1 Empirical Approach

1.1 Data

I employ a novel administrative dataset from the Canadian Intellectual Property Office, the IP Horizons Patent Researcher Datasets. The data identify patents in Canada from 1860 to 2023, with information on when the application for the patent was filed and granted as well as the parties involved in the application. Parties can be identified 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 January 2001 to June 2021, based on the application filing date. This period corresponds to the modern Canadian intellectual property institutional context, as reviewed on Section **??**. Further, limiting to a period smaller than twenty years reduces the possibility of double counting patent renewal applications. I assign patents to provinces based on where the majority of parties involved in a patent application report their location[1]. I only include the first two quarters of 2021 the last two quarters present an unusual downward trend for all provinces in late 2021, suggesting patent applications are yet to be updated. 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. Further, to allow for heterogeneity in treatment, I separate patents by their International Patent Classification (IPC) section. I then consider the count of patent applications for each IPC section as separate explained variables for some models.

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

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

reported by Innovation, Science and Economic Development Canada (2024) (2024) and the number of foreign parties involved in patent applications, which I obtain from the IP Horizons data. I aggregate data at the quarterly level by summing all variables except the consumer and new housing indices, for which I take arithmetic averages. Table 1 presents descriptive statistics for patents and all explanatory variables.

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 from 2001Q1 to 2021Q2. The sample includes all Canadian provinces except Newfoundland and Labrador, Prince Edward Island, Yukon and Nunavut.

1.2 Empirical Strategy

The AITC, as an investor tax credit, did not directly affect innovation inputs such as R&D expenditures. However, since it directly provided cheaper financing for innovative firms, it may have affected innovation output in the form of patent applications. To estimate the effect of the AITC on patent applications, I implement a two-way fixed effects (TWFE) difference-in-differences (DD) design, where I define treatment and control groups based on the period the program was passed, which was 2017Q1 (January 2017) (Alberta Economic Development and

Trade, 2017). While the first investment elegibility date was in 2016Q2 (April 2016), investors could only start funding businesses starting 2017Q1, hence any effect on patent applications would be observed after this period. 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 I believe the AITC affected Albertan patent applications. The DD design is implemented in a regression framework, according to the general specification below.

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

where P_{it} is the explained variable; in most specifications, P_{it} is the number of patents filed in a province i and period t. θ_i and τ_t are sets of province and period fixed effects. I use a natural logarithm transformation along with the addition of one to correct for provinces with small amounts of patent applications on some periods. The logarithm will aid give coefficients in the right hand side a percent interpretation. 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 average treatment effect of the AITC on the explained variable. $\mathbf{x}_{it}$ 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 stochastic error term, which varies between provinces and periods. I cluster standard errors at the province and period level, as the variance of the error term may be correlated both spatially and temporally.

Tables 2 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 the DD design is that, absent the AITC intervention, the trend of patent applications in Alberta would follow a similar pattern to that in control

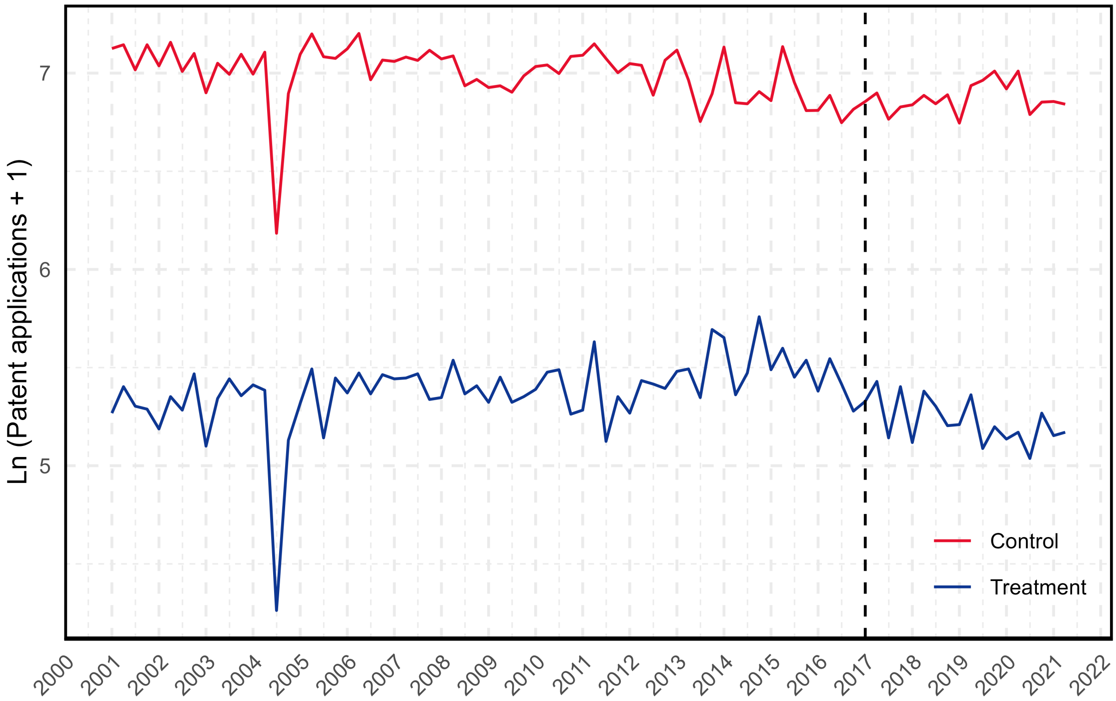
Table 2: Differences in means between treated and control provinces in province-quarter panel

Treatment		Pre	Post
Control	Ln +1 Patent applications	4.122	4.059
	Ln +1 Interested parties	5.683	5.464
	Ln +1 Inventors	301.511	292.143
	Ln +1 Applicants	158.663	139.122
	Ln +1 Owners	330.344	150.204
	Ln +1 Total population	8.632	8.731
	Ln +1 Section A applications	2.550	2.596
	Ln +1 Section B applications	2.238	2.037
	Ln +1 Section C applications	1.464	1.306
	Ln +1 Section D applications	0.349	0.185
	Ln +1 Section E applications	1.496	1.512
	Ln +1 Section F applications	1.618	1.417
	Ln +1 Section G applications	1.840	1.993
	Ln +1 Section H applications	1.620	1.380
	Ln +1 Multiple section applications	2.879	2.972
Treatment	Ln +1 Patent applications	5.379	5.254
	Ln +1 Interested parties	6.800	6.665
	Ln +1 Inventors	314.787	383.571
	Ln +1 Applicants	219.607	194.810
	Ln +1 Owners	378.721	211.762
	Ln +1 Total population	9.043	9.236
	Ln +1 Section A applications	2.704	2.896
	Ln +1 Section B applications	2.927	3.020
	Ln +1 Section C applications	2.632	2.633
	Ln +1 Section D applications	0.190	0.258
	Ln +1 Section E applications	4.073	3.761
	Ln +1 Section F applications	2.510	2.406
	Ln +1 Section G applications	3.075	2.853
	Ln +1 Section H applications	1.950	1.789
	Ln +1 Multiple section applications	4.100	4.111

Notes: Calculations based on a balanced panel of N = 656 province-monthly observations from 2001Q1 to 2021Q2. The sample includes all Canadian provinces except Newfoundland and Labrador, Prince Edward Island, Yukon and Nunavut. The treatment group is Alberta, and the control group is made from all remaining provinces. Post-intervention periods are those after April 2016 (2016Q2).

provinces. Figure 1 shows the quarterly time series of patent applications between Alberta and control provinces from 2001Q1 to 2021Q2. This visual representation of the trends shows that Alberta's patent applications follow a similar pattern to control provinces before the intervention, however, some deviations are present before 2016Q2.

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 from 2001Q1 to 2021Q2. The vertical line represents the start of the AITC intervention in 2017Q1. The treatment group is Alberta, and the control group is made from all remaining provinces except NL, PE, YT and NU.

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 $\hat{\beta}$.

$$\ln(P_{it} + 1) = \theta_i + \tau_t + \beta_t(t \cdot A_t) + \mathbf{x}_{it}^{'}\gamma + u_{it} \tag{2}$$

$\theta_i, \tau_t, \mathbf{x}_{it}, \gamma$ and u_{it} represent the same as in Equation 1. t is a set of binary variables for each of the periods for which there is data available, with the reference level set to one period before the AITC program start (2016Q4). A_t is a binary variable equal to unity if the observation is mapped to Alberta and zero otherwise. $t \cdot A_t$ is the interaction term between these two variables, and 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 their 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 interaction terms 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 average treatment effect of the AITC on patent applications. In section **??**, I justify the causal identification of the AITC effect by showing that the interaction terms before the intervention are not statistically significant. Further, I use event study regressions in the form of 2 to examine the effectiveness of the AITC by looking at post-treatment interaction terms, which should show statistically significant differences if the AITC affected Albertan patent applications.

1.3 Patent parties and province-month panel

I perform two main robustness checks on DD and event study analyses to ensure the validity of my results. First, to the extent that results may be driven by my 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 applicants[2].) to understand the effect of the AITC on the innovative capacity of different groups. Table 2 also presents the difference in means between treated and control provinces for parties 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. This panel allows for a more granular analysis of the effects of the AITC on patent applications, however, it is susceptible to noise due to the smaller number of observations, especially for the event study regressions. I present descriptive statistics of the monthly data in **??**.

[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.