Trends in complaints to the Fiji Independent Commission Against Corruption, 2007 – 2014

January 2018

Dr. Constantine Boussalis
Dr. Travis Coan
Dr. Caryn Peiffer
Dr. Grant Walton

Table of Contents

1. Introduction

This report statistically analyses complaints provided to Fiji's Independent Commission Against Corruption (FICAC) between 2007 (when the FICAC was established) and 2014. The report sheds light on the types of people who report, what they report about, how complaints relate to the FICAC's jurisdiction, and how those who reported defined corruption.

This research was conducted as a joint consultancy by researchers from the University of Birmingham, Australia National University, University of Exeter and Trinity College Dublin.

Each year FICAC now receives over 10,000 complaints about a range of issues from people across the country concerned about corruption, mismanagement, and abuse of power. FICAC summarises these complaints in their annual report, which provides basic, yet important, information about the number and type of complaints received. This report provides a more detailed and long term view of the type of complaints received by FICAC. It does this in two key ways. First it provides an analysis of trends over time and between provinces. This shows how the types of people who complain, and the issues they are concerned about, have changed over time and differ across space. Second, the report draws upon narratives of complaints, which provide a rich source of data on the variety of issues that those reporting corruption are concerned about. This analysis allows for the identification of key concerns, and allows us to better understand how respondents defined corruption. In turn, it is hoped that this analysis will help FICAC and other organisations better serve those who are concerned about corruption.

Key findings

To summarise, there are four key findings from this analysis. First, the volume and types of people making a complaint to FICAC has changed significantly over time, in particular:

- Complaints per month were relatively low across 2007 and 2008 (under 500 per month). In July 2009, FICAC, for the first time, received over 500 complaints for the month. Following this, the number of complaints per month slowly but steadily rose to May 2010 when it reached about 1,000 complaints received per month. A steady decline is then observed from May to March 2011, when the number of complaints received per month was lower than 500.
- The percentage of complaints initiated by females rose steadily from only 10% in 2007 to around 20% in 2012, and has after levelled off at around 25% of all complaints.
- Over the entire time period scrutinised, complainants from Ba province filed the largest number of complaints (15,664). However, very few complaints were received from Ba until 2010, when over 500 per month were received, and in April of 2011 there was a sharp increase in the complaints from Ba. At its peak, the number of complaints from Ba, per month, was close to 1,500.
- For a majority of complaints, the age of the complainant is not recorded, and as such, it is impossible to decipher trends with respect to the age group of complainants.

- A slight majority of complaints are received from urbanites, while almost 40% of all complaints are received from rural dwellers.

Second, the nature of complaints has also been shifting over time. A detailed look at the mode of reporting, dominant topic of complaint, and whether a complaint was within FICAC jurisdiction, reveals the following key trends:

- About half of all complaints received were reported in person, while just over a quarter of all complaints received were reported over the phone, and seven percent were received via the mail.
- Complaints involving the state started at close to 60%of all complaints received in 2007, but decreased steadily to be around only a third of all complaints received by 2014. Fairly steadily, across the time period scrutinised, private sector issues made up about 10% of all complaints received.
- The percentage of complaints falling under FICAC's mandate (which focuses on enforcing the penal code and investigating corruption in the public service) was relatively high (25% to 30%) to begin with (in 2007 and 2008 in particular). But the proportion of complaints falling under the mandate has decreased since this time. This indicates that, while there were fewer complaints in FICAC's early history, those who did complain early-on seemed to better understand its mandate.
- The vast majority of complaints received were not within FICAC jurisdiction and just over 15% of complaints received were coded as 'seeking assistance'. This means that between 2007 and 2014, less than 5% of the complaints received fell within FICAC jurisdiction.
- Overwhelmingly, the state is the primary subject of most of the complaints that FICAC receives. The next most popular subject is land.
- As employment issues usually do not involve the state, they make up the third most popular topic for those cases that do not fall under FICAC jurisdiction and only 3% of those cases that do fall under FICAC jurisdiction.

Third, regression analysis reveals that:

- Those who are making complaints that fall under FICAC's jurisdiction are mostly concerned about the public sector and transport; complaints are international in nature; respondents are most likely male; and these complaints tend to come from residents of Cakaudrove, Macuata, Rewa, or Rotuma.
- Females are significantly more likely to lodge a complaint about the family, the state, and the private sector, while men are more likely to lodge complaints about land, contracting, and transport-related issues.
- Rural dwellers are more likely to lodge a complaint related to land issues and the state.

Fourth, text analysis highlights how respondents defined corruption, and how these definitions have changed over time, and regression analysis using these definitions points to the factors that may contribute to different understandings of corruption. These analyses find that:

- When FICAC began, most complainants defined corruption as the 'abuse of public office for private gain' (this definition means that the state is involved in corruption).

- Most people making a complaint now are likely to define corruption as 'the abuse of power for private gain' (which emphasises corruption that involves people in positions of power but are not public officials).
- Females and rural dwellers are significantly more likely than males and urbanites, to lodge a complaint that defines corruption as 'an abuse of public office' or as 'decay'. Males and urbanites are more likely to lodge a complaint that defines corruption as 'an abuse of power'.

Map of report

This report first outlines the methodologies used for the analysis. It then, in section two, presents a profile of reporting over time and across the country by highlighting who complained, where complaints came from and how these factors have changed over time. Section three examines the nature of complaints over time, who was most likely to make a report that fit in with FICAC's jurisdiction, and what happened when a complaint was received. Section four examines the factors which influence whether a complaint will fall under FICAC's jurisdiction, and what factors influence the topic of a complaint. The penultimate section looks at how respondents defined corruption, and the report ends with a discussion about what these findings mean for improving FICAC operations.

2. Methodology

Four methodologies were used to produce this report.

Descriptive Frequency Distributions

First, descriptive frequency distributions were produced from the FICAC complaints data. These analyses allowed the research team to document the number of complaints in specific categories or the percentage of complaints that fell into a certain categories, over the entire period of scrutiny or over time.

Topic Modelling

Second, a topic model approach was employed to examine the text of the complaints. Topic modelling is a type of statistical model used for discovering the abstract "topics" that occur in a collection of documents. In this case, it was used to identify prominent or “key” topics in each of the complaints. As was the case with FICAC data, the methodology is able to identify and discern trends in topics raised. The topic modelling used in this report allowed the research team to 1) identify key topics that have been raised in FICAC complaints and categorise complaints by those key topics, 2) document and compare the number of complaints by key topic category over the entire period of scrutiny, and 3) track, over time, how FICAC complaints have varied with respect to the dominant topics they feature. A more detailed and technical description of this methodology can be found in Appendix A.

Regression Analyses

Third, regression analyses were also executed. Regression analysis was used to assess to what extent variations in a dependent variable are associated with variations in independent variables. For the purposes of this report, three regression analyses were run.

- The first provides a profile of factors that increase the likelihood that complaints filed fall under FICAC’s jurisdiction. Specifically, the regression analysis estimates how the gender of a complainant, location of complainant, and topic of the complaint is associated with whether a complaint lodged is within FICAC jurisdiction.
- The second regression examines how different factors influence whether a complaint is filed at all, by the main topic of complaint.
- The results of the final regression point to what factors contribute to how a complainant comes to define corruption. Specifically, the final regression examines how the gender of the complainant and their location is associated with whether the complaint defined corruption as being either ‘an abuse of public office,’ ‘an abuse of power’, or ‘decay.’

Supervised Classification Model

Finally, a supervised classification model was used. Supervised classification models are able to categorise a large dataset, on the basis of how a smaller, sub-sample of the data was

previously categorised. For this report, the supervised classification model was used to categorise the complaints into three categories, based on how the complainant defined corruption. Those categories are: complaints that involve the 1) abuse of public office, 2) abuse of other form of power (such as in the private sector), and 3) other types of decay (an ethically wrong act that does not involve any clear relationship of power). Categorising corruption in this way allows us to understand the assumptions that those who report to FICAC make about the nature of corruption, and the analysis allows us to see how definitions have changed over time. This type of information can be used by FICAC to understand how its mandate reflects (or does not) the definition of those making complaints. A more detailed and technical description of this methodology can be found in Appendix A.

3. Who, where and when:
Key demographic changes to reporting over time

This section examines the key demographics of those who complained to the FICAC across time, using descriptive frequency distributions. It starts with an analysis of changes to the gender of those making a complaint, before examining how reporting was different across different provinces and between urban and rural locations.[1]

Figure 3.1 shows the number of complaints over time on a quarterly basis. Overall, the number of complaints has risen significantly over this period, although there have been dips and surges along the way, and there was a lot of volatility in 2011 in particular.

Figure 3.1: Total number of complaints over time (2007-2014)

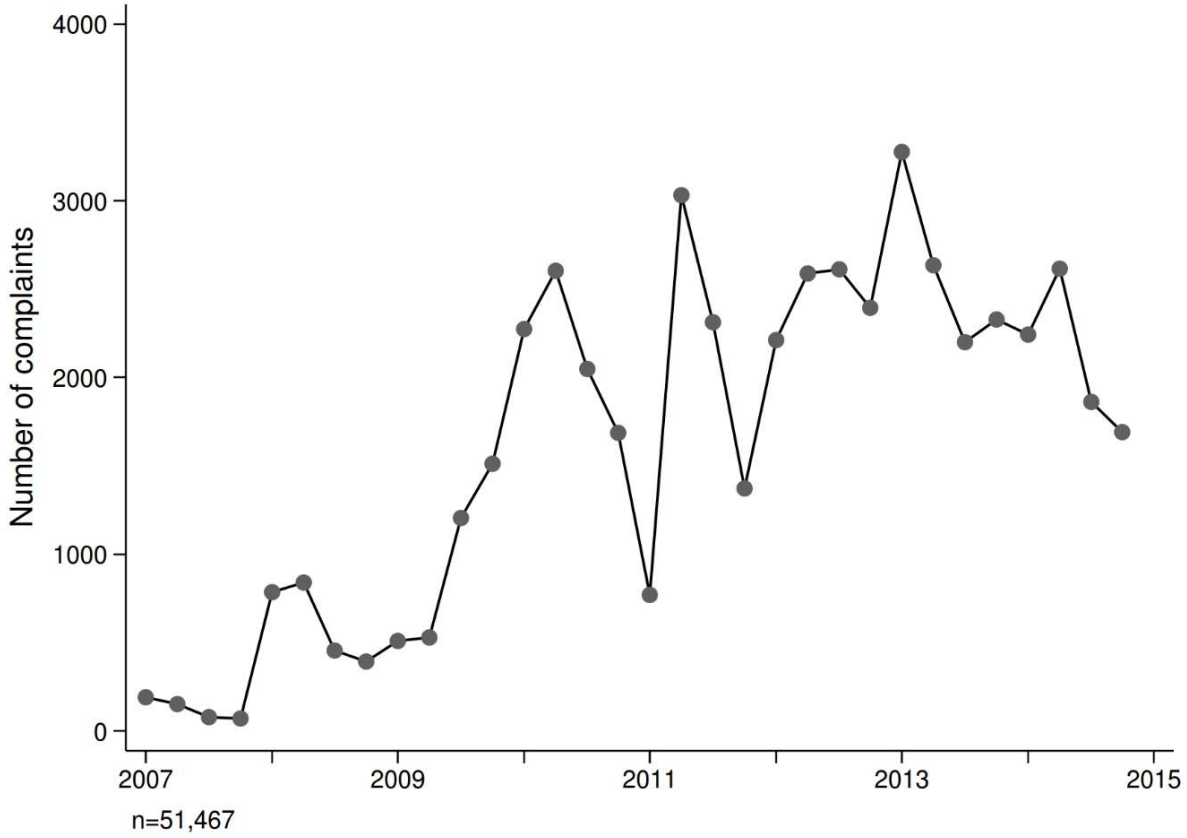
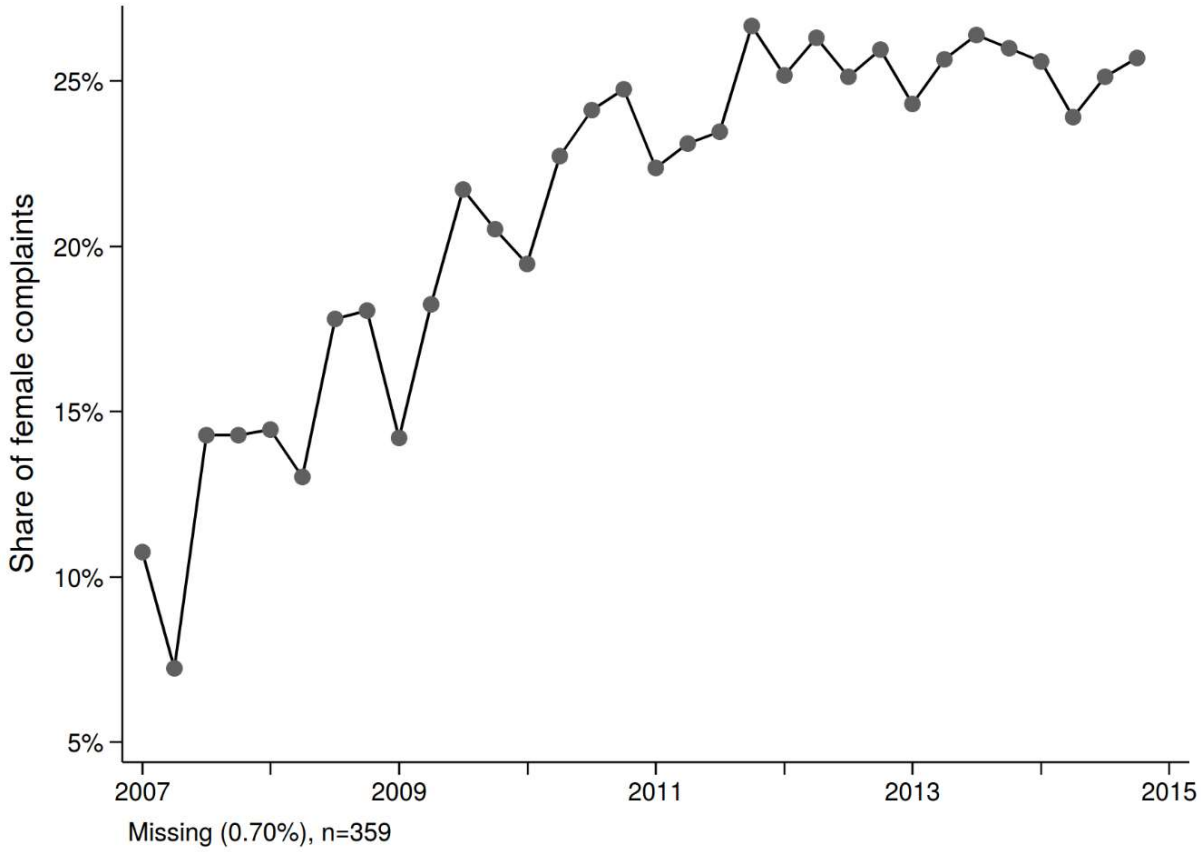


Figure 3.2 shows the share of complaints lodged by a female over time. Consistently, a majority of complaints are lodged by male complainants. At the start (2007), a little over 10% of all

[1] The demographic information provided in this section was limited by the information that was collected by FICAC over the period. While basic information was consistently collected by the agency between 2007 and 2014, some variables were not. This was particularly the case for the age category. Between 2007 and 2014, 76.9% of the complainants did not have their age recorded. Given this limitation, it remains unclear to what extent different age groups are more likely to register complaints with FICAC.

complaints were lodged by females. The percentage of female complainants rose steadily, however, reaching about 25% in 2012. From 2012 to 2015, with little fluctuation, about 25% of all complaints were lodged by female complainants. Although the increase in the rate of female complainants is significant, that females still make up only about one quarter of complainants may reflect gender norms that encourage men to engage with the state and on official matters on behalf of the household and/or the fact that men are more likely to be formally employed.

Figure 3.2: Gender of person making the complaint (2007 – 2014)



Examining the location of complaints, Figure 3.3 shows that from 2007 to 2014, people from Ba filed the largest number of complaints (15,664). In second and third place, respectively, complaints from Rewa were about two-thirds that of those filed from Ba (10,311), and complaints from Macuata were about half that of those from Ba (7,800). Together, complaints from these three locations make up about 70% of all complaints received.

Figure 3.3: Location of complainant (2007 – 2014)

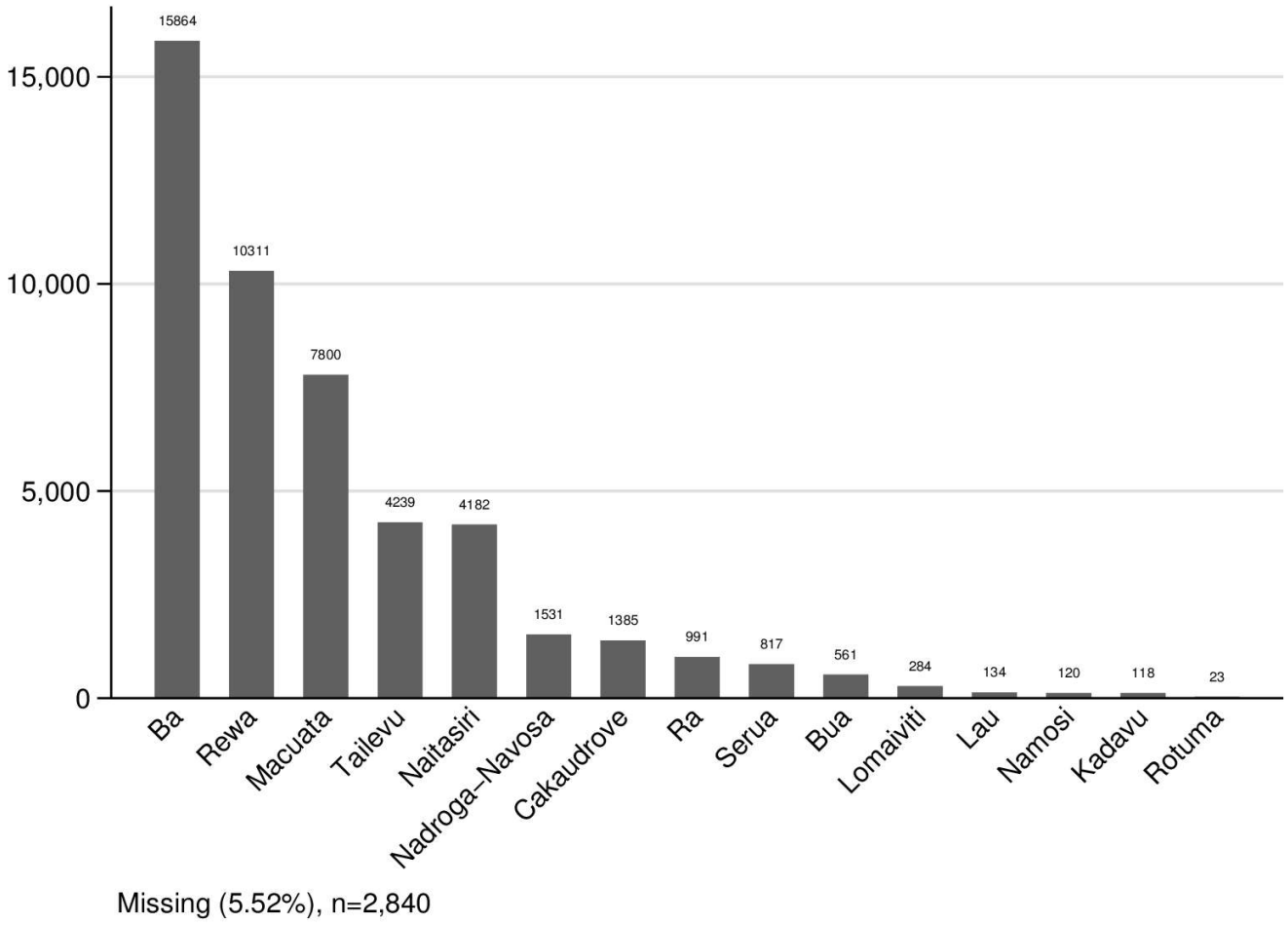


Figure 3.4 tracks the number of quarterly complaints from all of the provinces. From 2007 to 2014, though there have been fluctuations, the number of quarterly complaints from Rewa and Macuata seem to have increased fairly steadily over the time period scrutinised. From Rewa, at the start of 2007, the number of quarterly complaints was 45, and at the end of 2014 it was 294. From Macuata, at the start of 2007, the number of quarterly complaints was only 17, and at the final quarter in 2014 the number of quarter complaints was 204. With similar fluctuations, complaints from all other locations tended to steadily increase across the time period examined, with more than 600 complaints received per quarter by the end of 2014.

For complaints from Ba, the story is different. Less than 150 quarterly complaints were consistently received from Ba until late 2009. 652 quarterly complaints were received from Ba in early 2010. A sharp decline in complaints from Ba is noted in 2011 (only 76) and a more impressive sharp incline in complaints from Ba is registered for the second quarter of 2011: at its peak 1,381. Since 2012, with some fluctuations, the number of complaints from Ba seems to average at just below 800 per quarter.

Figure 3.4: Complaints from Ba, Macuata, Rewa, and other provinces over time (2007-2014)

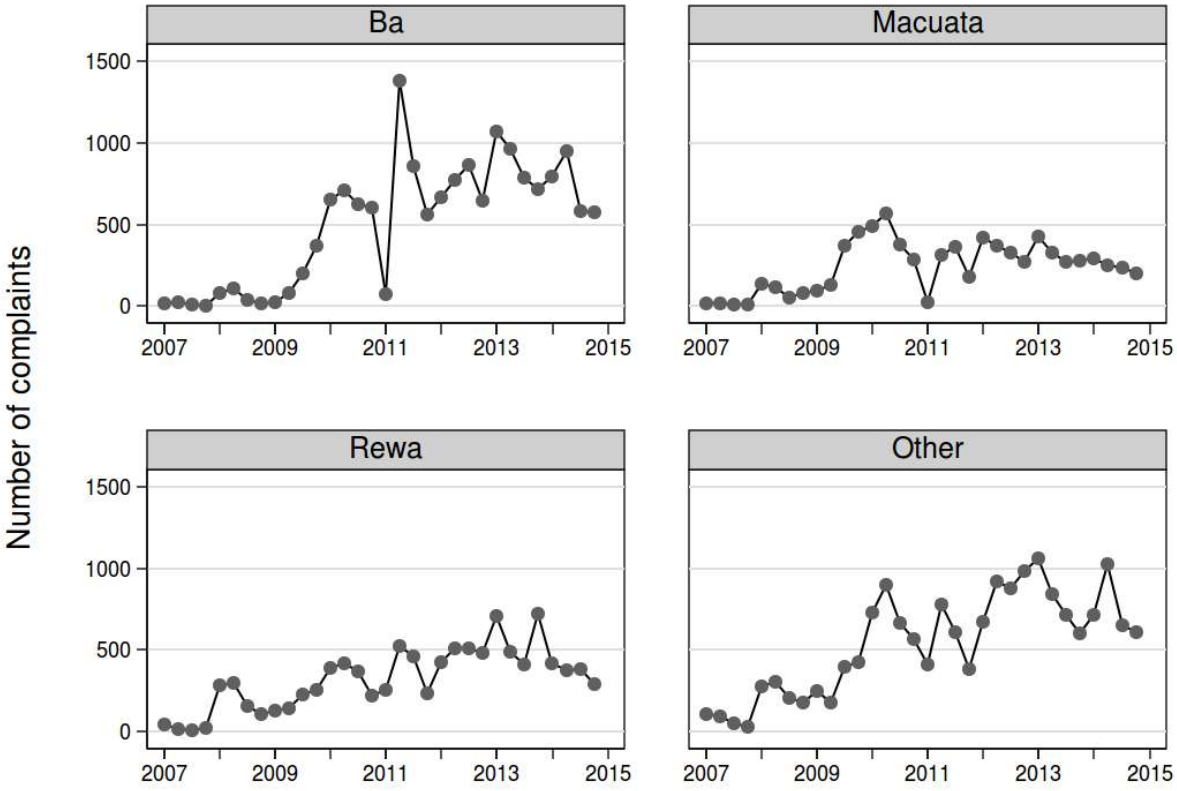


Figure 3.5 plots the number of complaints received per month over the entire time period. In total, between 2007 and 2014, the fewest number of complaints were made in January (3,640). April was the most popular month to make a complaint, at 5,460. In general, though, the number of complaints made by month is fairly steady, averaging at just over 4,200 complaints per month. We can also see from the graph that April, May, June and July are key months in which complaints to FICAC are made.

Figure 3.5: Seasonal analysis: month complaint made (2007-2014)

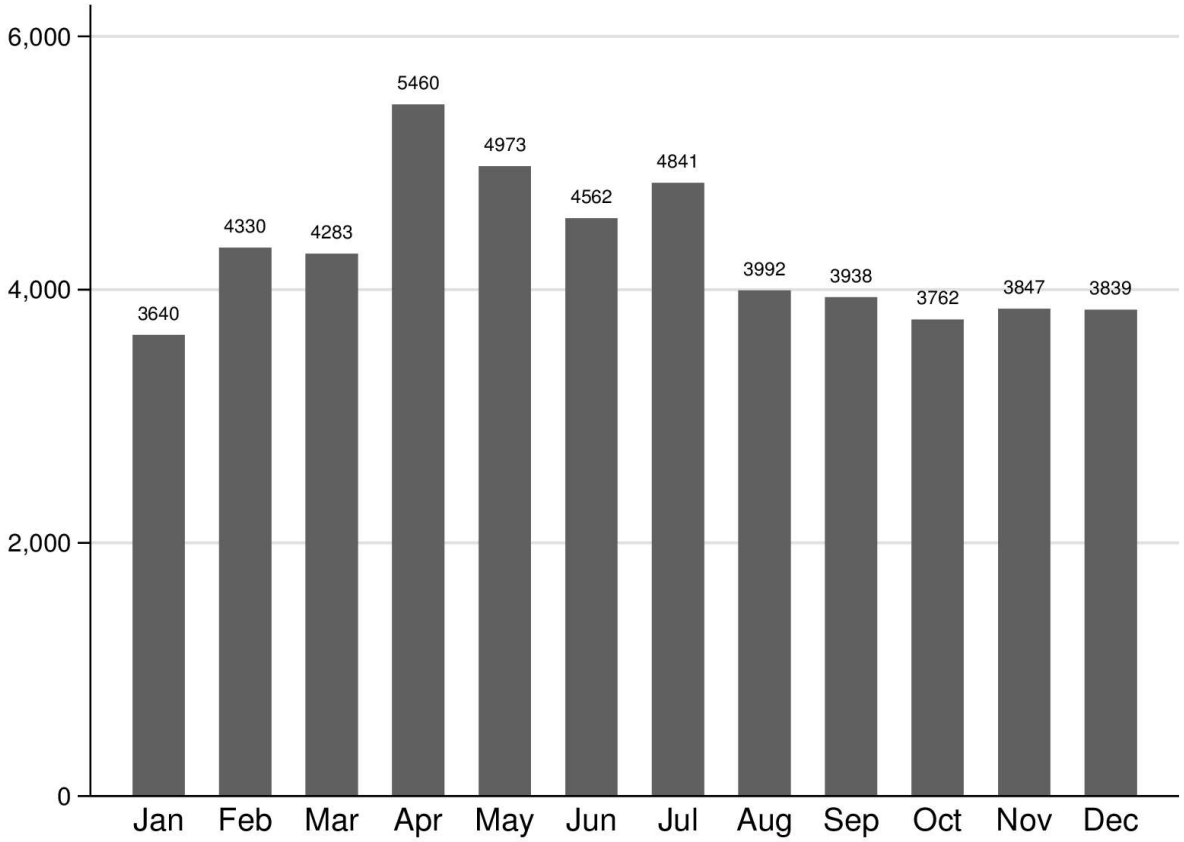


Figure 3.6 plots the number of complaints made by month, across years. It shows us that across months in 2007 and 2008, the number of complaints made per month were relatively and consistently low. In 2007, the average number of per month complaints were lower than in 2008. In 2007, the number of complaints per month did not exceed 100, and in 2008 the number of complaints per month did not exceed 350. The average number of complaints received per month in 2007 was 41, and it was 206 in 2008.

In July 2009 FICAC received over 450 complaints in one month (475) for the first time. The number of complaints per month slowly but steadily rose from then until May 2010, when it reaches just under 1,000 monthly complaints received (943). A steady decline is then observed from May to March 2011, when the per-month number of complaints received is again lower than 550.

The sharpest increase in complaints is observed in April 2011 when the number of complaints reached 1,343. We know from the Figure 4 that this sharp increase in complaints is mostly due to an increase in complaints from Ba.

With seemingly minor fluctuations, the number of per month complaints seems to level off at around 900 per month across 2012, 2013 and 2014. One notable increase is in February and March 2013, when the number of complaints received per month was around 1,200, and a dip is observed in August 2014 when the number of complaints received was around 500.

Figure 3.6: Seasonal analysis by individual years (2007-2014)

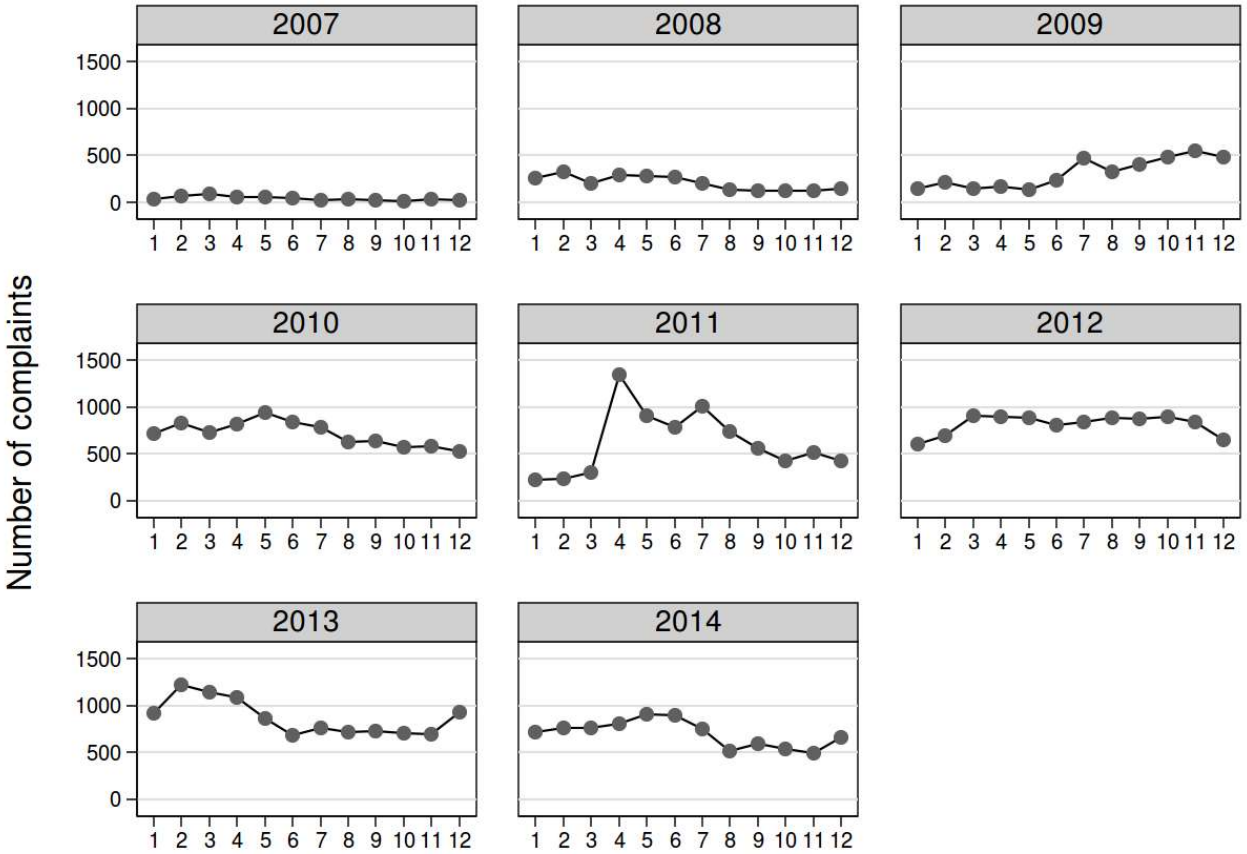
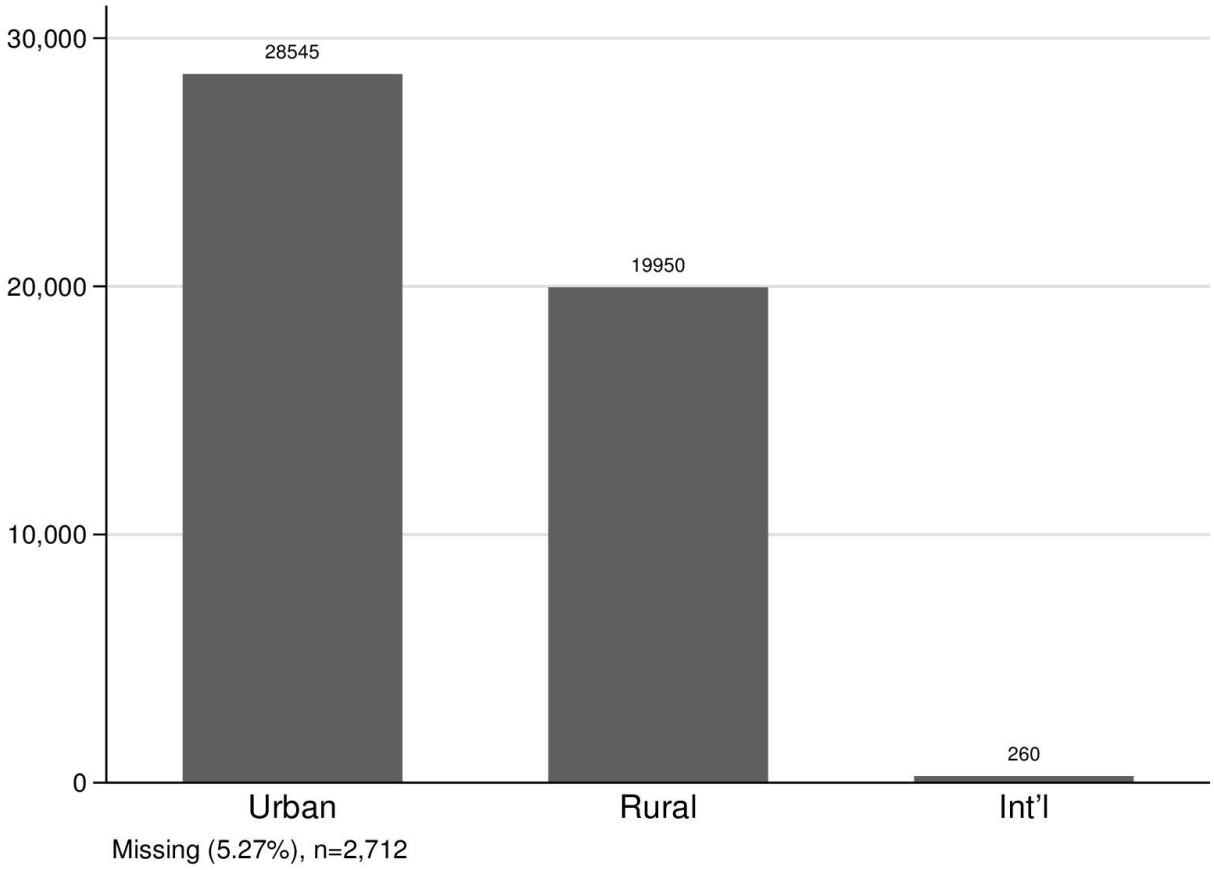


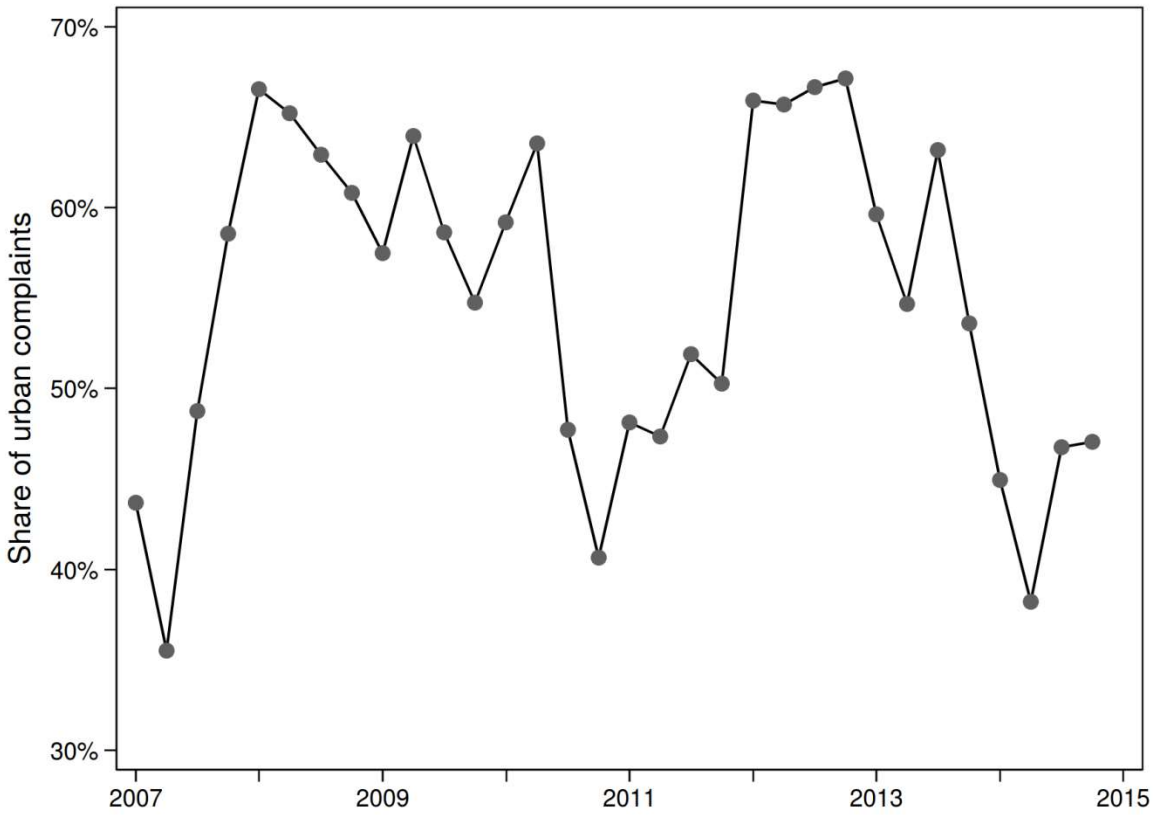
Figure 3.7 shows the number of complaints registered from urbanites, rural dwellers and from complainants residing outside of the country. Just over 55% of the complaints received were from urbanites, almost 39% were from rural dwellers, and for about 5% of the sample this information was missing. Only 260 of the complaints received were filed from complainants residing outside of the country.

Figure 3.7: Complaints from urban, rural & international locations (2007 – 2014)



The percentage of complaints from urban locations is shown in Figure 3.8. It shows that when it first began operations, FICAC was mostly taking complaints from rural areas, however this quickly changed. By 2008, two thirds of complaints were from urban locations. There was a significant dip in the share of complaints from urban areas in 2011, 2013 and 2014, which also indicates that the share of complaints from rural areas rose during these times.

Figure 3.8: Share of urban complaints (2007-2014)



4. Nature of complaints over time

This section examines the way in which complaints were made to FICAC, how the types of complaints reported changed over time, and how complaints relate to FICAC's mandate.

Figure 4.1 shows the number of complaints received via different modes of reporting. With respect to the top three modes, roughly half of all complaints received were reported in person (26,545), just over a quarter of all complaints received were reported over the phone (13,385), and seven percent (3,526) were received via the mail.

Figure 4.1: Mode of reporting (2007 – 2014)

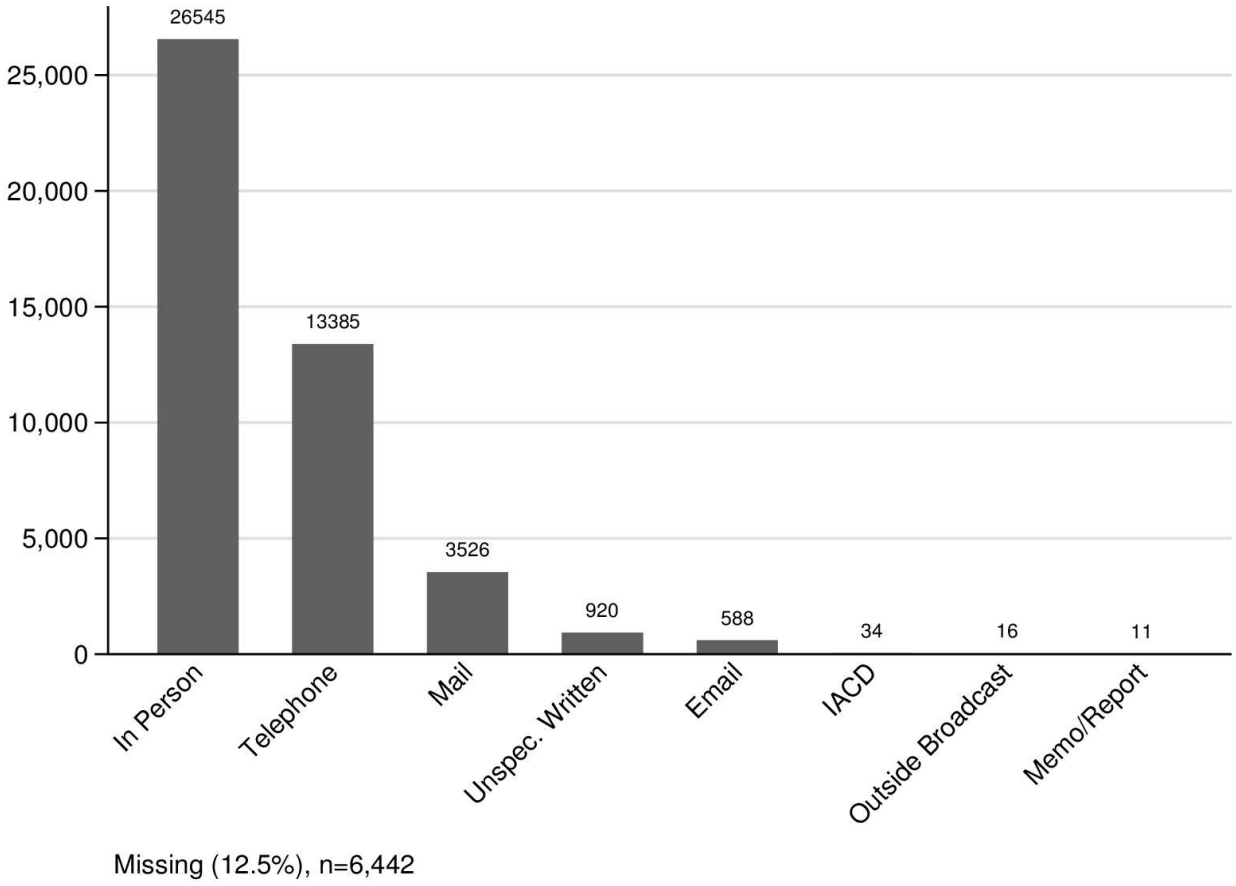
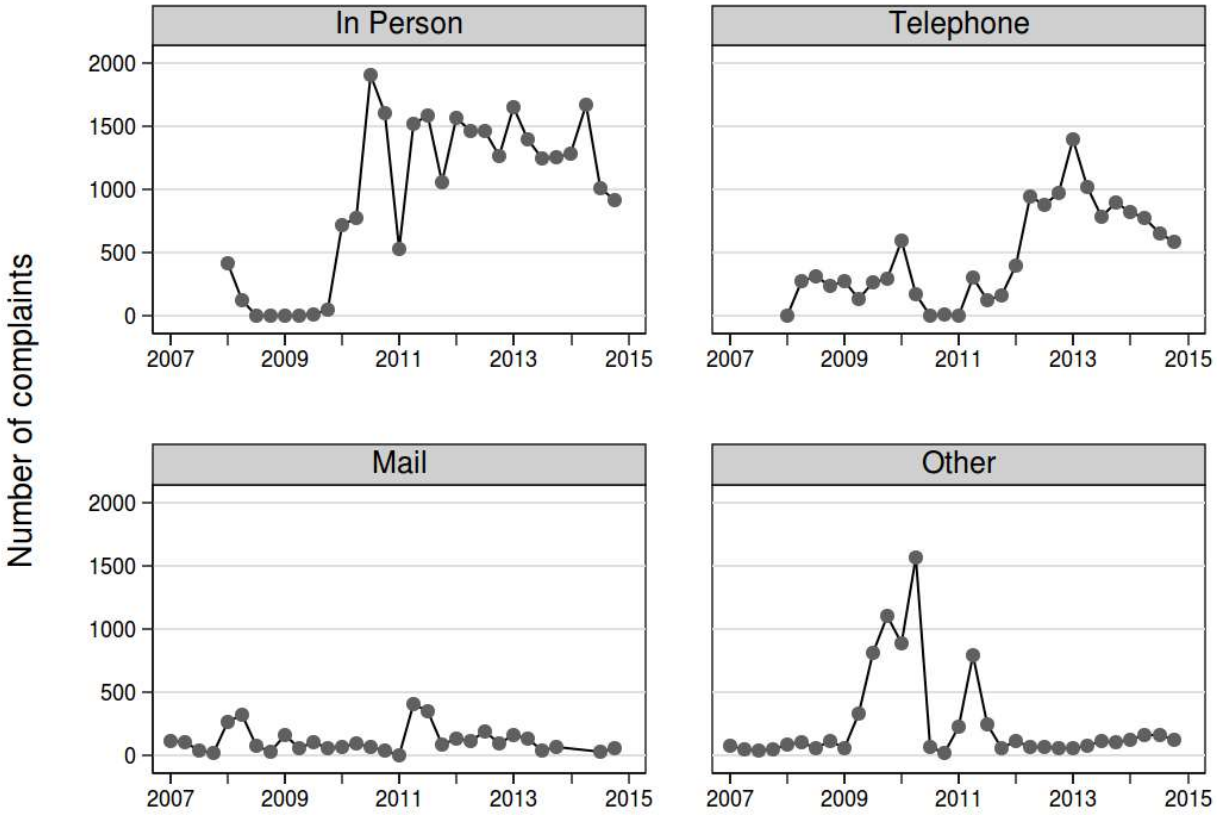


Figure 4.2 shows the number of complaints received over time via different modes of reporting. The number of complaints received in person increased dramatically in 2010: from 723 quarterly complaints in the first quarter to 1906 quarterly complaints during the third quarter. The number of quarterly complaints recorded at the beginning of 2011 dipped to 516, but then increased to over 1,000 complaints per quarter for much of the remaining period under scrutiny.

In comparison, quarterly telephone complaints were much lower from 2008 to 2012. Beginning in 2012, quarterly telephone complaints reached a peak of 1,401 in early 2013, and have remained fairly high (around 600) since. Complaints received per quarter via mail have stayed fairly steady and low throughout the period examined, at its lowest it was only 20 in late 2007, and at its highest it was 411 in the second quarter of 2011. Collectively, quarterly complaints received by any other mode remained relatively low (below 100 per quarter). Notable peaks were in late 2009 (1109 quarterly complaints), early 2010 (1562 quarterly complaints), and in mid-2011 (791 quarterly complaints).

Figure 4.2: Disaggregated methods of reporting (2007 – 2014)[2]



Moving away from the way in which people have complained, the following analysis examines what complaints were about. To do so, the research team applied a topic modelling approach to the text of the complaints. Topic modelling, a type of statistical model used for discovering the abstract "topics" that occur in a collection of documents, was used to identify prominent or "key" topics in each of the complaints (a detailed description of the methodology is found in the Appendix). Topic modelling found 32 key topics (i.e. reasons for reporting) within the data (See Appendix Table A.1). These topics were analysed and the relevant topics grouped into a

[2] There were no reported complaints by telephone or in person in 2007. We believe that this is a case of missing data, rather than no complaints occurring during this time.

total of seven different categories. Table 4.1 outlines the seven categories and highlights the types of issues they relate to.

Table 4.1: Topic categories and what they relate to

Category	Related issues captured
Employment	Issues relating to employment, including issues around payment of wages, leave, terminations, overtime, labour, and employee-employer relations.
Land	Family land disputes; tenancy issues and disputes; mataqali land leases; agriculture (including sugarcane, logging, etc.).
Society	Disputes within families (e.g. custody battles, maintenance); marital disputes; domestic violence; divorce; etc.
Contracting issues	Debt payments and disputes; failure to pay debts; permits and licences; consumer goods and services; contracts
State	Complaints about the police; legal process; public utilities; provincial development; management of public schools; abuse of office by public servants; health services; social welfare.
Private sector	Complaints predominately featuring the private sector, in particular financial institutions and travel agencies featured highly
Transport	Complaints related to marine activity, vehicles, and transport infrastructure (such as roads).

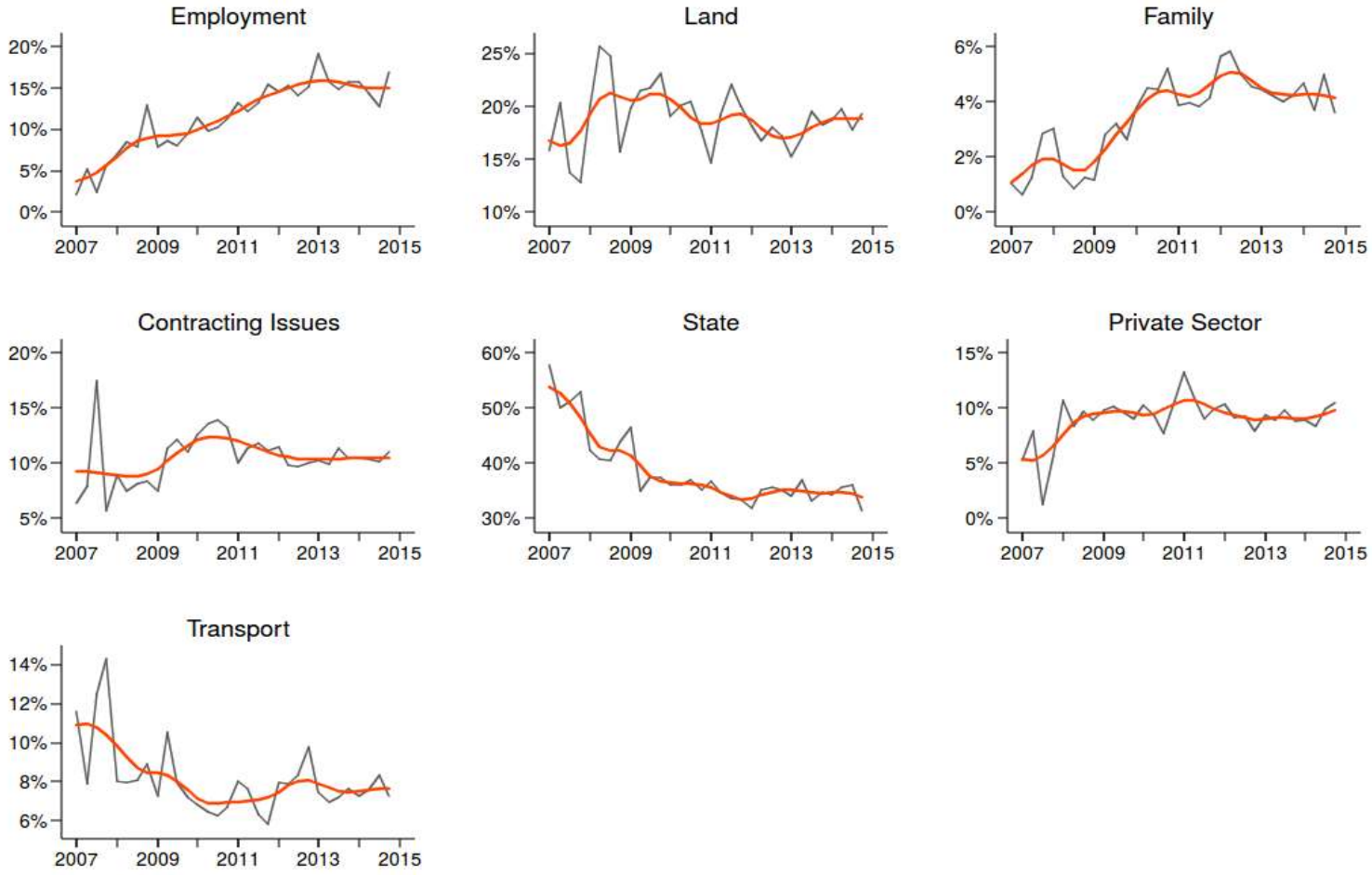
Using these categories, Figure 4.3 shows that between 2007 and 2014, the share of complaints involving employment and family rose, while issues mainly involving the state and transport fell. The share of complaints primarily involving employment started at less than 5% in 2007 and rose to its highest share of complaints (20%) in the first quarter of 2013. Complaints mostly involving the state were close to 60% of all complaints received in 2007, but decreased steadily to be around only a third of all complaints received by 2014.

Transport-focused complaints were about 11% of all complaints received in 2007, and with some fluctuation, its share steadily declined to about 7% of all complaints received in 2014. The share of complaints mostly concerning contracting, the family, land, and the private sector seems to stay fairly steady, as did the share of complaints primarily related to contracting, which varied between 9% and 15%. With some fluctuation, land issues were the main subject of about 17% of complaints; at a peak in the middle of 2008, land issues were the main topic of just over a quarter of all complaints. Also steady, private sector issues were the main topic in about 10% of all complaints received.

The share of complaints mainly involving the family remained very small throughout the period of scrutiny. Complaints focusing on the family were only about 1% of all complaints

received in 2007, but then rose to be about 4% of all complaints received in early 2008, and after a brief decline in the latter part of 2008, rose again to be primarily 3% to 4% of all complaints for the rest of the period scrutinised.

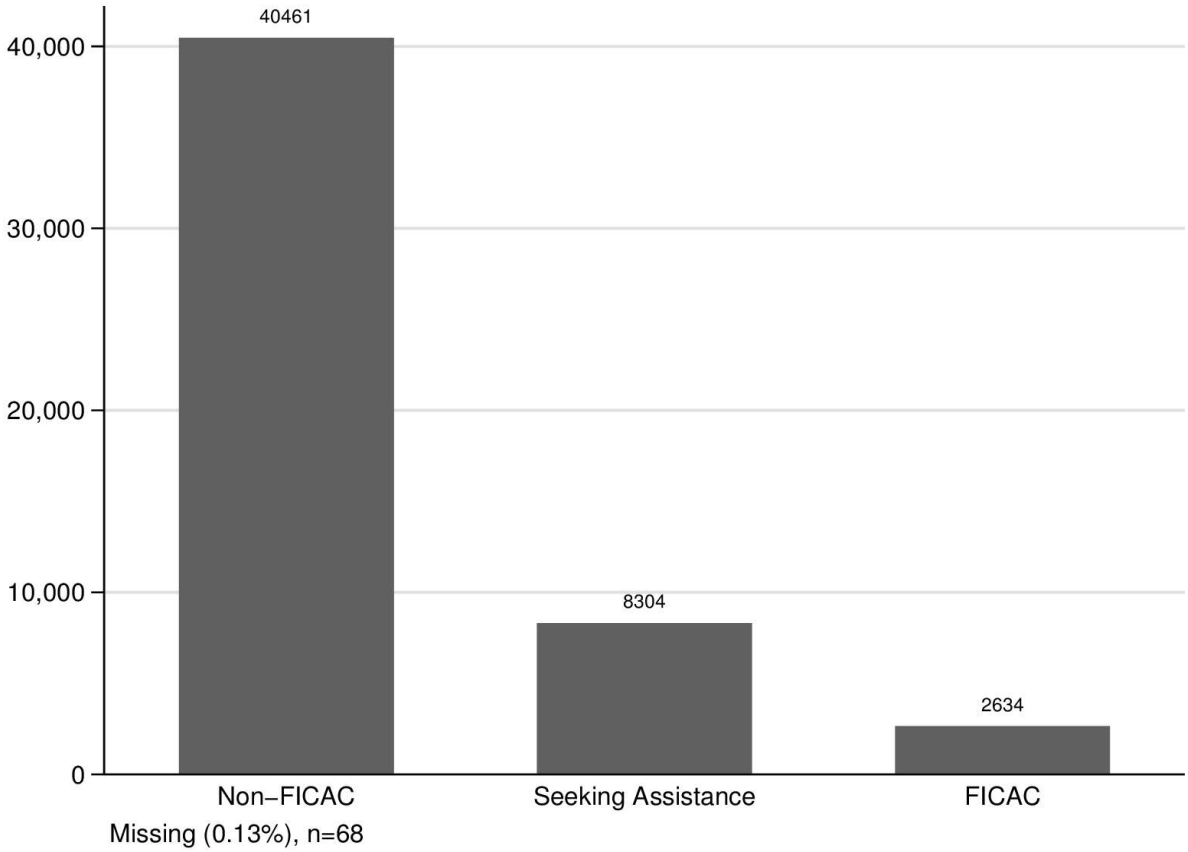
Figure 4.3: Frequency of key types of complaints (topic analysis)



Notes: Figure shows share of total complaints by dominant topic over entire sample period (2007-2014). Data are displayed at the quarterly level. A local polynomial line (orange) is displayed to aid interpretation. Note that y-axes are not on a common scale.

To understand FICAC's ability to address these different types of complaints, we now examine the percentage of complaints that fell within FICAC's jurisdiction (Figure 4.4). The vast majority of complaints (80%) received were not within FICAC jurisdiction. Just over 15% of complaints received were coded as 'seeking assistance'. This means that less than 5% of complaints received fell within FICAC's jurisdiction.

Figure 4.4: Number of complaints within and outside FICAC jurisdiction



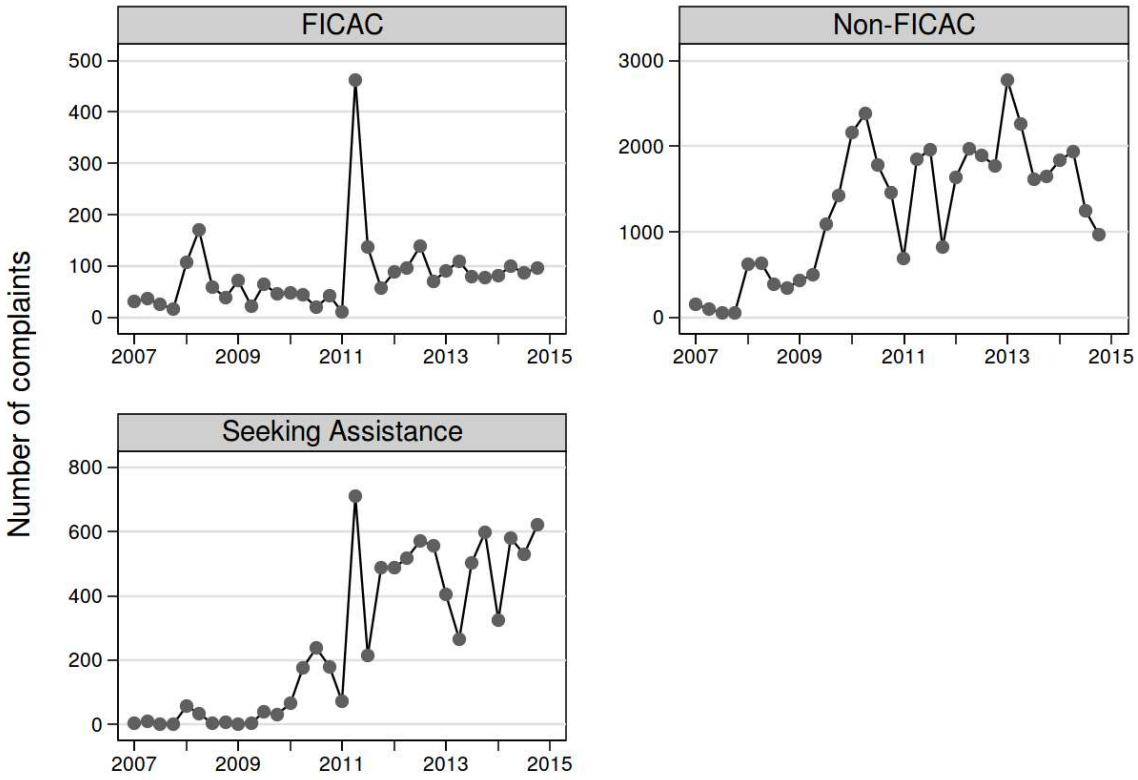
Notes: Number of complaints within the jurisdiction of FICAC, complaints outside the jurisdiction of FICAC, and submissions which sought advice or clarification on FICAC services (2007-2014).

Figure 4.5 shows how responses within and outside of FICAC's jurisdiction changed over time. Most notably, quarterly complaints within FICAC jurisdiction peaked in the second quarter of 2011 (462 complaints). This coincides with an increase in responses from complainants from Ba Province (see figure 3.4).

Much fluctuation is found amongst the complaints received that were not within the FICAC jurisdiction. Before 2010, quarterly complaints not within the FICAC jurisdiction did not exceed 700. Notable peaks include quarterly complaints not within FICAC jurisdiction reaching 2,380 in the second quarter of 2010, 1,954 in the 3rd quarter of 2011, 1,972 in the 2nd quarter of 2012, and 2,770 quarterly complaints recorded in the first quarter of 2013.

Cases coded as 'seeking assistance' remained, comparatively, much lower throughout the entire period. At its highest, 712 cases were coded as 'seeking assistance' in the 2nd quarter of 2011.

Figure 4.5: Changes to complaints within and outside FIAC jurisdiction over time



The above figure shows the number of complaints received (quarterly) that were either within the FICAC jurisdiction, not within the FICAC jurisdiction, or that were coded as 'seeking assistance' over the time period scrutinised.

Figure 4.6 shows the percentage of complaints that were either under FICAC jurisdiction, not within FICAC jurisdiction or were coded as 'seeking assistance'. The figures show that in 2007 to 2008, when the total number of complaints received by FICAC were the lowest, the majority of complaints received fell within FICAC jurisdiction. Less than 1% of all complaints received across most of the time period were under FICAC jurisdiction. In early 2011, about 1.5% of complaints received were under FICAC jurisdiction.

As is indicated in Figure 4.5, Figure 4.6 shows that there is a lot of fluctuation over time with respect to the percentage of complaints received that are not under FICAC jurisdiction. In 2010, this percentage peaks at about 95% of complaints received, but there are sharp decreases: in 2011 to around 60% of complaints received, and again in late 2014 to under 60% of complaints received. Most of the complaints coded as 'seeking assistance' were recorded post-2011. At its peaks, just over 3% of complaints received in late 2011 and late 2014 were coded as 'seeking

assistance'.

What is interesting here is that the percentage of complaints falling under FICAC's mandate was high to begin with, in 2007 and 2008 in particular. But the proportion of complaints falling under the mandate decreased over time. In other words, while there were fewer complaints in FICAC's early history, those who did complain were more likely to understand its mandate.

Figure 4.6: Changes to percentage of complaints within/outside FIAC jurisdiction

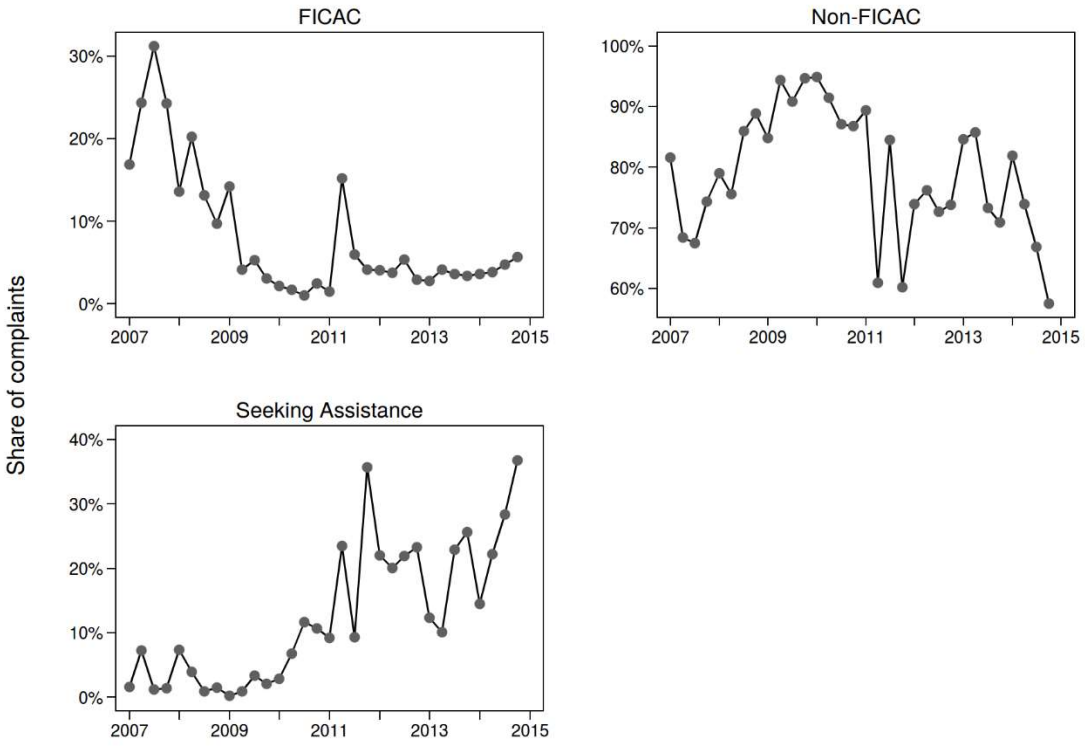


Figure 4.7 below shows the types of issues that were reported by people whose complaint fell within FICAC's jurisdiction. Figure 4.8 shows the types of complaints that fell outside FICAC's jurisdiction. Combined, the two figures show that overwhelmingly, the state is the subject of most of the complaints that FICAC receives. Complaints mainly concerning the state were almost double that of complaints concerning any other topic. 18,327 complaints primarily featured the state. The next most popular topic, irrespective of whether the complaint falls under FICAC jurisdiction, is land (9,728 complaints).

As employment issues are most usually not involving the state, they make up the third most popular topic for those cases that do not fall under FICAC jurisdiction, and only 3% of those cases that do fall under FICAC jurisdiction. 6,920 complaints, in total, were primarily concerned with employment issues. Transport issues are the third most popular topic which FICAC-jurisdiction complaints fall under, but feature in only 7% of non-FICAC jurisdiction complaints. 3,878 complaints, in total, primarily featured transport issues. In both cases, contracting issues make up about 10% of the complaints received (at 5,623 complaints, in

total), and complaints involving family members (2,144 complaints, in total) make up a very small fraction of cases received (<5%) in both cases.

Figure 4.7: Types of complaints within FICAC's jurisdiction (2007- 2014)

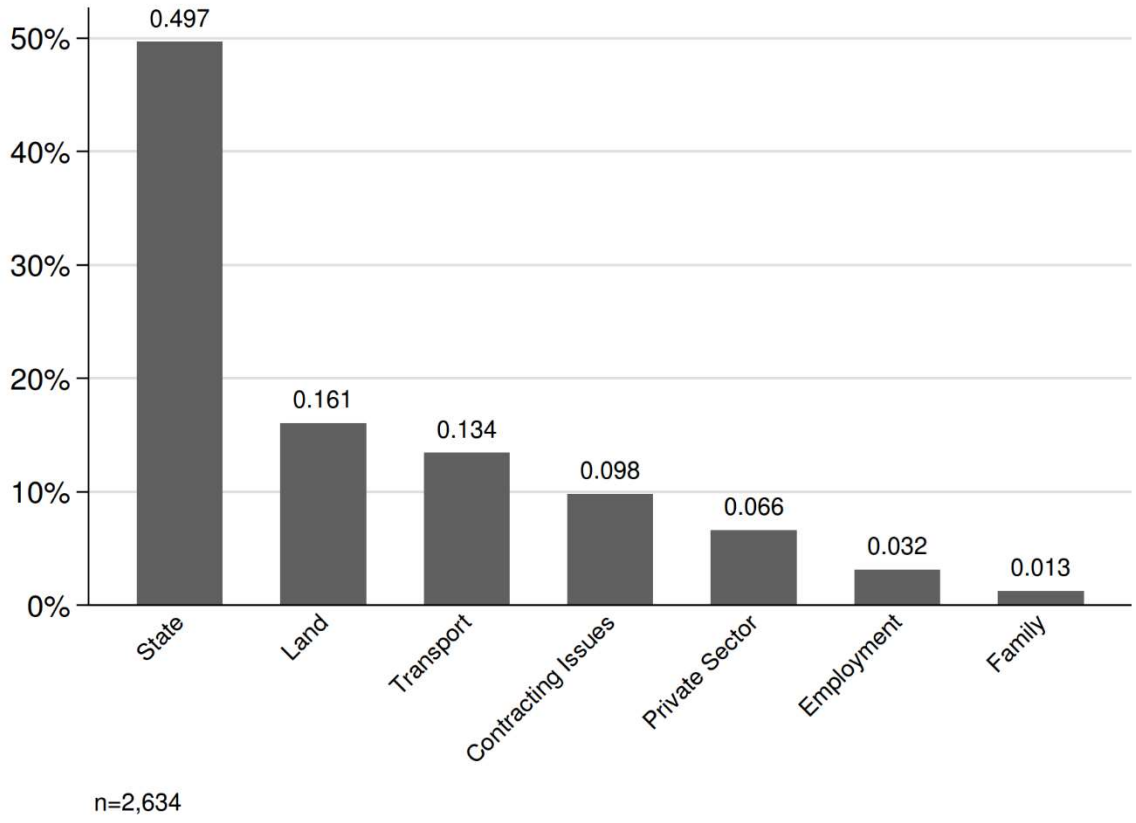
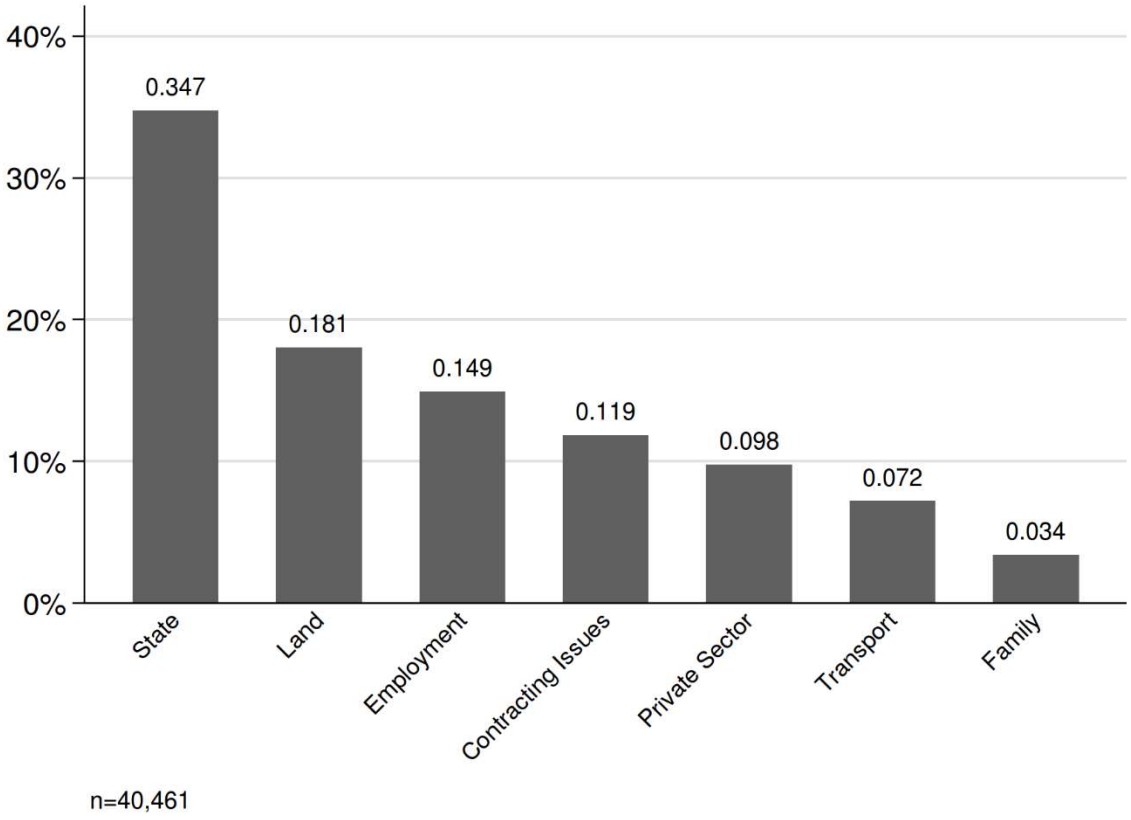


Figure 4.8: Types of reports not falling within FICAC's jurisdiction (2007-2014)



Action taken

What action was taken by FICAC staff once they had received a complaint? It is difficult to discern patterns in complaints based on the data provided. This is because many of the codes were not mutually exclusive. For example, a 'pending' code could also mean that it has been referred (which is another category). Other complaints are coded with just an institution, like 'Social Welfare', but it is not clear if this was the institution that the complaint was referred to or if it is the institution that is the topic of the complaint. With these caveats in mind, the following is true of the data collected on 'action taken':

- 33494 (70%) complaints are coded as being 'referred';
- 7501 (15.5%) are coded as 'no action taken';
- 3632 (7.5%) are coded as 'advised', 'closed' or 'resolved';
- 1308 (3%) are coded as 'forwarded for assessment' or 'investigation';
- 768 (1.5%) are coded as 'pending';
- 17 cases are coded as having been charged.

5. Key topics and FICAC's jurisdiction

The previous section examined the types of topics raised by people making a complaint. This section provides an examination of factors influencing whether a complaint will fall under FICAC's jurisdiction, and factors influencing the topic of a complaint. In other words, this section highlights who is most and least likely to make complaints that fall under FICAC's jurisdiction, and who is most and least likely to make a complaint that is focused on a specific topic. In both cases, probit regression analyses are used to examine these issues. Probit regression analyses estimate to what extent specific variables are associated with an identified, dichotomous dependent variable.

Table 5.1 displays the results from the first probit analyses. The dependent variable is whether or not the complaint falls within the FICAC jurisdiction. The coefficients are displayed, with the estimated influence of each of the variables listed in the far left column, and the stars signify the extent to which each of those variables are significantly associated with a complaint falling within FICAC jurisdiction (or not). For example, a negative and highly significant influence is shown as a negative coefficient with three stars – this means it is highly likely that increases in the associated variable will be associated with a decreased likelihood that the complaint falls under FICAC's jurisdiction.

Four trends stand out. First, complaints focusing on the state and on transport issues are positively and significantly associated with being under FICAC jurisdiction, while complaints focused on any other topic are negatively and significantly associated. This means that a complaint focused on transport or the state is more likely to fall under FICAC jurisdiction, while one that features any of the other topics is not likely to.

Second, being a female complainant is negatively and significantly associated with the complaint falling under FICAC jurisdiction. This means that females are less likely to lodge a complaint that falls within the FICAC jurisdiction, and males are significantly more likely to do so.

Third, rural dwelling complainants are less likely to lodge a complaint within FICAC's jurisdiction, but complaints lodged by complainants outside of the country are more likely to fall within FICAC's jurisdiction. These trends are expressed in the significant negative and positive coefficients associated with each of these variables.

Finally, the rest of the results in Table 5.1 relate to the influence of province location on complaint type. Ba is not included in the list, which was necessary for the analysis to be executed. One province had to be kept as the 'baseline' province. This means that the rest of the province results should be understood as showing how they compare to complaints from Ba. Ba was chosen as the 'baseline' province because, by far, more complaints were received from Ba than from elsewhere.

The results show that a complainant's region of origin sometimes matters. Complainants based in Cakaudrove, Macuata, Rewa, and Rotuma are significantly more likely to lodge complaints that fall under FICAC's jurisdiction, than those from Ba. All other regions are not significantly associated with the complaint falling under FICAC's jurisdiction, which means that

complainants from these other regions were not more or less likely to lodge a complaint that falls under FICAC's jurisdiction than complainants from Ba.

In sum, those who are making complaints that fall under FICAC's jurisdiction are concerned about the public sector and transport, complaints that are international in nature, complainants are most likely male and likely to reside in Cakaudrove, Macuata, Rewa, or Rotuma. In turn, FICAC may need to rethink its communication strategies to stress its mandate to other groups. This is particularly the case for women, who are making a significant number of complaints (see Figure 3.2), but whose complaints are not under FICAC's jurisdiction. In turn, FICAC may want to work with other groups (including NGOs and state agencies) to better communicate the array of services available to address complaints that are outside of FICAC's jurisdiction. This would also mean making sure that sufficient resources are available to other agencies to deal with any surge in complaints.

Table 5.1: Factors influencing whether complaint is within FICAC jurisdiction.

	Dependent variable: Complaint under FICAC jurisdiction (1) or not (0)						
Topic							
Employment	-1.681***						
Land		-0.379***					
Family			-1.559***				
Contracting				-0.544***			
State					0.123***		
Private sector						-0.830***	
Transport							0.597***
Profile of respondent							
Female	-0.292***	-0.270***	-0.209***	-0.272***	-0.268***	-0.262***	-0.254***
International	0.678**	0.843***	0.730**	0.708**	0.756**	0.718**	0.724**
Rural	-0.165***	-0.128***	-0.146***	-0.146***	-0.148***	-0.152***	-0.143***
Bua	0.190**	0.261***	0.245***	0.318***	0.252***	0.244***	0.265***
Cakaudrove	0.280***	0.309***	0.318***	0.357***	0.319***	0.315***	0.318***
Kadavu	0.072	0.107	0.124	0.146	0.117	0.160	0.135
Lau	-0.059	-0.047	-0.015	0.009	-0.023	0.028	-0.004
Lomaiviti	0.079	0.050	0.070	0.085	0.079	0.105	0.078
Macuata	0.219***	0.266***	0.271***	0.342***	0.276***	0.264***	0.277***
Nadroga-Navosa	0.075	0.101	0.099	0.101	0.102	0.104	0.096
Naitasiri	-0.060	-0.068	-0.059	-0.063	-0.062	-0.059	-0.061
Namosi	0.148	0.171	0.163	0.179	0.165	0.179	0.188
Other	-0.362	-0.438	-0.316	-0.304	-0.348	-0.287	-0.343
Ra	0.094	0.134*	0.124	0.122	0.124	0.127*	0.133*
Rewa	0.307***	0.275***	0.285***	0.286***	0.289***	0.293***	0.283***
Rotuma	1.027***	1.035***	1.057***	1.080***	1.057***	1.077***	1.073***
Serua	0.097	0.093	0.111	0.092	0.099	0.098	0.082
Tailevu	-0.039	-0.039	-0.027	-0.030	-0.029	-0.025	-0.030
N	48135	48135	48135	48135	48135	48135	48135
BIC	16969.9	17318.6	17256.6	17321.9	17358.8	17271.3	17303.2

Notes: Results are from probit regressions, with under FICAC jurisdiction or not as dependent variable. Coefficients are displayed with * p<0.10, ** p<0.05, *** p<0.010 indicating statistical significance. Dummy variables for the year of the complaint were also included in the analyses, but for ease of exposition, their effects are not included in the reported table.

Table 5.2 displays the results from the second set of probit analyses. The dependent variables are whether or not the complaint's main topic was employment, land, family, contracting, the state, private sector or transport, respectively. As above, the coefficients are displayed, which are the estimated influence of each of the variables listed in the far left column, and the stars signify to what extent each of those variables are significantly associated with the dependent variables.

Here, several trends deserve highlighting. First, the coefficients and significance stars in the first row indicates that females are significantly more likely to lodge a complaint about the family, the state, and the private sector, while men are more likely to lodge complaints about land, contracting, and transport related issues.

Second, though Table 5.1 (above) shows that international complaints are more likely to be within FICAC jurisdiction, Table 5.2 (below) shows that international complaints are most likely to be about land and transport issues. This is shown in the fact that the international variable is positively and significantly associated with these two dependent variables, but not with any other dependent variable. Third, the results in Table 5.2 show that rural dwellers are more likely to lodge a complaint related to land issues and the state but are less likely to lodge a complaint related to employment, contracting, the private sector, or transport.

The rest of the results in Table 5.2 relate to the influence of province location on complaint type. As with Table 5.1, Ba is not included in the list, and instead is used as the 'baseline' province. This means that the rest of the province results should be understood as showing how they compare to complaints from Ba. Ba was again chosen as the 'baseline' province because far more complaints were received from Ba than from elsewhere.

Three trends standout in the provincial results. First, compared to Ba residents, complainants from most other provinces were significantly less likely to lodge a complaint concerning employment, land, family and contracting. This trend is evidenced by the fact that most of the coefficients for those categories, associated with provinces other than Ba, are negative and significant. Second, compared to Ba residents, complainants from most other provinces (bar Macuata and 'other' provinces) were more likely to lodge complaints that primarily concerned the state. This trend is shown in the positive and significant coefficients for the state category, associated with most provinces on the left hand list.

Finally, less clear province patterns are picked up by the analyses for determining likelihood to lodge a private sector or transport focused complaint. In some instances, the coefficients associated with provinces for these categories are not significant, which signifies that complainants from these provinces are no more or less likely to lodge a complaint concerning these issues than complainants from Ba. In other cases the coefficients associated with provinces are positive and significant, and for others still they are negative and significant. For those that are negative and significant, this means that complainants from the associated provinces are less likely to lodge a complaint on the associated topic than residents from Ba, and a positive coefficient illustrates the opposite. For example, complainants from Bua, Cakaudrove, Macuata, Nadroga-Navosa and Naitasiri are less likely than complainants from Ba to report a complaint that focuses on the private sector, while those from Kadavu, Lau, and Lomaiviti are more likely to make such a complaint than complainants from Ba. And, complainants from Macuata are less likely than complainants from Ba to make a complaint that

focused on transport, while those from Cakaudrove, Lomaiviti, Nadroga-Navosa, Naitasiri, Rewa, Serua, and Tailevu are more likely to do so.

Table 5.2: Factors associated with main topic of complaint

	Employment	Land	Family	Contracting	State	Private Sector	Transport
Female	-0.024***	-0.006***	0.056***	-0.011***	0.007***	0.008***	-0.021***
International	-0.036	0.149***	0.004	-0.100	-0.076	0.021	0.046**
Rural	-0.019***	0.042***	0.001	-0.003**	0.024***	-0.010***	-0.006***
Bua	-0.069***	0.001	-0.010***	0.124***	0.009	-0.018***	-0.008
Cakaudrove	-0.056***	-0.053***	-0.005**	0.060***	0.048***	-0.015***	0.012***
Kadavu	-0.068***	-0.101***	-0.017***	-0.010	0.191***	0.030**	0.007
Lau	-0.060***	-0.123***	-0.011	-0.007	0.167***	0.042***	0.016
Lomaiviti	-0.031***	-0.109***	-0.016***	-0.006	0.102***	0.020**	0.015**
Macuata	-0.061***	-0.016***	0.003*	0.133***	-0.032***	-0.017***	-0.010***
Nadroga-Navosa	-0.034***	-0.021***	-0.006**	-0.014***	0.050***	-0.007*	0.015***
Naitasiri	-0.009***	-0.032***	-0.007***	-0.012***	0.053***	-0.005**	0.007***
Namosi	-0.028**	-0.035*	-0.011	-0.005	0.113***	0.003	-0.007
Other	-0.063**	-0.144***	-0.002	0.104	0.045	0.010	-0.015
Ra	-0.040***	0.010	-0.004	-0.012***	0.046***	0.001	-0.006
Rewa	0.004	-0.046***	-0.007***	-0.015***	0.036***	-0.003	0.012***
Rotuma	-0.071***	-0.144***	-0.031***	-0.003	0.177***	-0.008	0.028
Serua	-0.011*	-0.025***	0.009**	-0.022***	0.031***	-0.006	0.024***
Tailevu	-0.021***	-0.045***	-0.003*	-0.015***	0.052***	-0.004	0.010***
N	48176	48176	48176	48176	48176	48176	48176
BIC	-28392.0	-16892.7	-86197.0	-54240.0	-6996.3	-49030.8	-64175.2

Notes: Results are from probit regressions, with whether complaint fell under each respective topic, or not, as dependent variables. Coefficients are displayed with * p<0.10, ** p<0.05, *** p<0.010 indicating statistical significance. Dummy variables for the year of the complaint were also included in the analyses, but for ease of exposition their effects are not included in the reported table.

6. Definitions of corruption

Corruption is a contested concept which may be defined in many different ways. Academics debate the best meaning of corruption, but emerging from the literature are three key ways of defining the concept.

- First, corruption is defined as the **abuse of public office for private gain**. This definition relates corruption to public officials – which means that corruption must involve state officials or the state itself, even if the private sector is also involved. This is broadly how FICAC defines corruption, given it primarily focuses on the public sector.

- Corruption is also defined as **'the abuse of power for private gain'**. This definition draws attention away from the public sector and allows for corruption within the private sector and civil society. That is, it does not require state involvement.

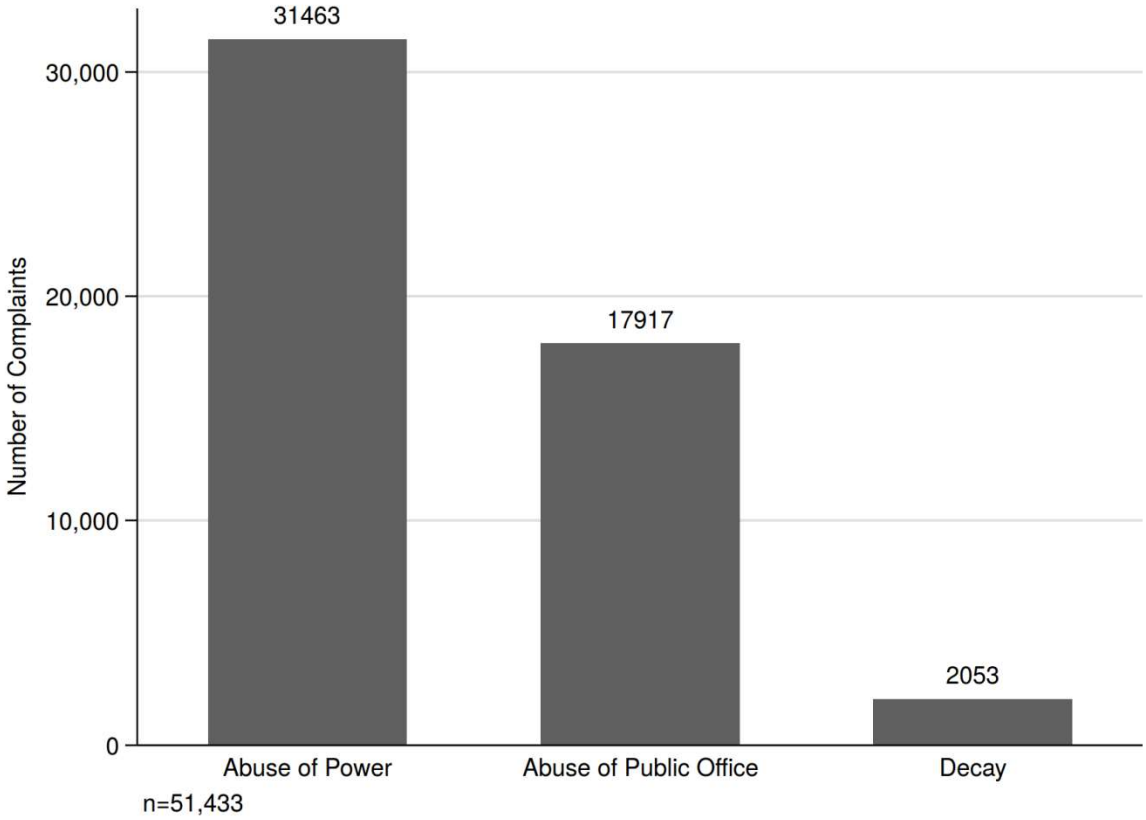
- Finally, corruption can be understood as a form of **decay** – a type of activity judged to be ethically and/or socially wrong, that does not necessitate people in power being involved. A complaint featuring a dispute within the family — perhaps one family member stealing from another — would be an example of this type of 'decay'.

To understand how all complaints reflected these definitions, two members of the research team first coded a sub-sample of complaints (1,000) into these three categories, and using that categorisation, a statistical program was trained to analyse the extent to which all of the complaints reflected these three definitions. This analytical approach is called supervised classification modelling.

It should be noted that complaints only received one code each. Complaints that clearly featured a member of the state were coded as being 'abuse of public office,' and this was even the case when the same complaint also featured a member from the private sector. Complaints were coded as 'abuse of power' if a member of the private sector was involved, a contract of any form was referenced, and/or another power dynamic was clearly present. In the decay category are all complaints where a clear power dynamic between the parties involved could not be discerned (like between friends or family members).

Figure 6.1 shows that most (61%) complaints reflected the abuse of non-state power definition. The analyses categorised 35% of the total number of complaints as concerning the abuse of public office. The remaining small minority of cases were classified as being a part of the *decay* category.

Figure 6.1: Number of complaints, by definition of corruption



The results suggest that the way complainants' defined corruption has changed over time, as Figure 6.2 demonstrates. Figure 6.2 shows the share of all cases that fall into each of the three categories, over yearly quarters.

Early on, most complaints reflected the abuse of public office definition: 67% at the beginning of 2007. This changed in 2008, after which most complaints were categorised as reflecting the abuse of power definition instead: 61% compared to 37% of complaints regarding the abuse of public office. The difference is greatest in the beginning of 2011, when only 25% of complaints were categorised by the analysis as having the abuse of public office definition and 72% of complaints were categorised as having the abuse of power definition. Post-2011, about 60% of cases are consistently categorised as having the abuse of power definition, while about a third of all cases are consistently categorised as having the abuse of public office definition.

In contrast to both of these trends, the share of complaints categorised by the analysis as 'decay' has consistently stayed low, not reaching over 10% in any of the quarters examined.

Figure 6.2: Share of complaints, by definition over time

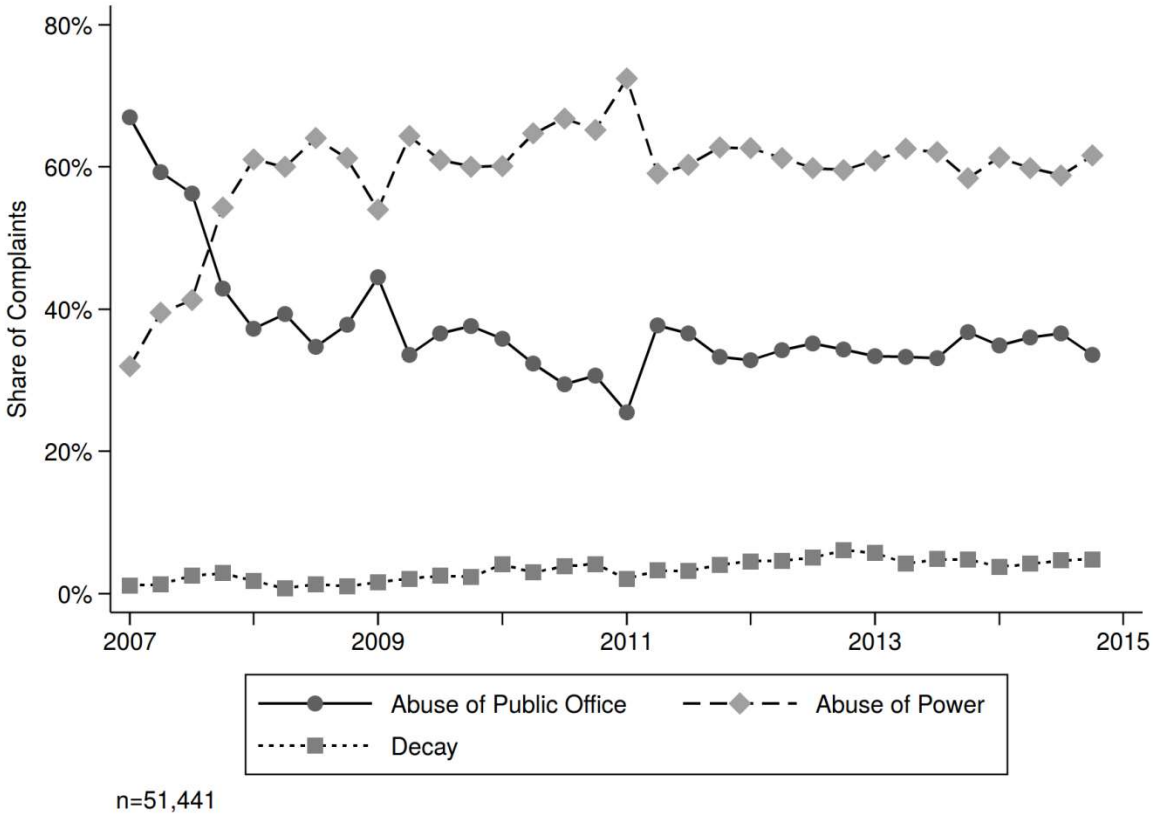


Table 6.1 displays the results from probit analyses that show how different factors influence the ways in which corruption was defined in the complaint. The dependent variables are whether or not the complaint defined corruption as being the 'abuse of public office', 'the abuse of power' or as 'decay.'

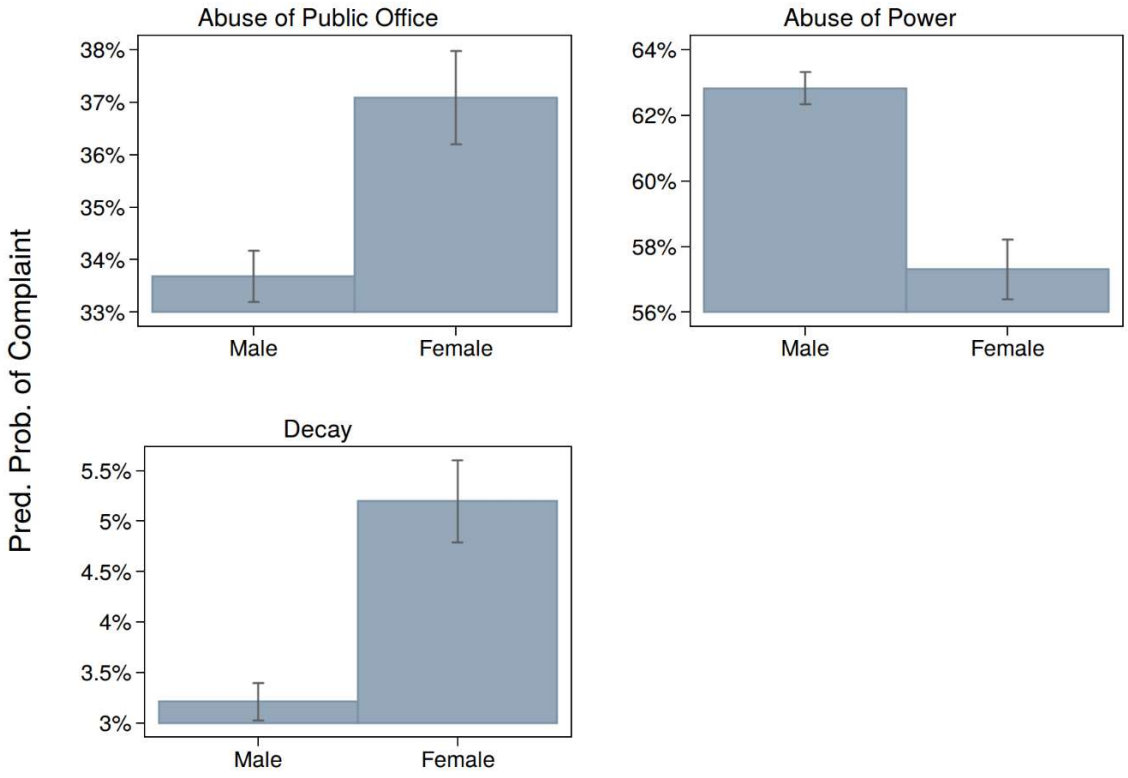
The results show that female complainants are significantly more likely than men to lodge a complaint that defines corruption as 'abuse of public office' or 'decay', and that females are less likely than men to lodge a complaint that defines corruption as 'the abuse of power'. Figure 6.3 focuses on these gendered trends. It shows the estimated probability, associated with both genders, of lodging a complaint using the three definitions. Holding the effects of all else constant, the model estimates that being a female complainant is associated with a 37% estimated probability of lodging a complaint under the 'abuse of public office definition', while male complainants only have about a 34% likelihood of doing so. It also shows that female complainants have a 5% probability of lodging a complaint under the definition of 'decay', while male complainants are estimated to only have a little over 3% probability of doing so. Finally, male complainants are predicted to have a 6% greater probability of lodging a complaint that falls under the decay category than female complainants do.

Table 6.1: Factors associated with definitions of corruption

	Abuse of Public Office	Abuse of Power	Decay
Female	0.092***	-0.143***	0.224***
International	-0.137	0.045	0.602**
Rural	0.041***	-0.051***	0.046*
Bua	-0.097*	0.076	0.085
Cakaudrove	0.122***	-0.143***	0.107*
Kadavu	0.220*	-0.191	-0.155
Lau	-0.209*	0.155	0.184
Lomaiviti	-0.047	0.035	0.052
Macuata	0.139***	-0.219***	0.326***
Nadroga-Navosa	0.105***	-0.090***	-0.064
Naitasiri	-0.174***	0.171***	-0.018
Namosi	-0.025	0.096	-0.604
Other	-0.005	0.077	-0.542*
Ra	0.086**	-0.082**	-0.005
Rewa	-0.129***	0.144***	-0.114***
Rotuma	0.464*	-0.349	--
Serua	-0.074	0.001	0.264***
Tailevu	-0.090***	0.085***	0.009
Pseudo R2	0.008	0.012	0.029
N	48144	48144	48122
BIC	61869.0	63746.9	15964.3

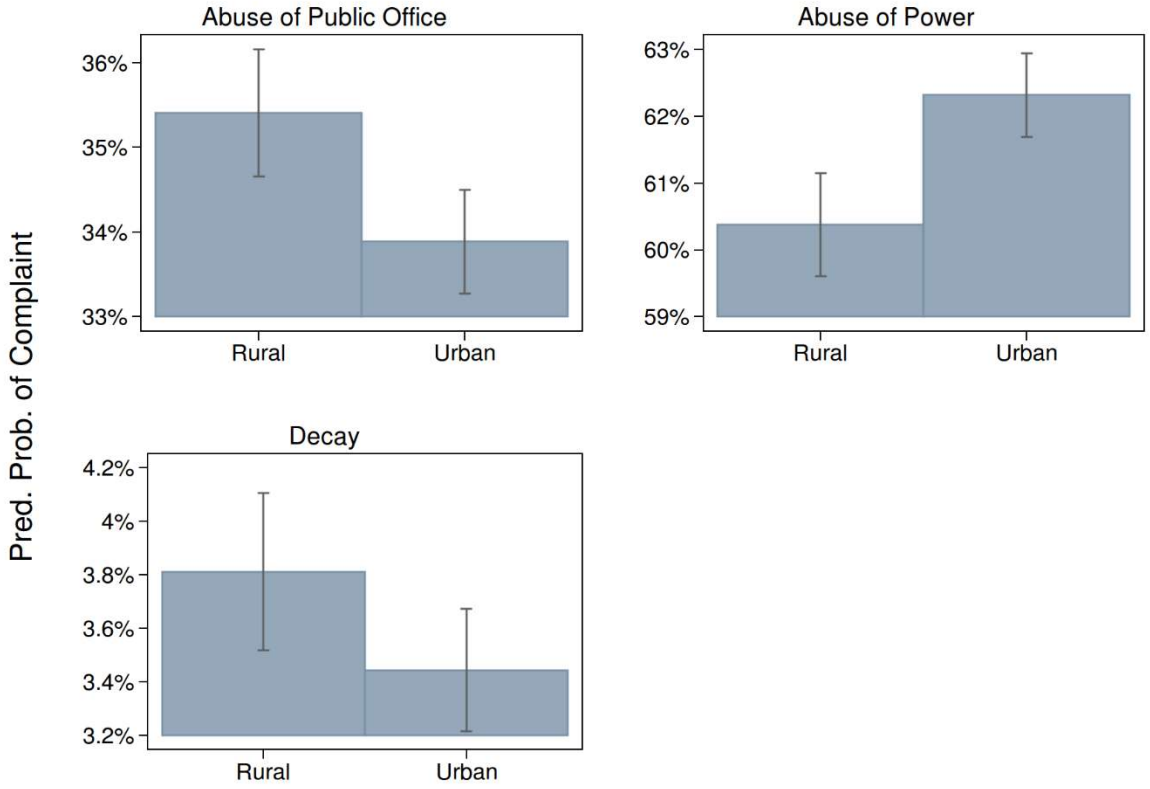
Notes: Results are from probit regressions, with whether complaint defined corruption as each respective definition as dependent variables. Coefficients are displayed with * p<0.10, ** p<0.05, *** p<0.010 indicating statistical significance. Dummy variables for the year of the complaint were also included in the analyses, but for ease of exposition their effects are not included in the reported table.

Figure 6.3: Definitions of corruption by gender



The results of Figure 6.4 and Table 6.1 also show that rural dwelling complainants are more likely than urbanites to define corruption in their complaints as the 'abuse of public office' or as 'decay,' and less likely than urbanites to define it as an 'abuse of power.'

Figure 6.4: Definitions of corruption by urban/rural



Few clear trends, across definitions, stand out strongly with respect to the influence of provincial location. Ba, again, is defined as the baseline category, and so the provincial results in Table 6.1 should be read as relative to complaints from Ba. It shows that complainants from Bua Lau, Naitasiri, Rewa, and Tailevu are significantly less likely than Ba complainants to define corruption as the 'abuse of public office', while those from Cakaudrove, Macuata, Nadroga-Navosa, Ra, and Rotuma are more likely than Ba complainants to do so. Complainants from Cakaudrove, Macuata, Nadroga-Navosa and Ra are significantly less likely than Ba complainants to define corruption as the 'abuse of power, while those from Naitasiri, Rewa, and Tailevu are more likely than Ba complainants to do so. Finally, complainants from Rewa and 'other' provinces are significantly less likely than Ba complainants to define corruption as 'decay', while those from Cakaudrove, Macuat, and Serua are more likely than Ba complainants to do so.

7. Discussion and Conclusions

This report has drawn upon complaints made to FICAC from its inception in 2007, to 2014. Analysing the available data has required the use of some sophisticated statistical approaches, which have yielded some important insights into the complaints the agency has received over the period. The analysis shows the extent to which complaints have increased between 2007 and 2014; the surge in complaints is a good news story. It shows that more and more citizens trust FICAC with their concerns about wrongdoing. However, this explosion in complaints has meant that only a very small percentage of complaints are responded to. Clearly there is a gap between citizen's concerns and FICAC's ability to respond. Addressing this gap should be a key part of the work FICAC does, because if left as it is, it is likely that complainants will become frustrated and possibly cynical about the agency's willingness to address their concerns. This report illuminates some of the factors FICAC should take into account when trying to address the gap between the demand for the service and its ability to respond.

The FICAC has been mandated with investigating corruption in Fiji's public sector. Although its role is expanding (in 2016 the FICAC Promulgation was amended to investigate and prosecute money laundering offences) the agency's focus is relatively narrow: it is concerned with corruption involving public officials. The analysis presented here suggests that complainants either do not understand this, or are unaware of other agencies who are better placed to respond to their concerns. Indeed, since 2008, most people who have complained are concerned with corruption that does *not* involve public officials (reflected by the rise of those defining corruption as the 'abuse of entrusted power for private gain'). If FICAC is to encourage these complainants to go to other agencies, it could focus on communicating that complaints made to the agency need to involve the public sector. It should be working with other government agencies to ensure citizens also know what other agencies might be better placed to address their concerns. This is not to say that citizens should be actively discouraged to call FICAC – indeed, its referrals to other agencies make it an important first port of call for many. However, working with other agencies to highlight the unique role FICAC and other agencies play to address disputes could help citizens find the right agency to lodge their complaint with more quickly. This should help address the 80% of complaints that currently fall outside of FICAC's jurisdiction.

It is encouraging to see that women have become more comfortable in using the service over time. However, women still only make up around one quarter of complainants. There are two possible reasons for this. First, women may have less engagement with the state - our research supports this hypothesis as it found that women are more concerned with family issues than men are, and their complaints were less likely to fall within FICAC's mandate. Given this, FICAC could better target its and other services to women, ensuring that services are tailored to women's experiences. Second, it may be the case that women do not want to be involved in disputes as much as men might. This suggests that more should be done to encourage women to report their dealings with officials, and that particular efforts be made to support women through the process. We suggest further research with women be undertaken to understand women's views on reporting corruption and that this research inform responses.

As people have been more willing to complain in person, it is not surprising that most complaints are made by urbanites. Telephone reports are increasingly a large percentage of complaints to FICAC, which provides an opportunity for FICAC to reach people in the more remote corners of the country. And there are signs that those in rural areas are particularly concerned about corruption involving state officials – rural respondents were more likely to define corruption as the 'abuse of public office for private gain'. However, there are signs that people in rural areas do not understand FICAC's mandate as well as their urban counterparts. People in rural areas are less likely to make a complaint falling within FICAC's mandate. In turn, it will be particularly important for FICAC to explain to this group their mandate, and the ways in which they can be assisted.

Appendix A
Detailed Methodological Note on Topic Modelling

This report seeks to classify 51,467 summaries of citizen complaints to FICAC over the period January 2007 – December 2014 according to overarching themes related to corruption. Given the scale of the corpus, a manual content analysis is infeasible due to the extremely high costs in terms of time and effort that such an approach requires. To accomplish the classification task at hand, therefore, we rely on computer-assisted methods from the fields of machine learning and natural language processing. The promise of computational methods is clear: they offer a reliable means to classify the primary topics or themes for a large corpus of text (Mikhaylov et al., 2012). Of course, these methods are not a silver bullet – considerable effort must go into ensuring model validity (Quinn et al., 2010; Grimmer and Stewart, 2013). Our approach holds the view that computational methods *assist* rather than *replace* the human researcher in classifying documents. As will be discussed below, we took a number of steps to manually validate the results of the topic modeling analysis.

1.1. Classification of complaints via an unsupervised learning algorithm

i. *Review of Latent Dirichlet Allocation*

While reading over 51,000 complaints is practically infeasible, carefully assessing 32 key "topics" is a much more attainable approach to comprehensively understanding the content of the corpus. As an initial step, therefore, we require a method to reduce our corpus of complaints to a core set of topics or themes. To this end, we employ the well-known latent Dirichlet allocation (LDA) model originally developed in Blei et al. (2003). LDA provides a statistical framework for exploring the presence of meaningful clusters of terms (i.e. "topics") that appear across the corpus of complaint summaries. These "topics" are detected by explicitly modeling the random process responsible for producing a complaint summary, assuming that each summary is composed of a mixture of "topics", and a mixture of words associated with each "topic". While the Bayesian methods used to produce "topics" are a bit involved, the important point is that the LDA has been shown to perform well in a wide range of areas, from population biology to information retrieval, and thus provides a suitable method for our data reduction task (see Blei, 2012).

ii. *Pre-processing procedures*

In an effort to economise computation, it is common when working with text data to perform a number of pre-processing steps before employing a particular analysis technique. To this end, we removed common "stop words" which appear quite frequently in the text but add little

substantive information (e.g. prepositions, articles, etc.). We also remove the affixes from each word and combine words with common *stems*, a process typically referred to as "stemming" (e.g. "complains" / "complained" / "complaining" = "complain").

iii. *Model estimation*

Using the pre-processed text, we then proceed with the topic model estimation. The LDA model requires that the researcher specify the number of topics a priori. We employed a data-driven approach (see Arun et al., 2010) to help us in formulating the topic selection criteria. After some experimentation, we determined the "correct" number of topics found in the corpus to be 32. Once the number of topics to be declared was finalised, we then proceeded with the model estimation with the use of a sparse Gibbs sampler (Yao et al., 2009; see Wallach et al., 2009 for details on our hyper-parameter optimisation routine). Each word in the corpus is assigned some probability of belonging to each of the 32 topics (or clusters, as discussed above). In turn, each topic has some probability of belonging to a given complaint summary. To ease interpretation of these topics – and with the assistance of employees at FICAC – we produced a descriptive label for each topic after evaluating the 19 most probable keywords.

Table A.1 provides a listing of these topic labels and the most probable keywords for each topic. The comparison of the topic labels to the most probable keywords provides a first pass at determining the level of *semantic validity* of the model – the extent to which each topic is coherent in terms of its meaning (see Quinn et al., 2010). To increase our confidence that the topic labels are accurate representations of the clusters produced by the model, we also manually reviewed the 20 most probable *complaint summaries* for each topic. Following this procedure, we were satisfied that the topic labels correspond to the true nature of the estimated latent topics.

iv. Generation of "super-topics"

Reviewing the topic labels presented in Table A.1, it becomes clear that many of the topics can be further grouped into higher-order themes. Relying on the topic keywords and labels, we were able to generate seven "super-topics": *employment, land, family, contracting issues, state, private sector,* and *transport.* Table A.2 provides a mapping of these "super-topic" labels to the underlying topics generated from the initial stage of the analysis. Note that some topics (e.g. "Complainant/Accused") did not fit well into any of these super-topics and were thus omitted.

2. Supervised classification of theoretical classes

The complaint summaries were further classified according to three theoretically meaningful groupings: *abuse of public office, abuse of power,* and *decay*. Given that these classes cut across the estimated topics discussed above, we employed a supervised classification approach to

measuring the prevalence of these classes within the corpus (see Grimmer and Stewart, 2013). In short, a random set of 500 complaints were coded by two researchers according to the three types of complaints discussed above. Several pilot studies were carried out to ensure adequate inter-coder reliability (Krippendorff's alpha = 0.8). Once a satisfactory level of inter-coder reliability was achieved, the coders proceeded to encode the remaining complaints in the set of 500. This coding was then used as a training set for stochastic gradient descent classifier (see Zhang, 2004). Having trained the model with these data, the algorithm was then employed to produce predicted classifications for the entire corpus (n=51,433). To determine the accuracy of the classification procedure, we calculate average F1 scores for the three classes via 10-fold cross validation. As is demonstrated in Table A.3, there is considerable variation in classification accuracy across the three theoretical classes, with the best performing category being *abuse of public office* (0.80) and the least being *decay* (0.36). Based on these validation metrics, we can have high confidence in the supervised classification of *abuse of public office* and *abuse of power*, while reserving scepticism regarding the classification of *decay.*

Table A.1. A full list of the estimated topics.

The table provides the unique topic ID, descriptive topic label, and the top 19 stemmed keywords.

Topic ID	**Topic Label**	**Most Probable Words**
0	Tenancy	landlord house rent notice vacate flat property renting tenant paying month pay bond agreement move tenants issued moved eviction
1	Police	police station officer reported officers inaction accused sigatoka action post nadi nausori lautoka charged statement informed force hrs valelevu
2	Permits & Licenses	taxi lta council town business permit license city base operating market suva officer nadi illegal driver nasinu driving scc
3	Marine Activities	company business boat tax resort frca limited fishing island customs director country return owner operating fish hotel sea vessel
4	Communication	caller mobile anonymous hrs received free toll line phone nadi fijian enquire enquiring enquired authority relevant grog calling vodafone
5	Mataqali Land Lease	mataqali lease land tltb trustees itltb money village tokatoka letter yavusa monies consent trust funds turaga offer paid distribution

6	Debts and Payments	money paid pay agreement sum payment amount accused payments refund balance total alleged purchase requested agreed promised loan cash
7	Consumer Goods/Services	items shop item selling phone repair purchased bought company machine courts food price purchase store sold hire suva missing
8	Financial Institutions	account bank money loan cheque union bsp deposited finance pay withdraw credit received westpac amount insurance housing payments fdb
9	Public Utilities	water meter fea waf authority bill electricity power supply bills paid disconnected pay arrears energy connection pipe received house
10	Post & Telecommunications	post taveuni email mail received address batteries david money subject bobby parcel master party times john send elizabeth james
11	Agriculture	cane farm sugar fsc sugarcane pine farmers cutting cut farmer trees harvest gang logging personally land mill limited payment
12	Family Land Disputes	land property lease father lands title brother transfer department house piece family transferred lawyer consent lot mother sold passed
13	Provincial Development	village church turaga koro project villagers island government yavusa funds meeting provincial money tui committee roko ratu kadavu development
14	Infrastructure & Transport	road roads pwd gravel health department neighbour house neighbor council town authority reported compound access damaged settlement government driveway
15	Health & Medical Services	hospital medical health doctor compensation lautoka injured nurse ministry admitted injuries cwm department sick injury center doctors staff nadi
16	Construction Contracting	house building company materials construction paid contractor job timber hired complete completed contract cost renovation plan private home build
17	Legal	court lawyer khan lautoka tribunal magistrate legal claim filed file hearing shah paid ali pay nadi sum law lawyers
18	Delaying Processes	received delaying payment reported feedback hasn suva months nausori month naitasiri village paid week receive follow tailevu process delay

19	False Promises	labasa personally money owing amounting pay savusavu paid northern promised macuata false seaqaqa telephone giving todate promises ministry calling
20	Higher Education	fnu tender ministry staff university company national fees manager students education post position officer contract awarded government interview school
21	Family1	husband family wife children father court mother passed maintenance late death married child left sister affairs birth daughter marital
22	Travel Agents	money travel agent paid visa obtaining deception sum naidu australia mukesh refund passport zealand agency false application immigration suva
23	Vehicles	vehicle car repair engine mechanic paid company registration lta garage bought motors owner accident van purchased truck motor repairs
24	Employee Entitlements	fnpf company pay employer wages deducted security paid deduction contributions contribution employed payment account deductions months week nonpayment deducting
25	Family2	son daughter brother law sister family wife home house mother father cousin husband left return nephew living wati lautoka
26	Public Service Vehicle	bus driver vehicle transport company driving hrs fare road lta buses government officer school morning carrier passengers registration truck
27	Employment Issues	labour department company employer pay wages leave employed paid termination terminated letter overtime unfair security reported officer week employment
28	Complainant/Accused	kumar chand prasad singh ram lautoka narayan lal sharma raj deo nand chandra reddy prakash nadi devi reported road
29	School Management	school funds teacher education committee ministry primary head association teachers students management principal government financial president manager college misuse
30	Social Welfare	welfare social department application family allowance receiving applied food elections voucher labasa suva received month officer scheme election mother
31	Government Referrals	letter officer lautoka received application government investigation attached department alleged officers suva documents manager information copy informed public ministry

Table A.2. A mapping of the seven "super-topics" onto the set of underlying topics.

Table A.2 shows which underlying topics made up the seven 'super-topics'. It is worthwhile noting that not all topics in the table above were used because some of them were considered 'junk topics' – that is topics that were not found to be meaningful enough to include. For example, the Post and Telecommunications category in the table above was about how communication was received, more than the nature of complaint. This topic was thus left out of the super topics.

"Super-topic" label	**Underlying topic labels [Topic ID]**
Employment	Employee Entitlements [24], Employment Issues [27]
Land	Tenancy [0], Mataqali Land Lease [5], Agriculture [11], Family Land Disputes [12]
Family	Family1 [21], Family2 [25]
Contracting Issues	Debts and Payments [6], Construction Contracting [16], False Promises [19]
State	Police [1], Permits & Licenses [2], Public Utilities [9], Provincial Development [13], Health & Medical Services [15], Delaying Processes [18], Public Service Vehicle [26], School Management [29], Social Welfare [30]
Private Sector	Consumer Goods/Services [7], Financial Institutions [8], Travel Agents [22]
Transport	Marine Activities [3], Infrastructure & Transport [14], Vehicles [23]

Table A.3. Average F1 score in 10-fold cross validation for theoretical classes.

Theoretical class	**F1 Score**
Abuse of Public Office	0.80
Abuse of Power	0.70
Decay	0.36

Works cited

Arun, R., Suresh, V., Veni Madhavan, C., & Narasimha Murthy, M. (2010). On finding the natural number of topics with latent dirichlet allocation: Some observations. *Advances in Knowledge Discovery and Data Mining*, 391-402.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993-1022.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*,*55*(4), 77-84.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, 267-297.

Mikhaylov, S., Laver, M., & Benoit, K. R. (2012). Coder reliability and misclassification in the human coding of party manifestos. *Political Analysis*,*20*(1), 78-91.

Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., & Radev, D. R. (2010). How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science*, *54*(1), 209-228.

Wallach, H. M., Mimno, D. M., & McCallum, A. (2009). Rethinking LDA: Why priors matter. In *Advances in neural information processing systems* (pp. 1973-1981).

Yao, L., Mimno, D., & McCallum, A. (2009). Streaming inference for latent dirichlet allocation. kdd, 2009. KDD.

Zhang, T. (2004, July). Solving large scale linear prediction problems using stochastic gradient descent algorithms. In *Proceedings of the twenty-first international conference on Machine learning* (p. 116). ACM.