

INTERNATIONAL MONETARY FUND

Public Perceptions of Canada's Investment Climate

A Natural Language Processing-Based Approach

Flora Lutz, Yuanchen Yang, Chengyu Huang

WP/24/165

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate.

The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

**2024
JUL**



WORKING PAPER

IMF Working Paper

Western Hemisphere Department

Public Perceptions of Canada's Investment Climate—A Natural Language Processing -Based Approach**Prepared by Flora Lutz, Yuanchen Yang, Chengyu Huang**

Authorized for distribution by Koshy Mathai

July 2024

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

ABSTRACT: Canada's muted productivity growth during recent years has sparked concerns about the country's investment climate. In this study, we develop a new natural language processing (NLP) based indicator, mining the richness of Twitter (now X) accounts to measure trends in the public perceptions of Canada's investment climate. We find that while the Canadian investment climate appears to be generally favorable, there are signs of slippage in some categories in recent periods, such as with respect to governance and infrastructure. This result is confirmed by both survey-based and NLP-based indicators. We also find that our NLP-based indicators would suggest that perceptions of Canada's investment climate are similar to perceptions of U.S. investment climate, except with respect to governance, where views of U.S. governance are notably more negative. Comparing our novel indicator relative to traditional survey-based indicators, we find that the NLP-based indicators are statistically significant in helping to predict investment flows, similar to survey-based measures. Meanwhile, the new NLP-based indicator offers insights into the nuances of data, allowing us to identify specific grievances. Finally, we construct a similar indicator for the U.S. and compare trends across countries.

RECOMMENDED CITATION: Y. Yang, F. Lutz and C. Huang, 2024. Public Perceptions of Canada's Investment Climate—A Natural Language-Based Approach, IMF Working Paper No. 165.

JEL Classification Numbers:	A10, C55, L82, Z13
Keywords:	Investment Climate; Canada; Machine Learning; Sentiment Analysis
Authors' E-Mail Address:	flutz@imf.org; yyang6@imf.org; chuang@imf.org

* The authors would like to thank Koshy Mathai, Lucy Liu, Tannous Kass-Hanna, Paola Morales Acevedo, Liam Masterson for their valuable comments.

WORKING PAPERS

Public Perceptions of Canada's Investment Climate

—A Natural Language Processing-Based Approach

Prepared by Flora Lutz, Yuanchen Yang, Chengyu Huang¹

¹ The authors would like to thank Koshy Mathai, Lucy Liu, Tannous Kass-Hanna, Paola Morales Acevedo, Liam Masterson, Alexander Amundsen and Samuel Jaime Pienknagura Looor for their valuable comments.

Contents

Contents	2
Introduction	3
What Do Survey Based Indicators Tell Us?	7
Regulatory Barriers to Firms Entry and Competition	8
Governance	9
Canada's Mining Sector	10
An NLP-Based Approach to Evaluating Public Perceptions	11
Data	11
Methodology	12
Results	14
Competition Policies	14
Governance	15
Infrastructure	17
Indigenous Communities and the Mining Sector	18
Canada and the United States Compared	20
Conclusions	23
References	24
Annex I. Survey Based Indicators	26
Annex II. NLP-Based Indicator	27
Annex III. Do Public Perceptions Matter for Investment Flows?	29
Regression Analysis	29
Random Forest	32

Introduction

Canada's muted productivity growth during recent years has echoed concerns that the Canadian investment climate is worsening relative to competitor countries. While effective policies and regulations can help attract productive and durable investments, create jobs, build stronger and more resilient value chains, and thereby lift productivity, a high regulatory burden on the industry can also cause bottlenecks to innovation, growth, and competitiveness. To limit unintended impacts of regulatory requirements on the competitiveness of Canadian businesses, the government has recently initiated a broader agenda towards regulatory modernization.¹ Against this background, this project aims to assess (i) how the public perception of the Canadian investment climate encourages or discourages private investment, (ii) how it has evolved over time and (iii) how it compares to international comparators.

The main novelty of this study is the development of a natural language processing (NLP) based indicator to measure trends in the public perceptions of Canada's investment climate. Specifically, we propose a new indicator based on a sentiment analysis mining the richness of Twitter accounts, one of the world's largest social networking platforms where 540 million users across the world share their views, opinions, and emotions in short texts called tweets. We focus on replies to tweets from official institutions such as the Bank of Canada and Natural Resources Canada, and adopt a keyword search strategy, based on a glossary containing the keywords of the survey-based indicators (see Annex II) to capture the ease of doing business. We then determine whether the public attitudes towards certain aspects of the investment climate are positive, negative, or neutral by performing a sentiment analysis on replies to the selected tweets.

We assess the performance of our novel indicator relative to traditional survey-based indicators by exploring a set of internationally comparable indicators related to regulatory barriers to firm entry and competition, governance, and political uncertainties. The selected indicators are largely survey-based, reflecting our focus on public perceptions. Moreover, compiled by well-established international organizations, such as the OECD and the World Bank, they are widely used and accepted in the literature, and focus on different dimensions of the investment climate such as the perception of governance, regulatory barriers in the product market, and risks.

Our new natural language processing (NLP hereafter) based measure of the investment climate offers three main advantages compared to the traditional survey-based indicators. First, it is less sensitive to biases in the views of experts and authorities, which are typically the main respondents of survey-based indicators. By analyzing a large set of replies to official tweets, our indicator is based on a broader and more diverse set of respondents, including experts as well as non-experts. Second, by using tweets up until the most recent period, our indicator provides a real-time indicator of the investment climate. This is clearly not the case for survey-based indicators which are

¹ In 2018, Canada introduced the [Targeted Regulatory reviews](#) to review existing regulations and identify novel approaches to support innovation and growth. In 2021, the Government of Canada committed to increase the Competition Bureau's budget by \$96 million over the following five years to enhance its enforcement capacity. In 2022, additional steps to strengthen the Competition Bureau's ability to protect and promote competitive markets were taken ([see 2022-2023 Annual Plan](#)).

usually updated only in relatively large intervals (e.g., every five years for the OECD Product Market Regulation Indicator) likely also due to the effort required to update survey-based indicators. Third, our indicator can easily be adjusted to cover additional and more nuanced dimensions of the investment climate.

It is important to note that our index, as well as several of the indices constructed by IFIs, focus on public perceptions of the investment climate, and do not necessarily reflect actual policy changes. We intentionally focus on public perceptions as they themselves are likely an important determinant of private investment decisions. It is further worth noting that in some cases an erosion of public perceptions could reflect an improvement in existing policies. For example, if perceptions are falling because of the unveiling of corruption cases, this could be a sign that the courts and the check and balances in place are functioning, despite of the perception of a worsening in institutions. These limitations are important to acknowledge when interpreting our findings.

The main findings of the analysis can be summarized as follows. First, survey-based indicators indicate that while the investment climate appears to be generally favorable in Canada, efforts in key areas such as competition policies and governance could help to boost private investment. Regulatory barriers appear to be above the OECD average, with weaknesses related to public procurement and SOE governance, as well as the licensing of new entrant businesses. Governance indicators are strong, but there are signs of slippage in recent years. Canada's mining sector continues to be an attractive investment destination, but policy uncertainty is an issue according to the annual survey of mining companies conducted by the Fraser Institute.

Second, we show that these findings are largely confirmed by our novel NLP-based sentiment analysis. Specifically, we find that over time, there has been a rising volume of grievances concerning Canada's competition policies and an increasing level of discontent with practices in this area, indicated by the increasing share of negative comments. Similarly, we observe a surge in the ratio of negative comments regarding governance, consistent with findings from the Worldwide Governance Indicators that indicated a decline in Canada's governance quality as of the most recent observation in 2021. Compared to existing survey-based indicators, our NLP-based indicator also yields several additional insights including the large public engagement with respect to indigenous policies, the heightened attention to competition policies in recent years as well as improvements in public perceptions related to governance and investment in 2022.

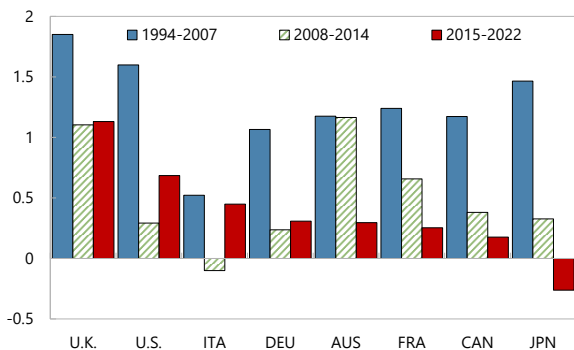
As a robustness check, we construct similar NLP-based indicators for the U.S. and compare the findings to the same set of survey-based indicators as well as across countries. Both indicators, survey-based as well as NLP-based, suggest that public perceptions with respect to governance are less favorable in the U.S. compared to Canada. While our NLP-based indicators point to similar public perceptions of issues related to competition policy in the two economies, survey-based indicators suggest a more competition friendly environment in the U.S. In the annex, we further assess if the public perception of the investment climate matters for actual investment outcomes in OECD countries using (i) a standard panel regression as well as (ii) a machine learning algorithm (random forest). In line with recent studies, we find evidence that public perceptions of the investment climate matter for actual investment outcomes (see e.g., Budina, 2023; IMF, 2023; and IMF, 2019).

Canada's weakening productivity and investment performance is evident from Figure 1 which plots different productivity measures and gross fixed capital formation relative to comparator countries. As can be seen, Canada's productivity growth has been muted during recent decades and weakened notably relative to the U.S. Specifically, labor productivity growth has remained slightly lower than in major comparator countries during most of the last three decades. During the early 2000s, Canada's labor productivity averaged 1.17 percent and was the third lowest among the G7 countries. From 2015–22, labor productivity growth dropped to 0.18 percent, clearly below the G7 average of 0.39 percent, and notably below the growth rates observed in leading economies including the U.S. (0.68) and U.K. (1.13). Similarly, multifactor productivity (MFP)² growth lagged, and remained below the G7 average. Relative to the U.S., labor productivity and gross fixed capital formation have weakened notably in recent years, especially since the collapse of oil prices in 2014. Gross fixed capital formation is now in the lowest quartile of OECD countries, while labor productivity too deteriorated significantly compared to U.S. and G7 levels, and is now close to the OECD average.³

Figure 1: Productivity Growth and Capital Formation in Canada and Comparator Countries

Canada: G7 Labor Productivity Growth

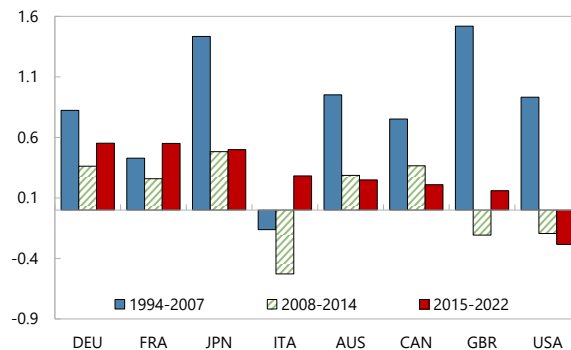
(Percent change, Period arithmetic average)



Sources: National Statistics, Haver Analytics and IMF staff calculations

Canada: Main Comparator Multifactor Productivity Growth

(Percent change, Period arithmetic average)



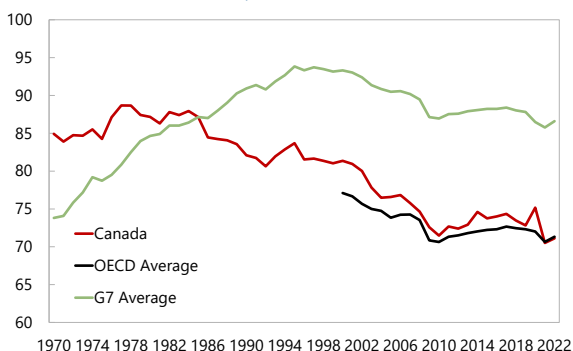
Sources: OECD and IMF staff calculations

² We recognize that MFP is a complicated measure, the outcome of which can vary due to methodological differences.

³ Labor productivity reflects GDP per hour worked. Relative productivities are evaluated in constant prices and adjusted for purchasing power.

Canada: Labor Productivity Relative to the US

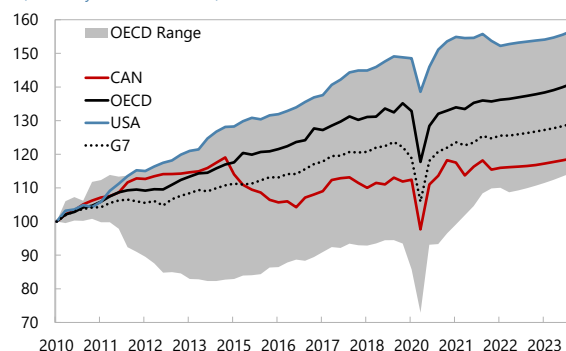
(Annual, Index, United States = 100)



Sources: OECD and IMF staff calculations

Canada: Gross Fixed Capital Formation

(Quarterly, Index 2010=100)



Sources: OECD and IMF staff calculations

The underlying drivers for the weakening productivity and investment performance are likely manifold, including an array of factors. Alvarez et al. (2019) highlight the importance of internal trade barriers in Canada estimating that complete liberalization of internal trade could increase GDP per capita by 4 percent. Similarly, the OECD (2023) highlights the unusually high inter-provincial non-tariff barriers in Canada. Amundsen et al. (2023) provide evidence for a significant increase in the share of zombie firms during the last two decades and argue that this zombification is increasingly lowering aggregate productivity growth. In the absence of rapid productivity growth, Canada has relied on skill-based immigration policies to support aggregate GDP growth, but recent studies indicate scope for improvement in addressing the country's specific labor needs. Liu, Lutz & Yang (forthcoming) further highlight capital misallocation during house price booms as another factor for the observed low productivity growth in Canada. Other factors that have been highlighted in the literature in different settings include financial frictions and high leverage (Kalemli-Özcan, Laeven, & Moreno, 2022), capital misallocation (Gopinath, Kalemli-Özcan, Karabarbounis, & Villegas-Sanchez, 2017), suboptimal infrastructure (e.g., OECD, 2015), economic uncertainty (e.g., Bloom et al., 2007 and Jarret et al., 2022), and, more recently, learning and productivity losses due to the pandemic (e.g., IMF, 2021).

In this study we focus on public perceptions as a complementary factor for the observed investment performance weakening in Canada. The most recent OECD survey (OECD 2023) highlights the need for better competition policy to help improve economic performance, including, in particular, challenges related to Big Tech and foreign ownership restrictions in Canada. Similarly, foreign ownership limits in some sectors, including airline, banking and cultural sectors are also highlighted by the U.S. Department of State (2022). In a cross-country setting, Budina et al. (2023) find that first generation reforms, including governance, business regulation and external sector reforms, can promote foreign investment and promote labor productivity in emerging market and developing economies, particularly when reforms are well sequenced. In a similar vein, recent work of the IMF has documented that reform pushes in areas such as governance, trade, labor, and product markets could deliver

sizeable output gains.⁴ Compared to these studies, we propose a novel NLP-based indicator measuring the public perception of the investment climate along several dimensions.

Finally, the study relates to a literature performing sentiment analysis in various settings. Sentiment analysis is an active field of research dedicated to improving the automated detection of sentiment within text, and machine learning-based sentiment analysis represents the state-of-the-art methodology (e.g., Barkema et al., 2021; Gonzalez-Garcia and Yang, 2022; Yang et al., 2023).⁵ It has been widely used to analyze market reactions (e.g., Baker et al., 2016, and Santi, 2023), product sales (Forman et al. 2008), and election outcomes (Tumasjan et al. 2010; O'Connor et al. 2010). While sentiment analysis has become a popular methodology due to its great flexibility, this is the first study to apply machine learning techniques to evaluate a country's investment climate in a comprehensive way, taking into account multiple dimensions. We take a supervised approach to analyzing vast quantities of Twitter data, which sheds light on various aspects of the investment sentiment in Canada.

What Do Survey Based Indicators Tell Us?

This section provides a comparison of the Canadian investment climate, both across countries and across time, based a set of internationally comparable, largely survey based, indicators. The selected indicators are compiled by well-established international organizations. Widely used and accepted in the literature, they focus on different dimensions of the investment climate. Specifically, the following four indicators are included: First, the [OECD indicators of product market concentration](#) (PMR) aim to measure the regulatory stance of an economy in an internationally comparable way by assessing barriers to firm entry and competition in a range of key policy areas (including licensing and public procurement, governance of SOEs, price controls, evaluation of new and existing regulations and foreign trade) and key individual sectors (network sectors, professional service and retail distribution). It is a de-jure indicator based on a questionnaire filled out by ministries and regulators. Second, the [Worldwide Governance indicators](#) compiled by the World Bank and the corruption perception indicators compiled by transparency international provide two survey based governance indicators covering different dimensions of governance including voice and accountability, regulatory quality, political stability, rule of law, government effectiveness and control of corruption. Third, the [International Country Risk Guide](#) provides monthly economic, financial, and political risk ratings for 141 countries combining survey-based indicators with other economic indicators. Finally, considering Canada's richness in critical "transition minerals" the findings of the [Annual Mining Survey](#) conducted by the Fraser Institute are also discussed.

⁴ See IMF World Economic Outlook (2019), chapter three as well as the IMF Regional Economic outlook, Middle East and Central Asia (2023).

⁵ Before the establishment of machine learning based algorithms, sentiment analysis largely relied on a lexicon approach counting the number of predefined happy and sad words. Compared to this traditional rule-based approach, machine learning approaches can train models on a massive amount of text using large trainings sets and can thereby learn to classify sentiment based on subtle cues on the text. One of the biggest hurdles for machine learning-based sentiment analysis is that it requires an extensive annotated training set to build a robust model.

Regulatory Barriers to Firms Entry and Competition

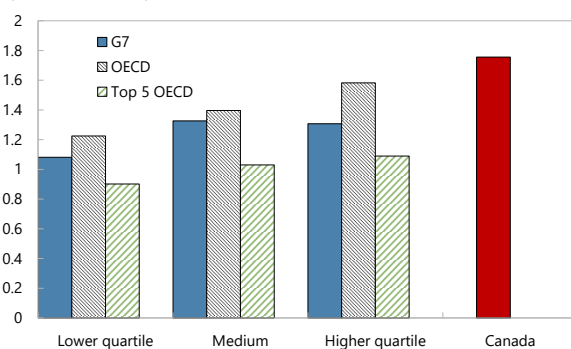
The OECD PMR Indicators aim to measure countries regulatory barriers to competition and relies on an extensive database including answers to questionnaires compiled by national authorities. The indicator is widely used by national governments, other international organizations (e.g., IMF, WB, and European Union), and international forums, such as the G20 and APEC, to determine areas in need for regulatory change. The left panel of Figure 2 plots the economy-wide PMR indicator, providing a general qualitative measure of a country's regulatory stance, for Canada as well as the average indicator for G7, OECD and the top-5 performing OECD economies. With Canada being in the highest quantile of G7 and OECD countries, regulatory barriers appear to be above the G7 and OECD average. Looking at PMR by network sector (right panel), a similar picture arises. Regulatory barriers to firm entry and competition appear to be above the G7 average in all sectors but remain below OECD averages.

Looking at the breakdown by sub-components of the economy-wide indicator reveals weaknesses related to public procurement and SOE governance (see Figure 3).⁶ Additionally, the licensing regime could better facilitate entry of new businesses and barriers to entry of foreign supplies are higher than the OECD average.⁷ On the positive side, regulatory procedures appear to be simple and administrative requirements necessary to set up new firms are low. Similarly, public ownership of large network operators is low.

Figure 2: Overall PMR Index

2018 PMR Economy Wide Indicator

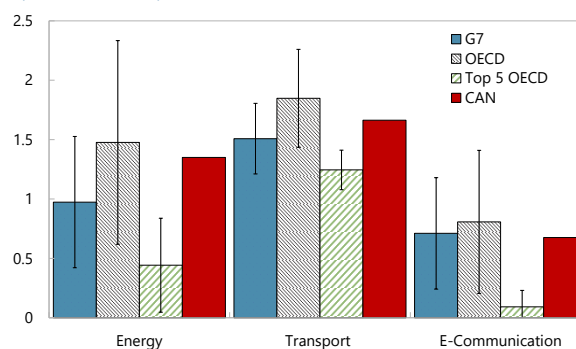
(Index, Min 0, Max 6)



Sources: OECD and IMF Staff Calculations

2018 PMR by Network Sector

(Index, Min 0, Max 6)



Sources: OECD and IMF Staff Calculations

Notes: A higher value indicates more statutory restrictions. The black lines indicate the confidence intervals around country group averages. The sector Energy includes Electricity and natural gas, Transport includes rail, road, water and air and E-communications includes the sectors fixed and mobile communications.

Against this background, authorities have recently started to set actions to improve Canada's competition policy. Following an increase in Competition Bureau Canada's budget in 2021, the government made targeted

⁶ According to the OECD (2018), the regulatory framework for the procurement of public work does not provide a level playing field for a bidder; state-owned enterprises benefit from exemptions from laws and benefit from better financing conditions in some states.

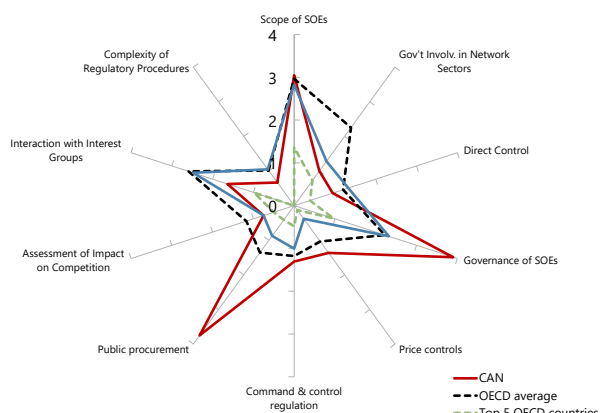
⁷ A one-stop shop for issuing and authorizing permits is not included in the licensing regime and there is no "silence is consent rule". (OECD, 2018)

improvements to Canada's competition act to strengthen the competition bureau's powers in June 2022 and launched the competition act review in November 2022.⁸

Figure 3: PMR Index by Sector

Distortions Induced by State Involvement

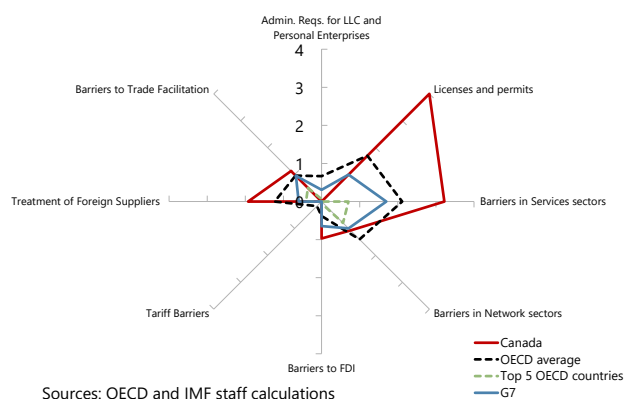
Index, Min=0, Max =6



Sources: OECD and IMF staff calculations

Barriers to Domestic and Foreign Entry

Index, Min=0, Max=6



Sources: OECD and IMF staff calculations

Governance

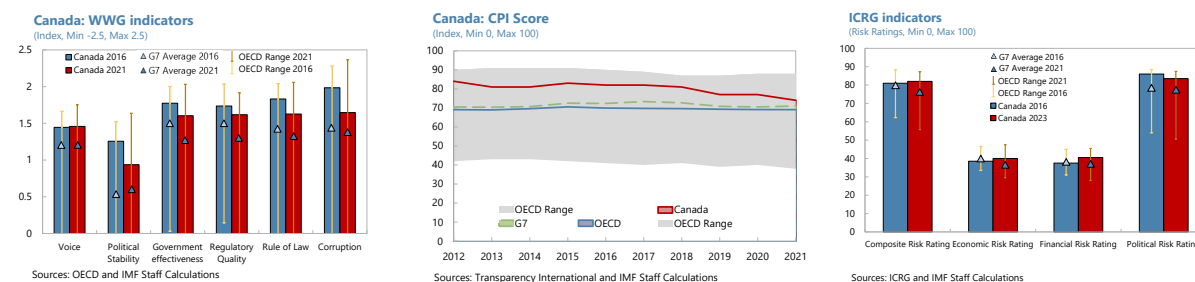
Another important dimension of the investment climate is governance. While governance is an equally broad concept as investment climate, the following discussion focuses on aspects covered by the *Worldwide Governance Indicators*: voice, political stability, government effectiveness, regulatory quality, rule of law and corruption. Canada's governance indicators are strong but there are signs of slippage in recent years. The countries' scores of the Worldwide governance indicators, compiled by the World Bank, report a decline in all categories except voice during the last five years. This decline is particularly notable in the category's government stability and corruption. A similar trend decline is reported by the Corruption Perception Index (CPI) and the political risk rating of the International Country Risk Guide (see Figure 4). The decline of Canada's CPI score can be largely explained by the money laundering problem that came into spotlight in recent years⁹, an issue, Canadian authorities are aware of and are currently working to address.¹⁰ Political risks seem to mainly arise from doubts about the durability of the current government.

⁸ [Minister Champagne maintains the Competition Act's merger notification threshold and extends the Competition Act consultation - Canada.ca](https://www.canada.ca/en/department-of-commerce-employment/news/2022/06/minister-champagne-maintains-the-competition-act-s-merger-notification-threshold-and-extends-the-competition-act-consultation.html)

⁹ [CPI 2019: Americas - News - Transparency.org, Cullen Commission Final Report Recommendations Mark Major Victory For Deterring Dirty Money From Entering Shell Companies and Shell Properties in Canada — Transparency International Canada \(transparencycanada.ca\)](https://www.transparencycanada.ca/en/news/cpi-2019-america-news-transparency.org-cullen-commission-final-report-recommendations-mark-major-victory-for-deterring-dirty-money-from-entering-shell-companies-and-shell-properties-in-canada)

¹⁰ Since 2019 the Government has provided \$319.9 million and \$48.8 million ongoing to strengthen the Anti-Money Laundering and Terrorist Financing Regime, including key investment areas such as the modernization of anti-money laundering laws, strengthening of inter-agency cooperation and information sharing, and equipping law enforcement with tools and expertise to support investigations.

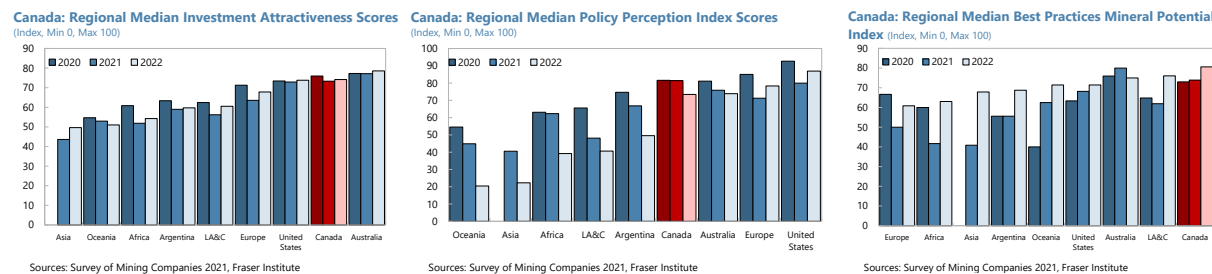
Figure 4: Governance Indicators



Canada's Mining Sector

Canada's mining sector continues to be an attractive investment destination, but policy uncertainty is an issue. The Fraser annual survey of mining companies assesses the perception of the attractiveness of mining jurisdictions based on government policies, geologic mineral potential, and overall attractiveness. Focusing on policy, although some provinces (e.g., Newfoundland & Labrador, Alberta, and New Brunswick) performed strongly according to the PPI, respondents continue to cite policy uncertainty—particularly around disputed land claims, protected areas, and environmental regulations—as a major deterrent to mining investment in Canada. Compared to 2021, respondents were less concerned about potential labor shortages, and government stability. Overall, Canada's median PPI decreased by almost 8 percentage points compared to 2021 from rank one in 2021 to rank four in 2022. The investment attractiveness index ranks Canada as the second most attractive region in the world in 2022, after Australia, reflecting the strong geologic attractiveness (ranked first according to the Best Practices Mineral Potential Index in 2022). The three jurisdictions ranked in the top 10 in terms of investment attractiveness remained unchanged compared to 2021, Saskatchewan (3rd), Newfoundland & Labrador (4th) and Quebec (8th).

Figure 5: Canada's Mining Sector Attractiveness



An NLP-Based Approach to Evaluating Public Perceptions

The primary aim of this section is to construct an indicator that systematically assesses public sentiment in relation to Canada's investment climate, based on user-generated content on Twitter, now known as X.¹¹ Sentiment analysis is an active area of research that seeks to enhance the automated recognition of sentiment in textual content. Contrary to news reports which tend to be predominantly neutral with relatively few expressing strong sentiment, positive or negative opinions are regularly expressed on the platform while neutral comments are comparatively less frequent. (see e.g., Sailunaz & Alhajj, 2019) The anonymity afforded to users on this platform further encourages the candid expression of their true thoughts. To conduct the sentiment analysis on Twitter, we draw on a state-of-the-art machine learning algorithm, trained on approximately 58 million tweets. Most importantly, we focus on retweets and replies to official government tweets, shedding light on people's genuine attitudes towards government policies. (Yang et al., 2023) In the following, we first describe the data and the empirical approach before we present the empirical findings.

Data

Twitter, or X, represents one of the world's largest and fastest-growing social networking platforms with over 540 million users who share their emotions, opinions, and views on a daily basis in short texts called tweets.¹² With one-quarter of Canadian internet users and 35.2% of social network users (compared with 19.5% and 27.9% of Americans) using the platform regularly it is even more popular in Canada than in the United States. User growth on the platform is predicted to rise by 2.7% in Canada and 1.9% in the United States in 2023. The number of active twitter users in Canada is summarized in Figure 7, underscoring the sharp increase in users during the last 5 years reflecting a broader global trend.

Twitter's large user base continuously produces diverse user-generated content, and offers a unique opportunity to access and understand the perspectives of users on different topics of interest. Several recent contributions relied on Twitter-based data to explain and predict various economic and financial phenomena, among which the most notable ones are stock market performance (Das and Chen 2007; Zimbra et al. 2015), political election outcomes (Tumasjan et al. 2010; O'Connor et al. 2010), and product sales (Liu, 2006; Forman et al. 2008). Like the current study, these analyses evaluate sentiments conveyed by users in their textual communications.

Our final sample comprises over 7500 official tweets, along with approximately 12500 replies or retweets corresponding to the official tweets. Therefore, each official tweet is, on average, associated with 1-2 replies. Our

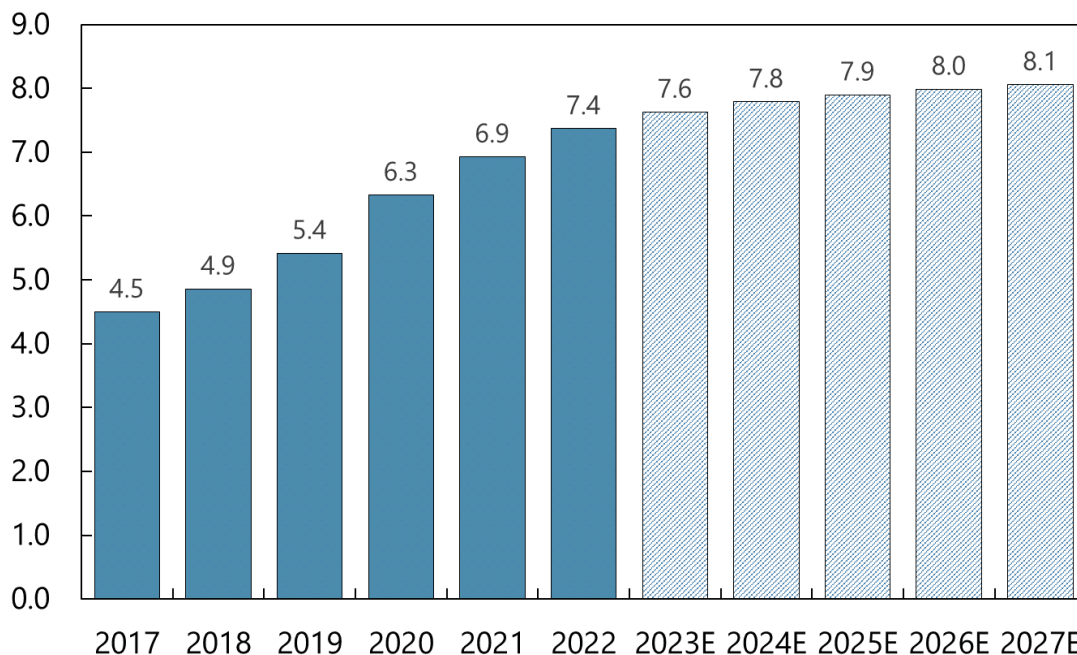
¹¹ At the time of data download, the official name of the platform was "Twitter" and we therefore use the term "Twitter" in the following analysis for consistency.

¹² <https://www.reuters.com/business/media-telecom/musk-says-x-monthly-users-reach-new-high-2023-07-28/>

discussion centers on three key dimensions of the investment climate—competition, governance, and infrastructure. Details about official accounts used for the search are provided in the Appendix.

Figure 6. Number of Active Twitter Users in Canada

Number of Twitter Users in Canada (in millions)



Sources: eMarketer, statista.

Methodology

The methodology is structured into four principal stages:

Stage 1: Data Collection

A list of verified Twitter¹³ accounts representing various government agencies (e.g., Department of Energy, Department of Commerce, Department of Labor; see Appendix for full list) was manually compiled. This dataset covers both Canada and the United States. Tweets, as well as retweets, were retrieved via Twitter's Full-Archive Search API, covering the period from early 2017 to the end of 2022.

Stage 2: Topic Identification

¹³ The methodology section maintains the use of Twitter as the data was collected and analyzed prior to the Twitter rebranding.

A refined list of keywords relevant to investment climate was developed, based on essential aspects of the investment climate summarized from survey-based indicators, to serve as topic identifiers. Only tweets containing at least one of these keywords were retained for subsequent analysis. The complete list of keywords used for the search are reported in Appendix II.

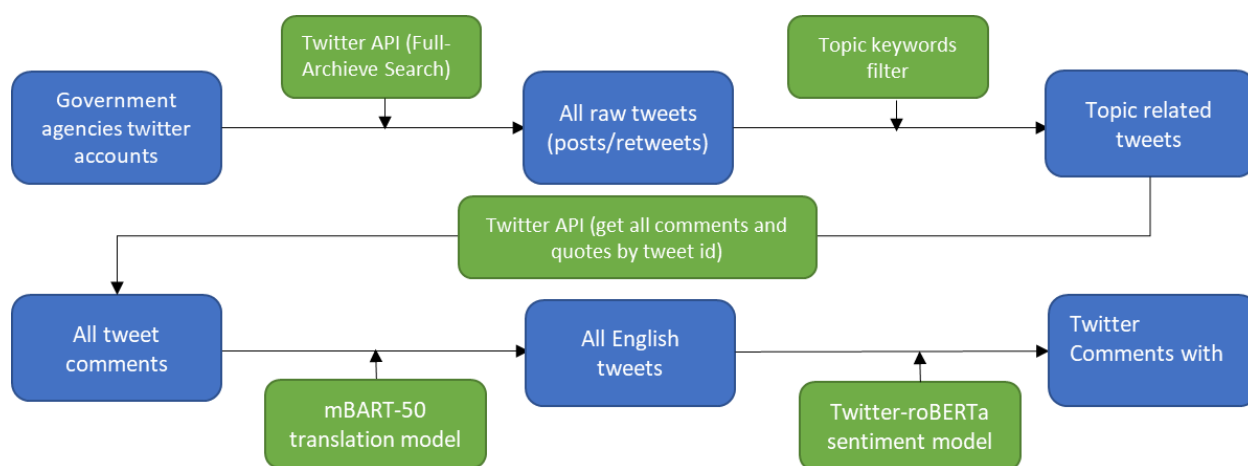
Stage 3: Comment Extraction

Following topic filtering, comments and quotes linked to each qualifying tweet were extracted using the same Full-Archive Search application programming interface (API), which enables bulk download of twitter data.

Stage 4: Sentiment Analysis

Sentiment analysis (also known as opinion mining, or emotion AI) is a method of analyzing text data to identify its intent. We conduct the sentiment analysis using the Twitter-roBERTa-base model¹⁴. Trained on more than 124 million tweets from January 2018 to December 2021, RoBERTa is a model specifically fine-tuned for Twitter data. Each comment was assigned a sentiment label (positive, negative, or neutral) along with a corresponding probability score. Overall, the Twitter-roBERTa model achieves an accuracy rate of about 72%.¹⁵ The collated data was then aggregated into an annual series for each country, facilitating a comprehensive analysis of public sentiment.

Figure 7: Flow Chart of Twitter Methodology



¹⁴ <https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest>

¹⁵ <https://paperswithcode.com/sota/sentiment-analysis-on-tweeteval?metric=ALL>

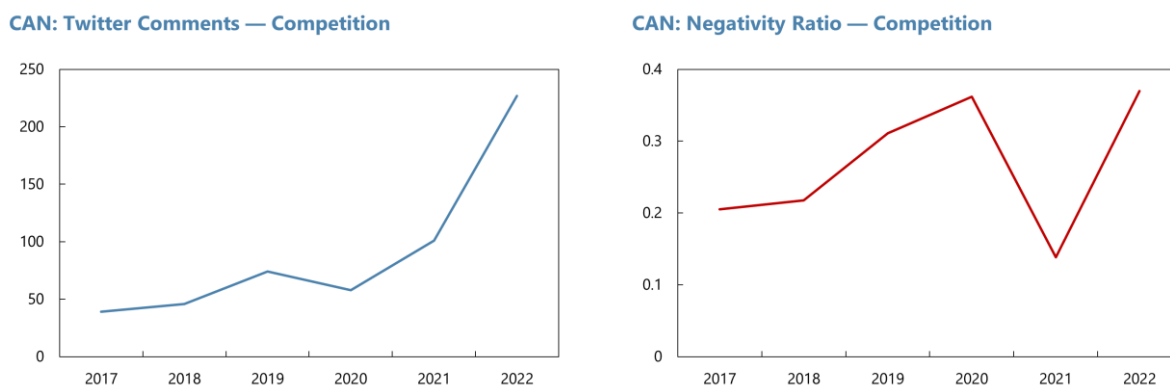
Results

In this section, we first present the findings of the NLP-based indicator related to the three main categories: competition, governance, and infrastructure.¹⁶ We then provide a sensitive analysis by compiling NLP-based indicators based on the same principles for the United States, Canada's by far most important trading partner and compare trends across countries.

Competition Policies

Applying natural-language processing techniques to the data, we find that competition may be an issue in Canada. There has been a growing focus on competition over the years, with the total number of tweets related to competition increasing nearly five-fold from January 2017 to December in 2022. This reflects both, an upward trend in official tweets by government agencies as well as in the replies by private users over the same period. At the same time, the negativity ratio, defined as the number of machine learning-identified negative comments over the total number of comments, has been on the rise and almost doubled over the same period (see Figure 8). We further observe a noticeable dip in the negativity ratio in 2021 likely at least partially related to temporal effects of the Covid-19 pandemic.

Figure 8: Selected Aspects of Twitter Investment Climate Analysis — Competition



Upon a more detailed examination of the comments identified as negative in the sentiment analysis, we find the discussions touch on some notable themes: discontent with the growing market power of large corporations, and complaints about the government support for conglomerates (see Table 1 for an illustrative example). Consistent with the results obtained using OECD's PMR survey, there are noticeable barriers to competition in Canada. Despite recent efforts by authorities to enhance Canada's competition policy, including a budget increase for the

¹⁶ While we also searched for other keywords such as political stability, product market regulation, and money laundering, we did not gather sufficient data to establish a meaningful trend. This, however, does not diminish the importance of these topics; it simply reflects a lower level of interest among Twitter users.

Competition Bureau Canada in 2021 and targeted amendments to Canada's competition act in June 2022, the negativity ratio continued to rise after a temporary decline in 2021.

Table 1. Illustrative Examples of Twitter Sentiment Labeling — Competition

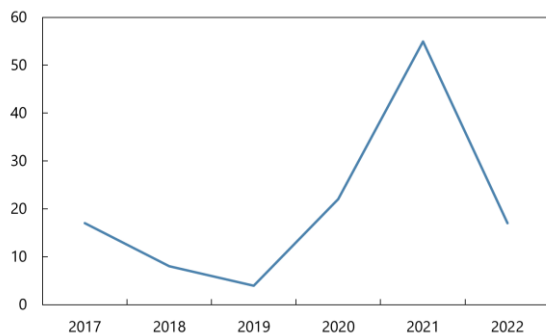
Type	Text	User Account	Date	Label
Original tweet	The Competition Bureau has reached an agreement today with Bell and Xplornet related to Bell's acquisition of MTS	Competition Bureau Canada @CompBureau	Feb 15, 2017	
Reply 1	could you clarify re "no penalties" for existing customers switching to Xplornet? #WirelessCode prohibits them anyway. ½			Neutral
Reply 2	Your own press release talks about how this deal will stifle competition and raise prices, but you still approved it? Useless			Negative
Reply 3	What is your mandate @CompBureau? Do you even know? You've destroyed competition in MB today. Way to go. #mbpoli			Negative

Governance

Turning next to the governance dimension, our Twitter content analysis suggests that prior to 2021, there was a rapid upward trend in the total number of tweets related to governance issues with the number of comments more than tripling since 2017 (see Figure 9). While we do observe a decline in the number of tweets in 2022, it is yet unclear if this reflects a short-term development or a longer-term trend. The ratio of negative comments on governance also peaked during the pandemic, in line with findings from the survey-based indicators that indicated a decline in the public perceptions of Canada's governance quality as of the latest available observation in 2021. Interestingly, the Twitter data show a significant improvement thereafter, as negative perceptions of governance declined substantially from its peak level at 0.27 in 2020 to 0.06 in 2022. Table 2 again provides two illustrative examples, this time relating to the broad concept of governance. It is also worth noting that the ML-based algorithm is able to adequately classify sentiment based on the subtle cues used in the replies.

Figure 9: Selected Aspects of Twitter Investment Climate Analysis — Governance

CAN: Twitter Comments — Governance



CAN: Negativity Ratio — Governance

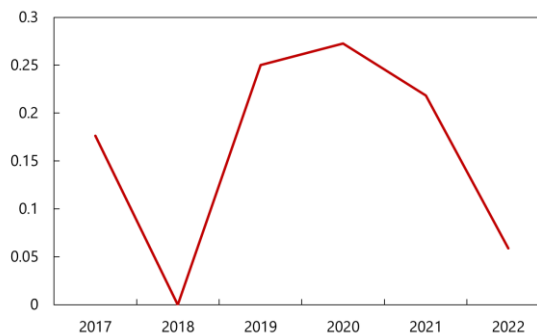


Table 2. Illustrative Examples of Twitter Sentiment Labeling — Governance

Type	Text	User Account	Date	Label
Example 1				
Original tweet	Today in Hamburg, PM Trudeau & President Peña Nieto discussed a number of issues including peace, security, good governance and trade. #G20	@CanadianPM	July 8, 2017	
Reply	I had to laugh at the "good governance" part. Justin could Cosplay Alice in Wonderland.			Negative
Example 2				
Original tweet	"Canada already leads the world on financial governance, standard-setting and sustainability." @Suncor CEO @itsMarkLittle joins @CPACanada CEO Charles-Antoine St-Jean in @GlobeBusiness to explain why Canada should host the global HQ of @IFRSFoundation 's ISSB. #MemberPerspective	Business Council of Canada @BizCouncilofCan	Oct. 1, 2021	
Reply	Canada leads the world on sustainability? Yikes. I didn't realize just how terribly the rest of the world is doing on sustainability.			Negative

Infrastructure

While the previous discussion of survey-based indicators did not explicitly include indicators related to the quality of infrastructure, mainly due to the lack of reliable data, we see one major appeal of the proposed NLP-based methodology in its great flexibility possibility as it can be easily extended to various aspects of the investment climate. Overall, we find that there is significant attention to infrastructure related discussions, and the total number of comments showed a significant uptick until 2020, suggesting heightened public engagement with the topic. (Figure 10) The negativity ratio is relatively stable but tends to be above the negativity ratios observed for other categories such as competition and governance.

Table 3 provides two illustrative example tweets that are extracted by the model when focusing on infrastructure. It is interesting to note that while the second original tweet clearly relates to the quality of infrastructure at a general level, reply 2 indicates concerns with respect to procurement policies and only indirectly relates to the quality of infrastructure. The negative labeled reply still enters the infrastructure negativity ratio as the algorithm categorizes tweets only based on the original tweets, highlighting one potential weakness of our NLP-based indicators.

Figure 10: Selected Aspects of Twitter Investment Climate Analysis — Infrastructure

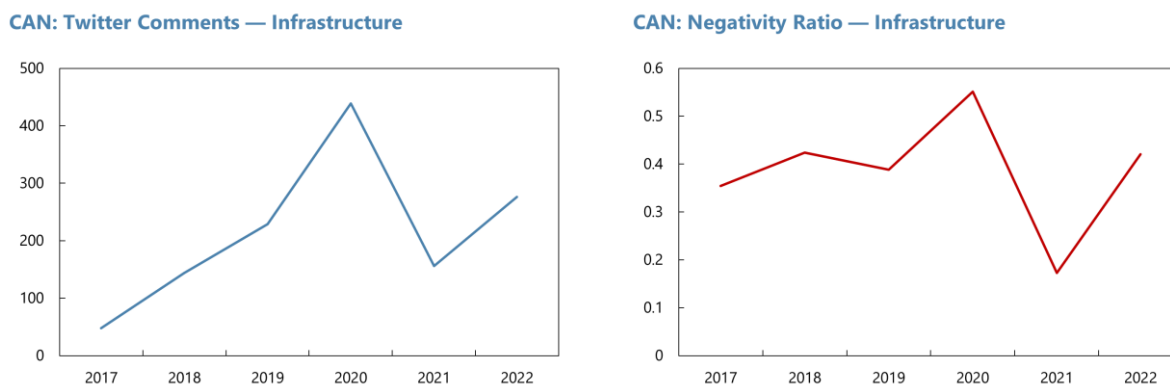


Table 3. Illustrative Examples of Twitter Sentiment Labeling — Infrastructure

Type	Text	User Account	Date	Label
Example 1				
Original tweet	Yesterday in the House of Commons, Prime Minister Justin Trudeau delivered a speech on the ongoing infrastructure disruptions caused by blockades across the country.	@CanadianPM	Feb 19, 2020	

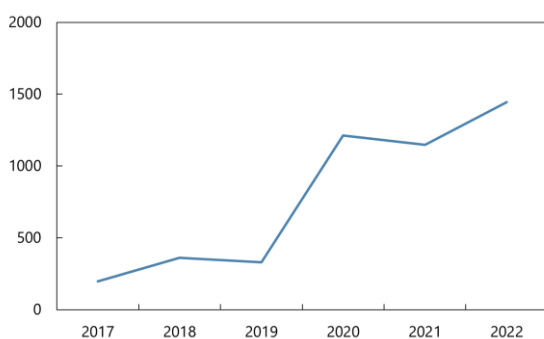
Reply 1	He delivered word salad! Even old pierre would be ashamed!		Negative
Reply 2	Cool...		Positive
Example 2			
Original tweet	Today, Prime Minister Justin Trudeau announced new major infrastructure initiatives to create jobs and economic growth. Learn more: http://ow.ly/xrpj50BGNsw	@CanadianPM	Oct 1, 2020
Reply 1	Good		Positive
Reply 2	How much of this will be allocated or sub contracted to Small businesses across Canada? Today's procurement policies make it very difficult for small to medium businesses to win a share of this business.		Negative

Indigenous Communities and the Mining Sector

In the following, we zoom into the public perceptions of the investment climate by specifically focusing on policies related to Indigenous groups. As part of a wider push towards more environmentally and socially responsible policies, shareholders, particularly those in the mining sector, are increasingly aware of indigenous rights and the critical need for indigenous participation. The findings are summarized in Figure 11.

Figure 11. Selected Aspects of Twitter Investment Climate Analysis — Indigenous Population

CAN: Twitter Comments — Indigenous



CAN: Negativity Ratio — Indigenous

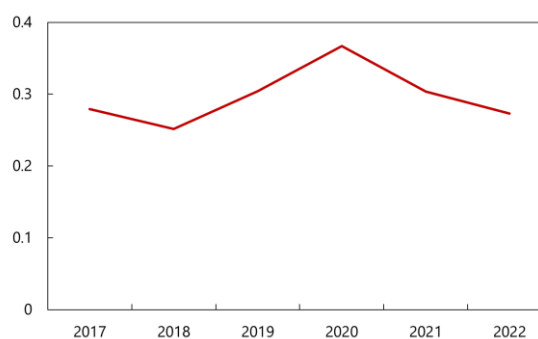


Table 4. Illustrative Examples of Twitter Sentiment Labeling — Indigenous

Type	Text	User Account	Date	Label
Example 1				
Original tweet	Minister Sohi announces funding for 3 #forestry projects in collaboration with Indigenous peoples in #AB to increase #ClimateResilience, create jobs and boost the local economy. #CEMMI2019 https://canada.ca/en/natural-resources-canada/news/2019/05/canada-invests-in-indigenous-participation-in-the-forest-sector.html...	Natural Resources Canada @NRCan	May 21, 2019	
Reply	You are the best for me. Vous commencez à agir			Positive
Example 2				
Original tweet	Today, PM Trudeau and Yukon Premier Silver discussed infrastructure investments and strengthening ties with Indigenous Peoples. #ytpoli	@CanadianPM	Feb 8, 2017	
Reply 1	time the indigenous people get smart and start to cooperation with government			Positive

Interestingly, it turns out that indigenous is the key word that obtained most attention among all analyzed key words during our sample period. Overall attention peaked in 2022 at almost 1500 tweets about three times the maximal number of tweets observed for the key words receiving the second most attention (infrastructure). We further find that more attention was paid to indigenous communities in recent years, likely reflecting the intention of the authorities to fully engaged them in the process of reviewing and monitoring major resource development projects. Unfavorable views of indigenous policies in Canada remained relatively unchanged from 2017 to 2022 and the level of the negativity ratio is similar to ratios observed for other categories. These findings align well with the continued attention of survey respondents to disputed land claims, protected areas, and environmental regulations as indicated by the Fraser Institute mining survey results.

While Twitter serves as a valuable tool for high-frequency indices of perceptions, there are important considerations when interpreting the results. First, there is a growing concern about the influence of Twitter bots in shaping political discussions on the platform. Second, despite the increasing number of Twitter users in Canada, their representativeness in relation to the average Canadian remains uncertain. For example, individuals with extreme views are more inclined to comment. Third, the continuous growth of users raises concerns about the evolving nature of the new user base. Unfortunately, due to data limitations, we cannot access the demographic characteristics, or trace the posting history of every Twitter user in the sample. However, we did manage to categorize the sample into existing users and new users who responded for the first time to official tweets. The

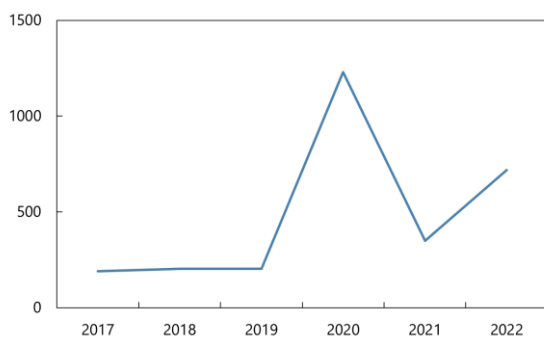
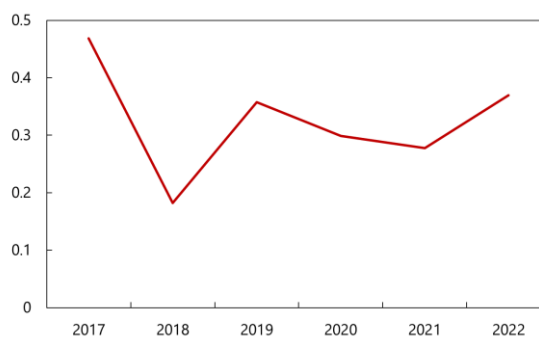
findings revealed no significant difference in the overall negativity ratio between these two groups (28.2% vs. 25.3%).

Canada and the United States Compared

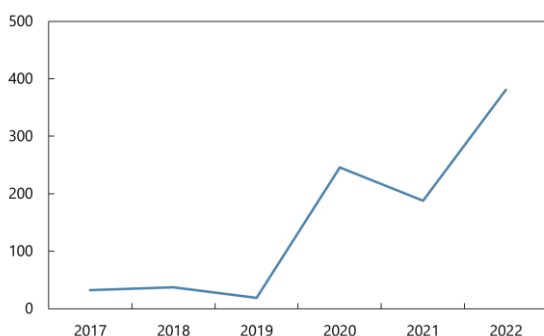
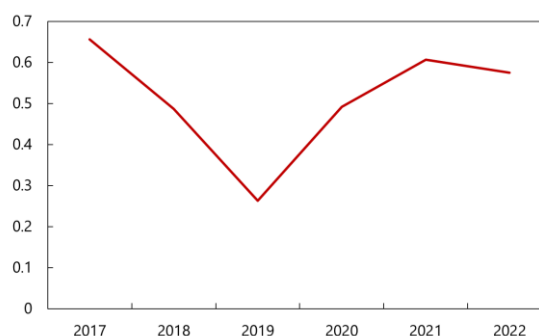
The earlier section has demonstrated that survey-based indicators indicate that Canada's regulatory and governance performances appear to be lagging compared to those of other G7 countries. In this section, we focus on one additional major advanced economy—the United States—Canada's by far most important trading partner and use Twitter data to compare investment sentiment across these countries. Nonetheless, it is crucial to interpret the findings with a measure of caution, considering that the pandemic may have varying impacts on different countries. Additionally, the total number of comments across countries may not be entirely comparable due to differences in Twitter's popularity in various regions or variations in usage by different official agencies. Results further crucially depend on the set of selected official institutions for which replies to official tweets are analyzed. To mitigate these concerns, we focus on the comparison of negativity ratios but do not compare overall numbers of tweets.

The U.S. sample comprises over 9000 tweets from official institutions including the U.S. Department of Commerce, U.S. Trade Representative, Federal Trade Commission, Washington International Trade Association, U.S. Chamber, U.S. Minority Business Development Agency, Internal Revenue Service, Federal Reserve, Department of Energy, Department of Labor, Department of Justice, and U.S. Environmental Protection Agency. A total of 28772 responses have been generated in reaction to tweets on selected topics.

While we find evidence for growing concerns over competition policies in Canada, reflected by the rising negativity ratio in the country, the situation has remained relatively stable in the United States. Absolute negativity ratios were similar across the two economies in 2022 at around 0.4 pointing to a similar public perception of issues related to competition policy across the two economies. Interestingly, we also find similar negativity ratios in both countries in 2018. This finding slightly contrasts the findings of the 2018 OECD PMR indicators suggesting a competition friendlier environment in the U.S with an economy wide PMR indicator in the medium quantile versus in the upper quantile for Canada (see Figure 12).

Figure 12: Selected Aspects of Twitter Investment Climate Robustness Check — Competition**US: Twitter Comments — Competition****US: Negativity Ratio — Competition**

As shown in the previous section, Twitter content analysis suggests that prior to the pandemic, there was an upward trend in negative perceptions of governance in Canada, which subsequently declined following the pandemic. In the United States, on the other hand, we observe a significant downward trend in the negativity ratio before the onset of the pandemic from 0.65 to less than 0.3 in 2019 before it reverted to 0.6 in 2021. The increase in the negativity ratio was accompanied by an increase in the number of related tweets. Interestingly, we find that the overall level of the negativity ratio, however, is significantly larger in the United States, particularly since the onset of the pandemic. This finding is in line with the evidence provided by the WWG indicators which show consistently lower scores for the United States across all six governance categories (see Figure 13).

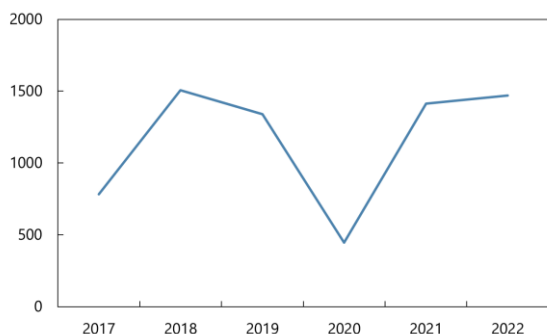
Figure 13: Selected Aspects of Twitter Investment Climate Robustness Check — Governance**US: Twitter Comments — Governance****US: Negativity Ratio — Governance**

Regarding infrastructure policies, we find evidence for a modestly declining negativity ratio in the United States while it has remained largely stable in Canada during recent years. The negativity ratios at a broadly similar levels in Canada and the United States but the recent downward trend in the U.S. has pushed the negativity ratio below

the one observed in Canada in 2022 (0.33 in the U.S. vs. 0.42 in Canada) providing some suggestive evidence that the public perceptions with respect to Canada's infrastructure have weakened slightly compared to the U.S. recently.

Figure 14: Selected Aspects of Twitter Investment Climate Robustness Check — Infrastructure

US: Twitter Comments — Infrastructure



US: Negativity Ratio — Infrastructure

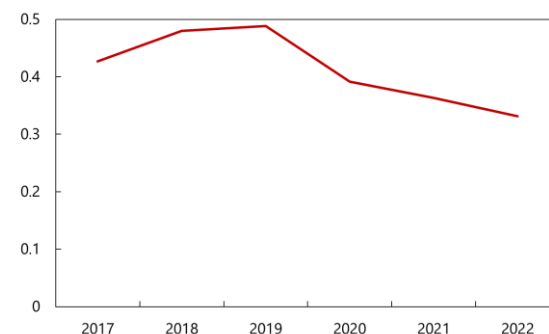
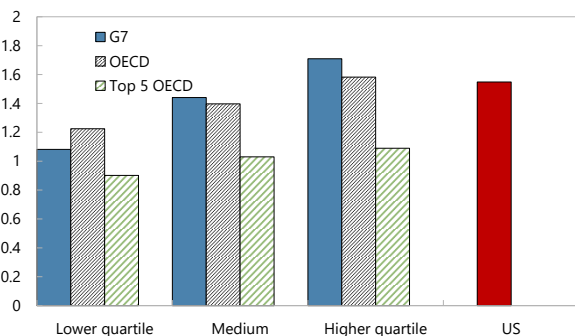


Figure 15: Survey-based indicators for the United States

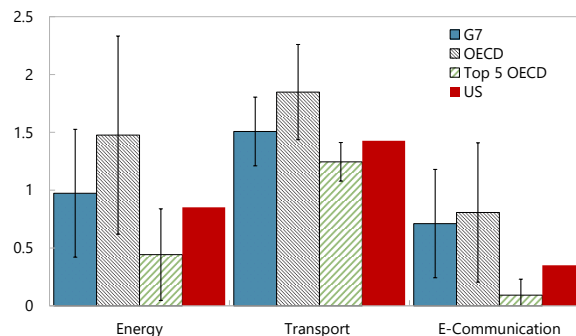
2018 PMR Economy Wide Indicator

(Index, Min 0, Max 6)



Sources: OECD and IMF Staff Calculations

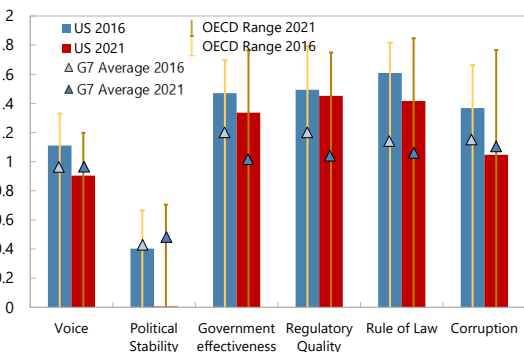
2018 PMR by Network Sector



Sources: OECD and IMF Staff Calculations

US: WWG indicators

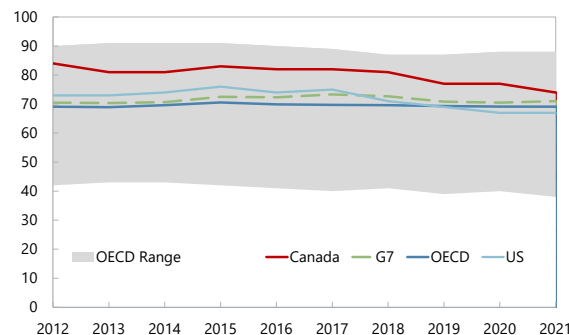
(Index, Min -2.5, Max 2.5)



Sources: OECD and IMF Staff Calculations

CPI Score

(Index, Min 0, Max 100)



Sources: Transparency International and IMF Staff Calculations

Conclusions

The slow productivity and investment growth observed in Canada during recent decades is attributed to various factors. In this paper, we focus on one potential factor, the public perceptions of the investment climate. Our main contribution is to develop a new NLP-based mining the richness of Twitter accounts to measure trends in the public perceptions of Canada's investment climate. Compared to existing indicators, our measure is (i) easier to update and adjustable to include additional countries or dimension of the investment climate and (ii) less likely to be biased by views among a certain group of individuals.

We find that while the Canadian investment climate appears to be generally favorable, there are signs of slipping performances in some categories in recent periods. This result is confirmed by both survey-based and NLP-based indicators. Compared to existing survey-based indicators, our NLP-based indicator yields several additional insights including the large public engagement with respect to indigenous policies, the heightened attention to competition policies in recent years as well as improvements in governance dimensions in 2022.

In summary, the novel measure of the investment climate proposed in this study provides valuable insights into the trends and patterns of public perceptions related to competition, governance, infrastructure, and policies related to the indigenous population in Canada. However, it is important to note that while the capacity of machine learning based models to process Twitter data facilitates a deeper understanding of user sentiments, they can be limited in handling evolving language or context-specific meanings. Sentiment analysis may categorize complex opinions into simplistic positive or negative labels, which can lead to overgeneralization and loss of nuanced understanding. For future research, there is scope to draw on the recently available large language models such as GPT to capture the full range of emotions and cultural nuances and enable more comprehensive insights into the investment climate in different countries. Finally, it is important to note our limited sample size (2017-2022) as well as the fact that some of the observed trends could be related to temporary developments driven by the Covid-19 pandemic.

References

- Amundsen, A., Lafrance_Cooke, A., and Leung, D., 2023. Winter is Coming? Zombie Firms and Ownership Type in Canada (June 12, 2023). Available at <http://dx.doi.org/10.2139/ssrn.4476595>
- Asteriou, Dimitrios, Simon Price, 2005. Uncertainty, Investment and Economic Growth: Evidence from a Dynamic Panel. *Review of Development Economics* 9(2): 277-288.
- Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring Economic Policy Uncertainty. *Quarterly Journal of Economics*, 113(4); 1593-1626.
- Barkema, J., Mircheva B., Mrkaic M., Yang Y., 2021. License to Spill: How Do We Discuss Spillovers in Article IV Staff Reports. IMF Working Paper No. 2021/134.
- Barrela, R., Paloma L.G., and Setzer R., 2022. Medium-term investment responses to activity shocks: the role of corporate debt, *ECB Working Paper* No. 2751.
- Budina, N., Ebeke, C., Jaumotte, F., Medici, A., Panton, A., Tavares, M., and Yao, B., 2023. Structural Reforms to Accelerate Growth, Ease Policy Trade-offs, and Support the Green Transition in Emerging Market and Developing Economies. *IMF Staff Discussion Note*.
- Das S., and Chen M., 2007. Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science*. 53 (9): 1375–1388.
- Forman, C., Ghose A., and Wiesenfeld B., 2008. Examining the relationships between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information System Research*. 19 (3).
- Furceri, D., and Li, G., 2017. The Macroeconomic (and Distributional) Effects of Public Investment in Developing Economies. *IMF Working Paper*.
- Geng, N., and N'Diaya, P., 2012. Determinants of Corporate Investment in China: Evidence from Cross-Country Firm Level Data. *IMF Working Paper*.
- Gonzales-Garcia J., and Yang Y., 2022. The effect of trade on market power-evidence from developing countries. *The Journal of International Trade & Economic Development* 31(6): 811-834.
- Gonzales-Garcia J., and Yang Y., 2020. International trade and corporate market power. *IMF Working Paper* 2020/131.
- Gopinath, G., Kalemli-Özcan, Ş., Karabarbounis, L., and Villegas-Sanchez, C., 2017. Capital Allocation and Productivity in South Europe. *Journal of Quarterly Economics*, 1915-1967.
- Guimaraes, R., and Unterberdoerster, O., 2006. What's Driving Private Investment in Malaysia? Aggregate Trends and Firm-Level Evidence. *IMF Working Paper*.
- Institute, F., 2022. Survey of Mining Companies.
- Jardet, C., Jude C., and Chinn M., 2022. Foreign direct investment under uncertainty evidence from a large panel of countries. *Review of International Economics* 31(3): 854-885.
- Kalemli-Özcan, S., Laeven, L., and Moreno, D., 2022. Debt Overhang, Rollover Risk, and Corporate Investment: Evidence from the European Crisis. *Journal of the European Economic Association*, 2353-2395.
- Liu, Y., 2006. Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of Marketing* 70, 74–89.
- Luca, O., and Spatafora, N., 2012. Capital Inflows, Financial Development and Domestic Investment: Determinants and Inter Relationships. IMF Working Paper.
- OECD. (2018). OECD Product Market Regulation (PMR) Indicators: How does Canada compare? 2018 PMR Country Note.
- OECD., 2023. OECD Economic Surveys: Canada 2023.

- O'Connor, B., Balasubramanyan R., Routledge B., and Smith N.. 2010. From tweets to polls: Linking text sentiment to public opinion time series. *In Proceedings of the AAAI Conference on Weblogs and Social Media*. 122–129.
- Rogoff, K., and Yang, Y., 2021. Has China's Housing Production Peaked? *China and the World Economy* 21 (1): 1-31.
- Rogoff, K., and Yang, Y., 2022. A Tale of Tier 3 Cities, *NBER Working Paper* No. 30519.
- Rossel, C., Dowsett, K., and Patterson, N. 2023. A Critical Juncture: Assessing Canada's Productivity Performance and Future Prospects. *International Productivity Monitor*, 45, 61-92.
- Sailunaz, K., & Alhaji, R., 2019. Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*.
- Santi, C., 2023. Investor Climate Sentiment and Financial Markets, *International Review of Financial Analysis*, 86(C).
- Simon M., 2023. Key Predictors for Climate Policy Support and Political Mobilization: The Role of Beliefs and Preferences, *Plos Climate* 2(8): e0000145.
- State, U. D. (n.d.). 2020. *Investment Climate Statements: Canada*.
- Tumasjan, A., T. Sprenger, P. Sandner, and I. Welpe. 2010. Predicting elections with Twitter: What 140 characters reveal about political sentiment. *In Proceedings of the AAAI Conference on Weblogs and Social Media*. 178–185.
- Wright M.N., Ziegler A., 2017. A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software* 77(1):1–17.
- Yang, Y., 2021. Are Passive Institutional Investors Engaged Monitors or Risk-Averse Owners? Both! *IMF Working Paper* No. 2021/158.
- Yang, Y., Huang, C., Zhang, Y., 2023. Decomposing Climate Risks in Stock Markets. *IMF Working Paper* No.2023/141.
- Yu, P., P., Zhang, H., Kang, H., Tang, W., Arnold, C. W., & Zhang, R., 2022. RPLHR-CT Dataset and Transformer Baseline for Volumetric Super-Resolution from CT Scans, *International Conference on Medical Image Computing and Computer-Assisted Intervention*: 344-353.
- Zhang, H., Polson, J., Nael, K., Salamon, N., Yoo, B., Speier, W. and Arnold, C., 2021. A Machine Learning Approach to Predict Acute Ischemic Stroke Thrombectomy Reperfusion Using Discriminative mr Image features. *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*.
- Zhang, H., Polson, J. S., Wang, Z., Nael, K., Rao, N. M., Speier, W. F., & Arnold, C. W., 2024. A Deep Learning Approach to Predict Recanalization First-Pass Effect following Mechanical Thrombectomy in Patients with Acute Ischemic Stroke. *American Journal of Neuroradiology*.
- Zimbra, D., Chen H., and Lusch R. F.. 2015. Stakeholder Analyses of Firm-related Web Forums: Applications in Stock Return Prediction. *ACM Transactions on Management Information Systems* 6 (1), 1–38.

Annex I. Survey Based Indicators

Appendix Table 1: Survey Based Indicators

Indicator	Compiled by	Methodology	Frequency	Data Source
PMR	OECD	Survey Based	Five-yearly	Indicators of Product Market Regulation - OECD
Worldwide Governance Indicators	World Bank	Survey Based	Annual	WGI 2022 Interactive > Home (worldbank.org)
International Country Risk Guide	PRS Group	Various Sources	Monthly	International Country Risk Guide (ICRG) The PRS Group (imf.org)
Corruption Perception Index	Transparency International	Survey Based	Annual	2021 Corruption Perceptions Index - Explore the... - Transparency.org
Fraser Annual Survey of Mining Companies	Fraser Institute	Survey Based	Annual	Annual Survey of Mining Companies, 2022 Fraser Institute

Annex II. NLP-Based Indicator

Appendix Table 2. Official Canadian Accounts Included in the Twitter Analysis

Agency	Twitter Account @
Bank of Canada	bankofcanada
Business Council of Canada	BizCouncilofCan
Canada Energy Regulator	CER_REC
Canada Revenue Agency	CanRevAgency
Canadian Intellectual Property Office	CIPO_Canada
CanadianPM	CanadianPM
<u>Competition Bureau Canada</u>	<u>CompBureau</u>
Conflict of Interest and Ethics Commissioner	EthicsCanada
Employment and Social Development Canada	ESDC_GC
<u>Finance Canada</u>	<u>FinanceCanada</u>
GCIIndigenous	GCIIndigenous
Impact Assessment Agency of Canada	iaac_aeic
Justice Canada	JusticeCanadaEN
Natural Resources	NRCan
Ontario Municipal Affairs	ONmunicipal
Superintendent of Financial Institutions	OSFICanada
The Official PDAC	the_PDAC

Appendix Table 3. Twitter Keyword List

<p>Competition barriers+entry barriers+non-tariff competition competition+law competition+policy competitiveness FDI+barriers FDI+restrictions foreign direct investment + barriers foreign direct investment + restrictions foreign suppliers price control tariffs taxes</p> <p>Governance accountability command and control corruption direct control governance government effectiveness government efficiency government involvement regulation regulatory quality rule of law</p>	<p>Infrastructure electricity infrastructure international shipments logistics network tracking timeliness utilities</p> <p>Indigenous indigenous first nations aboriginal</p> <p>Stability absence of terrorism absence of violence conflict uncertainty+economic uncertainty+economy uncertainty+political uncertainty+politics uncertainty+policy stability+government stability+political</p>	<p>Administrative burden administrative burden administrative requirement contracts license permit procedure</p> <p>Money laundering AML anti-money laundering CFT money launder money laundering</p> <p>Others business efficiency business operations credit environmental impact insolvency interaction with interest groups minority investors product market product market+regulation property rights protection public ownership public procurement SOE start a business</p>
---	---	---

Annex III. Do Public Perceptions Matter for Investment Flows?

In this section, we use various survey-based indicators of the investment climate including the worldwide governance indicators and the international country risk guide to shed additional light on the impact of the investment climate on investment flows. (Guimaraes and Unteroberdoerster, 2006; Geng and N'Diaya, 2012; Luca and Spatafora, 2012; Furceri and Li, 2017) Specifically, we apply two models, a standard econometric analysis and machine learning algorithm (random forest) to study the determinants of investment flows in OECD countries. The models relate real gross fixed capital formation to domestic factors such as changes in real interest rates, real GDP growth, inflation, general government debt, external variables such as the terms of trade, the real exchange rate, the current account, and global factors such as the federal funds rate and the VIX index. To measure the investment climate, we look at the following dimensions—voice and accountability, political stability, government effectiveness, regulatory quality, rule of law, control of corruption, size of government, legal system and property rights, sound money, freedom to trade internationally, and economic regulation. Our sample includes all OECD countries and ranges from 2000-2022.

Regression Analysis

We first estimate a panel regression to evaluate important drivers of investment flows. Specifically, we estimate the following regression.

$$\Delta \ln(\text{GFCE})_{it} = \alpha_{it} + \beta_{it} X_{it-1} + \delta_{it} Y_{it-1} + U_i + \varepsilon_{it}, \quad (1)$$

where the dependent variable is the percentage change in gross fixed capital formation, X captures the standard economic factors (see Table 4), Y captures the variables related to the investment climate (see Table 5), u captures the unobserved fixed effects and ε the error term. It is important to note that our results should be interpreted as an indication of correlation rather than causation as the specification likely involves endogeneity (variables such as e.g., GDP and investment are determined simultaneously). While the concern might be partially mitigated by using lags of the explanatory variables, fully addressing the concern would require an instrumental variable approach or an event study. Instead, we focus on a descriptive analysis highlighting the importance of the public perception of the investment climate for actual investment outcomes and show that these results are surprisingly robust to different specifications and modelling techniques.

The results of the baseline specification, including only national aggregate accounts variables, are summarized in Table 4. Overall, the results are broadly in line with our expectations. First, global factors seem to be a major factor impacting real gross fixed capital formation in OECD countries. Both, tighter US monetary conditions and higher global risk aversion are adversely correlated with investment. Second, among the external variables, an improvement in the terms of trade tends to increase gross fixed capital formation while an exchange rate

depreciation has the opposite effect.¹⁷ Finally, domestic factors, in general, seem to be of lower importance. While the estimated coefficients of real GDP and general government debt have the expected sign, they are only significant in the last two specifications. We find that these results are robust to alternative specifications excluding post-Covid-19 periods or non-G7 countries.

Appendix Table 4: Baseline Regression Results

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Short-term interest rate	0.0 (0.7)	0.0512 (0.814)	0.815 (0.709)	0.747 (0.660)
Real GDP	1.3 (2.7)	3.487 (2.335)	5.233** (2.318)	7.305*** (1.825)
Consumer prices, period average	-0.2 (0.5)	0.0664 (0.291)	0.102 (0.332)	0.179 (0.190)
General government debt, % of GDP	-0.1 (0.1)	-0.0873 (0.0732)	-0.147* (0.0783)	-0.125* (0.0676)
Total Terms of Trade	0.3*** (0.1)	0.302*** (0.0769)	0.194*** (0.0467)	0.236*** (0.0409)
Real Exchange Rate	-0.0*** (0.0)	-0.0373*** (0.0112)	-0.0367*** (0.0130)	-0.0330*** (0.0107)
Current Account, % of GDP	0.8 (1.9)	0.482 (1.771)	0.543 (1.883)	0.322 (1.732)
VIX Index	-0.0*** (0.0)	-0.00460*** (0.000843)	0.00815*** (0.00270)	0.00998*** (0.00279)
Federal Funds Rate	-0.0*** (0.0)	-0.0220*** (0.00492)	-0.0513*** (0.0137)	-0.0566*** (0.0152)
Constant	0.1*** (0.0)	0.0995*** (0.0176)	-0.236*** (0.0755)	-0.295*** (0.0757)
Observations	668	668	668	668
R-squared	0.1		0.348	
Number of c	32	32	32	32
Country FE	YES	NO	YES	NO
Year FE	YES	NO	NO	YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: All variables are lagged by one period and represented percentage changes. The sample includes all OCED countries and ranges from 2020-2022. Data source: WEO Live.

In the next step we add third-party indicators to the regression to shed light on how certain dimensions of the investment climate can impact actual investment flows. The results are summarized in Table 5. The coefficient of economic policy uncertainty is negative and significant in specifications (2) and (3), indicating a negative impact of policy uncertainty on capital formation but is insignificant if both fixed effects (country and year) are included. For the worldwide governance indicators, we find that a higher score (indicating a better outcome) in the categories government effectiveness and control of corruption tend to be associated with an increase in capital formation. Regulatory quality, on the other hand, seems to have the opposite effect with a negative and significant

¹⁷ An increase in the real exchange rate indicates an exchange rate depreciation.

coefficient in most specifications. Finally, the risk ratings from the ICRG are insignificant in all specifications. These results are broadly robust to alternative specifications excluding post Covid episodes or focusing on G7 countries.

Appendix Table 5: Regression Results Including Survey-Based Indicators

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Short-term interest rate	0.4 (10.9)	-3.105 (8.849)	-4.160 (9.486)	0.978 (10.36)
Real GDP	0.0** (0.0)	0.00689*** (0.00183)	0.00356 (0.00266)	0.00875*** (0.00153)
Consumer prices, period average	-0.8 (0.5)	-1.017*** (0.334)	-1.834*** (0.183)	-0.238 (0.290)
General government debt, % of GDP	-0.1** (0.0)	-0.0523 (0.0424)	-0.120* (0.0644)	-0.0702*** (0.0259)
Total Terms of Trade	0.1 (0.1)	0.254*** (0.0959)	0.135 (0.0977)	0.178** (0.0800)
Real Exchange Rate	0.0 (0.0)	0.00557 (0.00659)	0.0100* (0.00492)	0.00351 (0.00703)
Current Account, % of GDP	2.9 (2.1)	1.969 (1.925)	2.207 (2.042)	2.506 (1.990)
VIX Index	0.0*** (0.0)	0.00416 (0.00308)	0.00568* (0.00315)	0.0202*** (0.00521)
Federal Funds Rate	0.0 (0.0)	0.0240 (0.0174)	0.0318* (0.0176)	0.0272 (0.0257)
Economic Policy Uncertainty Index	0.0 (0.0)	-0.181*** (0.0569)	-0.197*** (0.0578)	0.0684 (0.0494)
Voice and Accountability (WWG)	-0.0 (0.0)	-0.000279 (0.000204)	-9.39e-05 (0.000243)	-0.000202 (0.000175)
Political Stability and Absence of Violence (WWG)	-0.0 (0.0)	0.000175 (0.00297)	3.91e-05 (0.00283)	-0.00121 (0.00203)
Government Effectiveness (WWG)	0.0** (0.0)	0.00711* (0.00427)	0.00903** (0.00365)	0.0121** (0.00566)
Regulatory Quality (WWG)	-0.0* (0.0)	-0.000576* (0.000348)	-0.000363 (0.000325)	-0.000492** (0.000234)
Rule of Law (WWG)	0.0 (0.0)	0.0180 (0.0277)	0.0314 (0.0274)	0.0123 (0.0238)
Control of Corruption (WWG)	0.0* (0.0)	0.0126* (0.00734)	0.0125* (0.00668)	0.0128* (0.00746)
Economic Risk Rating (ICRG)	0.2 (0.2)	-0.0984 (0.121)	-0.0261 (0.127)	0.236 (0.194)
Financial Risk Rating (ICRG)	-0.1 (0.1)	0.0793 (0.0846)	0.0219 (0.0812)	-0.0464 (0.0854)
Political Risk Rating (ICRG)	0.3 (0.3)	0.274 (0.349)	0.263 (0.403)	0.242 (0.250)
Constant	-0.4*** (0.1)	-0.0443 (0.0510)	-0.0473 (0.0521)	-0.403*** (0.103)
Observations	303	303	303	303
R-squared	0.4		0.240	
Number of c	16	16	16	16
Country FE	YES	NO	YES	NO
Year FE	YES	NO	NO	YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: All variables are lagged by one period and represented in percentage changes. The sample includes all OCED countries and ranges from 2020-2022. Data source: WEO Live, WWG Indicators (World Bank), ICRG and Economic Policy Uncertainty.

Random Forest

This section presents findings generated from machine-learning algorithms. Machine-learning models are well-suited for exploring the relative importance of established explanations because of their tremendous flexibility in identifying non-linear patterns and interactions within complex datasets that feature many different predictors. (Zhang et al., 2021; 2024) By comparison, traditional approaches such as ordinary least squares models, assume linear relationships among variables, necessitating the explicit specification of potential interactions.¹⁸

Out of the diverse machine-learning algorithms available, we have opted for the random forest model, one of the most widely used classification algorithm. (Gonzalez-Garcia and Yang, 2020) This choice is supported by a large strand of literature that discusses the strengths and weaknesses of various variable importance scoring methods for random forest models. Moreover, in our analysis, the random forest model demonstrates slightly superior performance in modeling investment climate when compared to the XGBoost. Nonetheless, in various robustness assessments, the results obtained using the XGBoost model did not present significant deviations from our baseline findings.

Model Specification

As before, we focus on various investment climate measurements such as the worldwide governance indicators, the international country risk guide, and the economic freedom index and include several additional predictor variables such as the real interest rate, real GDP growth, inflation, general government debt, external variables such as the terms of trade, the real exchange rate, the current account, and global factors such as the federal funds rate, the VIX index, and the economic policy uncertainty index are also included as predictor variables. The dependent variable is a binary variable equal to one in periods of positive real gross capital formation and equal to zero otherwise.

Model Evaluation

Machine-learning models undergo evaluation using data that had not been utilized during the model's learning phase. (Yu et al., 2022) This ensures that the model does not overfit, i.e., overly adapt to specific data patterns that might not be relevant in other datasets. The evaluation process, involving data not employed during training, is commonly referred to as out-of-sample performance assessment. The resulting scores are then employed to select the most effective model for interpretation. Following standard practice, we allocate 80% of the data to the training set, while the remaining 20% is reserved for assessing out-of-sample performance. We apply tenfold cross-validation, which entails randomly dividing the training data into ten subsets, each omitted during the respective training phase. This approach, combining cross-validation and out-of-sample prediction, aims to mitigate overfitting of the training set.

We evaluate the model performance for various specifications of the random forest through two common performance metrics: the F1 score and the receiver operating characteristic (ROC) area under the curve (AUC).

¹⁸ It is worth noting that while machine learning models provide a comprehensive assessment of a large number of explanatory variables, understanding causal relationships would require additional analysis such as a structural econometric model.

The F1 score balances precision and recall, measuring the accuracy of positive classifications and the extent to which positive class predictions are missed for a given decision threshold.¹⁹ Higher scores indicate superior performance. The ROC AUC is a metric that evaluates model performance across all possible decision thresholds. An ideal prediction model would correctly classify all outcomes, and the AUC quantifies overall performance regardless of marginal uncertainty in ROC curves.

We find that all models (whole sample, OECD, non-OECD) achieve relatively high F1 scores, ranging from 0.80 to 0.90. Following existing literature on variable importance scores, we fine-tune the model's hyperparameters, exploring different hyperparameter combinations for factors including the number of trees, the number of randomly sampled predictors at each tree split, the minimum data points required to split nodes into additional leaves etc. In the following the study feature importance using the hyperparameter setting that maximizes AUC-ROC score for the three different samples.

Feature Importance

Annex figure 1 shows variable importance scores using the fine-tuned random forest model for each individual predictor, for a sample of 163 countries including both OECD (36) and non-OECD (127) members.

On a scale from 0 to 100, GDP growth is the single most important factor predicting investment flows, which is consistent with our empirical finding that GDP growth is a statistically significant explanation for the change in fixed capital formation. Clearly, investors are more likely to commit funds when they perceive an environment that fosters growth and profitability.

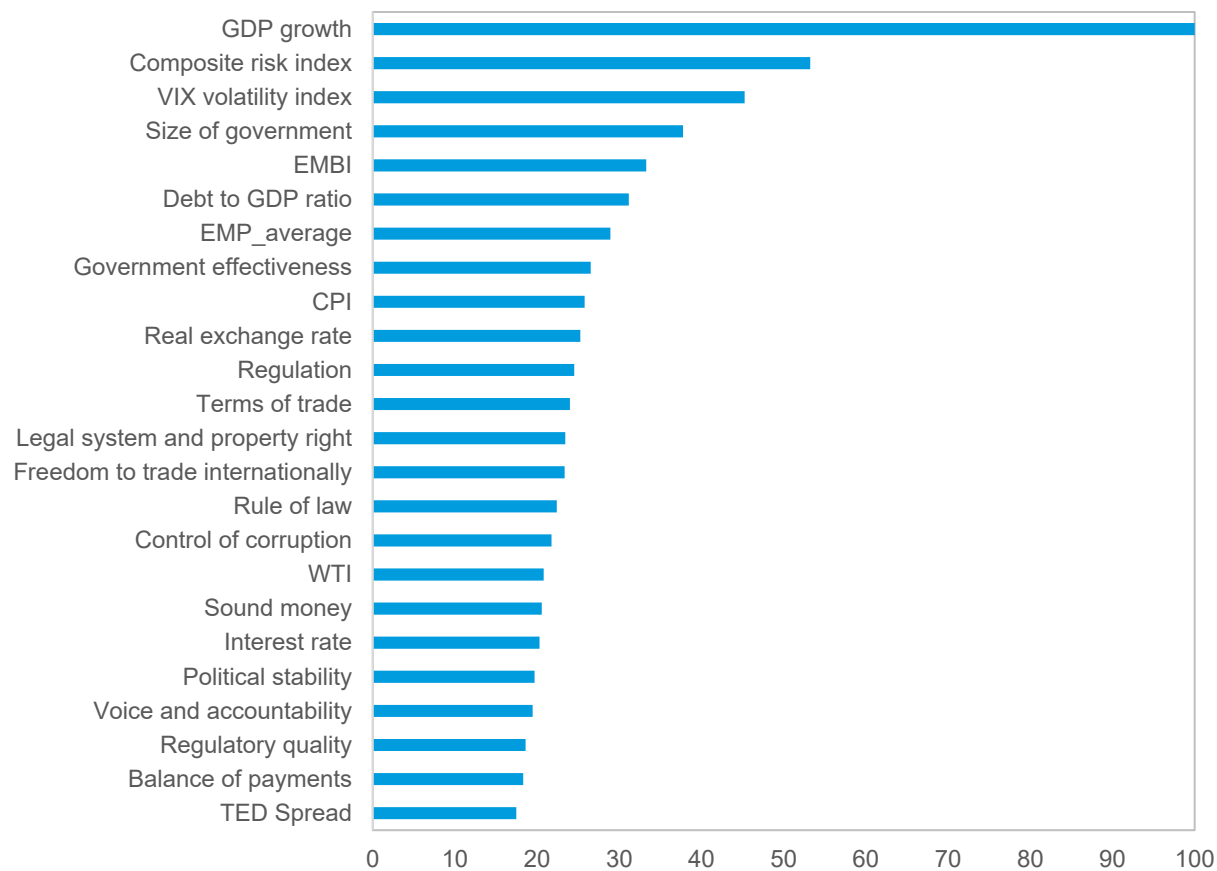
The volatility index (VIX) comes second in the whole sample and the sample that includes only OECD countries. There is also abundant evidence suggesting that uncertainty, particularly economic policy uncertainty, reduces both investment and growth. (see e.g., Asteriou and Price, 2005) Previous studies also show that uncertainty has discernable and substantial effects on capital flows. The primary driver of such effect is not the uncertainty within the host country, but rather the level of global uncertainty. (Jardet et al., 2022) In comparison, geopolitical risk, measured by the composite risk index, plays a major role in predicting investment flows for non-OECD countries.

In addition to volatility, other factors related to investment climate, such as the size of government, government effectiveness, rule of law, regulation, and political stability, are also associated with high predictive power for investment. Predictable regulations, political stability, and legal protections give investors confidence that their investments will not be subject to sudden adverse changes. Overall, a favorable investment climate sets the stage for attracting and facilitating investment flows.

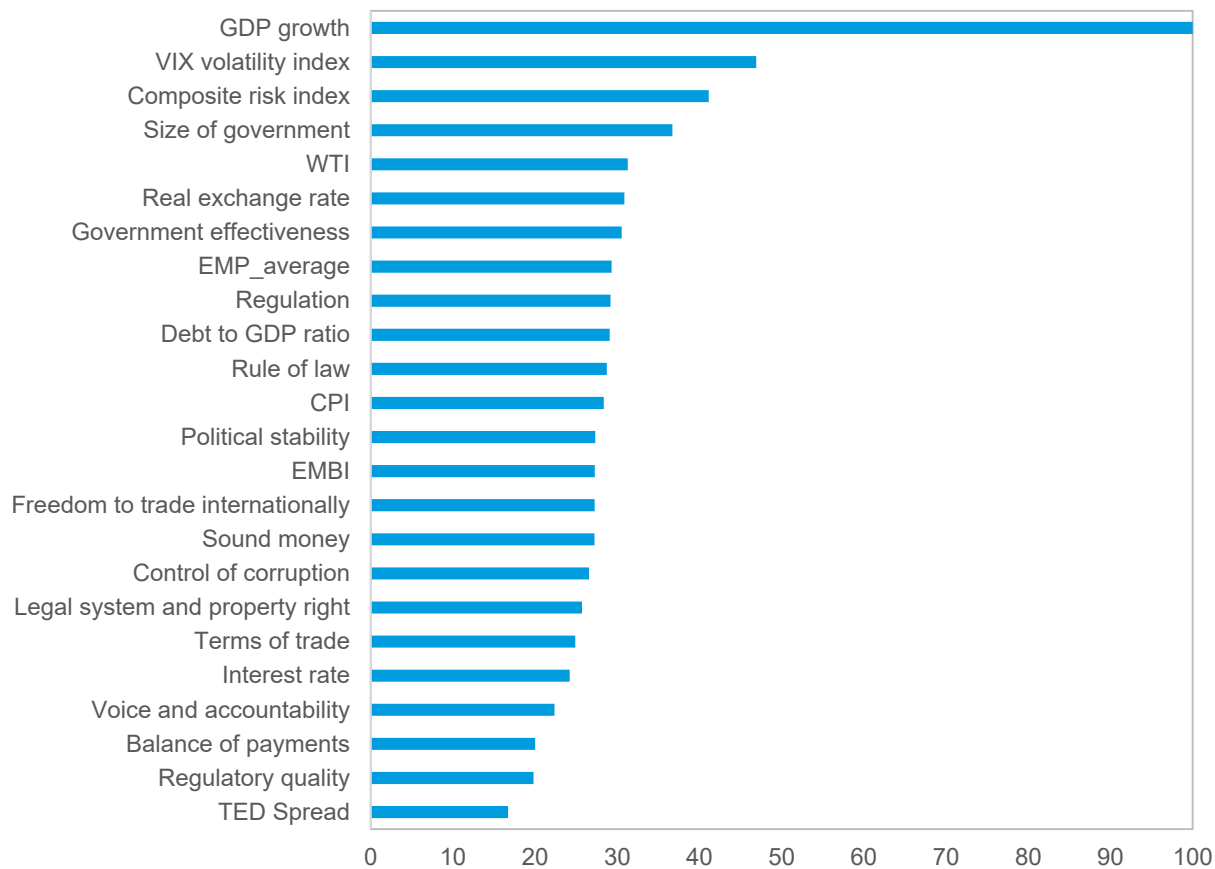
¹⁹ Note that the data exhibits class imbalance in the data, with about three quarters of the observations showing positive capital inflows. We adopt a five-fold cross-validation model and the F1 score, which is effective in dealing with class imbalance problems, shows decent performance.

Interestingly, we find that for non-OECD members, debt-to-GDP ratio is an important factor, following composite risk and GDP growth, that explains a country's investment flows. The negative link between elevated debt levels and reduced investment has been firmly established in the literature with several channels at work. (e.g., Rogoff and Yang, 2021; 2022) Heightened corporate indebtedness translates to increased interest expenses, resulting in a decreased pool of available funds for investment. Companies with overhang debt also encounter greater difficulties in securing new capital from external sources, due to their increased default risk. In addition, the desire to repair balance sheets often forces firms to trim their debt load, consequently causing them to forego potential investment opportunities. (Barrela et al., 2022)

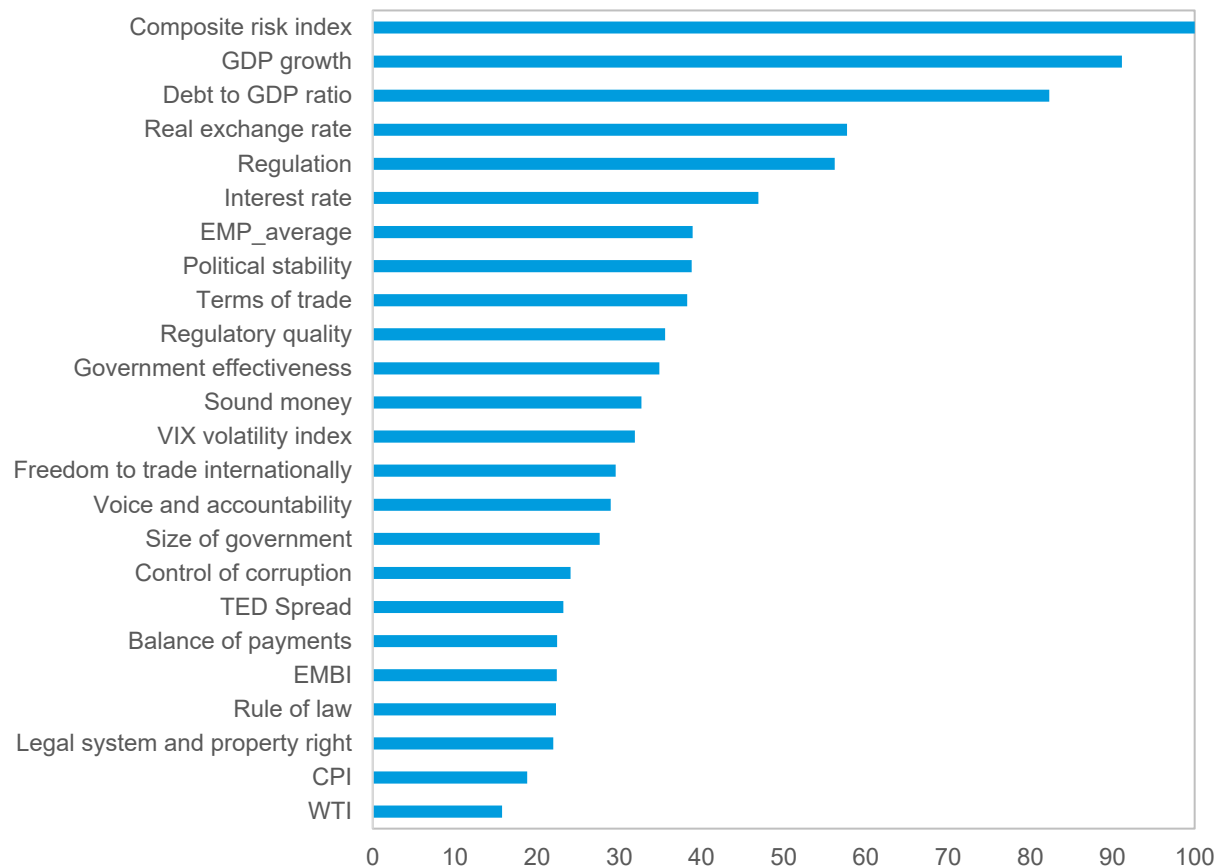
Appendix Figure 1: Importance Scores Using the Random Forest Machine Learning Model



Appendix Figure 2: Importance Scores Using the Random Forest Machine Learning Model (OECD Countries)



Appendix Figure 3: Importance Scores Using the Random Forest Machine Learning Model (Non-OECD Countries)

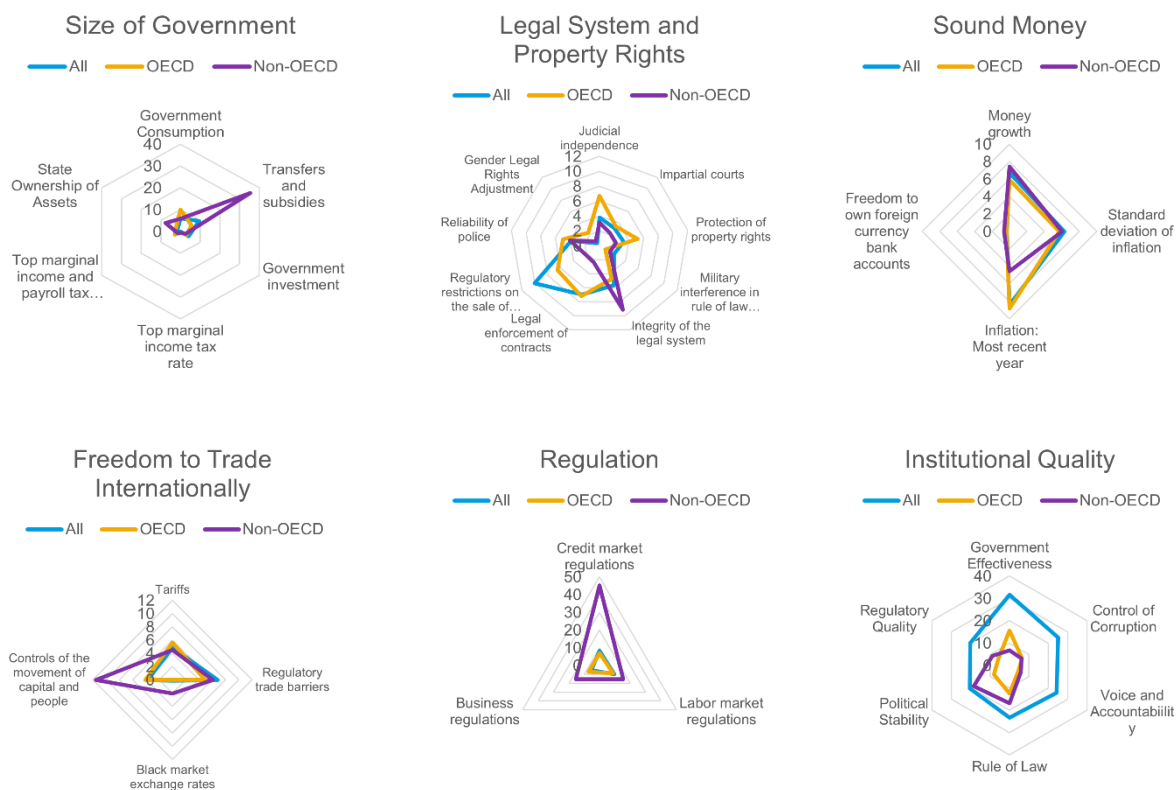


Subsequently, we disaggregate the economic freedom index into two additional tiers, with each tier adding more granularity. In the first tier, the size of government indicator is further decomposed into government consumption, transfers and subsidies, government investment, top marginal tax rate, and state ownership of assets. Similarly, the legal system and property rights indicator is disaggregated into areas of judicial independence, impartial courts, protection of property rights, military interference in rule of law and politics, integrity of the legal system, legal enforcement of contracts, regulatory restriction on the sale of real property, reliability of police, and gender legal rights adjustment. Sound money comprises money growth, inflation, standard deviation of inflation, and the freedom to own foreign currency bank accounts. The freedom to trade internationally indicator includes subcategories like tariffs, regulatory trade barriers, black market exchange rates, and controls of the movement of capital and people. And finally, the regulation indicator is broken down into credit market regulations, labor market regulations, and business regulations.

Figure 14 presents results about which specific subcategories best explain the outcomes. For OECD countries, inflation, particularly in most recent years, play a critical role in explaining investment flows, whereas for non-OECD countries, transfers and subsidies, integrity of the legal system, and capital flows are key factors. (Yang,

2021) This set of complementary tests help deepen the analysis and permit more granular insights into the relative importance of various dimensions of investment climate.

Appendix Figure 4: Decomposing the Investment Climate Measurement





PUBLICATIONS

Public Perceptions of Canada's Investment Climate—A Natural Language Processing-Based Approach
Working Paper No. WP/2024/165