1 Empirical Strategy

1.1 Data

I employ a novel administrative dataset from the Canadian Intellectual Property Office, the IP Horizons Patent Researcher Datasets . The dataset identifies patents in Canada from 1860 to 2023. All parties involved in a patent application can be identified along with their registered province of origin if the party comes from Canada and the state if the party is recorded in the United States.

With these data, I compute quarterly and monthly patent application counts at the province level from January 2001 to June 2021, based on the application filing date. This period corresponds to the modern Canadian intellectual property institutional context, which I reviewed on Section ??. I assign patents to provinces based on where the majority of parties involved in a patent application report their location. I only include 2021 as a partial year since the data present a downward trend for all provinces in late 2021, suggesting patent applications are yet to be updated for the most recent periods of the IP Horizons data. Further, I drop Newfoundland and Labrador, Prince Edward Island, Yukon and Nunavut due to missing observations on most explanatory variables.

The explained variable of interest is the count of patent applications. To allow for heterogeneity in treatment, I separate patents by their International Patent Classification (IPC) section, which defines a broad classification of the technology being patented. For robustness checks, I also consider the number of parties involved in a patent application as explained variables for robustness checks, separating by party type (inventors, owners and applicants).

For my explanatory variables, I extract province-level data at the monthly frequency from Statistics Canada. These include data from the Labour Force Survey (LFS), such as labour force characteristics, employment wages, among others .