.

Aesthetically, this plot is very concise. It does not include many elements. Maybe some people would like to have a plot with more colors. In this plot, we could distinct the lines by color. But we didn't think it would work better than this. We discussed this issue and we concluded that if we just have two lines (i.e. two types), the full line and dash line should be the best match. For it is really clear and really easy to read.

This plot actually provides the audience with a clear first impression of the topic we are going to work on. It works well.

(b)
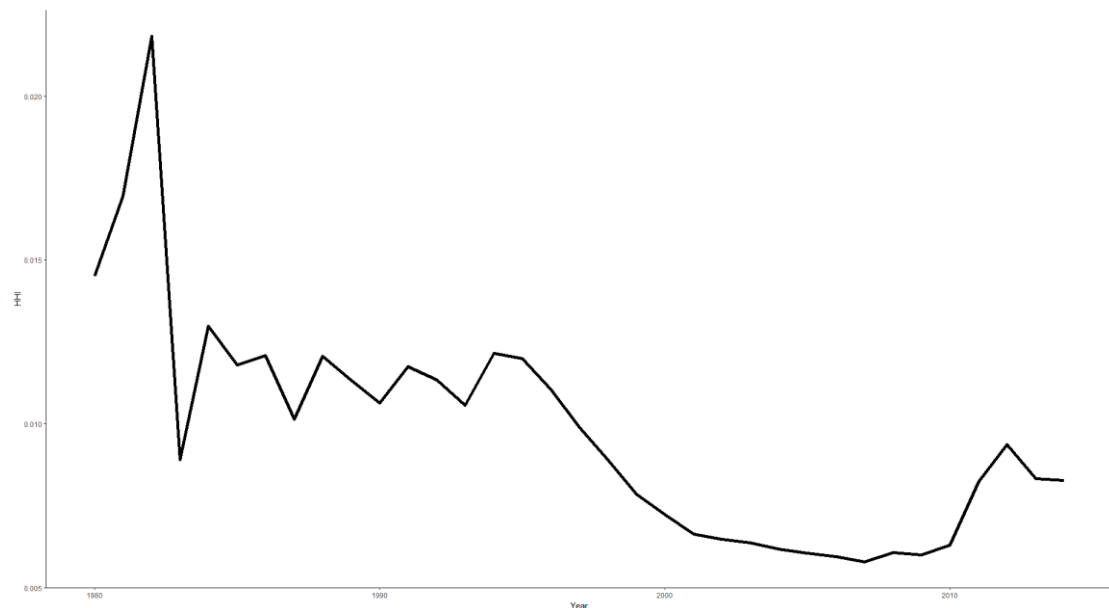
This second plot shows New patents are not concentrated among a few firms.

Same as the first one, this plot is absolutely truthful for it just shows the objective variables. As this is quite a simple plot, we wouldn't use the word "beautiful" to describe this one. But it is functional, for it is really insightful and enlightening.

In the first plot, we show that Computer patents during the dot-com bubble display extremely high level of concentration. In this one, we are focused on the new patents.

This plot is very insightful and enlightening. Most people pay more attention to the overall situation but overlook some potential subtopics. In this case, most literature focused on the overall situation of patents. They discovered that among the period of dot-com bubble, most Computers patents belong to several big companies. However, when we look at the new patents assigned each year, especially during the early 2000s, we can discover that the HHI is actually very low. It tells us the fact that the new patents assigned (during the period of dot-com bubble) are not only applied by those a few

monopoly companies. Truly insightful and enlightening.

(c)
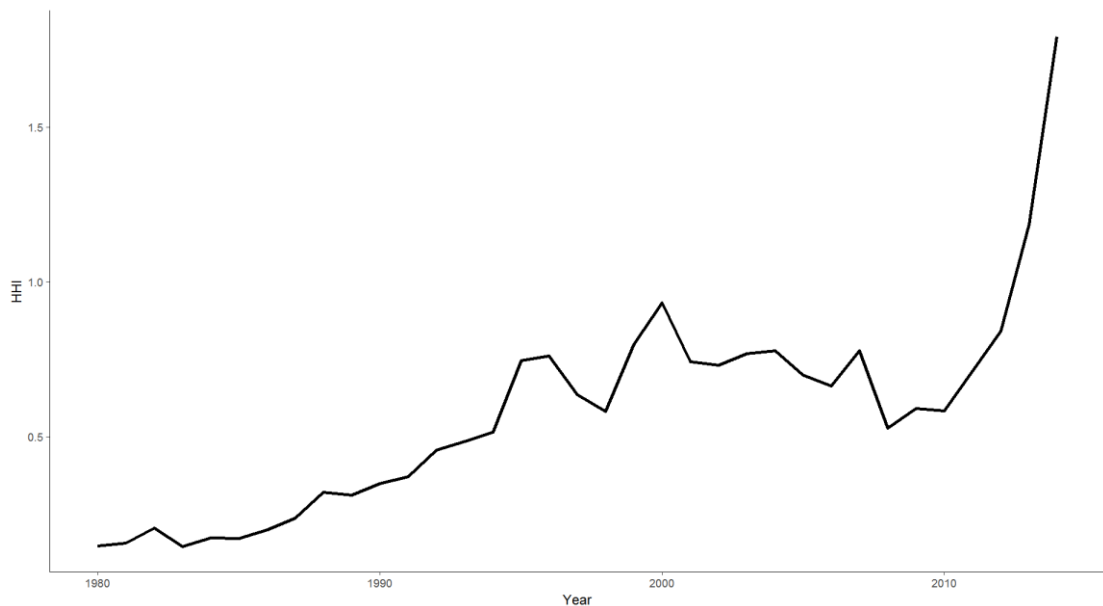
Again, this plot is concise but really insightful and enlightening. Aesthetically, this plot is basically same as the previous one.

The figure above reports the HHI of citations made by new patents to patents owned by different firms. It tells the story that citations made by new patents are concentrated in the patents owned by a select few firms. We are going deeper in this plot. What we learn from this one is that the new patents are highly connected with those patents owned by selected a few companies (mainly the monopolies in the industry). In other words, new patents, though not highly concentrated in some companies themselves, are based on those patents belonging to several selected companies.

(d)
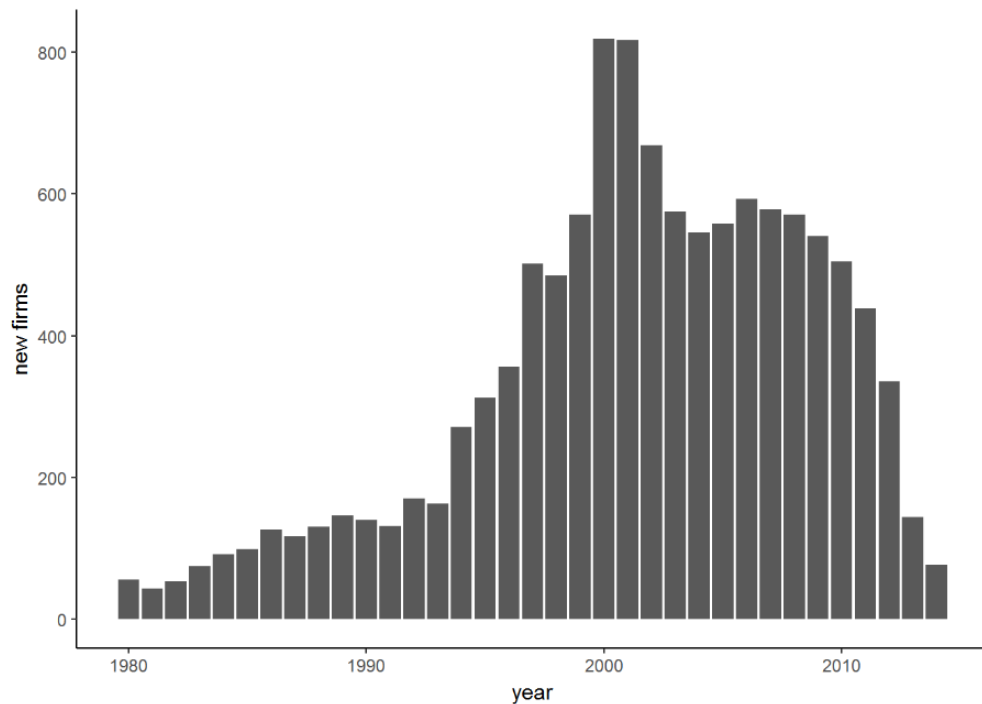
This is one of the bar charts in the project. The original data we had didn't include the amount of companies. So we used some tricks (tidying and mutating the dataset) to get that data.

To certain extents, this chart is even more truthful than the previous charts. HHI is an index proposed by economists to illustrate concentration ratio of an industry. Some people trust this indicator while others don't. However, the amount of companies which get a patent for the first time in a single year is an absolutely objective variable.

Frankly, in most cases, line chart is better at showing the change of a variable. We chose the bar chart for two reasons: firstly, we already have three line charts so we would like to have something new. We shouldn't make others feel bored reading our

project. It might be a good way to draw the attention of audience by using something new. Secondly, the bar chart can make a visual impact. We want to emphasize the large amount during the dot-com bubble. The bar chart can give people this impression and it is functional.

(e)
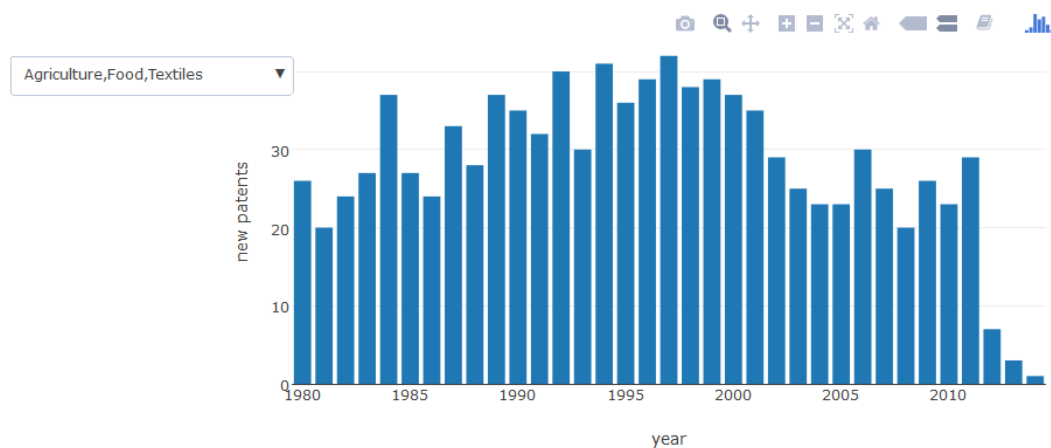
The interactive plot is truthful as it just shows the number of new patents under each sub-category in each year.

It is really functional. We want to give the audience an opportunity to get access to the change through years of each subcategory. It is impossible to put the charts every subcategory in a report for there are too many subcategories. The interactive chart fits well in this case. The audience can choose whatever subcategory they would like to see. Furthermore, we think the color is pretty good.

(f)

The amount applied in terms of category
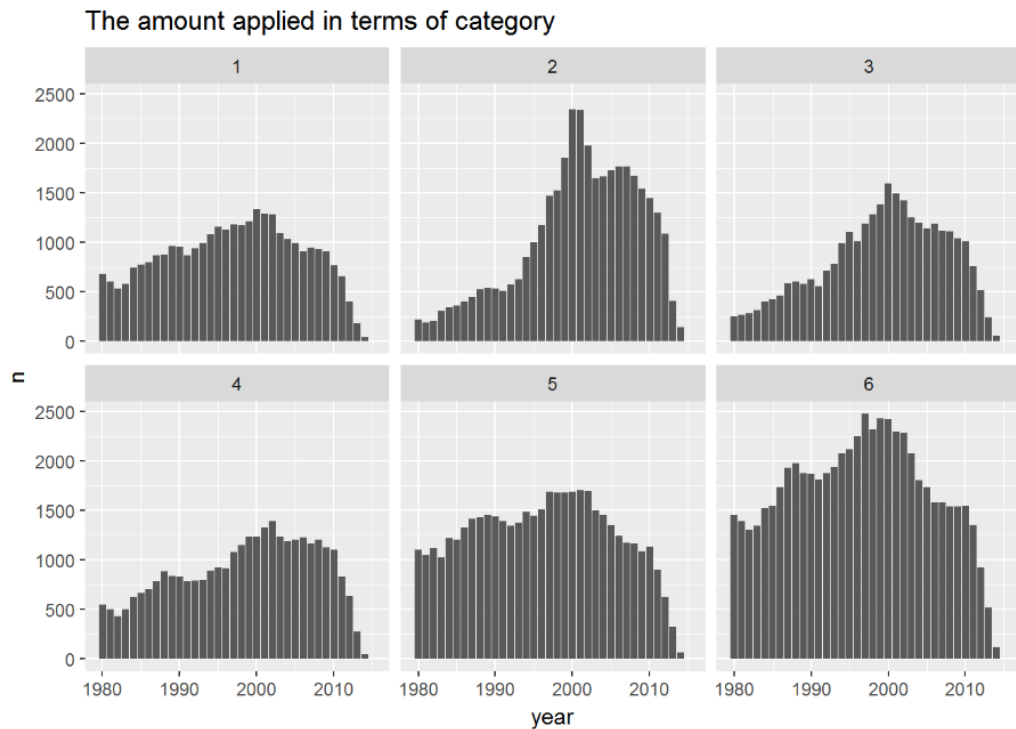
This is an addition to the previous interactive chart. As there are only six categories (more than sixty subcategories in total), it is feasible to put them in one chart by using the facet function.

It makes it quite clear that there is some similarity and difference among patents under different categories. So it is functional. The pattern of category 2 is basically same with the pattern in the bar chart. So we can conclude that the dot-com bubble can be the main reason leads to the pattern shown in the bar chart. Thus, it is insightful and enlightening.