



INSTITUT DE HAUTES
ÉTUDES INTERNATIONALES
ET DU DÉVELOPPEMENT
GRADUATE INSTITUTE
OF INTERNATIONAL AND
DEVELOPMENT STUDIES

1

2 **When central-peripheral relations intersect with the**
3 **asylum-seeking regime: how media outlets from different**
4 **political camps report on asylum seekers in Hong Kong**

5 **THESIS**

6 submitted at the Graduate Institute
7 in fulfillment of the requirements of the
8 The International Relations/Political Science Department

9 by

10 Chin Man Kwan

11 Thesis No. 12345

12 **Supervisor: Prof. Sungmin Rho**
13 **Second Reader: Prof. Ravinder Bhavnani**

14 **Geneva**
15 **2022**

16
17

INSTITUT DE HAUTES ETUDES INTERNATIONALES ET DU DEVELOPPEMENT
GRADUATE INSTITUTE OF INTERNATIONAL AND DEVELOPMENT STUDIES

18

RESUME / ABSTRACT

19
20
21
22
23
24
25
26
27

Non-refoulement claimants (or asylum seekers) has become a salient issue in Hong Kong's public debate since early 2016 after a senior official from the Immigration Department claimed that the city's screening mechanism of asylum seekers was being abused. When Hong Kong embroiled into one of its largest conflict between pro-democracy protesters and the HKSAR Government in 2019, how might the portrayal of asylum seekers by media outlets affiliated with different political camps be? This thesis aims to investigate this question by using topic modelling from Natural Language Processing to discover the topics of news articles reporting on asylum seekers in 2019 by printed news media. It is found that ... (to be completed). With the

For my loved ones

Acknowledgements

I would like to take this opportunity to thank all the people who have offered enormous support to me throughout my study at the Graduate Institute. First of all, my parents, Mr. Lam Cheong KWAN and Ms. Suk Yee TAM, have always supported me in every moment of my life. Without them I would not have been able to be who I am today.

Next, my friends have also brought a lot of joy throughout my study at the Institute over the last two years. My cohort in the Department definitely

Chin Man KWAN
Graduate Institute, Geneva
10 January 2022

Contents

40	Preface	1
41	1 Introduction	2
42	2 Literature Review	4
43	2.1 Theories	4
44	2.1.1 What is national identity?	4
45	2.1.2 Impact of conceptions of national identity on attitudes towards immigrants	6
46	2.1.3 Chineseness in Hong Kong	8
47	2.2 Empirics	10
48	2.2.1 How do different political camps in Hong Kong perceive national identity? .	10
49	The pro-Beijing camp	11
50	The pro-Democracy camp (or pan-democrats)	13
51	The localist camp	15
52	2.2.2 Asylum seekers in Hong Kong	18
53	Hong Kong's policy regime	18
54	Responses by political camps	20
55	3 Methods	24
56	3.1 Hypotheses	24
57	3.2 Data collection	25
58	3.3 Modelling	28
59	3.4 Operationalisation	30
60	3.4.1 Main independent variable	30
61	3.4.2 Dependent variable	31
62	3.4.3 Control variables	32

CONTENTS

63	4 Results	34
64	4.1 Exploratory data analysis (EDA)	34
65	4.1.1 How does the number of news articles vary by political camps and month?	34
66	4.1.2 Polarities of the news articles	37
67	4.1.3 Presence of racial labels	37
68	4.1.4 Character lengths of news articles and titles	39
69	4.2 Sentiment analysis	39
70	4.2.1 Preprocessing	39
71	4.2.2 Training the model	42
72	4.2.3 Is the pro-Beijing camp more likely to portray asylum seekers in 2019 more	
73	negatively than other outlets?	43
74	5 Conclusion	48
75	5.1 How might the instigation of the National Security Law affect the public discourse	
76	on asylum seekers in Hong Kong?	48
77	Appendix: The Echoes of the Code	50
78	Appendix: The Echoes of the Code redux	51
79	References	54

80

List of Figures

81	2.1	Number of non-refoulement claim cases (left) and refugees (right) in Hong Kong .	19
82	2.2	Number of non-refoulement claimants by nationality in 2021	21
83	4.1	News articles on asylum seekers in 2019 by news outlet (left) and political camp	
84		(right)	35
85	4.2	Temporal patterns of the publication of news articles about asylum seekers in Hong	
86		Kong in 2019	36
87	4.3	Presence of racial labels in the news articles by sentiment	38
88	4.4	Distributions of the word counts of the articles' titles (left) and main texts (right)	39
89	4.5	Elbow plot of the reconstruction error of NMF as a function of the number of	
90		pre-specified latent topics	41
91	4.6	Overall SHAP values of the features in the tuned XGBoost model	44
92	4.7	The SHAP values of the features in the prediction of whether an article has a	
93		negative polarity	45
94	4.8	The SHAP values of the features in the prediction of whether an article has a	
95		neutral polarity	46
96	4.9	The SHAP values of the features in the prediction of whether an article has a	
97		positive polarity	46
98	5.1	Word list of each of the topics generated by the NMF model	53

List of Tables

100	2.1	Comparison of the three camps on key dimensions of Hong-Kong-China relations .	17
101	3.1	Translation of the keywords used to search for media reports on WiseNews	28
102	3.2	Classification of the associated political camp of media outlets in Hong Kong in-	
103		cluded into the sample, source: Fong (2017a); Lee (2018); author's own research .	31
104	4.1	Polarities of the news articles on asylum seekers in Hong Kong in 2019	37
105	4.2	Summary statistics of the word counts of the news articles' titles and main texts .	40
106	4.3	Log loss on 5-fold cv and test set for the 4 baseline models	42
107	4.4	Classification report of the baseline XGBoost model	43
108	4.5	Classification report of the tuned XGBoost model	43

List of Abbreviations

110	ASPDMC	the Alliance for the Support of the Patriotic and Democratic Movement
111	NATO	North Atlantic Treaty Organization
112	UN	United Nations
113	UNICEF	United Nations Children's Fund
114	WHO	World Health Organization
115	WTO	World Trade Organization

Preface

The issue of refugees and asylum seekers have ignited a lot of attention and debates in the 21st century, particularly so after the 2015 European refugee crisis. In the academia, a great deal of effort has been contributed to understanding how the public's and/or policy-makers' perceptions towards asylum seekers and

Chapter 1

Introduction

Asylum seekers, refugees and these two groups of vulnerable population's movements across national borders have become some of the most salient issues in the 21st Century, especially after the outbreak of the 2015 Refugee Crisis in Europe. A lot of academic researches have subsequently investigated how politics in the .

Since the transfer of sovereignty from Britain to China in 1997, Hong Kong has witnessed increasing tension with its new sovereign owner. Until July 2020 when the HKSAR¹ Government passed the National Security Law and the subsequent commencement of the large-scale suppression of the opposition camp such as the mass arrest of 53 pro-democracy activists in January 2021 (BBC 2021), protests against the authority of the HKSAR and Beijing Governments had occurred. Apart from the annual July 1 protest since 2003, the two most prominent political movements would be the 2014 Umbrella Movement and the more recent 2019 anti-extradition law protest which created considerable attention from other countries as well.

This thesis thus aims to investigate *how may the framing of asylum seekers in Hong Kong vary by news media outlets associated with different political camps, given each of them has considerably different definition of the city's national identity*. The academic contribution of this thesis .

Before moving on, it should be mentioned that I will use the words "asylum seekers", "refugees" and "non-refoulement claimants" interchangeably. It is true that difference exists between the definitions of asylum seekers and refugees². Nevertheless, since the above three terms are generally

¹A list of the 30 most prominent words in each of the topic and the code to generate it are available in the appendix.

²For the complete documentation of the default parameters of the models used in this thesis, refer to the websites of scikit-learn and XGBoost Documentation — xgboost 1.5.1 documentation.

1. Introduction

142 used to refer to the same group of population in Hong Kong who flee from their own countries
143 to escape from persecution, treating these terms as synonyms in this context should better cover
144 how this group of population is perceived by media outlets in different political camps. In fact,
145 scholars researching on this issue often use the terms I listed without strictly differentiating their
146 meanings (e.g. Vecchio (2014), Ng, Choi, et al. (2019)). Moreover, to increase the transparency
147 of this study, all the source codes will be uploaded into this

148 This thesis will proceed as follow. In section 2, I will review the literature on the theories
149 about how the perception of national identity might affect attitudes towards immigration, how
150 each major political camp in Hong Kong generally perceives the national identity of Hong Kong,
151 and how each camp perceives asylum seekers. Section 3 provides an overview of the data collection
152 process and methodology used to answer the research questions. Section 4 presents the result
153 of the analysis, and section 5 concludes the findings as well as mention the implications of this
154 thesis.

Chapter 2

Literature Review

To understand how the different conceptions of the Hong Kong identity may affect attitude towards asylum seekers, it is important to first review the literature on how national identity may affect attitudes towards immigrants and how Chineseness may be articulated on theoretical level, as well as how political camps in Hong Kong may define the city's national identity and the situation of asylum seekers on empirical level.

2.1 Theories

2.1.1 What is national identity?

I will adopt the definition proposed by Greenfeld and Eastwood (2007), which is how an individual relates to a broader societal order called the nation and is central to modern political and social organisations. Lewin-Epstein and Levanon (2005) point out modern nationalism's conception that people with similar cultures (e.g. language, myths, historic territory) should congregate together as the basis of political community called the nation-state. A closely related but not identical concept to nationalism is ethnicity, defined by Varshney (2007) as sense of belonging shared by a group in ancestry, language, history, culture and/or history. Whereas nations have their own territory-based political authority, this is not necessarily the case for ethnic groups.

Greenfeld and Eastwood (2007) also elaborate on some peculiar features of nationalism. For instance, the secular world where nations are located is deemed the most important to human beings and their life. But the more well-known characteristic is that the world is perceived to be divided into distinct communities (i.e. nations) occupying their own territories, and each member of "the people" within a nation is treated as equal to one another. It is also "the people" who

2. Literature Review

ultimately control political legitimacy and authority. These features together contribute to two fundamental implications. Socially, the fact that nationals are considered as being equal and that status is achieved instead of ascribed promotes social mobility. Politically, nationalism has made modern (nation-)states “impersonal” (p.261) because whoever holds the government office does not affect whatsoever the nation-state’s existence and legitimacy which ultimately rest in popular sovereignty (Greenfeld and Eastwood 2007). But for cultures that had been existing long before nationalism’s entry into these societies, the constructions of novel national identities may often rebrand pre-existing cultures despite the latter’s internal heterogeneity, of which Chineseness is an example which I will look into later.

Lastly, Greenfeld and Eastwood (2007) propose a typology of nationalism with two criteria. The first one is how the nation-state relates itself to its members, from the individualist end of treating a nation as merely a massive group of individuals to the collectivist end of which individuals are subservient to the nation. The other criterion is the well-known “ethnic” versus “civic” dimension which is extensively studied, including its relation with the perception towards immigrants which I will discuss later. In general, ethnic nationalism defines a nation’s members based on ascriptive and rather objective standards such as descent, language and customs which are largely decided at birth, whereas civic nationalism focuses less on the members’ cultural backgrounds than their voluntary allegiance to a political community and its legal institutions within a political territory (Greenfeld and Eastwood 2007; Lecours 2000; Lewin-Epstein and Levanon 2005; Heath and Tilley 2005).

A note on the ethnic-civic dimension is that these two aspects are not mutually exclusive since Medrano and Koenig (2005) highlight that nationalism in one country can often incorporate both civic and ethnic features, which is quite frequently observed from empirical. Accordingly, scholars often create other dimensions of describing nationalism theoretically. For example, Hjern (1998) proposes to additionally gauge whether identifications with ethnic and civic features of nationalism are high (“multiple national identity”) or low (“pluralist”) at the same time (p. 453). Meanwhile, some scholars prefer more fine-grained categories of citizenship criteria. Medrano (2005) breaks the ethnic and civic dimensions further into five categories while describing the importance of criteria for citizenship perceived by the Spanish population, namely (p.137),

- the civic dimension being broken down into:

2. Literature Review

- Pure republican criteria: Long-term residence, feeling of belonging, and being a citizen of that state
- Political criteria: Respect for laws and political institutions
- and the ethnic dimension into:
 - Territorial criteria: Place of birth
 - Cultural criteria: Language, religion, customs and traditions and
 - Descent

2.1.2 Impact of conceptions of national identity on attitudes towards immigrants

Perhaps one of the most prominent consequences of how national identity is perceived by individuals is their attitudes towards immigrants. This is because national identity simultaneously defines members and outsiders of a nation (Hjerm 1998; Lewin-Epstein and Levanon 2005; Heath and Tilley 2005). For simplicity's sake, I will mainly focus on how tendency towards defining national identity in ethnic and civic terms may be correlated with perceptions towards immigrants.

To begin with, although it is generally agreed that characteristics used to demarcate membership in a nation can be classified as ethnic- or civic-based on a macro-level, additional categories are often needed to capture the general patterns of how national identity is defined within a country and how such definitions may affect citizen's attitudes towards immigrants. A potential reason for such a need is that although individuals may hold a rather "pure" civic conception of identity mostly devoid of ethnic components, it is quite unlikely that individuals would consider ethnic components as standalone enough for constituting a nation's citizenship without any civic components. Indeed, empirical evidence from Hjerm (1998), Heath and Tilley (2005) and Medrano (2005) demonstrate that it is very rare for individuals to perceive national identity solely in ethnic terms, since those who attribute importance to ethnic-based components are also likely to do so to civic-based ones simultaneously. Often, scholars may even challenge the ethnic-civic-dimension in classifying the perception of national identity and propose alternative schema. Drawing on the Weberian idea of social closure which measures the number of obstacles imposed by a group for outsiders to join, Medrano (2005) proposes the "postnationalist" versus "credentialist" (p.136)

2. Literature Review

dimension which differentiates how restrictive people consider acquiring citizenship to their nation should be. Specifically, postnationalists impose few limits on foreigners to become citizens, whereas credentialists require a long list of criteria being fulfilled. That said, even though there exist multiple and equally valid alternatives for categorising how national identity may be defined, the ethnic-vs-civic dimension is quite broadly applicable to explain the typologies of nationalism in countries with some modifications and thus is a decent starting point.

As for the potential impact of national identity on attitudes towards immigrants, the literature tends to agree that the more an individual holds ethnic-based criteria of membership in a nation, the more likely (s)he will be averse against immigrants (i.e. xenophobic). It seems that the effect of ethnic-based national identity conception on xenophobic attitudes is consistent across different immigration policy regimes differing on the ease of naturalisation and the mode of integration. Hjerm (1998) discovers that despite Germany, Australia and Sweden having considerably different immigration policy regimes, individuals who agree that having both ethnic and civic characteristics are crucial for being a member of their nations are more likely to be xenophobic than others who only use civic components to demarcate nationhood or ascribe little importance to nationalism. Similar findings are observed by Heath and Tilley (2005) in the UK context, as British who place more emphasis on ethnic-based nationhood criteria will be more inclined to approve reducing immigrants inflows and more stringent measures against illegal immigrants. Interestingly, Heath also discovers that after controlling for how national identity is conceived, the extent to which an individual is attached to the nation is no longer significant in affecting his/her attitude towards immigrants.

In some cases, however, the effect of national identity on perceptions towards immigrants may be less due to whether individuals hold ethnic-based definitions of national identity than how many criteria individuals deem to be indispensable for obtaining a nationality. For instance, Medrano (2005) argues that even though there is an association between including ethnic-based (or in his terminology “ethno-biological republican”) (p.148) criteria for obtaining Spanish citizenship and holding negative attitudes towards immigrants, it is ultimately how many criteria a Spanish regards as important for being a Spanish citizen that better predicts attitudes towards immigrants, evidenced by the stronger correlation between these two items found in his study.

2. Literature Review

Furthermore, depending on the ethnic composition of a nation-state, not only may different ethnic groups have their own definitions of what constitute to the ethnic and civic components of national identity, but also the relations between perceptions of national identity and attitudes towards immigrants may go into rather unexpected directions. Lewin-Epstein and Levanon (2005) discover that in the Israeli society which can be broadly separated into three ethnic groups (namely, veterans Jews who have lived in Israel for a long time, Olim who are recent Jewish immigrants from former Soviet Union regions, and ethnic Arabs citizens in Israel), there are only moderate similarities in terms of which components should be included into the ethnic and civic dimensions of national identity, as they disagree on whether the ability to speak Hebrew and whether being born in Israel should be labelled as ethnic or civic components. Moreover, contrary to the expectation in the literature, veteran Jews who consider the Israeli nationality as more ethnic-based are actually *less* averse towards immigrants, possibly because they are likely to perceive immigrants as mostly overseas Jews who share the same ethnicity (Lewin-Epstein and Levanon 2005). Nevertheless, the surprising case of Israel is likely to be more due to its specific socio-political context rather the theoretical pitfalls of how ethnic-based conceptions of national identity may be related to xenophobia per se.

2.1.3 Chineseness in Hong Kong

Given over 90% of the population being ethnically Han Chinese by 2016 (Home_Affairs_Department 2018) and also large-scale pro-democracy political movements in the 2010s challenging the authority of the HKSAR and Chinese Governments (particularly the 2014 Umbrella Movement and the 2019 Anti-Extradition Law Protest), Chineseness indeed occupies a salient position on the discourse of the Hong Kong identity. It is then important to understand what Chineseness may mean in Hong Kong, a city under British colonial rule for over a century until 1997.

Although Chineseness is usually connoted with a homogeneous image of a group of people with the same ethnic origin, place of residence or cultural characteristics (e.g. language, customs etc.), Chun (1996) problematises this conception not least because expressions used to describe different aspects of Chineseness actually contains heterogeneity. For example, despite the contemporary use of Han Chinese to denote ethnic homogeneity, people under the Han empire were anything but ethnically homogeneous. It was not until the 1911 Revolution then China became a nation

2. Literature Review

inhabited by the single ethnic group called *Zhonghua Minzu*, and a homogeneous national culture was accordingly invented and legitimised by the state. The efforts to articulate a Chinese national identity continued after the foundation of the People's Republic of China (PRC) in 1949, as both the Kuomintang (KMT) regime in Taiwan and the PRC (after the Maoist period) extensively used Chinese history, political ideology and/or values to construct their national cultures and identities to claim themselves being the orthodox Chinese nation.

Whereas Chineseness was constructed by Chinese regimes holding political authority in Taiwan and mainland China, Chineseness manifested quite differently in Hong Kong where it was a Chinese society ruled by the British Government before 1997. According to Chun (1996), most inhabitants in Hong Kong had identified as Chinese until the establishment of the PRC in 1949. So (2015) also mentions that before 1949, ethnic Chinese in Hong Kong did not intend to permanently reside in the city and identified with the broader Chinese nation, evidenced by their participation in the anti-colonial Hong Kong-Canton general strikes in the 1920s. When the Chinese Communist Party (CCP) took power in 1949, however, the identity of ethnic Chinese in Hong Kong started to gradually depart from that on the mainland. As Hong Kong had become the venue to the KMT and the CCP's contest of the Chinese identity in the 1950s, the British Government then attempted to pull Hong Kong away from this nationalist conflict by actively promoting economic transformation from a Chinese entrepot to exported-oriented economy, and the colonial government started to create a new Hong Kong identity to further distance the citizens from Chinese nationalism following the outbreak of leftist riots in the late 1960s (Chun 1996; So 2015).

With intellectuals in Hong Kong paying more and more attention to local instead of Chinese affairs and the huge influx of illegal immigrants from China in the 1970s, the conflict between the old and newly arrived residents eventually led to the rise of a Hongkonger identity (So 2015). By mid-1980s, Hong Kong eventually developed its own culture which was indifferent to politics while also fusing habits from Chinese and Western cultures together (Chun 1996). As the transfer of the city's sovereignty approached in the 1980s, however, the Hong Kong identity became more complicated. On one hand, the investment opportunities in and relocation of factories to mainland China due to the reform and opening up policy helped resurge Chinese nationalism among Hong Kong business elites and working class who married mainland Chinese (So 2015). Meanwhile, the

2. Literature Review

uncertainty associated with the transition of sovereignty, especially after the Tiananmen Massacre in 1989, caused the public to be more conscious of values respected in Hong Kong such as the rule of law and civil liberties and thus increased their perceived difference between Hong Kong and China (So 2015; Fong 2017b; Lin and Jackson 2021) .

Amidst the fear of Hong Kong losing its unique after the handover to China in the 1990s, the *One Country, Two Systems* (OCTS) framework which stipulates Hong Kong's autonomy and continuity of the capitalist system for 50 years was accordingly put in place to secure confidence of the public regarding the handover (So 2015; Fong 2017b) . Indeed, the first decade after 1997, So (2015) and Fong (2017b) report that overall more and more Hongkongers identified as Chinese despite Beijing's increasing intervention (more on this later) after the mass protest against Article 23 in 2003, and such sentiment peaked in 2008 when China was the host of that year's Olympic Games. But the Hong Kong identity has once again become prominent since 2009 as Hongkongers become aware of Beijing's increasing political interference and the social disruption brought by the large influx of mainland Chinese, and Fong (2017b) highlights how the Hong Kong identity has emerged in a mentality of resisting against "invaders" from the mainland during the period (p.19).

In short, Chineseness as how closely Hong Kong people identify themselves as members of the Chinese nation have evolved throughout the course of history, and one could argue that following the transferral of Hong Kong's sovereignty to China, the Chineseness of Hong Kong identity has become more and more contested. This phenomenon is best reflected by how each prominent political camp in Hong Kong envisions the Hong Kong identity.

2.2 Empirics

2.2.1 How do different political camps in Hong Kong perceive national identity?

With Hong Kong's major political cleavage being demarcated by the city's relations with Beijing and the pace of democratisation, different political camps may adopt their own vision of what Hong Kong's national identity should be. Indeed, Lecours (2000) highlights the importance of political goals in affecting how political actors with different ideologies on a region's relations with the national government may formulate national identity through the case study of Quebec.

2. Literature Review

352 He argues that the secessionist Parti Quebecois (PQ) is motivated to adopt an ethnic-based
353 nationalism which emphasises on French being the Quebecois' mother tongue because the party
354 needs a more salient cultural marker to convince Quebecois that they are indeed qualitatively
355 different from the Anglophone Canada and thus warrants a separate nation. By contrast, since
356 the autonomist Parti Liberal du Quebec's (PLQ) goal is to demand more political autonomy
357 instead of secession, the party can adopt more civic-based framing of nationalism which does not
358 restrict membership to the Quebecois society exclusively to Francophones while not explicitly
359 relying on ethnic markers to distinguish between Quebecois and Canadians.

360 In the following paragraphs, I will discuss the political goals of the three prominent political
361 camps in Hong Kong regarding the city's political regime and relations with the Beijing Gov-
362 ernment, namely, the pro-Beijing camp, the pro-Democracy camp and the newly arising localist
363 camp in the 2010s, and how their political goals may affect their articulation of Hong Kong's
364 national identity.

365 **The pro-Beijing camp**

366 As the politically conservative force governing Hong Kong, the pro-Beijing camp's objective can
367 be summarised as legitimising China's sovereignty over the city and protecting the political in-
368 terests of China. Members of the camp include the HKSAR Government itself, political parties
369 and societal actors who lean towards the Beijing Government etc. Perhaps the most prominent
370 manoeuvre by the pro-Beijing camp to achieve the above objective is counteracting against de-
371 mands for more rapid democratisation in both the executive and legislative branches to prevent
372 the opposition from gaining control over the political institutions (Ma 2012) .

373 Apart from controlling high-level political institutions in the HKSAR, the pro-Beijing also
374 invests in efforts to foster patriotism towards China in different societal sectors, of which education
375 is one. Bottom up endeavours to instil the Chinese identity into students' minds are organising
376 trips to China as well as pedagogy of pro-China ideologies and Chinese cultures, by pro-Beijing
377 schools and regular schools receiving sponsorship from the Government or pro-Beijing actors,
378 whereas top-down endeavours from the Government mainly revolves around policies on national
379 education such as teaching Putonghua in primary school since 1998, encouraging schools to raise
380 the Chinese flag and sing the Chinese anthem in significant occasions and the failed attempt to

2. Literature Review

381 make national education itself a mandatory subject which avoids discussion on sensitive issues
382 in 2012 (Lau 2013; Fong 2017b) . The pro-Beijing camp also sets up support organisations in
383 communities for electoral mobilisation, as Loh (2010) and Fong (2017b) out how the largest pro-
384 Beijing parties in Hong Kong have numerous subsidiaries grassroots organisations which can serve
385 to mobilise for votes during elections under the coordination of the Liaison Office (LOCPG), and
386 Fong (2017a) reports an alleged vote-rigging incident in 2015 during which pro-Beijing-affiliated
387 elderly homes were spotted sending elderlies, often of low cognitive abilities, to cast their votes.

388 With the pro-Beijing camp's goal of securing the ultimate authority and sovereignty of the
389 Beijing Government over the HKSAR, its vision on how the identity of Hong Kong should be
390 is then closely intertwined with the Chinese identity. In other words, pro-Beijing forces view
391 the Chinese and Hong Kong identities as a whole in which Hong Kong is subordinate to the
392 Chinese nation, and identification with the Chinese nation is translated into support towards
393 the incumbent PRC regime, which self-proclaims as the “vanguard of the Chinese people and
394 the Chinese nation” (*Full Text of Constitution of Communist Party of China - Xinhua* 2017)
395 . Indeed, Lin and Jackson (2021) argue that at least during 2012-19, the HKSAR government
396 primarily portrays the Chineseness of Hong Kong citizens in an essentialist and narrow ethnic-
397 based manner which emphasises on the shared biological inheritance from the ancient Peking
398 Man between Hong Kong citizens and mainland Chinese while downplaying other aspects such as
399 values, cultural heritage and lifestyle. For instance, former Chief Executive CY Leung and pro-
400 Beijing politician Ronny Tong claimed that Hong Kong citizens were obliged to identify themselves
401 as Chinese due to their ascribed biological similarities with mainland Chinese. Ethnic markers
402 were also co-mentioned while HKSAR officials were describing the PRC, as former Chief Secretary
403 Matthew Cheung stated that he was proud as “a son of the Chinese” while witnessing the progress
404 made by mainland China (Lin and Jackson 2021) (p.914).

405 This view echoes with that held by the Beijing Government, as Chinese President Xi Jinping
406 once claimed that Hong Kong's fate was closely linked to mainland China, and the city was crucial
407 for achieving the Chinese dream (Lin and Jackson 2021) . In essence, as the ruling government
408 of post-1997 Hong Kong is now under the sovereignty of the PRC, it is likely that the HKSAR
409 Government and its political ally, the pro-Beijing camp, actively promote Chineseness as an ethnic
410 feature that both Hong Kong citizens and mainland Chinese share closely to realign the Hong

2. Literature Review

Kong identity, which is deemed to have developed separately in British Hong Kong (Chun 1996; So 2015), under the umbrella of Chinese so that Beijing could exert more control over the city's ideology. By contrast, civic values are less salient in the official discourse in promoting Chineseness in Hong Kong (except for the value of patriotism which is ultimately another manifestation of the ethnic conception of Chineseness), but are often used by the pro-Democracy camp to portray Hong Kong's distinct (but **not** independent) identity from mainland China.

The pro-Democracy camp (or pan-democrats)

According to Ma (2012), even though several pro-democracy political groups were founded in the 1980s after the commencement of the city's democratisation and elections as well as the signing of the Sino-British Joint Declaration in 1984, it was not until the late 1980s when these groups congregated together for better mobilisation power in large-scale political movements, especially during the 1989 Tiananmen Movement when the Alliance for the Support of the Patriotic and Democratic Movement (ASPDMC) was formed to support students protesting for democracy in Beijing. Shortly afterwards, the first pro-democracy party, *the United Democrats of Hong Kong* (the predecessor of the present Democratic Party) was created in 1990 for next year's first direct election in the legislature (Ma 2012). After the handover in 1997, the pro-democracy camp remains as one of the significant political forces in the HKSAR, including the provision of an alternative perception of the Hong Kong identity vis-a-vis the official rhetoric.

Compared to the pro-Beijing camp which considers protecting the interests of the Chinese nation (currently under the PRC regime) as paramount over other concerns, the pro-democracy camp instead prioritises values such as the rule of law and civil liberties which are regarded to be constitutive to the civic elements of Hong Kong's identity (Kwan 2016). This is evidenced by the pro-democracy camp deeming the Article 23 legislation back in 2003 as contravening personal freedom under the pretext of national security and thus openly opposed to and organised the 1 July rally against the law's promulgation (Ma 2012), and the camp also supported the Occupy Central Movement (a precursor to the 2014 Umbrella Movement) to protest against Beijing's proposal of only allowing pre-approved candidates to run for Chief Executive elections in August 2014 as violating democratic principles of universal suffrage (Kwan 2016). It can thus be argued that the pro-Democracy camp puts more importance on protecting the civic values that they

2. Literature Review

deem crucial for the identity of Hong Kong, even if this means often clashing with the interests of Beijing.

When it comes to identifying which nation Hong Kong belongs to, however, it should be noted that the pro-democracy camp in general still identify themselves as Chinese, as Kwong (2016) states that especially for the older generation pro-democratic politicians, they are in favour of Hong Kong's sovereignty being transferred from Britain to China in 1997, and they claim to share affinity with China's history, culture and ethnicity despite opposing the CCP's one-party dictatorship. In fact, the identification with the Chinese nation by the pro-democracy camp is well-inscribed in the manifesto of the largest and most historic pro-democracy party, the *Democratic Party*, as the very first article clearly claims that "Hong Kong is an indivisible part of China" (Manifesto 2021). The pro-democracy camp's identification as members of the Chinese nation may very well explain how they contextualise Hong Kong's democratisation movement and why they do not think Hong Kong identity is independent of the Chinese one. Kwong (2016) mentions that the pro-Democracy camp generally considers Hong Kong's democratisation as integral to that of mainland China, and the annual candlelight vigil by the ASPDMC is a totem of this mentality since some organisers perceive a democratic China, if realised, might subsequently bring democracy to the city as well.

Admittedly, the pro-democracy camp attempted to appeal to localist supporters when the ideology gained huge popularity around the time of the 2016 Legislative Council election, such as calling for the right of self-determination for Hong Kong (Kaeding 2017; Fong 2017b). Nevertheless, this does not mean the pro-democracy camp giving up on caring about political issues in China while relating them to the situation in Hong Kong. Even in the 2021 Tiananmen Massacre candlelight vigil, the ASPDMC stated that Hong Kong people shared "a common fate with those oppressed and deprived of freedom in China", and the organisation would struggle until both Hong Kong and China became democratic (ASPDMC 2021).

In short, the pro-democracy camp considers civic values in Hong Kong such as the rule of law and civil liberties to be constituting to the distinctiveness of Hong Kong and is averse towards Beijing's political interventions which often run against these values. Nevertheless, the camp still regards Hong Kong as belonging to the same nation with mainland China and is even willing to help their Chinese compatriots in political movements which advocate freedom and democracy.

2. Literature Review

The geographical scope of how wide Hong Kong's political movement to democratisation should concern, then, is partly what makes the localist camp different from the pro-democracy camp even though both favour a democratic regime in Hong Kong.

The localist camp

Broadly speaking, the rise of localism as a third prominent political ideology in the 2010s was due to both the increasing integration with mainland China both politically and socio-economically as well as the perceived incompetence of the pro-Democracy camp to push forward substantive democratisation in Hong Kong. Although the term "localist camp" in reality represents an aggregation of groups which vary considerably in choosing violent or non-violent means of resisting against the Beijing Government, whether Hong Kong should abandon OCTS and pursue independence, and are much less coordinated in their actions compared to the pro-Beijing and pro-democracy camps, organisations belonging to this camp hold strong aversive sentiment against China **both** in terms of the PRC regime and mainland Chinese who are deemed to be depriving the locals of already scarce public resources or disrupting the way of life of ordinary Hong Kong citizens, and they (Kwong 2016) . The first substantive discourse to materialise localism as a political ideology was "Hong Kong as a city-state" published in 2011 by scholar *Chin Wan-kan* which, among many other ideas, advocates separating the integration between Hong Kong and mainland China as well as prioritising the needs of Hongkongers (So 2015) . The following paragraphs will outline how the rise of localism can be treated as a backlash against both long-existing political camps in Hong Kong while also exploring how localists perceive Hong Kong's national identity.

Starting with the so-called "China factor", the localist camp obviously rejects the pro-Beijing camp's (and the Chinese Government) claim that Hong Kong's identity should be treated as one with that of Chinese. Fong (2017b) illustrates how the China factor has facilitated the rise of localism under the framework of "state-building" versus "peripheral" nationalism (p.2), where Hong Kong is attempting to consolidate its unique identity against Beijing's incorporation of the city into the hegemonic national culture. Specifically, after 500,000 Hongkongers had rallied against the Article 23 legislation in 2003, Beijing then started to increase its interventions in

2. Literature Review

multiple areas, such as asserting that its authority overrides the city's autonomy, sponsoring pro-Beijing candidates during elections and fostering further economic integration between Hong Kong and China in finance, tourism and trade.(So 2015; Fong 2017b) . In turn, the high proportion of respondents in a 2016 survey identifying themselves as Hongkongese amidst the foundings of more radical localist organisations which aspire to make Hong Kong a separate nation can be evidence of how the PRC's attempts to assimilate Hong Kong into the hegemonic Chinese identity has risen the consciousness of Hong Kong people about their uniqueness vis-a-vis China and thus foster the increasing popularity of localism (Fong 2017b) . Apart from Chin's foundational discourse on localism as mentioned before, overt advocacy of Hong Kong becoming an independent nation from China is often proposed by some localist organisations, most signified by the establishment of the disbanded *Hong Kong National Party* in 2016 by Andy Chan which aimed to build a "Hong Kong Republic" (Kaeding 2017) (p.165). Some organisations even dare openly insult the PRC to challenge its authority and legitimacy, as two former legislators from the self-determination-advocating *Youngspiration* had taken their office oaths by pronouncing "China" derogatorily in 2016, which caused them to be disqualified from their seats shortly afterwards (Kaeding 2017) .

Due to the localist camp's dissociation of Hong Kong with mainland China concerning both Beijing's assimilation and which nation Hong Kong belongs to, they also reject the pro-Democracy camp's idea that political development of Hong Kong should be linked to that of China. In effect, the indifference of the localist camp to China's political situation is largely due to the "transition fatigue" which witnesses increasing dissatisfaction with the traditional pro-democracy camp has started to accumulate due to the lack of substantive progress in democratisation by 2008 (Kwong 2016) (p.63). The Democratic Party's decision to negotiate with the Beijing Government for the political reform package and its subsequent support in May 2010 created profound division within the pro-democracy camp as certain parties believed that more radical measures were needed to fend off Beijing's interference into the city, but the more significant turning point for localism to gain its popularity was the 2014 Umbrella Movement when some youngsters blamed the pro-democracy camp for the Movement's failure (Kwong 2016; Kwan 2016) .

The disagreement between the pro-democracy and localist camp lies less in what kinds of values Hong Kong should uphold (as both camps aim at fighting for democracy and autonomy from Beijing) than whether Hong Kong should involve in China's democratisation. Specifically,

2. Literature Review

the pro-democracy camp's insistence on Hong Kong's responsibility to help build democracy in China is met with strong criticism from the localists, who consider that Hong Kong should focus only on its political development since binding Hong Kong's future to a democratic China would hamper the city's own interests (Kwong 2016; Kwan 2016) . Thus, the different conceptions of which nation Hong Kong belongs to between the two camps and how much Hong Kong should dissociate from China have caused both camps to have rather contentious relations, as some pan-democrats also consider the localist camp's hostility towards mainland Chinese as xenophobic (Kwong 2016) .

In short, the localist camp does not just treat the Hong Kong identity as distinctive but ultimately belongs to the Chinese nation like the pro-democracy camp. Instead, Hong Kong is a separate political entity that should avoid close integration with mainland China in order to protect its own interests. This causes the localist camp to become a third force in Hong Kong's political arena in the 2010s confronting against both the pro-Beijing and pro-Democracy camps.

Before moving onto hypothesising how the three camp's different views on Hong Kong's national identity may impact their views on asylum seekers within the city, table 2.1 summarises how they differ on the vision of the Hong-Kong-China relations.

Table 2.1: Comparison of the three camps on key dimensions of Hong-Kong-China relations

	Pro-Beijing camp China	Pro-Democracy camp China	Localist camp Hong Kong
The nation that Hong Kong belongs to			
Attitude towards political development in China	Political ally of the ruling PRC	Opposes one-party rule of the CCP, sympathetic to Chinese people	Indifference, only focuses on Hong Kong's situation
Prioritised political values	Patriotism and nationalism over individual rights	civil liberties, the rule of law and political autonomy	those of the pro-democracy camp plus prioritising Hong Kong's interests
Means of contention	N/A (as an ally of the Government)	Non-violent	Both violent and non-violent
Degree of within-camp coordination	Strong	Moderate	Weak

2.2.2 Asylum seekers in Hong Kong

Hong Kong's policy regime

Regarding the legal framework and policies of assessing asylum claims and refugee recognition, the Hong Kong Government has a strong position of not making the city a destination for refugee resettlement. Although Hong Kong is a signatory of the 1992 United Nations Conventions Against Torture (UNCAT) which requires the Government to temporarily house asylum seekers and process torture-related asylum applications, the city has never signed the 1951 Refugee Convention and thus has no legal obligations to permanently resettle asylum seekers and refugees within its territory (Ng 2020) . This remains true even after the implementation of the Unified Screening Mechanism (USM) in 2014 which made the Government become the sole handler of all the non-refoulement claims, including those which had been previously handled by the UNHCR for refugee status (Ng 2020; *Unified Screening Mechanism Procedures* 2021) . Instead, the Government reiterates its lack of intention of determining individuals' refugee status and allowing refugees to permanently settle in the city, while also claiming that the recognition and resettlement of refugees should still be referred to the UNHCR "in accordance with its mandate" (*Making a Claim for Non-refoulement Protection in Hong Kong* 2021) , even though the UN agency states on its website that it would no longer process new refugee and asylum claims following the implementation of the USM (*Unified Screening Mechanism Procedures* 2021) . Simply put, the official institutions responsible for assessing asylum applications in Hong Kong hardly have intention of making the city a destination of permanent resettlement for individuals fleeing from their countries because of persecutions.

Even if the HKSAR Government is obliged to adjudicate non-refoulement applications as per the UNCAT, the legal and social status of the asylum seekers are quite precarious while awaiting their application results. To begin with, individuals must wait until their visas expire in order to be eligible for non-refoulement applications and then have to report regularly to the Immigration Department (*Unified Screening Mechanism Procedures* 2021) , an arrangement that Ng (2020) considers as the Government's attempt to exclude asylum seekers from the rights granted to legal residents in the city (e.g. working,) and receiving welfare supports through criminalisation. Furthermore, the Immigration Department has the final authority to prosecute the asylum seekers for having illegally overstayed beyond the validity of their visas, and officials frequently call

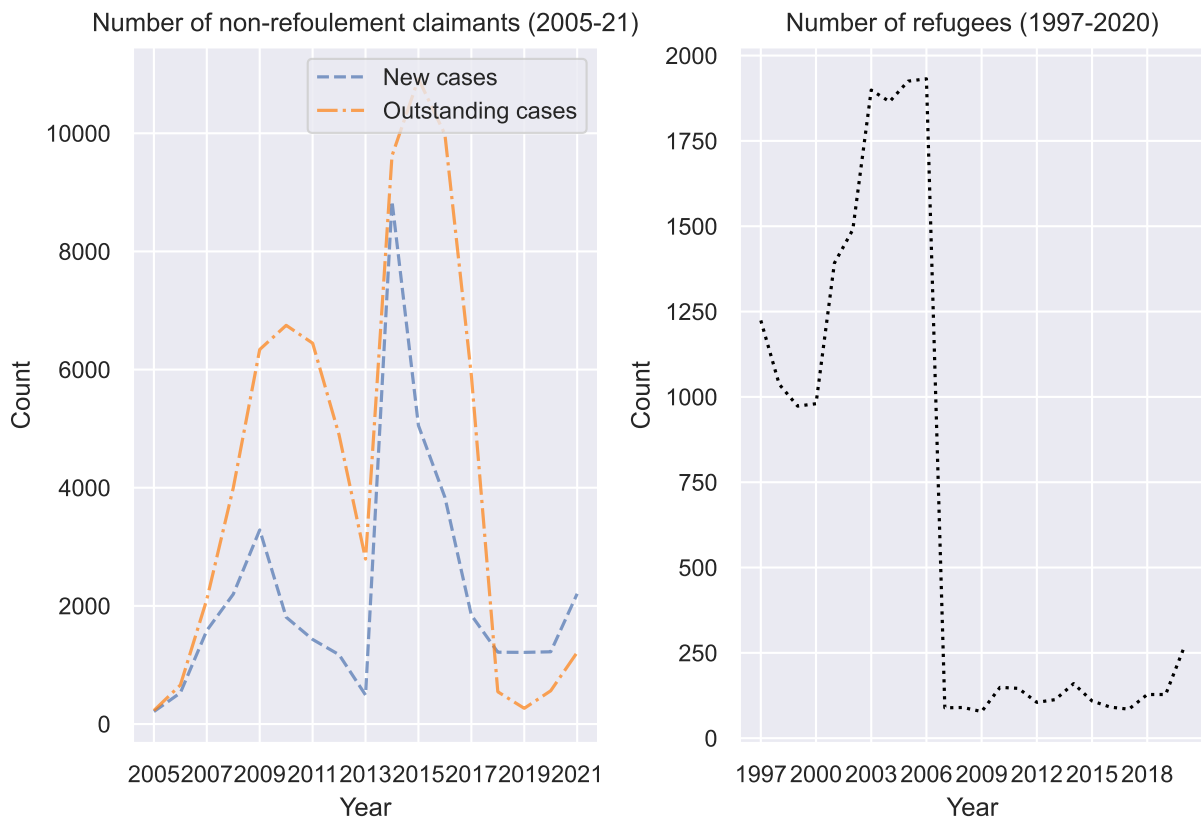


Figure 2.1: Number of non-refoulement claim cases (left) and refugees (right) in Hong Kong

these populations as *illegal immigrants* (often with racial labels) (Ng 2020) . Coupled with the Government's refusal to recognise and resettle refugees in its territory, the institutional framework of asylum seeking in Hong Kong can thus be described as unwelcoming towards those who are fleeing from persecutions in their countries of origin, since these applicants not only have to first become illegal immigrants for eligibility to request non-refoulement, but also they are under the constant threat of being charged for their immigration status.

The HKSAR Government's unwelcoming stance towards asylum seekers and refugees is likely to be one of the major reasons for the low numbers of non-refoulement claims and refugees in the city since 2005. For instance, the left plot in figure 2.1 (*Enforcement / Statistics on Non-refoulement Claim 2021*) shows that the number of *new* non-refoulement claims submitted each year from 2005 to 2021 on average was lower than 4000, and the sudden rise of new applications to around 9000 instances in 2014 was more due to the implementation of the USM requiring the Government to also assess claims previously handled by the UNHCR than the commencement of

2. Literature Review

a more liberal asylum-seeking regime (*Enforcement / Statistics on Non-refoulement Claim 2021*) . Likewise, the number of *outstanding* non-refoulement cases by year during the same period generally surpassed that of new cases except from 2018, a few years after the USM had been implemented. By 2018, non-refoulement claimants constituted to a mere 0.2% of Hong Kong's total population of over 7 million, a figure considerably smaller when compared to a total of 3.5 million refugees in Asia as Ng (2020) contrasts. The right plot in figure 2.1 (*Refugee Population by Country or Territory of Asylum - Hong Kong SAR, China / Data 2021*) shows how the number of *refugees* in Hong Kong has dramatically reduced to no more than 250 persons each year (save for 2020) since 2007 after the Government had closed the last refugee camp hosting Vietnamese who fled from the Vietnam War in the 1970s (Ng 2020) . Lastly, figure 2.2 (*Enforcement / Statistics on Non-refoulement Claim 2021*) indicates the nationalities of non-refoulement claimants, who mostly are from South and Southeast Asia, with a few from Africa as well. One particularly interesting insight from this graph is that despite the frequent racial labelling of non-refoulement claimants as South Asians by some media outlets (Ng, Choi, et al. 2019) , it is actually **Southeast Asians** who constitute the largest proportion of non-refoulement claimants.

Responses by political camps

Given the extremely small number of non-refoulement claimants, the issue of asylum seeking regime in Hong Kong hadn't gained much attention until early 2016, after a senior Immigration Department official's accusation of the USM being abused during a TV interview in October 2015 and the subsequent propagation of two fabricated WhatsApp audios alleging asylum seekers as perpetrators of robberies and homicides in the New Territories district (Ng, Choi, et al. 2019) . Soon afterwards, coverage of asylum seekers by newspaper media has increased, with two pro-Beijing newspapers named *The Sun* (which ceased publication in April 2016) and *Oriental Daily* contributing to the largest proportion of articles which depict this group of vulnerable population as a source of social problems (e.g. committing crimes) which warrant stringent measures like detention camps as solutions (Ng, Choi, et al. 2019) .

Different opinions of how the Government should deal with the asylum seekers in the political debate could also be classified according to political camps at the same time. On one side, echoing with the HKSAR Government's already restrictive policies for assessing non-refoulement

2. Literature Review

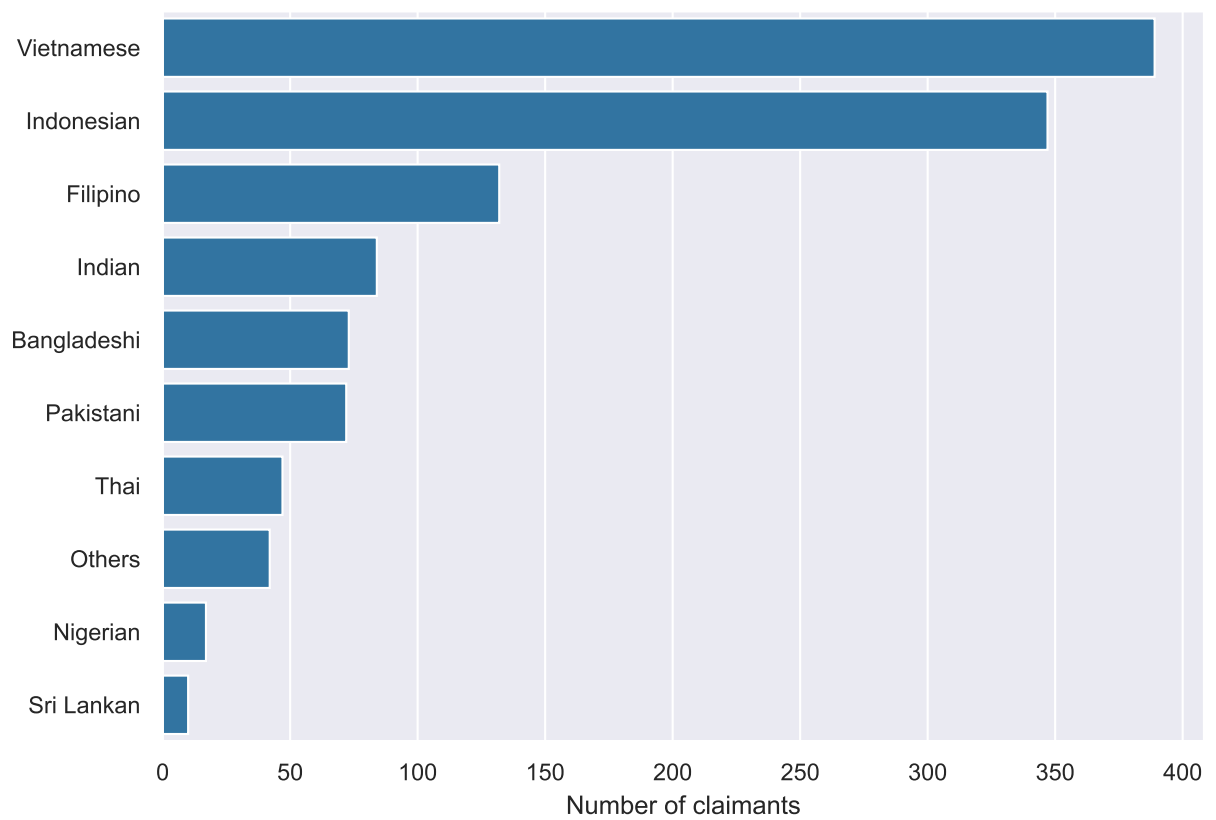


Figure 2.2: Number of non-refoulement claimants by nationality in 2021

2. Literature Review

claims, pro-Beijing camp politicians also create narratives that non-refoulement claimants and asylum seekers are anything but genuine in the legislature. For instance, less than a year after asylum seekers had gained salience in public debate, legislator Holden Chow submitted a motion in November 2016 under the banner of “combating bogus refugees” who, among other sins, were working illegally and/or committing crimes while calling for the revision of the USM and carrying out measures against illegal immigrants adopted by other countries (Yuen 2016) . Some pro-Beijing parties also reach out to the public to sway the latter’s opinion, such as the DAB¹ setting up a booth in the ethnic-minority-concentrated Tsim Sha Tsui district to hand out pamphlets about measures against “fake refugees” (Cheung and Grundy 2016) . Perhaps the most dramatic proposal of dealing with the so-called “fake refugee” issue by the pro-Beijing camp was the former Chief Executive *CY Leung*’s suggestion to withdraw from the UNCAT in 2016 so that Hong Kong would not be obliged to assess non-refoulement claims anymore (Ng, Choi, et al. 2019) . Coupled with the high frequency of pro-Beijing media outlets in pejoratively portraying asylum seekers in Hong Kong as mentioned earlier, the pro-Beijing camp promotes an unwelcoming rhetoric against asylum seekers that is consistent with the Government’s refusal to permanently resettle refugees and asylum seekers.

By contrast, the pro-democracy camp is less hostile towards non-refoulement claimants and even advocates better treatment of these vulnerable individuals. When Holden Chow’s motion was tabled for voting in the legislature in December 2016, it was barely rejected due to the pro-democracy camp’s opposition of Chow’s motion as “hypocrisy and discrimination in disguise” by indiscriminately applying the proposed “draconian measures” on all asylum seekers (Cheng 2016) . Among all pro-democracy legislators, Dr Fernando Cheung was perhaps the most vocal in fighting for better treatments of asylum seekers in Hong Kong, as he met with non-refoulement claimants directly on several occasions while receiving approval from this group of population (Justice Centre Hong Kong 2014; Vision First 2015; Refugee Union 2018) . Indeed, legislator Cheung’s prominent advocacy for the rights of asylum seekers had indirectly caused him to become a victim of political smearing a few months before the 2016 Legislative Council election, as political banners imposing legislator Cheung’s party were seen hanging in streets which maliciously called the legislator “the father of refugees” (Ngo 2016) . It should also be noted that, nevertheless, a

¹A list of the 30 most prominent words in each of the topic and the code to generate it are available in the appendix.

2. Literature Review

more lenient attitude towards asylum seekers does not mean that the pro-democracy camp totally disregards the perceived issue of bogus refugees. Even though legislator Charles Mok condemned Holden Chow’s motion against “bogus” refugees was “disturbing”, he also attributed the most responsibility of “bogus non-refoulement claims” to the Government due to the latter’s inefficiency and lack of knowledge in handling the issue (Cheng 2016) . But in general, the pro-democracy camp supports a more humane treatment towards non-refoulement claimants who are genuinely in need of fleeing from their countries due to prosecution.

Lastly, the localist camp’s stance on dealing with asylum seekers is less confrontational than that towards mainland Chinese. As one of the few localist lawmakers who were not disqualified in the oath-taking controversy by December 2016, Dr Cheng Chung-tai’s speech on Holden Chow’s anti-fake-refugees motion had two points which worth highlighting² (*Hui4yi4 Guo4cheng2 Zheng4shi4 Ji4lu4* 2016) (p.1232-33). Firstly, he shared the pro-democracy camp’s view of the main reason for the so-called “bogus refugees” issue being the Government not allocating adequate resources to assess the claimants’ applications. But the more intriguing point of view is how he framed Hong Kong’s signatory status of the UNCAT as a proof of the city having its own *de-facto* sovereignty from mainland China, and thus he opposed CY Leung’s suggestion of withdrawing from the treaty since it would signify Hong Kong’s regression in civilisation to the world. Interestingly, the localist camp’s reservation of adopting more stringent measures against asylum seekers has drawn smear from some pro-Beijing medias, such as an editorial from the Sun (2016) accused the localists of “remaining silent” on non-refoulement claimants wasting public resources of Hong Kong.

²For the complete documentation of the default parameters of the models used in this thesis, refer to the websites of scikit-learn and XGBoost Documentation — xgboost 1.5.1 documentation.

Chapter 3

Methods

After reviewing both the theories of the impact of national identity on xenophobia as well as how political camps in Hong Kong imagine the city's national identity, how could we know the possible associations between each camp's definition of the city's national identity and its attitude towards asylum seekers who are mainly non-ethnic Chinese? This section will discuss the strategy of answering this question.

3.1 Hypotheses

Based on the aforementioned literature in the previous section, I will formulate hypotheses to be tested and explain why I expect them to go in certain directions. The main hypothesis to be tested will be:

H₁: Pro-Beijing media outlets are more likely to negatively portray asylum seekers than media outlets from other political camps.

I expect the pro-Beijing camp news media to frame asylum seekers more negatively based on the literature on the relationship between national identity and xenophobia (e.g. Hjerme (1998) and Heath and Tilley (2005)) which states that a more ethnic-based conception of national identity is generally associated with a higher degree of xenophobia towards immigrants, of which asylum seekers is one of the (most vulnerable) subgroups of immigrants. Accordingly, since pro-Beijing politicians and officials (including the Chinese President *Xi Jinping* himself) (Lin and Jackson 2021) resort heavily to ethnic-based discourses for justifying Chineseness of the Hong Kong identity, pro-Beijing media outlets can be expected to hold more negative attitudes towards

3. Methods

the mostly non-ethnic Chinese non-refoulement claimants. In particular, the pro-Beijing camp's emphasis on how Hong Kong citizens belong to the greater Chinese national identity based on shared ethnicity with mainland Chinese might lead to its affiliated media outlets to portray asylum seekers as a threat to society since their ethnic distinctiveness marks them as ever-being strangers of the family of the majority Chinese population as Erni (2012) describes.

By contrast, since the pro-democracy camp focuses more on civic values (e.g. civil liberties, the rule of law) while defining the Hong Kong identity, the non-Chineseness of asylum seekers should appear less threatening to the Hong Kong identity from the pro-democracy camp's point of view. Lastly, the localist camp's primary focus on the mainland-Hong-Kong division regarding the city's identity again prevents a clear direction for H_1 from being formulated beforehand.

3.2 Data collection

I will study news media reports on asylum seekers and non-refoulement claimants in Hong Kong by local mass media during the year of 2019. Specifically, media publishing in **Chinese** in printed format will be included in the data¹. For the retrieval of the relevant newspaper articles, I will rely on the *WiseNews* database which can fetch news articles with search keywords published in Hong Kong². While it would have been more ideal to also include online news media into the dataset as Lee (2018) mentions that they have become quite prominent in the 2010s for providing counter-narratives to the mainstream media which are often co-opted by the Chinese Government³, WiseNews does not directly provide the texts of online media news article and thus requires additional steps of web scrapping which are, unfortunately, beyond the author's capacity by the time of writing. That being said, including only the printed news media during 2019 would still cover 16 of the well-known news media with associations to different political camps according to the survey on media credibility by the Chinese University of Hong Kong (CUHK 2020). In other words, just searching for printed news media outlets still allows me to discover whether the attitudes of reporting on asylum seekers by some of the most well-known and widely circulated media outlets in Hong Kong might be related to their political associations.

¹A list of the 30 most prominent words in each of the topic and the code to generate it are available in the appendix.

²For the complete documentation of the default parameters of the models used in this thesis, refer to the websites of scikit-learn and XGBoost Documentation — xgboost 1.5.1 documentation.

³The code for tuning the model can be found in the appendix.

3. Methods

713 The reason for choosing to collect media reports published in 2019 is because it was when
714 the Anti-Extradition Law Amendment protest first broke out in June and then lasted for several
715 months, a period when one can argue the manifestation of political ideologies should be the most
716 salient for all the political camps. This should indeed be the case for both the pro-democracy
717 camp whose supporters participated in several large-scale protests and sometimes even engaged
718 in numerous violent clashes with the Hong Kong police to oppose a law that would have permitted
719 extradition of fugitives to mainland China, and the pro-Beijing camp whose authority was
720 once again immensely challenged after the 2014 Umbrella Movement. Accordingly, it would be
721 interesting to investigate whether the issue of asylum seekers still remained its salience and was
722 reported in similar manners given the co-occurrence of another arguably more salient political
723 event (i.e. the anti-Extradition Law protests), as compared to when the issue first had become
724 prominent back in 2016. Specifically, analysing newspaper articles in 2019 provides an opportunity
725 to observe whether and how asylum seekers in Hong Kong might be represented by the
726 media in tandem with major political events.

727 It is worth discussing some of the limitations of using media reports as the data for this
728 thesis. For starters, media reports do not necessarily reflect how individuals identifying with
729 each political camp might perceive asylum seekers directly and might not be fit for investigating
730 the causal mechanism of how definition of Hong Kong's national identity may cause a political
731 camp to frame asylum seekers in particular ways. Despite these shortcomings, media reports
732 can still adequately help prompt the hypothesis of whether depictions of asylum seekers in Hong
733 Kong might be associated with their political camp's definition of the city's national identity. In
734 fact, previous research about the framing of asylum seekers by mass media (Cooper et al. 2017;
735 McKay et al. 2011; Ng, Choi, et al. 2019; O'Doherty and Lecouteur 2007) point to its potential
736 impact on affecting how the public may perceive the issue via being a major provider of related
737 information, and sometimes these media reports may even serve to justify policymaking or reflect
738 the government's position on the issue. Drawing from Fong (2017a) and Lee (2018), the media
739 industry in Hong Kong is quite remarked for their political stances, and quite a number of outlets
740 have their owners being affiliated with the Chinese and/or HKSAR Governments. Therefore,
741 reports published by media outlets should be able to reflect political camps' attitudes towards

3. Methods

742 asylum seekers and, perhaps more importantly, the version of framing that they would like to per-
743 suade their readers to accept, thereby making them a suitable data source to test the hypotheses
744 proposed above.

745 Language-wise, including only Chinese-language news media omits how asylum seekers may
746 be represented by English-language media in Hong Kong (e.g. the South China Morning Post) and
747 thus cannot investigate whether the representation of asylum seekers by media outlets to potential
748 non-Chinese audience may also vary based on each outlet’s political affiliation. Moreover, solely
749 focusing on textual content of the media reports means discarding graphical and (if also published
750 online) audio-visual materials accompanied in a given news report, which may contain additional
751 information that are not observed in the texts of news articles. Nevertheless, analysing the textual
752 content of the media reports should still yield a considerable amount of information about the
753 representation of asylum seekers in Hong Kong by media outlets linked to different political camps,
754 since after all a large proportion of the content in news reports are represented textually.

755 To search for relevant media reports on WiseNews, I will refer to the keywords used by
756 Ng, Choi, et al. (2019) which refer to the group of population-in-interest. These terms can
757 either be directly describing this group of population (e.g. refugees, illegal immigrants) or policies
758 related to asylum seeking (e.g. non-refoulement claim, immigration paper). *Table 3.1* lists all
759 the Chinese terms used for searching the relevant news articles on WiseNews and their meanings
760 in English. Additional notes are used for indicating whether some articles were omitted and
761 why. The author also omitted articles only mentioning the keywords briefly throughout the main
762 text which discusses issues not directly related to asylum seekers. As for the keywords “South
763 Asian”, “illegal immigration” and “human snake”, the author only included news entries fetched
764 from these keywords if at least one of the other keywords directly referring to asylum seekers
765 (i.e. non-refoulement claim(ant), torture claim(ant), asylum seeking, refugee and immigration
766 paper) co-appeared in either the title or main text. This is to ensure that the news articles in
767 the dataset are obviously describing non-refoulement claimants in Hong Kong.

3. Methods

Table 3.1: Translation of the keywords used to search for media reports on WiseNews

Original terms in Chinese pinyin	Literal meaning in English	Notes
mian3 qian3fan3 sheng1qing3	Non-refoulement claim(ant)	
ku4xing2 sheng1qing3	Torture claim(ant)	
fei1fa3 ru4jing4, ren2she2	Illegal immigration, human snake (alias of illegal immigrants)	Excluded illegal immigration from mainland China
xun2qiu2 bi4hu4	Asylum seeking	Excluded asylum seekers <i>outside of</i> Hong Kong
nan4min2	Refugee	Excluded refugees <i>outside of</i> Hong Kong and Hong Kong activists seeking refuge overseas
nan2ya4	South Asian	Articles about the South Asian region are excluded.
xing2 jie1 zhi3	Immigration paper	Documents issued by the Immigration Department to non-refoulement claimants.

3.3 Modelling

Since the period under study spans for the year 2019, the dataset contains almost 700 articles in total which would be extremely time- and resource-consuming to qualitatively analyse each article line by line, whereas only focusing on a small subset of the articles may risk losing generalisability. On the other hand, while it is true that using human coding (whether it be done by trained experts or crowd coding platforms) for sentiment analysis will be likely to yield the most accurate results as van Atteveldt et al. (2021) have noted, doing so will be too resource-intensive for a master-level thesis, and this method may also fall short of reliability, i.e. how other researchers may replicate the study if wanted. Despite some of the caveats listed by van Atteveldt et al. (2021) (e.g. accidentally identifying spurious relationships between variables, problem of overfitting) which can be remedied with corresponding solutions, I will use supervised machine learning to test the hypothesis. This means H_1 will be treated as a sentiment analysis problem which classifies the *polarity* of the news articles (i.e. whether they reported on asylum seekers *positively*, *neutrally* or *negatively*). How the independent (both the media outlet’s political camp and other controls) and dependent variables are operationalised and coded will be discussed in the next section, and for now I will talk about what kinds of tools and models will be used for the analysis.

3. Methods

Starting with the models that will be used for the analysis, I will rely on some commonly used classification models. Candidate models include logistic regression, support vector machine (SVM), random forest and gradient boosting. Although van Atteveldt et al. (2021) also mention that deep learning is another popular choice of machine learning model for sentiment analysis tasks, I did not use such a model for this paper due to the relatively small size of the dataset at hand as well as the difficulty of interpreting the features' impact on the dependent variable directly from neural network models. I will first run some baseline models with the default parameter settings of the models, and then select the one for further tuning based on their performance on the *log loss* of the validation set (or test set). To decide whether I should use the tuned or baseline model for interpretation later, I will also look at the *f1 score* (or the harmonic mean of *precision* and *recall*) between the two versions. The model version which overall obtains a higher f1 score across all the classes in the dependent variable will then be used for calculating the *SHAP values* of the features (Lundberg 2022) which evaluate how much impact each feature has on the model prediction when the features is at certain values versus at its baseline value. In essence, the higher the SHAP value of a feature, the higher its impact of the model's prediction.

In order to pre-process the articles into suitable formats as machine learning model inputs, I will also transform the news articles into a term-document matrix which records the frequency of each word (i.e. token) appearing in each article (i.e. document). Instead of using a simple bag-of-word approach, I will use the term-frequency inverse-document-frequency (TF-IDF) matrix transformation which also considers in how many documents a word appears out of all the documents in the corpus. This is because words that appear too frequently across all the documents in the corpus may not convey much meaning of each document and thus should be treated as stop words. Due to the relatively small size of the dataset (with only around 550 articles), I will also adopt non-negative matrix factorisation (NMF) to reduce the dimensionality of the dataset to prevent overfitting. According to Stevens et al. (2012) (p.953), the matrix denoted as H which captures the weight of each topic (as columns) in each document (as rows) of the corpus can help summarise the information of the articles in terms of which topic(s) they primarily focus on. Thus, I will use NMF to transform the TF-IDF matrix into another which summarises the news articles into a certain number of pre-defined latent topics (the exact number will be figured

3. Methods

813 out in the *preprocessing* section later), and then use this matrix as part of the inputs for the
814 models.

The majority of the implementation from data preprocessing, modelling to validation will be done via the `scikit-learn` package in Python (Pedregosa et al. 2011), whereas the tokenisation of the Chinese-language news articles will be done by the `jieba` package which is designed for natural language processing in Chinese (Junyi 2021). For better tokenisation of the articles, I will also remove stop words, punctuations and hyper-links of the texts and add words related to Hong Kong politics (downloaded from elgarteo (2021)) and asylum seekers into the dictionary. As for the gradient boosting model, the `XGBoost` package

@chenXgboostExtremeGradient2015

815 will be used since it is one of the most renowned machine learning models in data science, and it
816 also offers a lot of options for tuning the model for better performance.

817 3.4 Operationalisation

818 3.4.1 Main independent variable

819 Starting from the main independent variable which classifies the associated political camp of the
820 media outlets, I will mainly refer to the literature exploring how Hong Kong’s media industry is
821 intertwined with politics and perform additional research when needed. I consulted the articles
822 by Fong (2017a) and Lee (2018) on the development of post-1997 Hong Kong’s media industry,
823 of which they have documented the co-optation of Hong Kong media outlets by the Chinese
824 Government (e.g. the major shareholders holding positions in such as Chinese People’s Political
825 Consultative Conference). Building on the concept of “allocative control” (p.11) by Lee (2018)
826 about media outlet owners having the ultimate decision on the personnel and resources at the
827 newsroom’s disposal for reporting news which may lead to self-censorship of the frontline journal-
828 ists, I will code media outlets whose owners are affiliated with the official Chinese and/or Hong
829 Kong Government organisations as belonging to the pro-Beijing camp. As for those media owners
830 with less apparent links with the Chinese or Hong Kong Governments, I will decide the classifi-
831 cation according to whether the owners are known to be affiliated with a certain political camp
832 and how a given media outlet is perceived for its political stance. Media outlets without apparent

3. Methods

political stances are labelled as “neutral”. Table 3.2 is the resulting classification scheme. Note that due to the medium of publication of localist news media being almost exclusively online, none of the outlets in that camp was included in the dataset. Therefore, it would not be possible to examine how the localist media outlets might report on asylum seekers and non-refoulement claimants in news articles.

Table 3.2: Classification of the associated political camp of media outlets in Hong Kong included into the sample, source: Fong (2017a); Lee (2018); author’s own research

Associated political camp	Media outlets
Pro-Beijing camp (directly owned by the Liaison Office or the HKSAR Government)	Wen Wei Po, Hong Kong Government News, Hong Kong Commercial Daily, Ta Kung Pao
Pro-Beijing Camp (owners affiliated with political institutions on mainland China/ HKSAR Government)	Oriental Daily News, Sing Tao Daily, Sing Pao, Headline Daily, Hong Kong Economic Journal, Hong Kong Economic Times, am730, Sky Post, HK01 Newspaper
Pro-Democracy camp	Apple Daily, Kung Kao Po
Localist camp	N/A
Neutral	Metro Daily, Ming Pao Daily News

3.4.2 Dependent variable

As mentioned before, the dependent variable of H_1 will be the sentiment of the news articles towards asylum seekers in the polarities of *positive*, *neutral* and *negative*. But how exactly should these three polarities mean in this context? Again referring to the literature can help devise a sound coding scheme. Drawing McKay et al. (2011), Cooper et al. (2017) and Ng, Choi, et al. (2019), *positive* articles usually acclaim the person or event in concern, describe personal stories, report expert opinions without derogatory terms or depict refugees favourably, *neutral* articles simply report an incident without an overt sentiment or opinion towards particular stakeholders and with little interpretation by the journalists, and *negative* articles frame asylum seekers either as bringing problems to the host society or use inaccurate terms to delegitimise this group of population. I will follow these coding criteria for the polarities of the news articles in the dataset since these definitions suit the purpose of this study. The numerical representation of the polarities in the dataset will be as follows, namely, positive articles will be coded as 2, neutral articles will be coded as 1, and negative articles will be coded as 0.

For the context of Hong Kong about the “incorrect” terms and information used for describing asylum seekers, I will refer to expert opinions about whether descriptions of asylum seekers in

3. Methods

Hong Kong are founded⁴. For example, Ng, Choi, et al. (2019) point out that the term “fake/bogus refugees” (*jia3 nan4min2*) should be considered as both derogatory and inaccurate because those applying for non-refoulement claims are not automatically qualified as refugees while awaiting the final decisions of their applications. *Therefore, the polarity of news articles in which the news media directly refers non-refoulement claimants as “fake refugees” in its reportage will be coded as negative, but the report should be coded at least as neutral instead if the news outlet is merely quoting someone who describes non-refoulement claimants as “fake refugees”.* Furthermore, Ng (2020) mentions that non-refoulement claimants are forced to be illegal immigrants by the policy design of the HKSAR Government so that this group of population can be excluded from social rights that legal residents enjoy. News articles equating non-refoulement claimants as illegal immigrants will thus also be coded as negative since such reportage is justifying the exclusion of non-refoulement claimants by resorting to their illegality that had been forced upon by the HKSAR Government.

3.4.3 Control variables

I will also incorporate some control variables to see if the association between political camps and polarities of the news articles still remains significant afterwards. For starters, the time when the articles were published will be included as a control, though I will bin the twelve months into four evenly split quarters (*Quarter*) with the first quarter beginning in January and ending in March. Since 2019 was also the year when one of the most large-scale and long-lasting pro-democracy protests broke out, It would be interesting to see how asylum seekers may be portrayed during the peak of the protest. Moreover, the presence of racial labels will be included as a control, since it is possible that more overtly racialised discourses may be associated with more negative sentiments towards asylum seekers in general according to the literature. This variable will be named *Racial_label* in the dataset and will be binary, with 1 meaning there is presence of racial labels in the main text and 0 otherwise. Racial labels referring to both regional (e.g. South Asia, Africa) and national (e.g. the nationalities included in figure 2.2) will be counted. Lastly, the metadata of the news articles, namely the character counts of titles and main texts, will also be

⁴A detailed discussion of the validity of the depiction of asylum seekers in Hong Kong can be found in Ng (2020) and Ng, Choi, et al. (2019).

3. Methods

881 added to see if the lengths of the articles may be associated with the sentiment towards asylum
882 seekers.

Chapter 4

Results

After the data collection process, there were in total 557 articles published in 2019 by 16 newspapers which reported on non-refoulement claimants residing in Hong Kong. In this section, I will first explore the data set, and then move onto sentiment analysis with machine learning models to find out whether the political camp of media outlets is associated with the polarity of the news articles towards asylum seekers.

4.1 Exploratory data analysis (EDA)

4.1.1 How does the number of news articles vary by political camps and month?

Starting with the number of articles by media outlets as shown in the left plot of figure 4.1, consistent with the study by Ng, Choi, et al. (2019), Oriental Daily News continues to be the media outlet covering the most frequently on asylum seekers with 384 (or 68.94%) articles throughout 2019. By contrast, the second-most frequent publisher *Sing Tao Daily* only had 45 entries (or 8.08%) of the total number of articles published. Each of the other newspaper outlets only constituted to a small portion of news articles about non-refoulement claimants in 2019. Therefore, the issue of asylum seekers in Hong Kong still appeared to be the most salient for Oriental Daily News by 2019, evidenced by its unmatched volume of articles related to this topic vis-a-vis other media outlets.

On a higher level of political stance, the right plot of figure 4.1 indicates that largely due to the huge volume of articles by Oriental Daily News, the pro-Beijing camp dominated the coverage of asylum seekers in Hong Kong in 2019. Meanwhile, both neutral and pro-democracy newspaper outlets published similar amounts of articles throughout 2019, and both camps constituted to

4. Results

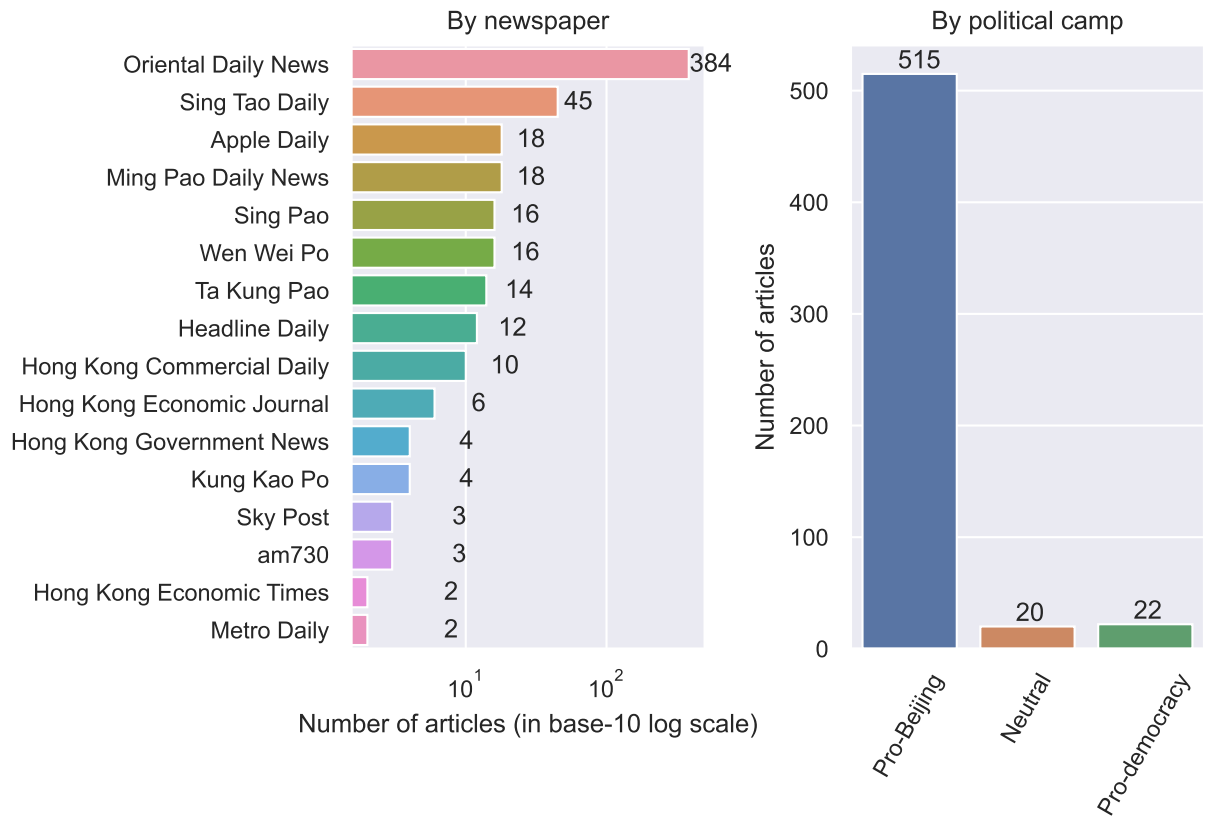


Figure 4.1: News articles on asylum seekers in 2019 by news outlet (left) and political camp (right)

small proportions of the share of articles during the year. Even if we omitted the sheer volume of articles published by Oriental Daily News, the pro-Beijing media would still have 173 articles published altogether which was still considerably more than the quantity of articles by neutral and pro-democracy media outlets.

It will also be intriguing to see how the number of articles might vary by month in 2019. As noted before, the anti-extradition law protest lasted mostly from June to November when numerous large-scale clashes between protesters and the police occurred. From figure 4.2, it appears that coincidentally, there were the fewest amounts of articles about asylum seekers published between August and November when some of the most intense clashes (notably the siege of the Hong Kong Polytechnic University in November 2019) took place.

In short, the majority of news articles about non-refoulement claimants in Hong Kong in 2019 were published by pro-Beijing media outlets, of which a huge proportion was from Oriental Daily News. Moreover, the number of articles by month was the lowest from August to November when

4. Results

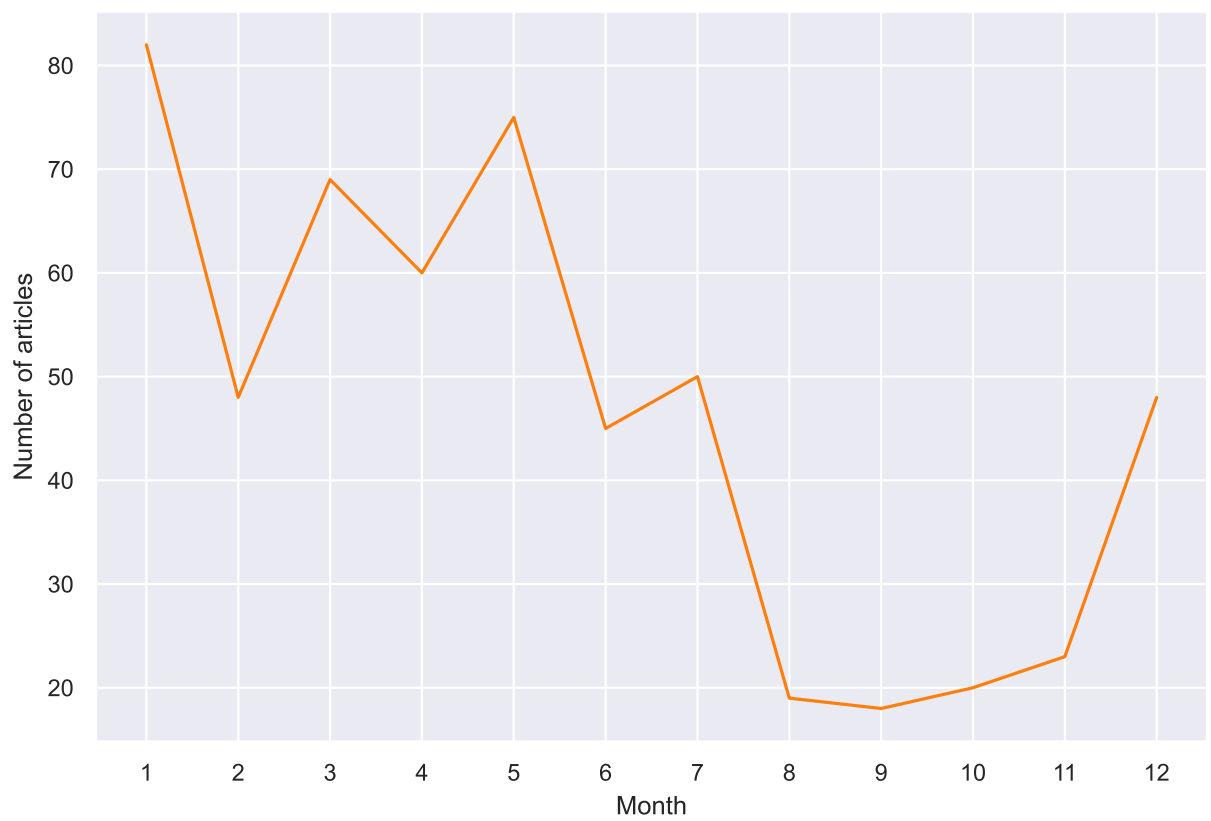


Figure 4.2: Temporal patterns of the publication of news articles about asylum seekers in Hong Kong in 2019

4. Results

Table 4.1: Polarities of the news articles on asylum seekers in Hong Kong in 2019

	Negative	Neutral	Positive	All
Neutral	0	12	8	20
Pro-Beijing	402	108	5	515
Pro-democracy	0	11	11	22
All	402	131	24	557

the anti-extradition law witnessed some of the most large-scale and intense clashes.

4.1.2 Polarities of the news articles

According to table 4.1, the polarity of the news articles about asylum seekers in Hong Kong in 2019 tilted towards negative, since only around 4.3% and 23.5% of articles respectively depicted asylum seekers positively and neutrally. The fact that the sentiment of the news articles in 2019 was skewed towards negativity implies that I will need to take class imbalance into account for modelling later. Political-camp-wise, pro-Beijing media outlets had over 70% of its articles depicting asylum seekers in Hong Kong in negative lights, whereas neutral and pro-democracy media outlets had their reportage evenly spread between neutral and positive articles (albeit they altogether constituted to only a small proportion of the total number of articles in 2019). While H_1 shall be tested formally with machine learning models after including other control variables later, preliminary evidence suggests that the polarities of the news articles vary with the political camp that the outlets belong to.

4.1.3 Presence of racial labels

Given the majority of asylum seekers in Hong Kong being non-ethnic Chinese, it will also be worth glimpsing whether the presence of racial labels for describing asylum seekers is associated with the sentiment of the news articles. Judging from figure 4.3 preliminarily, however, it appears that the patterns of the polarities are quite similar whether news articles contain racial labels or not, namely, most of the articles framed non-refoulement claimants negatively, some reported on events about this group of population neutrally, and only a small amount of articles were favourable towards asylum seekers residing in the city. In any case, the machine learning models can add the presence of racial labels as a control variable to test this potential association more formally later.

4. Results

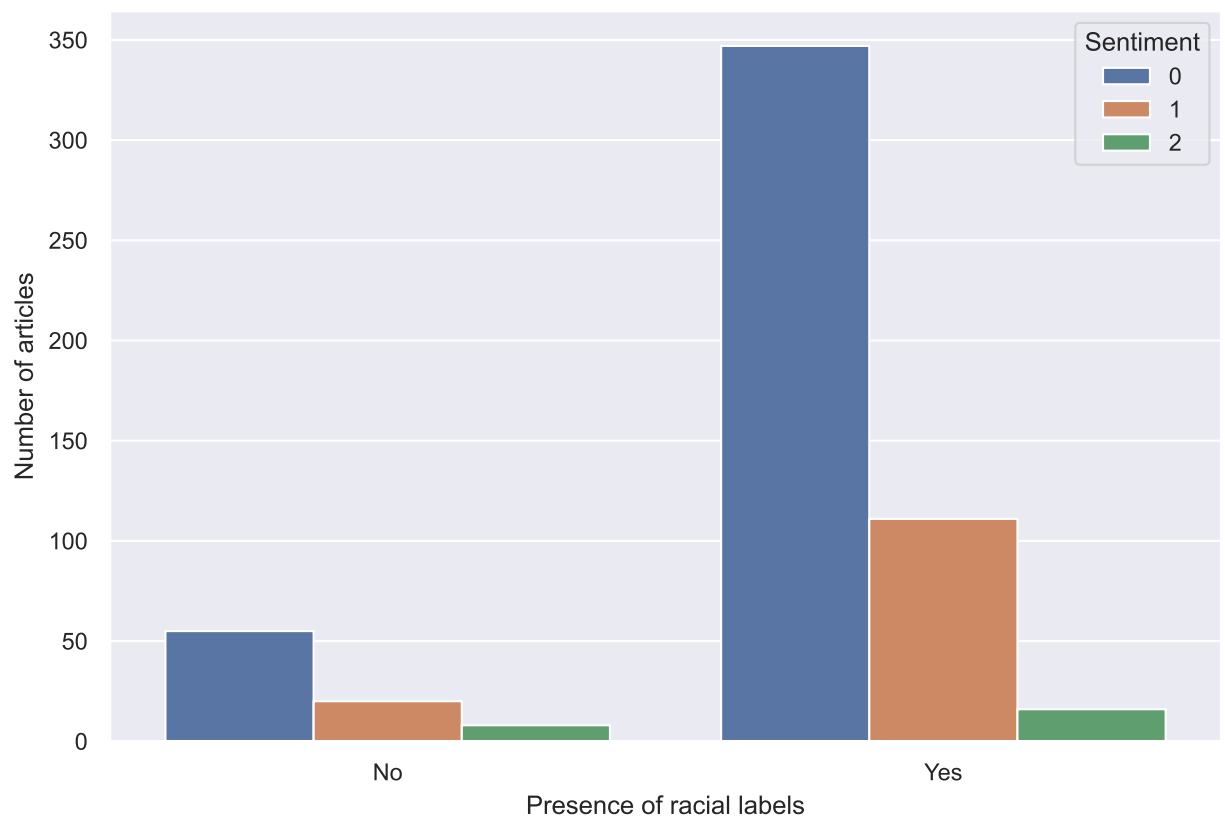


Figure 4.3: Presence of racial labels in the news articles by sentiment

4. Results

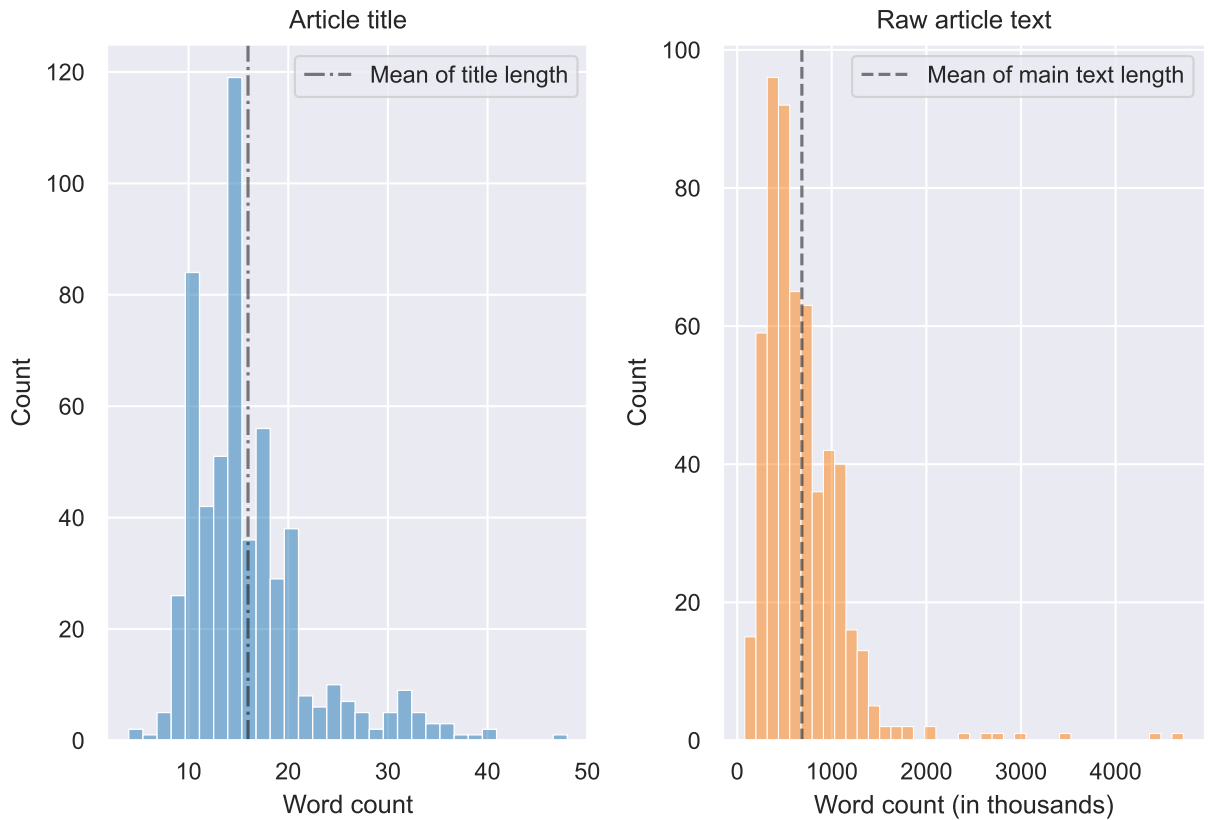


Figure 4.4: Distributions of the word counts of the articles' titles (left) and main texts (right)

4.1.4 Character lengths of news articles and titles

Lastly, let's look at the distribution of the character lengths of the titles and main texts of the news articles. According to figure 4.4 and table 4.2, it appears that both the title and main text lengths have right-skewed distributions. In other words, while most of the news articles on asylum seekers in Hong Kong in 2019 had relatively short titles and/or main texts, a few of them were considerably more verbose than the rest of the articles.

4.2 Sentiment analysis

4.2.1 Preprocessing

After making sense of the dataset with EDA, it is time to build the sentiment analysis model to see whether the political affiliation of news media outlets is associated with the polarities of the news articles after controlling for other variables. But first there are some preprocessing steps to be done so that the data are transformed into suitable formats as inputs for machine learning

4. Results

Table 4.2: Summary statistics of the word counts of the news articles' titles and main texts

	Title	Raw main text
count	557.000000	557.0000
mean	15.965889	683.9264
std	5.993154	453.9607
min	4.000000	80.0000
25%	12.000000	404.0000
50%	15.000000	581.0000
75%	18.000000	893.0000
max	48.000000	4715.0000

models. For starters, columns of the metadata should be excluded for being the inputs of the models. Note that I have also removed the `Newspaper` column since H_1 is more interested in whether newspaper outlets of the pro-Beijing camp *as a whole* may hold more negative attitudes towards asylum seekers in Hong Kong vis-a-vis media outlets with other political stances. The removed metadata columns are: `Index`, `Date`, `Category`, `Page_number` and `Newspaper`.

Furthermore, I have binned `Month` into four even split yearly quarters (`Quarter`) to reduce the dimensionality of the dataset. A further note on the categorical features is that they will need to be transformed via one-hot encoding, meaning that each of them will be transformed into n variables, with n being the number of the original distinct values. Meanwhile, it would also be better to standardise the numerical features (i.e. other than `Political_camp` and `Quarter`) by centering their means at 0 for better model convergence, but the standardiser should only be fitted on the training set after splitting the data into the training and validation sets in order to avoid data leakage (the same is also true for creating the TF-IDF matrix).

The next step is to transform both the titles and main texts of the articles into a TF-IDF term-document matrix. Apart from joining the `Title` and `Text` columns together as the complete `Article`, I will also add additional words into the dictionary and remove stop words as well as punctuation for better tokenisation so that the NMF model can better discover the latent topics.

To avoid data leakage as mentioned before, I will only fit the `TfidfVectorizer` and the NMF models on the train set (i.e. `X_train`) and then use the fitted instances to transform both the train and test sets. I set the number of latent topics (`n_components`) as 10 for the NMF model, and this is decided based on figure 4.5 which plots the reconstruction error measuring the difference of the values between the original TF-IDF matrix and the reconstructed version after NMF. Although there are certainly other valid choices of the number of latent topics to be discovered by NMF, 10

4. Results

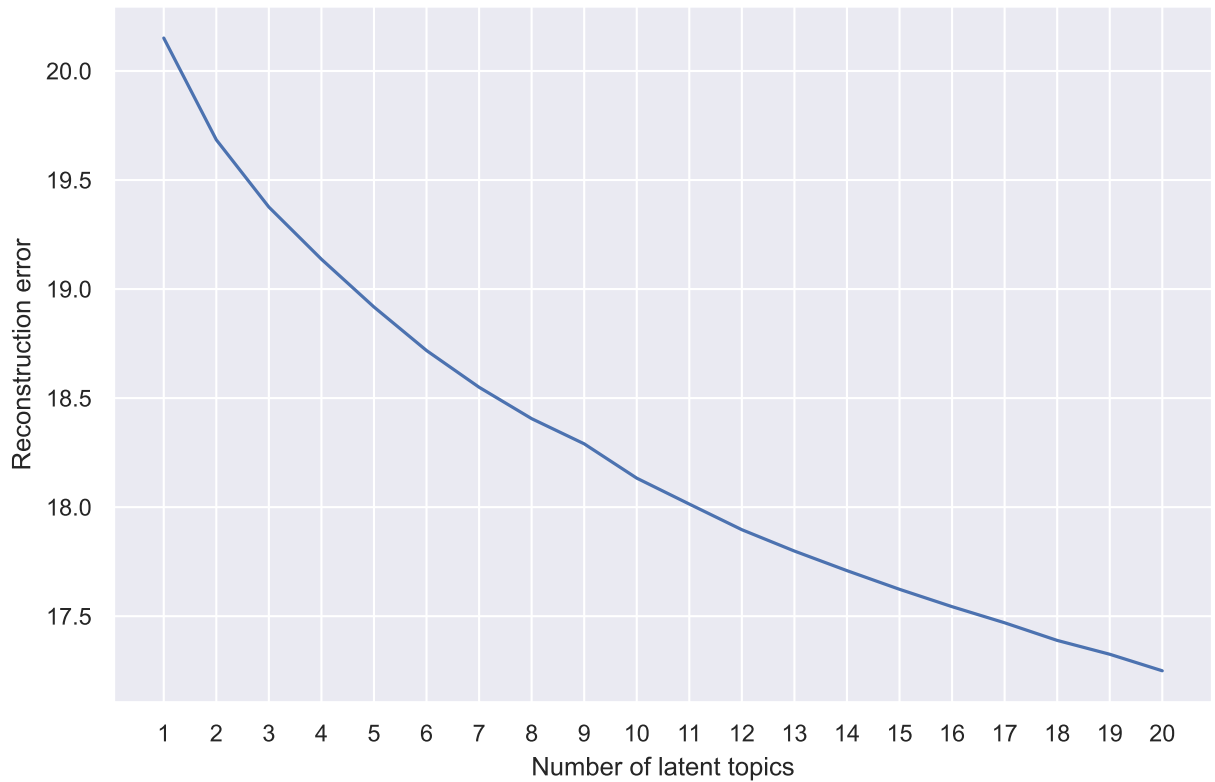


Figure 4.5: Elbow plot of the reconstruction error of NMF as a function of the number of pre-specified latent topics

appears to be a reasonable choice as a compromise between finding out a wide variety of topics in the corpus and not fitting too much into the noise of the data.

In order to make the latent topics generated by the NMF model be named more intuitive, I will inspect the 30 most prominent words of each latent topic and then summarise each topic¹. Overall, the ten topics generated by NMF are more or less semantically coherent and can be summed up concisely. Finally, I will transform the validation set's articles with the fitted instance of the TF-IDF and NMF models on the training set. The ten latent topics that were discovered by the NMF model are: *crimes, non-refoulement policy, illegal labours, illegal gambling, drugs, illegal immigration, murder, robbery, South Asian settlements* and *public security*.

¹A list of the 30 most prominent words in each of the topic and the code to generate it are available in the appendix.

4. Results

Table 4.3: Log loss on 5-fold cv and test set for the 4 baseline models

	5-fold cv log loss	Test set log loss
Logistic regression	0.4648889	0.3495389
SVM	0.4416793	0.3371468
Random forest	0.3805292	0.3466180
XGBoost classifier	0.3863108	0.2859553

4.2.2 Training the model

After the above preprocessing steps, it is time to train a model that adequately predicts the relations between the features and the sentiment of the articles before finding out the importance of the political camp as the main independent variable. To facilitate the decision of which model to use and model tuning, I will first run some baseline models with the default hyper-parameters, except that I have adjusted the weights of each class in the dependent variable due to class imbalance and also tweaked the `n_estimators` of the XGBoost model to 15 instead of the default value of 200 to prevent the model from being too big². Moreover, tree-based models (i.e. random forest and xgboost) do not necessarily need to have the numerical features standardised, and thus only the categorical columns need to be one-hot encoded. The baseline models will be compared based on their performance on log loss in 5-fold cross validation and on the testing set.

Table 4.3 contains the performance of the log loss on both the 5-fold cross validation and test set log loss scores for the four baseline models. Judging by the performance of log loss on the test set, it seems that the XGBoost model performs the best out of all candidates. I will then perform hyper-parameter tuning of the XGBoost model to see if there could be any improvements of its performance³. Surprisingly, the tuned XGBoost model has a higher log loss on the test data than the pre-tuned one. If we look at the per-class f1 score in tables 4.4 and 4.5, however, we can see that the tuned XGBoost model performs considerably better in the f1 score for predicting whether a news article’s polarity is positive or not (i.e., class 2). Therefore, I will use the tuned model as the basis of interpreting the impact of the features on predicting the sentiments of the news articles about asylum seekers in Hong Kong in 2019.

²For the complete documentation of the default parameters of the models used in this thesis, refer to the websites of scikit-learn and XGBoost Documentation — xgboost 1.5.1 documentation.

³The code for tuning the model can be found in the appendix.

Table 4.4: Classification report of the baseline XGBoost model

	precision	recall	f1-score	support
0	0.8953	0.9506	0.9222	81.000
1	0.7826	0.6923	0.7347	26.000
2	1.0000	0.6000	0.7500	5.000
accuracy	0.8750	0.8750	0.8750	0.875
macro avg	0.8927	0.7476	0.8023	112.000
weighted avg	0.8738	0.8750	0.8710	112.000

Table 4.5: Classification report of the tuned XGBoost model

	precision	recall	f1-score	support
0	0.9259	0.9259	0.9259	81.0000
1	0.7407	0.7692	0.7547	26.0000
2	1.0000	0.8000	0.8889	5.0000
accuracy	0.8839	0.8839	0.8839	0.8839
macro avg	0.8889	0.8317	0.8565	112.0000
weighted avg	0.8862	0.8839	0.8845	112.0000

4.2.3 Is the pro-Beijing camp more likely to portray asylum seekers in 2019 more negatively than other outlets?

With the trained model at hand, we can now answer whether H_1 is supported by the model's results using SHAP values. According to figure 4.6, we can see that on the level of the whole XGBoost model, whether a media belongs to the pro-Beijing camp or not (`Political_camp_Pro-Beijing`) is the fourth most important features in predicting the sentiment of a news article, only being lower than the magnitudes of three themes of the news articles about asylum seekers (i.e., public security, murder and crimes). Specifically, pro-Beijing affiliation has the second largest magnitude in affecting the prediction of whether a news article has a negative narrative against non-refoulement claimants, as one can see the pink bar representing the impact of a feature on predicting class 0 in the model is the second longest for the `Political_camp_Pro-Beijing` feature (only behind that of the theme `Crimes` in the articles). Let's zoom into the SHAP values plot for predicting class 0.

The beeswarm plot in figure 4.7 zooms into the importance of each feature in predicting whether a news article on asylum seekers has a negative (coded as 0) polarity, and features appearing at the top of the y-axis are deemed more important than those at the lower end of the y-axis. In a SHAP value beeswarm plot, dots in red mean the value of a feature is high (or present in case of a binary feature, e.g. one-hot-encoded columns), whereas those in blue mean the value of a feature is low (or absent in the case of a binary feature). As one can see, not only is

4. Results

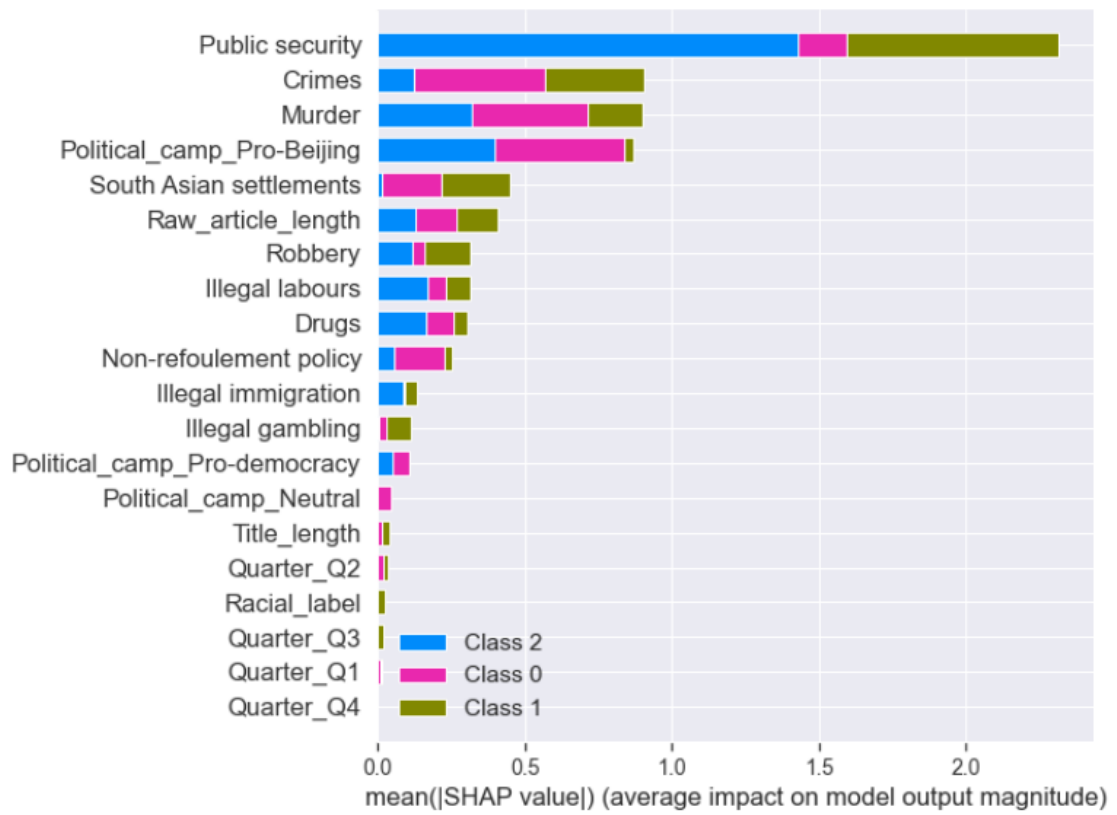


Figure 4.6: Overall SHAP values of the features in the tuned XGBoost model

4. Results

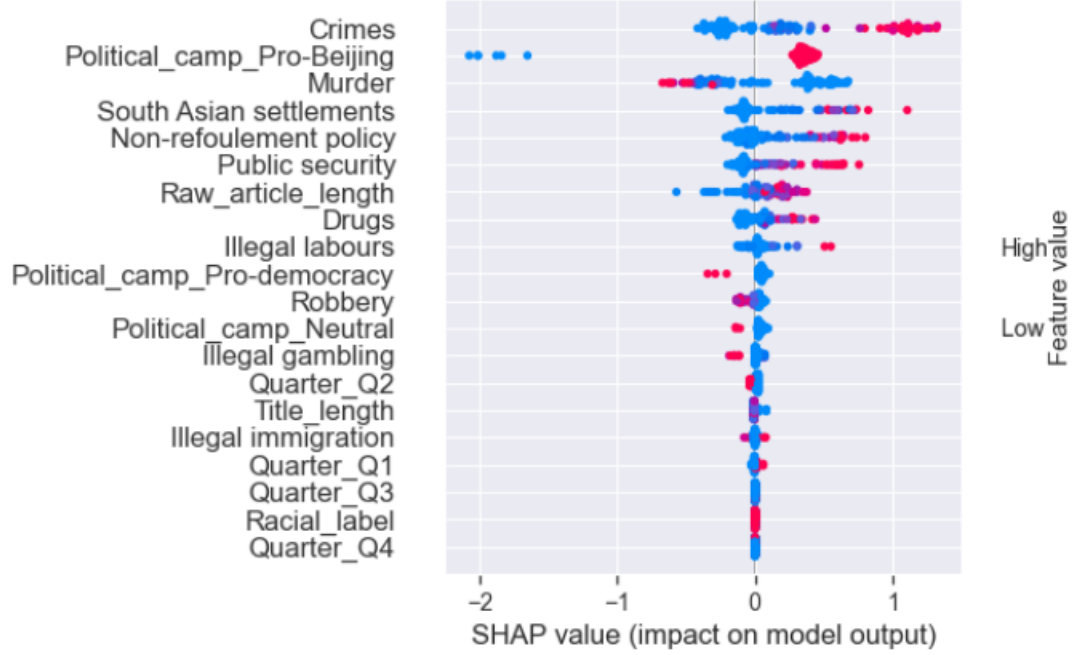


Figure 4.7: The SHAP values of the features in the prediction of whether an article has a negative polarity

the affiliation to pro-Beijing camp the second most determinant feature in predicting whether a news article has negative polarity, but such affiliation will also increase the probability of a news articles portraying asylum seekers in negative lights.

If we inquire further in figures 4.8 and 4.9 about the impact of the features on the predictions of neutral and positive articles respectively, then some interesting observations arise. Firstly, pro-Beijing affiliation is not a very prominent feature in affecting the prediction of whether a news article is simply reporting on news related to asylum seekers objectively without much added sentiment and interpretation by the journalists, as figure 4.8 shows that the `Political_camp_Pro-Beijing` only occupies the middle layer of the y-axis and does not have a significant magnitude in affecting the prediction. By contrast, pro-Beijing affiliation of a media outlet once again is the second most important feature for the prediction of whether an article depicts asylum seekers in Hong Kong favourably. In particular, pro-Beijing media outlets are less likely to have positive reportage on non-refoulement claimants vis-a-vis media outlets from other camps.

Based on the SHAP values of the features as shown in figures 4.6-4.9, we can therefore conclude that H_1 is supported by the evidence produced from the above sentiment analysis on the articles

4. Results

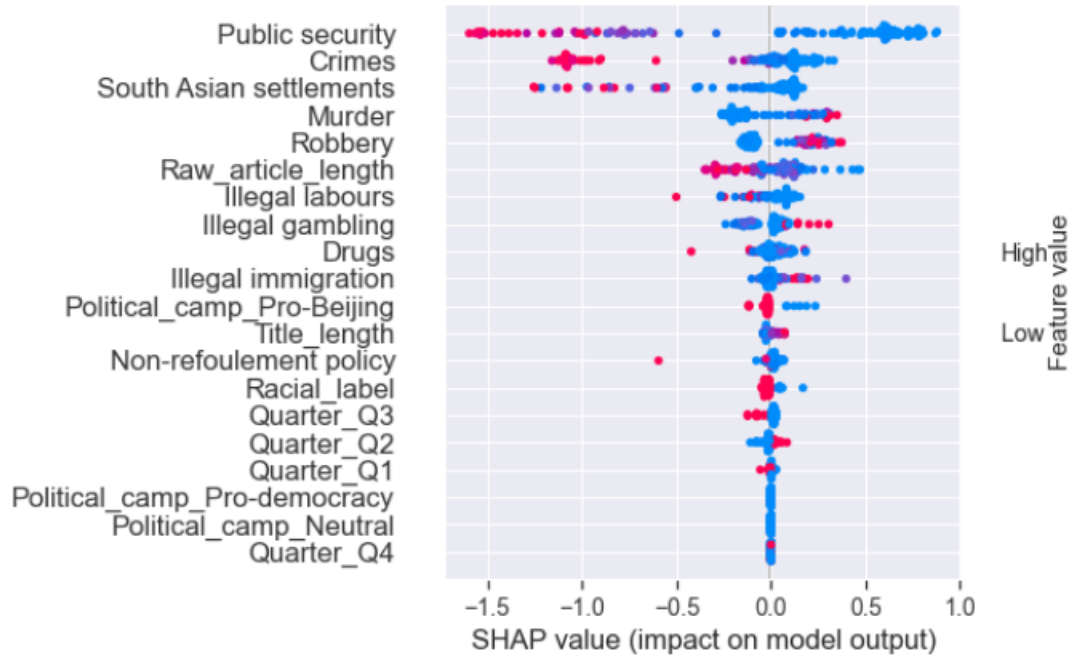


Figure 4.8: The SHAP values of the features in the prediction of whether an article has a neutral polarity

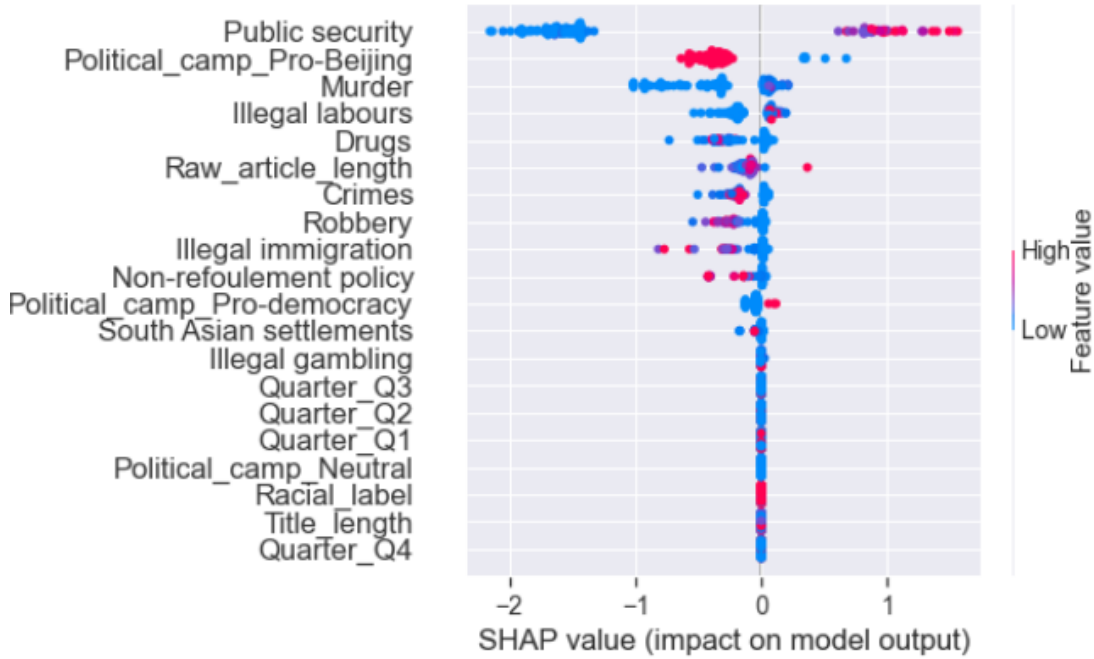


Figure 4.9: The SHAP values of the features in the prediction of whether an article has a positive polarity

4. Results

1042 about asylum seekers in Hong Kong in 2019. Specifically, pro-Beijing media outlets are more
1043 likely to frame asylum seekers negatively by using incorrect terms such as “fake refugees” to refer
1044 to this group of population or criticising them as social ills, and these media outlets are quite
1045 unlikely to have favourable reportage on non-refoulement claimants.

1046 Here are also some notable observations about other features in the dataset. For starters, it
1047 seems that when news articles are related to the topic of **Public security**, then their polarities are
1048 less likely to be neutral, meaning that such reportage is likely to be loaded with interpretations by
1049 the media outlets. Referring to figures 4.7 and 4.9, in general, articles which link non-refoulement
1050 claimants with the frame of public security are less likely to have positive polarity. Relatedly,
1051 articles reporting on crimes with involvement by non-refoulement claimants are likely to have
1052 negative tones.

1053 Surprisingly, racial labels was not an informative feature in predicting the sentiment of the
1054 articles. One potential reason may be that the connotation between non-ethnic Chineseness and
1055 non-refoulement claimants is quite salient that these two categories may to some extent be used
1056 interchangeably, or the media find it difficult to not mention both at the same time no matter
1057 the sentiment or attitudes towards asylum seekers in Hong Kong.

Chapter 5

Conclusion

In conclusion, it is found that at least in 2019, the political camp of media outlets was associated with their attitudes towards asylum seekers in Hong Kong. Specifically,

- I will sum up the findings after editing the main sections

5.1 How might the instigation of the National Security Law affect the public discourse on asylum seekers in Hong Kong?

Just a year after the anti-extradition law protest had started and once again mobilised a huge section of Hong Kong's society against the authority, the HKSAR Government promulgated the National Security Law in July 2020 which aims to tighten the control over the political landscape of Hong Kong. As a result, non-pro-Beijing forces have witnessed their political influence being severely restricted, both in official institutions and the society. Even though the Legislative Council had been one of the few avenues where pro-democracy activists and politicians could run for office (despite lacking substantive authority to participate in policymaking by proposing bills), with the conclusion of the recent 2021 Legislative Council election after an overhaul of the electoral system which essentially permits candidacy only to the "patriots" loyal to the Beijing and HKSAR Governments (Lau and Yam 2021), the door for pro-democracy camp to advocate alternative political discourses has been shut. This means when pro-Beijing lawmakers propose non-binding bills to combat the so-called "fake refugee" problems as Holden Chow did back in 2016 (Cheng 2016), such bills will very unlikely be opposed by others, and the voting results

5. Conclusion

1079 on the however unbinding bills may then be used by the Government to justify more stringent
1080 measures against non-refoulement claimants.

1081 Even the presence of alternative media which strive to provide an alternative narrative of
1082 news contrary to the state-coopted mainstream media as Lee (2018) describes may not be the
1083 case anymore. In June 2021, the flagship pro-democracy newspaper *Apple Daily* was forced to
1084 shut down after the HKSAR Government had frozen the media outlet's asset under the pretext
1085 of violating the National Security Law (Tan 2021). Just days before the year 2021 ended, the
1086 popular pro-democracy online media *Stand News* had to shut down as well after the Hong Kong
1087 police raided its office and arrested senior staff (Ng and Pomfret 2021), followed by the closure of
1088 another online pro-democracy news outlets days after ("Hong Kong" 2022). It is likely that more
1089 alternative media may be pressured to defunct in the future as well due to the threat from the
1090 National Security Law, and this implies that alternative discourses will exist in the media industry
1091 of Hong Kong. Consequently, since the pro-Beijing camp media in general are more hostile to
1092 non-refoulement claimants, it is likely that Hong Kong's media landscape will be dominated by
1093 more negative rhetoric towards asylum seekers in the city given that pro-democracy outlets are
1094 now finding increasing difficulties to operate without running into legal repercussion.

1095 **Appendix: The Echoes of the Code**

1096 The goal of this appendix is to echo the code you used in your thesis for a greater sense of
1097 transparency and replicability of your research. Note that `ref.labels` can be set to any label.
1098 Hence, you can filter the code you want replicated in the appendix by setting labels to the desired
1099 code chunks in the various chapters. See this excellent resource for more information.

1100 This might be particularly useful when you perform model selection to output intermediary
1101 steps here instead of in the code to avoid cluttering your report.

Appendix: The Echoes of the Code redux

Add as many appendices as you like.

```
from hyperopt import fmin, hp, tpe, Trials

xgb_space = {
    "n_estimators": hp.quniform("n_estimators", 10, 50, 5),
    "max_depth": hp.quniform("max_depth", 2, 6, 1),
    "learning_rate": hp.quniform("learning_rate", 0.01, 0.3, 0.01),
    "gamma": hp.quniform("gamma", 0.1, 10, 0.1),
    "min_child_weight": hp.quniform("min_child_weight", 1, 10, 1),
    "subsample": hp.quniform("subsample", 0.5, 0.9, 0.1),
    "colsample_bytree": hp.quniform("colsample_bytree", 0.5, 0.9, 0.1),
    "reg_lambda": hp.quniform("reg_lambda", 1, 100, 1)
}

def xgb_objective(params):
    xgboost_clf = xgb.XGBClassifier(objective="multi:softmax",
                                    eval_metric="mlogloss",
                                    random_state=1,
                                    use_label_encoder=False)

    xgb_params = {
        "n_estimators": int(params["n_estimators"]),
```

References

```
"max_depth": int(params["max_depth"]),
"learning_rate": params["learning_rate"],
"gamma": params["gamma"],
"min_child_weight": int(params["min_child_weight"]),
"subsample": params["subsample"],
"colsample_bytree": params["colsample_bytree"],
"reg_lambda": params["reg_lambda"]
}

loss = np.mean(- cross_val_score(xgboost_clf.set_params(**xgb_params), X_train_final, y_train, cv=fi

return loss

xgb_trials = Trials()

best_xgb_params = fmin(xgb_objective, xgb_space, algo=tpe.suggest, max_evals=500, rstate=np.random.see
```

```
{python inspecting_topic_words}

# Defining a function to extract the most prominent words in each topic

def topic_words(model, vectorizer, top_n_words):

    vocabulary = vectorizer.get_feature_names()

    for idx, topic in enumerate(model.components_):

        print(f"\nTopic {idx}: ")

        print(" ".join([vocabulary[i] for i in topic.argsort()[::-top_n_words - 1: -1]]))

# Extracting the 30 most prominent words in each topic

_ = topic_words(nmf_10, tfidf_vec, 30) # the output is shown in the figure below
```

References

Topic 0:
南亞 事主 漢 現場 兵團 一名 受傷 刑事 報警 調查隊 警方 犯案 調查 附近 案件 凌晨 兩名 男子 事件簿 送院 警區 治理 追緝 逃 初步 接報 進 求助 交由
Topic 1:
聲請 處理 司法覆核 個案 免遣返聲請 提出 上訴 立法會 宗 法律 入境處 政府 審核 程序 假難民 保安局 遣返 修例 修訂 建議 議員 相關 問題 年度 當局 提交 考慮 指 申請 司法機構
Topic 2:
入境處 非法 黑工 勞工 監禁 被判 工作 僱傭 聘用 介乎 被捕 年齡 行動 定罪 最高 僱主 一經 離境 十五 罰款 工行 月 餐廳 擔保書 拘捕 一名 入境者 涉嫌 女子 身份
Topic 3:
遊戲機 賭檔 中心 黑幫 賭博 仔 賭博場 行動 賭客 警方 持牌 非法 包括 魚機 遊戲 經營 釣 機 戲機 積分 被捕 釣魚機 集團 南亞 現金 三合 機及 人士 舖 萬元
Topic 4:
毒品 檢獲 約 販毒 海關 大麻 單位 警方 行動 懷疑 男子 分銷 可卡因 被捕 市值 拘捕 鈔 涉嫌 調查 冰毒 職務隊 搜查 探員 追查 來源 特別 南亞 公斤 製 報
Topic 5:
人蛇 水警 偷渡 蛇頭 偷渡客 來港 南亞 搶灘 非法入境 西貢 截獲 警方 一艘 巴基斯坦 船 集團 屯門 八名 可疑 發現 水域 深圳 安排 拘捕 船上 窟 西分區 昨晨 籍 海里
Topic 6:
死者 房 割 妻子 男子 殺 警方 單位 非洲 行街紙 死亡 同鄉 傷痕 住客 殺案 謀 謀殺 調查 印度 印度籍 發現 爭執 土瓜灣 海防道 九龍城 被捕 昏迷 床上 被告
Topic 7:
事主 背囊 賊 搶 休班 匪徒 黃埔 兩名 換店 紅磡 刀 警長 一名 現金 巨款 找 男子 萬元 匪 制服 牛肉 逃 截劫 疑匪 劫 劫案 南亞 合力 疑犯 德安街
Topic 8:
南亞 村 拆 木板 新村 橋底 屋 滑 露宿 村民 政府 通州街 深水埗 居民 區議員 公園 部門 天橋底 衛生 街坊 搭建 渡船街 雜物 油麻地 治安 更 假難民 垃圾 吸毒
Topic 9:
港府 假難民 香港 政府 說 難民 問題 社會 市民 治安 南亞 做 派 更 政策 港人 沒 政治 港 退出 長 毒癮 禁閉營 移民 加拿大 解決 法治 錢 美國 億元

Figure 5.1: Word list of each of the topics generated by the NMF model

References

- 1105
- 1106 *Basic Law - Basic Law - Chapter I (EN)* (Apr. 2021).
- 1107 <https://www.basiclaw.gov.hk/en/basiclaw/chapter1.html>.
- 1108 BBC (Jan. 2021). *National Security Law: Hong Kong Rounds up 53 pro-Democracy Activists*.
- 1109 <https://www.bbc.com/news/world-asia-china-55555299>.
- 1110 Chen, Tianqi et al. (2015). “Xgboost: Extreme Gradient Boosting”. In: *R package version 0.4-2*
- 1111 1.4, pp. 1–4.
- 1112 Cheng, Kris (Dec. 2016). *Legislature Rejects Lawmaker’s Non-Binding Motion to Combat ‘bogus’*
- 1113 *Refugees*. [https://hongkongfp.com/2016/12/01/legislature-rejects-lawmakers-non-binding-](https://hongkongfp.com/2016/12/01/legislature-rejects-lawmakers-non-binding-motion-combat-bogus-refugees/)
- 1114 [motion-combat-bogus-refugees/](https://hongkongfp.com/2016/12/01/legislature-rejects-lawmakers-non-binding-motion-combat-bogus-refugees/).
- 1115 Cheung, Karen and Tom Grundy (Apr. 2016). *DAB Party Sets up in Tsim Sha Tsui Calling for*
- 1116 *‘fake’ Refugees to Be Put in Camp and Deported*. [https://hongkongfp.com/2016/04/11/dab-](https://hongkongfp.com/2016/04/11/dab-party-sets-up-in-tsim-sha-tsui-calling-for-fake-refugees-to-be-put-in-camp-and-deported/)
- 1117 [party-sets-up-in-tsim-sha-tsui-calling-for-fake-refugees-to-be-put-in-camp-and-deported/](https://hongkongfp.com/2016/04/11/dab-party-sets-up-in-tsim-sha-tsui-calling-for-fake-refugees-to-be-put-in-camp-and-deported/).
- 1118 Chun, Allen (1996). “Fuck Chineseness: On the Ambiguities of Ethnicity as Culture as Identity”.
- 1119 In: *boundary 2* 23.2, p. 111. ISSN: 01903659. DOI: 10.2307/303809.
- 1120 Cooper, Samantha et al. (2017). “Media Coverage of Refugees and Asylum Seekers in Regional
- 1121 Australia: A Critical Discourse Analysis”. In: *Media International Australia* 162.1, pp. 78–89.
- 1122 CUHK (2020). *Public Evaluation on Media Credibility*.
- 1123 <https://ccpos.com.cuhk.edu.hk/public-evaluation-on-media-credibility/>.
- 1124 elgarteo (2021). *HKPolDict: Hong Kong Politics Dictionary for Text Mining*.
- 1125 <https://github.com/elgarteo/HKPolDict>.
- 1126 *Enforcement / Statistics on Non-refoulement Claim* (2021).
- 1127 <https://www.immd.gov.hk/eng/facts/enforcement.html>.
- 1128 Erni, John Nguyet (2012). “Who Needs Strangers? Un-imagining Hong Kong Chineseness”. In:
- 1129 *Chinese Journal of Communication* 5.1, pp. 78–87.

References

- Fong, Brian (2017a). “In-between Liberal Authoritarianism and Electoral Authoritarianism: Hong Kong’s Democratization under Chinese Sovereignty, 1997-2016”. In: *Democratization* 24.4, pp. 724–750. ISSN: 1351-0347, 1743-890X. DOI: 10.1080/13510347.2016.1232249.
- Fong, Brian CH (2017b). “One Country, Two Nationalisms: Center-periphery Relations between Mainland China and Hong Kong, 1997-2016”. In: *Modern China* 43.5, pp. 523–556.
- Full Text of Constitution of Communist Party of China - Xinhua* (Nov. 2017).
http://www.xinhuanet.com/english/special/2017-11/03/c_136725945.htm.
- Greenfeld, Liah and Jonathan Eastwood (2007). “Nationalism and National Identity: Definitional Issues”. In: *The Oxford Handbook of Comparative Politics*. Ed. by Carles Boix and Susan Stokes. 1st. Oxford University Press, pp. 256–273.
- Heath, Anthony F. and James R. Tilley (2005). “British National Identity and Attitudes towards Immigration”. In: *International Journal on Multicultural Societies* 7.2, pp. 119–132.
- Hjerm, Mikael (1998). “National Identity: A Comparison of Sweden, Germany and Australia”. In: *Journal of Ethnic and Migration Studies* 24.3, pp. 451–469.
- Home_Affairs_Department (Jan. 2018). *The Demographics: Ethnic Groups*.
https://www.had.gov.hk/rru/english/info/info_dem.html.
- “Hong Kong” (Jan. 2022). “Hong Kong: Pro-democracy Website Citizen News Is Latest to Close”. In: *BBC News*.
- Hui4yi4 Guo4cheng2 Zheng4shi4 Ji4lu4* (Dec. 2016).
<https://webcache.googleusercontent.com/search?q=cache:zyvAUu3wlSoJ:https://www.legco.gov.hk/yr16-17/chinese/counmtg/hansard/cm20161201-translate-c.pdf+&cd=8&hl=zh-TW&ct=clnk&gl=ch>.
- Junyi, Sun (Dec. 2021). *Jieba*.
- Kaeding, Malte Philipp (2017). “The Rise of "Localism" in Hong Kong”. In: *Journal of Democracy* 28.1, pp. 157–171.
- Kwan, Justin P. (2016). “The Rise of Civic Nationalism: Shifting Identities in Hong Kong and Taiwan”. In: *Contemporary Chinese Political Economy and Strategic Relations: An International Journal* 2.2, pp. 941–973.
- Kwong, Ying-ho (2016). “The Growth of “Localism” in Hong Kong. A New Path for the Democracy Movement?” In: *China Perspectives* 2016.2016/3, pp. 63–68.

References

- 1160 Lau, Jessie and Shui-yin Sharon Yam (Dec. 2021). *‘Patriots Only’: Hong Kong’s New Election*
1161 *System in Action*.
1162 <https://thedi diplomat.com/2021/11/patriots-only-hong-kongs-new-election-system-in-action/>.
- 1163 Lau, Tracy (2013). “State Formation and Education in Hong Kong: Pro-Beijing Schools and
1164 National Education”. In: *Asian Survey* 53.4, pp. 728–753.
- 1165 Laver, Michael, Kenneth Benoit, and John Garry (2003). “Extracting Policy Positions from
1166 Political Texts Using Words as Data”. In: *American political science review* 97.2,
1167 pp. 311–331.
- 1168 Law, Kam-Yee and Kim-Ming Lee (Mar. 2012). “The Myth of Multiculturalism in ‘Asia’s World
1169 City’: Incomprehensive Policies for Ethnic Minorities in Hong Kong”. In: *Journal of Asian*
1170 *Public Policy* 5.1, pp. 117–134. ISSN: 1751-6234. DOI: 10.1080/17516234.2012.662353.
- 1171 Lecours, André (2000). “Ethnic and Civic Nationalism: Towards a New Dimension”. In: *Space*
1172 *and Polity* 4.2, pp. 153–166.
- 1173 Lee, Francis LF (2018). “Changing Political Economy of the Hong Kong Media”. In: *China*
1174 *Perspectives* 2018.2018/3, pp. 9–18.
- 1175 Lewin-Epstein, Noah and Asaf Levanon (2005). “National Identity and Xenophobia in an
1176 Ethnically Divided Society”. In: *International Journal on Multicultural Societies* 7.2,
1177 pp. 90–118.
- 1178 Lin, Cong and Liz Jackson (2021). “Make China Great Again: The Blood-Based View of
1179 Chineseness in Hong Kong”. In: *Educational Philosophy and Theory* 53.9, pp. 907–919.
- 1180 Loh, Christine (2010). *Underground Front: The Chinese Communist Party in Hong Kong*. Hong
1181 Kong University Press.
- 1182 Lundberg, Scott (Jan. 2022). *Slundberg/Shap*.
- 1183 Ma, Ngok (2012). “Political Parties and Elections”. In: *Contemporary Hong Kong Government*
1184 *and Politics*. Ed. by Wai-man Lam, Percy Luen-tim Lui, and Wilson Wong. Second. Hong
1185 Kong University Press, pp. 159–176.
- 1186 *Making a Claim for Non-refoulement Protection in Hong Kong* (2021).
1187 https://www.immd.gov.hk/eng/useful_information/non-refoulement-making-claim.html.
- 1188 *Manifesto* (2021). https://www.dphk.org/index.php?route=eng/about_us/detail&id=1.

References

- Mathews, Gordon (2018). “Asylum Seekers as Symbols of Hong Kong’s Non-Chineseness”. In:
China Perspectives 2018.2018/3, pp. 51–58.
- McKay, Fiona H., Samantha L. Thomas, and R. Warwick Blood (2011). “‘Any One of These
 Boat People Could Be a Terrorist for All We Know!’Media Representations and Public
 Perceptions of ‘Boat People’Arrivals in Australia”. In: *Journalism* 12.5, pp. 607–626.
- Medrano, Juan Diez (2005). “Nation, Citizenship and Immigration in Contemporary Spain”. In:
International Journal on Multicultural Societies 7.2, pp. 133–156.
- Medrano, Juan Diez and Matthias Koenig (2005). “Nationalism, Citizenship and Immigration in
 Social Science Research”. In: *International Journal on Multicultural Societies* 7.2, pp. 82–89.
- Ng, Edmond and James Pomfret (Dec. 2021). “Hong Kong Pro-Democracy Stand News Closes
 after Police Raids Condemned by U.N., Germany”. In: *Reuters*.
- Ng, Isabella (2020). “Criminalizing the Innocents: Social Exclusion of Asylum-Seekers and
 Refugees in Hong Kong”. In: *Journal of Asian Public Policy* 13.3, pp. 319–332.
- Ng, Isabella, Sharice Fungyee Choi, and Alex Lihshing Chan (May 2019). “Framing the Issue of
 Asylum Seekers and Refugees for Tougher Refugee Policya Study of the Media’s Portrayal in
 Post-colonial Hong Kong”. In: *Journal of International Migration and Integration* 20.2,
 pp. 593–617. ISSN: 1874-6365. DOI: 10.1007/s12134-018-0624-7.
- Ngo, Jennifer (Apr. 2016). *Political Dark Arts Live on in Hong Kong with Bogus Banners
 Designed to Shame Lawmakers*. [https://www.scmp.com/news/hong-](https://www.scmp.com/news/hong-kong/politics/article/1933095/political-dark-arts-live-hong-kong-bogus-banners-designed)
[kong/politics/article/1933095/political-dark-arts-live-hong-kong-bogus-banners-designed](https://www.scmp.com/news/hong-kong/politics/article/1933095/political-dark-arts-live-hong-kong-bogus-banners-designed).
- O’Doherty, Kieran and Amanda Lecouteur (2007). “‘Asylum Seekers’, ‘Boat People’ and
 ‘Illegal Immigrants’: Social Categorisation in the Media”. In: *Australian journal of
 psychology* 59.1, pp. 1–12.
- OCTOBER ROUNDUP AT THE REFUGEE UNION (Nov. 2018).
<http://refugeeunion.org/9739/october-roundup-refugee-union-office/>.
- Pedregosa, Fabian et al. (2011). “Scikit-Learn: Machine Learning in Python”. In: *the Journal of
 machine Learning research* 12, pp. 2825–2830.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Manual. R
 Foundation for Statistical Computing. Vienna, Austria.

References

- 1218 *Refugee Community Supports Lawmaker Fernando Cheung Chiu-hung* (Mar. 2015).
1219 <http://www.vfnw.org/7401/vision-first-supports-lawmaker-fernando-cheung-chiu-hung/>.
- 1220 *Refugee Population by Country or Territory of Asylum - Hong Kong SAR, China / Data* (2021).
1221 <https://data.worldbank.org/indicator/SM.POP.REFG?locations=HK>.
- 1222 So, Alvin Y. (2015). “The Making of Hong Kong Nationalism”. In: *Asian Nationalisms*
1223 *Reconsidered*. Ed. by Jeff Kingston. 1st. Routledge, pp. 151–162.
- 1224 Soltoff, Benjamin (Sept. 2021). *Topic Modeling*.
1225 <https://cfss.uchicago.edu/notes/topic-modeling/>.
- 1226 Stevens, Keith et al. (2012). “Exploring Topic Coherence over Many Models and Many Topics”.
1227 In: *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language*
1228 *Processing and Computational Natural Language Learning*, pp. 952–961.
- 1229 Tan, Yvette (June 2021). “Apple Daily: The Hong Kong Newspaper That Pushed the
1230 Boundary”. In: *BBC News*.
- 1231 the Sun (Mar. 2016). *jiu4shi4lun4shi4: ben3tu3pai4 shi1sheng1*.
1232 http://the-sun.on.cc/cnt/news/20160312/00418_009.html.
- 1233 *Theme of the 32nd Anni. of June 4 “For Freedom, a Common Fate, a Shared Struggle”* (2021).
1234 <https://web.archive.org/web/20210501051630/https://hka8964.wordpress.com/6432-2/>.
- 1235 *Unified Screening Mechanism Procedures* (2021).
1236 <https://www.unhcr.org/hk/en/2675-usm-procedures-2.html>.
- 1237 van Atteveldt, Wouter, Mariken A. C. G. van der Velden, and Mark Boukes (Apr. 2021). “The
1238 Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary
1239 Approaches, and Machine Learning Algorithms”. In: *Communication Methods and Measures*
1240 15.2, pp. 121–140. ISSN: 1931-2458. DOI: 10.1080/19312458.2020.1869198.
- 1241 Varshney, Ashutosh (2007). “Ethnicity and Ethnic Conflict”. In: *The Oxford Handbook of*
1242 *Comparative Politics*. Ed. by Carles Boix and Susan Stokes. 1st. Oxford University Press,
1243 pp. 274–294.
- 1244 Vecchio, Francesco (2014). *Asylum Seeking and the Global City*. Routledge.
- 1245 *Voices for Protection Go to LegCo* (July 2014).
- 1246 Xie, Yihui (2016). *Bookdown: Authoring Books and Technical Documents with R Markdown*.
1247 Boca Raton, Florida: Chapman and Hall/CRC.

References

- 1248 Xie, Yihui (2021). *Bookdown: Authoring Books and Technical Documents with r Markdown*.
1249 Manual.
- 1250 Yuen, Chantal (Nov. 2016). *Pro-Beijing Lawmaker Submits Motion to Stop 'Bogus Refugees'*.
1251 [https://hongkongfp.com/2016/11/18/pro-establishment-lawmaker-submits-motion-to-stop-](https://hongkongfp.com/2016/11/18/pro-establishment-lawmaker-submits-motion-to-stop-bogus-refugees/)
1252 [bogus-refugees/](https://hongkongfp.com/2016/11/18/pro-establishment-lawmaker-submits-motion-to-stop-bogus-refugees/).