

1 Empirical Approach

1.1 Data

I employ a novel administrative dataset from the Canadian Intellectual Property Office, the IP Horizons Patent researcher datasets (2023). The data identify patent applications in Canada from 1869 to August 2021, including all involved parties and the application filing data. Parties can be mapped to provinces based on their location, which can be in Canada or other countries.

With these data, I compute quarterly patent application counts at the province level from 2001Q1 to 2021Q2. This period, spanning January 2001 to June 2021, corresponds to the modern Canadian institutional context, as reviewed in Section ??, when most provinces had implemented their SR&ED programs. I assign patents to provinces based on where the majority of parties from patent applications report their location¹. Further, I drop Newfoundland and Labrador (NL), Prince Edward Island (PE), Yukon (YU) and Nunavut (NU) due to missing observations on explanatory variables.

Patent application counts are my main explained variable. Further, I separate patents by their International Patent Classification (IPC) section as separate explained variables for some models. For explanatory variables, I extract province-level data at the monthly frequency from Statistics Canada and aggregate it at the quarterly frequency. These include data from the Labour Force Survey (LFS), such as labour force characteristics, employment wages, among others. I also consider the consumer price index, international merchandise exports and imports, retail, wholesale and manufacturing trade sales, food services receipts, the new housing price index and electric power generation. I also include the number of business insolvencies as reported by Innovation, Science and Economic Development Canada (2024) and the number of foreign parties involved in patent applications from the IP Horizons data. I aggregate data at the quarterly level by summing all variables except the consumer and new housing indices, which I average over months. Table ?? in Appendix ?? provides a detailed account of my data sources. Table 1 presents descriptive statistics for the main explained variable and all explanatory variables.

1. Patent applications without information of party locations or with an equal number of interested parties from two provinces are dropped from the sample.

Table 1: Descriptive statistics for the province-quarter sample

	Mean	SD	Min	Median	Max
Ln +1 Patent applications	4.261	1.405	1.099	4.107	6.691
Ln Full-time employment	8.026	1.034	6.726	7.831	9.814
Ln Median wage	2.949	0.192	2.523	2.956	3.395
CPI	119.145	12.668	95.400	119.400	148.900
Ln +1 Business insolvencies	4.403	1.396	0.693	4.197	6.957
Ln Intl. exports	15.810	1.139	13.694	15.848	17.804
Ln Intl. imports	15.646	1.198	13.715	15.369	18.372
Ln Retail sales	15.963	1.028	14.424	15.774	17.913
Ln Wholesale sales	15.910	1.292	13.907	15.892	18.490
Ln Manufacturing sales	16.027	1.179	14.398	15.729	18.213
Ln International travellers	12.470	1.779	4.344	12.387	15.929
Ln Arriving vehicles	11.944	3.562	0.000	12.516	15.801
Ln Electric power generation	16.213	0.997	14.344	16.219	17.990
Ln Average actual hours	3.545	0.050	3.311	3.550	3.676
New housing price index	88.064	16.987	42.900	94.250	129.500
Ln Food services receipts	13.737	1.108	12.255	13.575	15.857
Ln Average job tenure	4.636	0.088	4.399	4.653	4.830
Ln +1 Foreign patent parties	3.609	1.918	0.000	3.842	6.671

Notes: All statistics based on a balanced panel of $N = 656$ province-quarter observations. The sample includes all Canadian provinces except NL, PE, YU and NU.

1.2 Identification Strategy

The AITC, as an investor tax credit, did not directly affect innovation inputs such as R&D expenditures. However, since it provided access to financing for innovative firms, it may have affected innovation in the form of patent applications. To estimate this effect, I implement

a two-way fixed effects (TWFE) difference-in-differences (DD) design, where I define treatment and control groups based on when the program was passed (2017Q1). While the first investment eligibility date was in 2016, businesses only started receiving AITC funding after 2017Q1, hence any effect would only be observed then. The treatment group is Alberta, and the treatment period is composed of all periods after 2017Q1. The control group is all remaining Canadian provinces in the sample. Treated observations are those from Alberta after 2017Q1, where the AITC may have affected Albertan patent applications. The DD design is implemented in a regression framework according to Equation 1 below.

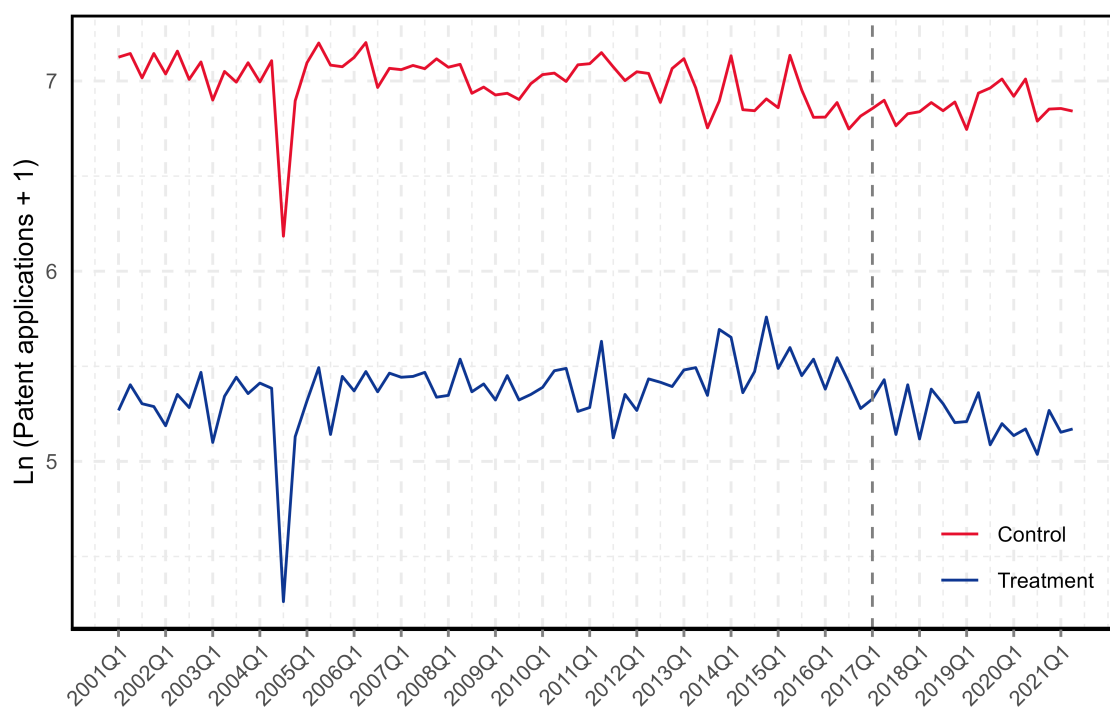
$$\ln(P_{it} + 1) = \theta_i + \theta_t + \beta T_{it} + \mathbf{x}_{it}'\gamma + u_{it} \quad (1)$$

where P_{it} is patent applications in a province i and period t . θ_i and θ_t are sets of province and period fixed effects. I use a natural logarithm transformation with the addition of one to correct for provinces with small amounts of patent applications on some periods. The logarithm will give percent interpretations to the coefficients on the right-hand side. T_{it} is a binary variable equal to unity for observations for treated observations and zero otherwise. Hence, the estimated parameter $\hat{\beta}$ is the coefficient of interest, which is my estimate for the effect of the AITC on P_{it} . \mathbf{x}_{it} is a vector of time and province-varying controls, as described in the previous subsection, and γ is the associated vector of parameters. u_{it} is a province and time-varying error term. I cluster standard errors at the province and period level, as the variance of the error term may be spatially and temporally correlated.

Table ?? in Appendix ?? presents the difference in means between treated and control provinces for all explained variables. This presents the simplest version of the DD estimate, where I compare the average number of patent applications between Alberta and the control provinces before and after the AITC intervention. This simple comparison suggests a small or null effect; the regression analysis described above will provide a more robust estimate, controlling for other factors that may affect patent applications.

The key identifying assumption of DD is that absent the intervention, the trend of patent applications in Alberta would follow a similar pattern to that in control provinces. Figure 1 shows the time series of patent applications between Alberta and control provinces. Alberta's

Figure 1: Quarterly time series of patent applications between treatment and control groups



Notes: The figure shows the quarterly time series of patent applications between the treatment and control groups. The vertical line represents the start of the AITC intervention in 2017Q1. The treatment group is Alberta, and the control group consists of all remaining provinces except NL, PE, YT and NU.

patent applications follow a similar pattern to control provinces before the intervention, however, some deviations are present before 2016Q2.

To allay the concern of unobservable factors impacting patent application trends across provinces, I estimate event study regressions following Equation 2 below and provide supporting evidence for causal identification of $\hat{\beta}$.

$$\ln(P_{it} + 1) = \theta_i + \tau_t + \beta_t(\tau_t \cdot A_t) + \mathbf{x}_{it}'\gamma + u_{it} \quad (2)$$

θ_i , τ_t , \mathbf{x}_{it} , γ and u_{it} represent the same as in Equation 1. τ_t has its reference level set to one period before the treatment start period (2016Q4). A_t is a binary variable equal to unity for Alberta observations and zero otherwise. $\tau_t \cdot A_t$ is the interaction term between these two variables. β_t is the associated vector of coefficients, which will show the difference between the treatment and control groups in the explained variable for all t . For these regressions, I show the values of the interaction terms in event study plots, along with 95% confidence intervals. I cluster standard errors at the province and period level.

Evidence in favour of the identifying assumption will be observed if the β_t before 2016Q4 are not statistically significant. This supports the idea that Alberta had no significant differences in the trend of patent applications to other provinces before the intervention. Thus, I use the event study regressions to provide evidence of the causal identification of the effect of the AITC. Further, I use event study regressions in the form of Equation 2 to examine the effectiveness of the AITC by looking at post-treatment interaction terms.

1.3 Patent parties and province-month panel

I perform two robustness checks on DD and event study analyses to ensure the validity of my results. First, to the extent that results could be driven by the patent-province mapping, I consider the number of Canadian parties involved in a patent application as an alternative explained variable. I separate parties by type (all parties, inventors, owners and applicants²). Table ?? also presents the difference in means between treated and control provinces for par-

2. I do not consider agents as a separate category due to them being hired professionals, which is not informative about the patent application team.

ties involved in patent applications.

Second, I reestimate models on a province-month panel, to ensure that the results are not driven by the aggregation of data at the quarterly level. I present descriptive statistics of the monthly data in Appendix ??.