The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Social Determinants of Police Corruption: Towards Public Policies for the Prevention of Police Corruption

Strategies for the prevention of police corruption, e.g., bribery, commonly neglects its social dimension in spite of the fact that police corruption has societal causes and undertaking a reform of the police requires, to some extent, reforming society. In this paper, we built a decision tree from socio-economic profiles of 103 countries classified according to their level of police corruption using data from the United Nations Statistics Division and Transparency International. From the rules of the resultant decision tree, we identified and analyzed social determinants of police corruption to assist policymakers in designing societal level strategies to control police corruption by improving socio-economic conditions. We found that school life expectancy, involvement of women in society, economic development, and work-related indicators are relevant to police corruption. Moreover, empirical results indicate that countries should gradually improve social indicators to reduce police corruption.

Keywords: police corruption; corruption; public policies; social determinants of police corruption; police anti-corruption strategies

J. Octavio Gutierrez-Garcia*
Department of Computer Science
Instituto Tecnológico Autónomo de México
Mexico City, DF 01080, Mexico
octavio.gutierrez@itam.mx

Luis-Felipe Rodríguez
Department of Computer Science
Instituto Tecnológico de Sonora
Ciudad Obregon, Sonora 85000, Mexico
luis.rodriguez@itson.edu.mx

*Corresponding author's contact details:
Tel. +52 (55) 5628 4000 Ext. 3645, octavio.gutierrez@itam.mx
1 Rio Hondo St., Progreso Tizapan, Mexico City, DF 01080, Mexico.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

1. Introduction

Corruption has been recognized as one of the major obstacles for economic growth and development (Jain 2001). Among governmental institutions, *the police* is commonly identified as one of the most corrupt institutions according to Transparency International (2014). In addition, police corruption may even foster a perception of impunity for crimes (Newham 2000), undermining the credibility of the entire criminal justice system (Bruce 2008).

Police corruption, and particularly police bribery (as a prototypical form of police corruption (Newburn 1999)) is a social phenomenon (Carvajal 1999). Dimant (2013) points out that corruption has profound roots in social aspects such as education, gender, religion, urbanization, and ethnical separation. In the same vein, Botero et al. (2012) present a statistical analysis indicating that educated countries are less prone to corruption. Similarly, Dutta et al. (2011) found a relationship between corruption and the informal labor force. Whereas the research efforts of Botero et al. (2012), Dimant (2013), and Dutta et al. (2011) emphasize the social aspects of corruption in general, there is no analysis focused on the social determinants of police corruption.

Strategies for reducing police corruption are mostly focused on reforms at the level of police departments by inducing changes into police culture, recruitment, and training (Bayley 2011). Other strategies for controlling police corruption are police integrity tests (Prenzler 2001) and complaint profiles (Prenzler 2003). Nonetheless, the social dimension of police corruption is commonly neglected in spite of the fact that *the reform of the police becomes, then, part of the reform of the society* (Punch 2000) and that corruption has societal causes (Botero et al. 2012; De Graaf 2007; Dimant 2013; Dutta et al. 2011). Moreover, as stated by McCusker (2006), effective anti-corruption strategies require a potentially long-term cultural reform at a societal level. Then,

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

strategies aimed at controlling police corruption should take into account the society which the police serves.

The purpose of this paper is to determine the societal factors on police corruption in order to provide policymakers with guiding principles for the design of long-term anti-corruption strategies focused on changing the social environment. In this paper, we present a quantitative analysis of socio-economic profiles of 103 countries supported by decision trees (Quinlan 1993) to assist in the design of public policies for reducing police corruption. Unlike previous related work (Bayley 2011; Prenzler 2001; Punch 2000) that focused on anti-corruption strategies at the level of police departments, this paper adopts a societal perspective. We determine the social indicators that are relevant for classifying police institutions of 103 countries according to their police corruption score (PCS). Socio-economic profiles of the countries were created from survey results and statistics reported by the United Nations Statistics Division (2014). The PCSs of the countries were obtained from survey results provided by Transparency International (2014) in its global corruption barometer.

The major contribution of this paper is that, from a quantitative analysis, we propose a set of guiding principles to assist policymakers in designing societal level strategies to control police corruption by improving social indicators. In addition, by analyzing quantitative evidence, we also contribute with a list of social determinants of police corruption. Moreover, we provide empirical results suggesting that countries should gradually improve social indicators to reduce police corruption.

2. Strategies for the Prevention and Control of Police Corruption

According to Newburn (1999), strategies for the prevention and control of police corruption can be categorized into (i) human resource management, (ii) anti-corruption

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

policies, (iii) internal controls, and (iv) external environment and external controls.

Strategies involving *human resource management* make use of specialized recruitment techniques, e.g., psychological preemployment screening (Arrigo 2003; Dantzker 2011), whose objective is to detect personality factors that may affect police integrity. In addition, training practices for police personnel have been designed in order to control police corruption by teaching what is wrong and what is right to police officers (Lamboo 2008). Another technique of human resource management to control police corruption is developing pride (Newburn 1999), under the assumption that police officers that are proud to be in charge of law enforcement are less prone to corruption. A more drastic approach to control police corruption is making police supervisors responsible and accountable for the integrity of their subordinate police officers (Klaver 2013).

Strategies involving *anti-corruption policies* aim at designing and promoting the adoption of ethical codes, e.g., the standards of police conduct proposed by the International Association of Chiefs of Police (2014). Another anti-corruption policy, as pointed out by Newburn (1999) and Punch (1994), is creating (i) police supervisors in charge of promoting the ethical code and/or (ii) ethics commissions which police officers may confidentially consult when dealing with ethical issues. In addition, as indicated by Small (2003), corruption can also be prevented by rewarding the ethical behavior of police officers.

Strategies involving *internal controls* aim at improving (or removing) internal procedures of police organizations, e.g., eradicating corruption-prone procedures as suggested by Newburn (1999). In addition, anti-corruption internal controls may involve the creation of supervision programs for police officers. For instance, the Australian Victoria Police developed an early warning system that profiles police

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

officers by using citizens' reports (Macintyre 2008). In doing so, the accumulation of several minor complaints (reported by citizens) about a given police officer are used to issue warnings about potential and more severe misbehaviors of the police officer. Another successful implementation of anti-corruption internal controls is the police reform in Georgia (Di Puppo 2010), which involved salary increases, the redesign of the public image of the police agency, and the introduction of a new regulatory framework. A common form of anti-corruption internal control is the implementation of internal affairs departments, which may even use undercover officers to monitor and detect potential acts of corruption (Girodo 1998). Other forms of anti-corruption internal controls are integrity tests and polygraph tests (Prenzler 2009).

Strategies involving the *external environment and external controls*, instead of aiming at changing police agencies, aim at changing the (social) environment. This type of anti-corruption strategy is based on the idea that corruption can also result from external pressures on police officers to become corrupt (Newburn 1999). In this regard, Sherman (1978b) states that a police agency may become corrupt as a result of an external pressure from a corrupt political environment. Williams (2002) argues that the cultural traditions of countries are a core factor of police corruption, and thus, an external pressure on police officers to become corrupt. Moreover, in some countries, citizens see police corruption as an acceptable everyday inconvenience and share negative attitudes towards the law with police officers (Williams 2002). Williams (2002) also documented that citizens, according to their culture, may prefer to pay a small bribe rather than a large fine. Following this trend, police corruption has been recognized as a social phenomenon given that *it takes two to bribe* (Carvajal 1999; Ivković 2003; Newburn 1999; Polinsky 2001; Punch 2000; Sherman 1978a). In response to this, some anti-corruption strategies aimed at changing the environment

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

have been designed (Kies 2011; Rugege 2006). For instance, Rwanda's government introduced a new constitutional law that severely punishes bribe-givers (Rugege 2006). Kies (2011) highlights the proliferation of citizen journalism as an external control of police misconduct by video recording police activities and uploading the videos to public video streaming services on the web.

In spite of the numerous research efforts focused on developing strategies for the prevention and control of police corruption, there is still a long journey ahead before eradicating police corruption. In fact, there is not a single anti-corruption strategy that has been sufficient to fully address police corruption. Nonetheless, as concluded by Punch (2000), the combination of multiple anti-corruption strategies has helped to prevent and control police corruption. In this paper, we identify and analyze the social determinants of police corruption. By determining the societal factors on police corruption, we aim at (i) complementing existing strategies and (ii) providing guiding principles for the design of long-term anti-corruption strategies focused on changing the social environment.

3. Data on Police Corruption and Countries' Social Indicators

In order to identify the social determinants of police corruption, we created socio-economic profiles for 103 countries, which were classified according to their levels of police corruption. The data on police corruption (see Table 1) and social indicators (see Table 2) of countries was obtained from survey results and statistics reported by Transparency International (2014) in the global corruption barometer and the United Nations Statistics Division (2014), respectively. We collected data of only 103 countries because they were the only countries whose information about police corruption and social indicators is available in both Transparency International (2014) and the United

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Nations Statistics Division (2014).

From the global corruption barometer of Transparency International (2014), we obtained a police corruption score (PCS) for each country. The PCS was determined by asking those who are involved in acts of corruption with the police, e.g., citizens (Ivković 2003). Transparency International (2014) uses a real number between 1.0 (not at all corrupt) to 5.0 (extremely corrupt) to indicate the PCS of a country. Nonetheless, we used integer-valued PCSs that resulted from rounding the real-valued PCSs. In doing so, we reduced the number of potential classes from forty one (from 1.0 to 5.0 in 0.1 increments) to five classes: 1, 2, 3, 4, and 5. This made the analysis tractable given that, as indicated by Quinlan (1993), decision trees work better when there are considerably more instances (e.g., socio-economic profiles of countries) than classes (e.g., PCSs).

Countries	**Police Corruption Score (PCS)**
None	1 (not at all corrupt)
Denmark, Finland, Norway, Rwanda, and Switzerland.	2
Afghanistan, Australia, Azerbaijan, Belgium, Cambodia, Canada, Estonia, Ethiopia, Fiji, France, Georgia, Germany, Hungary, Iraq, Italy, South Korea, Libya, Luxembourg, Maldives, New Zealand, Palestine, Portugal, Slovenia, Spain, Turkey, The United Kingdom, The United States, and Uruguay.	3
Albania, Algeria, Argentina, Armenia, Bangladesh, Bosnia and Herzegovina, Brazil, Bulgaria, Burundi, Cameroon, Chile, Colombia, Croatia, Cyprus, Czech Republic, Democratic Republic of the Congo, Egypt, Macedonia, Greece, India, Israel, Japan, Kazakhstan, Latvia, Lebanon, Lithuania, Malaysia, Moldova, Mongolia, Morocco, Mozambique, Nepal, Pakistan, Papua New Guinea, Paraguay, Peru, Philippines, Romania, Senegal, Serbia, Sierra Leone, Slovakia, Solomon Islands, South Africa, Sri Lanka, Sudan, Thailand, Tunisia, Ukraine, Vanuatu, Venezuela, Vietnam, and Yemen.	4
Bolivia, El Salvador, Ghana, Indonesia, Jamaica, Kenya, Kyrgyzstan, Liberia, Madagascar, Malawi, Mexico, Nigeria, Russia, Tanzania, Uganda, Zambia, and Zimbabwe.	5 (extremely corrupt)

Table 1. Police corruption scores

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

The number and type of social indicators (Table 2) were selected because they represent the minimum list of social indicators for monitoring and following up social development of countries proposed by the United Nations Working Group on International Statistical Programmes and Coordination (2014). Thus, by using the list of social indicators, we created complete but compact socio-economic profiles for each country. The social indicators are classified into five categories: population, health, housing, education, and work. From these categories, we analyzed 68 social indicators for each of the 103 countries. Please see the documents provided by the United Nations Statistics Division (2014) for a detailed definition of the social indicators and how they were measured.

It should be noted that all the social indicators were normalized with respect to the maximum value of each indicator to range from 0.0 to 1.0 in order to prepare the data to create the decision tree shown in Fig. 1.

It is worth mentioning that for some countries, some of their social indicators were missing. The missing value of a given social indicator (of a country) was replaced with the mean value of the available values (of the other countries) for the social indicator. This data cleaning process is commonly used to deal with missing values in real-world data sets when applying machine learning techniques (Saar-Tsechansky 2007).

Social indicator	**Category**
Population size	Population
Total percentage of population under 15 years	
Percentage of male population over 60 years	
Percentage of female population over 60 years	
Sex ratio in 60+ age group	
Annual population growth rate	
Percentage of urban population	

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Sex ratio of international migrants	
Life expectancy at birth of women	Health
Life expectancy at birth of men	
Maternal mortality ratio	
Infant mortality rate	
Under 5 mortality rate	
Adolescent fertility rate	
Total fertility rate	
Contraceptive prevalence for any method	
Contraceptive prevalence for modern methods	
Estimated number of adults living with HIV/AIDS	
Women's share of adults living with HIV/AIDS	
Average number of people per room	Housing
Average number of people per room in urban areas	
Average number of people per room in rural areas	
Population distribution in urban areas	
Population distribution in rural areas	
Annual rate of population change in urban areas	
Annual rate of population change in rural areas	
Total percentage of improved drinking water coverage	
Percentage of improved drinking water coverage in urban areas	
Percentage of improved drinking water coverage in rural areas	
Total percentage of improved sanitation coverage	
Percentage of improved sanitation coverage in urban areas	
Percentage of improved sanitation coverage in rural areas	
Total adult literacy rate	Education
Male adult literacy rate	
Female adult literacy rate	
Total youth literacy rate	
Male youth literacy rate	
Female youth literacy rate	
Girls' net enrolment ratio in primary education	
Boys' net enrolment ratio in primary education	
Girls' share of primary enrolment	
Girls' net enrolment ratio in secondary education	
Boys' net enrolment ratio in secondary education	
Girls' share of secondary enrolment	
Women's net enrolment ratio in tertiary education	
Men's net enrolment ratio in tertiary education	
Women's share of tertiary enrolment	
Total school life expectancy	
Male school life expectancy	
Female school life expectancy	
Per capita gross domestic product	Work
Total adult economic activity rate	
Male adult economic activity rate	
Female adult economic activity rate	
Total percentage of part-time employment	
Percentage of male part-time employment	
Women's share of part-time employment	

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Percentage of female employees
Percentage of male employees
Percentage of female employers
Percentage of male employers
Percentage of female own-account workers
Percentage of male own-account workers
Percentage of female contributing family workers
Percentage of male contributing family workers
Total adult unemployment rate
Male adult unemployment rate
Female adult unemployment rate

Table 2. Social indicators for defining socio-economic profiles of countries.

4. Method

We conducted the data analysis using decision trees (Fig. 1). A decision tree is a supervised machine learning technique for classification that allows for the generation of decision rules.

A decision-tree analysis was selected because, as indicated by Stevenson et al. (2009), decision trees can better identity subgroups (e.g., police institutions with a relatively high PCS) than some statistical analyses, e.g., logistic regression approaches. In addition, there is evidence (see Zurada and Lonial (2011)) that decision trees provide better classification rate than some statistical analyses, e.g., regression analysis, when nonlinearities are observed. Moreover, decision trees are easy to understand and intelligible to policymakers (Quinlan 1993).

The structure of a decision tree is composed of decision nodes and leaf nodes represented by rounded rectangles and circles in Fig. 1, respectively. A decision node represents a test on a feature (e.g., a social indicator) that determines the branch traversed and as a consequence the next decision node to be evaluated until reaching a leaf node. A leaf node indicates the class (e.g., highly corrupt) to which a feature vector belongs.

To build the decision tree, we used the J48 algorithm (which is an open source implementation of the C4.5 algorithm proposed by Quinlan (1993)) implemented in Weka (Hall 2009), a data mining software package. The input of the J48 algorithm was the 103 feature vectors (described in Section 3), which are composed of the 68 normalized social indicators of a given country (reported in Table 2). In addition, given that decision trees are a supervised machine learning technique, each feature vector was labeled with their corresponding class (namely 1, 2, 3, 4, and 5) representing the possible PCSs of a country.

The J48 algorithm recursively splits the feature vectors into two subsets by using tests until each subset contains feature vectors of a single class.

Given that the features of the countries (analyzed in this work) are numerical values, a test (i.e., a decision node) is composed of a social indicator, a relational operator, and a threshold value. For instance, a test that separates feature vectors is as follows: *annual population growth rate* less than or equal to a threshold value of 0.84.

The J48 algorithm makes use of information entropy, namely information gain ratio, to select the most informative feature for each decision node at each recursive step. In order to define information gain ratio, we firstly define entropy and information entropy. Entropy is a measure of uncertainty. As described by Weaver (1949), when "*a situation is highly organized, it is not characterized by a large degree of randomness or of choice - that is to say, the information (or the entropy) is low*". Conversely, when a situation is characterized by a large degree of randomness or of choice, the information (or the entropy) is high. Information entropy is a measure of uncertainty of information (Bashkirov 2000). The higher the information entropy of a feature, the most informative

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

it is. Information gain ratio represents the proportion of information generated by splitting a set of instances using a test (Quinlan 1993), i.e., a decision node. By selecting features based on information gain ratio, the branches of decision trees are composed of the most informative features, i.e., the features that better classify the instances. Please see (Shannon 1946) and (Quinlan 1993) for mathematical definitions of entropy and the information gain ratio criterion, respectively. Once a feature f has been selected for a decision node, the J48 algorithm sorts feature vectors based on f in ascending order. Then, the information gain ratio is measured for each available value of feature f, and the threshold value with the highest information gain ratio is selected.

It should be noted that we did not make use of a test set, i.e., a set of feature vectors to test the capacity of the decision tree to classify unseen feature vectors. This is because we are using decision trees as a decision support tool, as in (Welsh 2014), and not as a classifier of unseen feature vectors.

At the expense of generality, the decision tree was not pruned and the minimum number of instances per leaf was set to 1 in order to keep all the features that were found relevant according to the information gain ratio.

A rule group is a branch of the decision tree that classifies countries into one of the PCSs. For instance, rule groups 2a, 2b, and 2c (or just rule group 2) classify countries with a PCS of 2, see Figs. 1 and 2. Rule groups that classify countries with better (i.e., lower) PCSs, can be used to assist in the decision-making process (Savage 2003) to determine what social indicators should be improved by other countries with worse (i.e., higher) PCSs to reduce police corruption.

In order to measure how distant the features (i.e., the social indicators) of a given country were from the decision node values that classify countries based on a given PCS, we designed three metrics: separation distance, rule compliance, and combined metric.

(1) The *separation distance* metric measures the absolute differences of non-complying features of a country with respect to the values of the decision nodes of a given rule group (see Figs. 1 and 2). For instance, if a country has a normalized *male school life expectancy* of 0.78, the absolute difference with respect to the numeric value of the first decision node of rule group 2a (i.e., *male school life expectancy > 0.81*) is 0.03. Once all the non-complying features of the country have been measured, the sum of the absolute differences is divided by the number of rules of the rule group from which the country's features are being compared to. The lower the value of the *separation distance* metric, the better the fit of the features (of a country) to comply with the rule group. The value of the *separation distance* metric ranges from 0.0 to 1.0.

(2) The *rule compliance* metric measures the percentage of rules (of a given rule group) that a country complies with. The higher the value of the *rule compliance* metric, the better the fit of the features (of a country) to comply with the rule group. The value of the *rule compliance* metric ranges from 0.0 to 1.0.

(3) The *combined* metric is a combination of both *separation distance* and *rule compliance* metrics. The combined metric of a country for a given rule group is defined as 1 *minus* the value of its *separation distance* metric *plus* the value of its *rule compliance* metric divided by 2. The higher the value of the *combined* metric, the better the fit of the features (of a country) to comply with the rule group. The value of the *combined* metric ranges from 0.0 to 1.0.

It should be noted that neither the *separation distance* metric, nor the *rule compliance* metric, nor the *combined* metric take into account that rules are nested, see

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Fig. 2. This is because we designed the metrics to give the same priority to all the features of a rule group.

5. Data Analysis and Results

We constructed a decision tree (Fig. 1) to (i) determine the social indicators that are relevant for classifying police institutions according to their PCSs and (ii) give insights into the causes of police corruption at a societal level. The analysis and results are divided into two categories:

(1) Analysis of the relationship between the social indicators and the level of police corruption of a country (Section 5.1).

(2) Analysis of the most relevant rule groups and how these can be used to give insights to address police corruption according to the current socio-economic situation of each country (Section 5.2).

5.1. Analysis of the Relationship between Social Indicators and Police Corruption

From the analysis of the structure of the decision tree, we drew seven observations.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

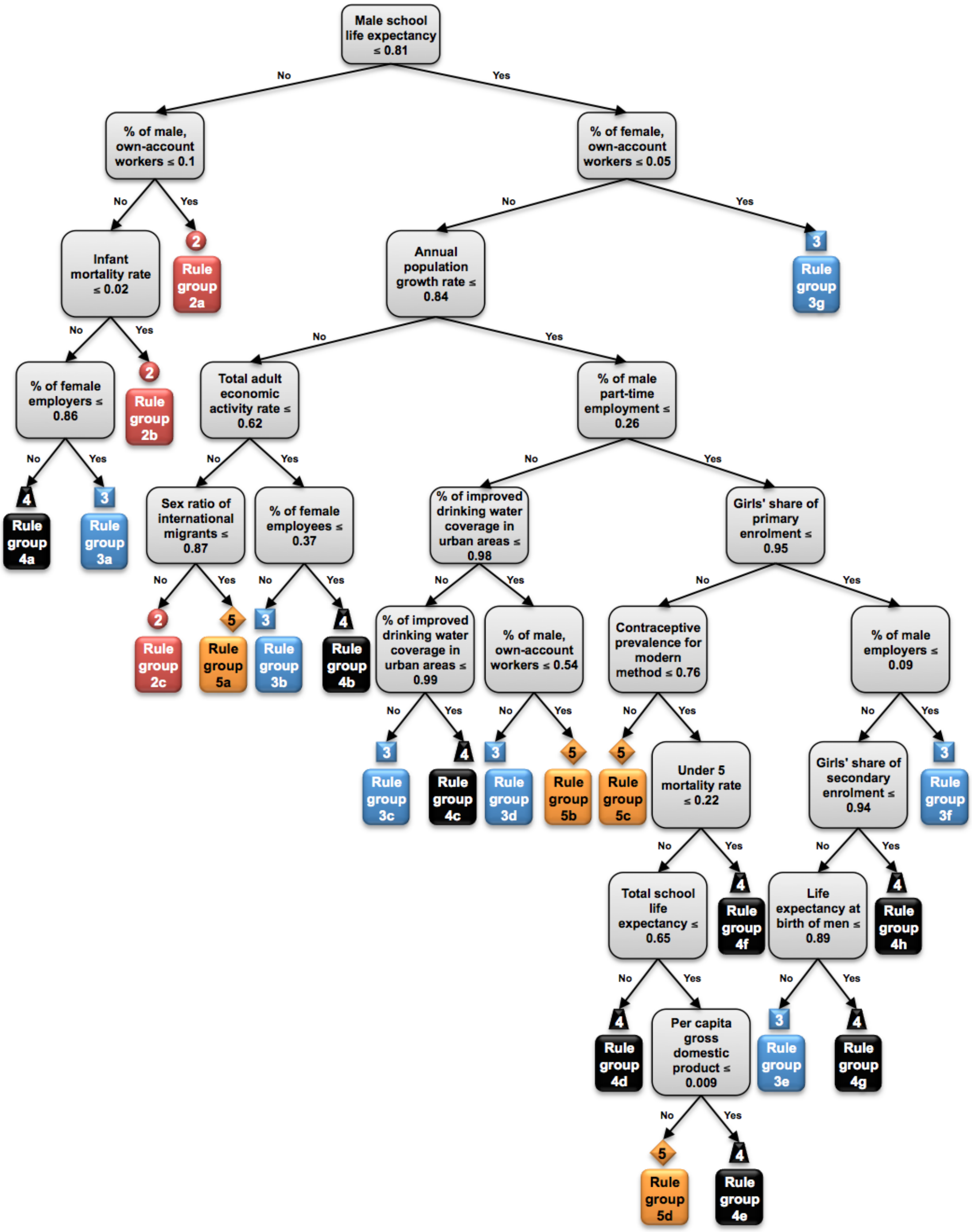

Fig. 1. Decision tree for classification of countries by police corruption score.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Rule group 2a:

If Men's school life expectancy > 0.81 **Then**
  If Percentage of male, own-account workers ≤ 0.1 **Then**
    The police corruption score of the country is 2

Rule group 2b:

If Men's school life expectancy > 0.81 **Then**
  If Percentage of male, own-account workers > 0.1 **Then**
    If Infant mortality rate ≤ 0.02 **Then**
      The police corruption score of the country is 2

Rule group 2c:

If Men's school life expectancy ≤ 0.81 **Then**
  If Percentage of female, own-account workers > 0.05 **Then**
    If Annual population growth rate > 0.84 **Then**
      If Total adult economic activity rate > 0.62 **Then**
        If Sex ratio of international migrants > 0.87 **Then**
          The police corruption score of the country is 2

Fig. 2. Rule group 2: decision tree's branches that classify countries with a PCS of 2.

Observation 1. The level of police corruption of a country can be characterized by nested *if-then* rules involving features used for monitoring social development proposed by the United Nations.

Analysis. From the data described in Section 3, we obtained the decision tree shown in Fig. 1 that correctly classifies with a 0% classification error rate all the 103 countries according to their PCS by using their social indicators. The decision tree (Fig. 1) is composed of 20 out of 68 social indicators, which were the ones that provided the most information gain and were sufficient to classify all the countries. It should be noted that according to Transparency International (2014), no country was considered (by its inhabitants) to be free of police corruption, i.e., with a PCS of 1. Hence, the countries were classified into only four classes: 2, 3, 4, and 5.

Observation 2. The most informative social indicator (from the list of indicators for monitoring social development proposed by the United Nations) with respect to police corruption is *male school life expectancy*.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Analysis. The most informative feature of rule group 2a (and of all the other rule groups) is *male school life expectancy*, see Fig. 1. As indicated in rule group 2a (Fig. 2), countries (e.g., Denmark, Norway, and Switzerland) with a *male school life expectancy* greater than a normalized value of 0.81 tend to have less police corruption than countries with a lower *male school life expectancy*. This is consistent with data reported by Botero (2012) whose statistical analysis suggests that the more educated a country is, the less corrupt the country, under the assumption that corruption is reduced because educated people tend to report it. Moreover, Dreher and Herzfeld (2005) in their analysis of the costs of corruption state that a high level of corruption reduces expenditures on education of a country significantly affecting school enrollment. Whereas no causal relationship can be established between *male school life expectancy* and the level of police corruption of countries, the decision tree (Fig. 1), Botero (2012), and Drexer and Herzfeld (2005) agree that the level of education is relevant to the corruption level of a country.

Observation 3. Countries with a well-educated population should also have a low *percentage of male own-account workers* in order to have a low level of police corruption, e.g., a PCS of 2.

Analysis. The second rule of rule group 2a (Fig. 2) indicates that in addition to having a well-educated population, a country should have a low *percentage of male own-account workers* (less than or equal to a normalized value of 0.1). Own-account workers are commonly located in the informal sector of the labor force (Saunders 2003). In this regard, Dutta et al. (2011) found a direct and proportional relationship between corruption and the informal labor force. The larger the share of the informal labor force, the higher the corruption (Dutta et al. 2011). In addition, Tonoyan (2004) found that only the wealthier self-employed are less inclined to corruption.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

By taking into account the results reported in (Saunders 2003; Dutta et al. 2011; Tonoyan 2004) and the information provided by the decision tree, we can conclude that (i) the *percentage of male own-account workers* is relevant to police corruption and (ii) in general, a large formal labor force prevents police corruption.

Observation 4. *Infant mortality rate* was selected as a social indicator relevant to police corruption. However, its relevance may derive from its correlation with *male school life expectancy.*

Analysis. Rule group 2b (Fig. 2) included a healthcare-related decision node, namely *infant mortality rate*. Whereas, as far as we know, no direct relationship between healthcare and police corruption has been found, as stated by Husted (1999), education *is likely to improve health, not only of those receiving it, but of their children*. In fact, according to data reported by the United Nations Statistics Division (2014), there is a significant correlation between *male school life expectancy* and *infant mortality rate* of -0.58. Thus, using *infant mortality rate* as a decision node of rule group 2b seems to be due to its correlation with *male school life expectancy.* It should be noted that given that the features of the decision nodes are selected based on information gain ratio, the features are good at classifying countries into the different corruption levels, but that does not impede some correlation between other decision nodes.

In addition, it should be pointed out that only Finland was classified by rule group 2b. However, Finland was very close to be classified by using rule group 2a, with a *male school life expectancy* greater than a normalized value of 0.81 and a *percentage of male own-account workers* of 0.15, see Fig. 2.

Observation 5. The masculinity of a country is relevant to police corruption.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Analysis. The feature of the second decision node of rule group 2c (Fig. 2) is *percentage of female own-account workers*, which is also a feature of 17 out of the 22 rule groups that compose the decision tree. In addition, the decision tree (Fig. 1) is composed of five decision nodes related to the active participation of women in a country, namely (i) *percentage of female own-account workers*, (ii) *percentage of female employers*, (iii) *percentage of female employees*, (iv) *girls' share of primary enrolment*, and (v) *girls' share of secondary enrolment*. In this regard, as demonstrated by Husted (1999), the masculinity of a country is positively correlated with its corruption level. In addition, Swamy et al. (2001) found that women are less prone to corruption and that countries where women comprised a relatively large share of the labor force are less corrupt. Both Husted (1999) and Swamy et al. (2001) establish a relationship between the involvement of women in society and corruption in general, whereas the present work indicates that such relationship holds for police corruption as well. This observation suggests that the active involvement of women in society as self-employed workers, employees, employers, and students is relevant to differentiate countries with respect to the level of police corruption.

Observation 6. Economic development and work-related social indicators are relevant to police corruption.

Analysis. The decision tree (Fig. 1) is mostly composed of work-related social indicators, 9 out of 21 decision nodes make use of a social indicator that belongs to the *work* category (see Table 2). Among the work-related social indicators are *percentage of female employees* and *total adult economic activity rate*, which are related to the overall employment and economic activity of a country.

With respect to economic development, Husted (1999) hypothesizes that the higher the level of economic development, the lower the level of corruption. With

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

respect to work-related social indicators, Newham (2012) indicates that unemployment contributes to police corruption. Both Husted (1999) and Newham (2002) suggest a direct relationship between economic development and work-related social indicators and the (police) corruption level of a country. However, whereas the work-related features included in the decision tree suggest that economic development has a relationship with police corruption, the decision tree does not provide a clear indication that the higher the level of economic development, the lower the level of police corruption. Nevertheless, the decision node involving the *per capita gross domestic product* indicates that (after evaluating eight decision nodes) countries with a very small per capita gross domestic product (i.e., less than or equal to a normalized value of 0.009) have a PCS of 5. In contrast, countries with a higher per capita gross domestic product have a better PCS of 4.

In addition to the above analysis, it should be noted that the relationship between corruption in general and economic growth has been widely studied (Bardhan 1997; Johnson 2004; Rose-Ackerman 2007). In this regard, *Observation 6* suggests that police corruption is not the exception, but the norm.

Observation 7. Rwanda is a special case that generated its own rule group.

Analysis. Only Rwanda was classified by rule group 2c (Fig. 2). Rwanda is the only country with a PCS of 2 with a relatively low *male school life expectancy* with a normalized value of 0.57. In fact, rule group 2c required five rules to classify Rwanda, in contrast to the two and three rules of rule groups 2a and 2b, respectively. This suggests that Rwanda could be a special case.

In 2003, Rwanda implemented many reforms to reduce corruption in its entire judicial system (Bachors 2009; Rugege 2006). Rwanda's government introduced a new constitutional law with special emphasis on anti-corruption measures including severe punishments for bribe-givers (Rugege 2006). The successful application of such anti-corruption measures (Bachors 2009) may have considerably reduced corruption of many government institutions, including the police.

In addition, *Observation 7* may have also resulted from overfitting, i.e., when the decision tree perfectly fits the training data that classifies, at the expense of losing its capability to generalize (Rokach 2008). Another instance of overfitting presented in the decision tree is rule group 3c, which has two decisions nodes using *percentage of improved drinking water coverage in urban areas* with very similar values (≤ 0.98 and ≤ 0.99). Rule group 3c perfectly fits the training data, but it has lost its generalization capability. The overfitting can be prevented by pruning the trees' branches, which may increase its generalization capability, but it may decrease its classification accuracy (Rokach 2008). However, in this present work, the decision tree is not used as a classifier, but as a decision support tool, which can be used to discover features relevant to police corruption that can assist in the design of strategies to reduce police corruption. In this regard, the overfitting represented by the nodes involving *percentage of improved drinking water coverage in urban areas* should be discarded from the analysis or consideration.

5.2. Analysis of the Rule Groups

Given that police corruption is a multi-factorial social phenomenon (Carvajal 1999), there is not a one-size-fits-all strategy (McCusker 2006), and the best strategy to be adopted by a country depends on its current social and economic situation (McCusker 2006). However, general guiding principles to prevent police corruption can be identified. Then, we conducted an evaluation of the most relevant rule groups to determine what rules can be used to assist in the design of strategies for countries with

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

different PCSs.

It is acknowledged that the analysis is performed under the following two assumptions:

a) Given that there is no exemplar country with a PCS of 1, we assumed that the best rules to be followed are those provided by rule group 2, see Fig. 2. Then, the social indicators of all the countries (except for the exemplar countries with a PCS of 2) were evaluated against the best rules of rule group 2. In addition, the social indicators of the countries with a PCS of 4 were also evaluated against the best rules of rule group 3, and the countries with a PCS of 5 were also evaluated against the best rules of rule groups 3 and 4.

b) Since some branches of the decision tree may have resulted from exceptional cases, e.g., rule group 2c that only classifies Rwanda, we selected the most representative rule group of each PCS category. The most representative rule group for each category was determined by counting the number of countries that were classified using a given rule group. In doing so, we assumed that the more countries a rule group classifies into a PCS category, the more representative the rule group is for that particular PCS category. The most representative rule groups are 2a, 3a, and 4h for countries with PCSs of 2, 3, and 4, respectively. Rule groups 2a, 3a, and 4h classified 60%, 42.8%, and 52.8% of all the countries belonging to their corresponding category. The social indicators of rule groups 2a, 3a, and 4h (see Fig. 1) are as follows.

(1) Rule group 2a: (i) male school life expectancy and (ii) percentage of male own-account workers.

(2) Rule group 3a: (i) male school life expectancy, (ii) percentage of male own-account workers, (iii) infant mortality rate, and (iv) percentage of female employers.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

(3) Rule group 4h: (i) male school life expectancy, (ii) percentage of female own-account workers, (iii) annual population growth rate, (iv) percentage of male part-time employment, (v) girls' share of primary enrolment, (vi) percentage of male employers, and (vii) girls' share of secondary enrolment.

From the evaluation of the most relevant rule groups against the social indicators of all the countries with a PCS of 3, 4, and 5, we drew three observations.

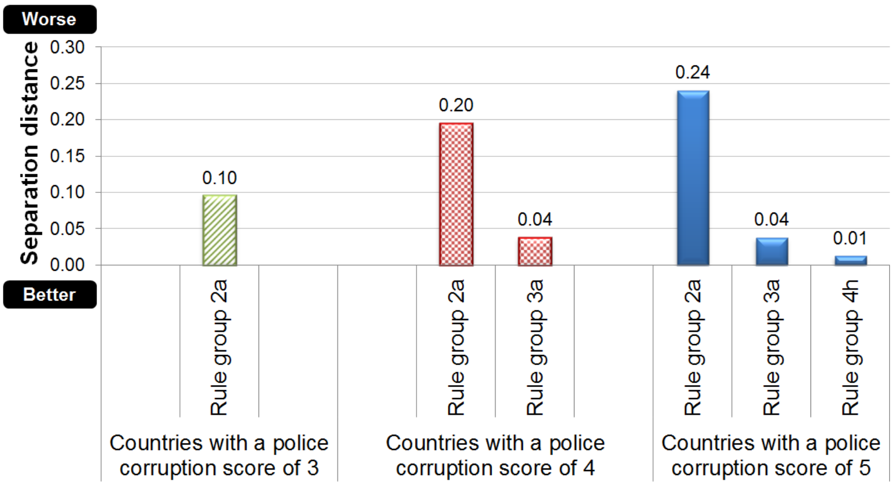

Fig. 3. Evaluation of rule groups based on the separation distance metric.

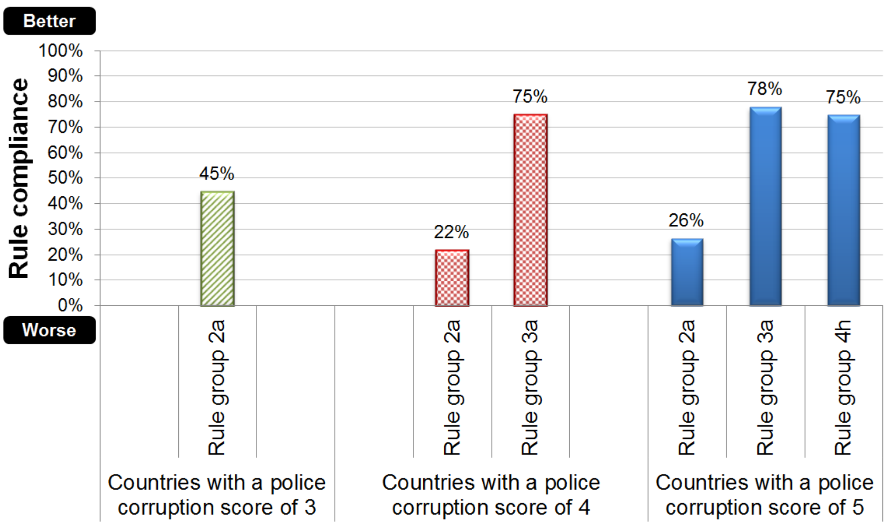

Fig. 4. Evaluation of rule groups based on the rule compliance metric.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

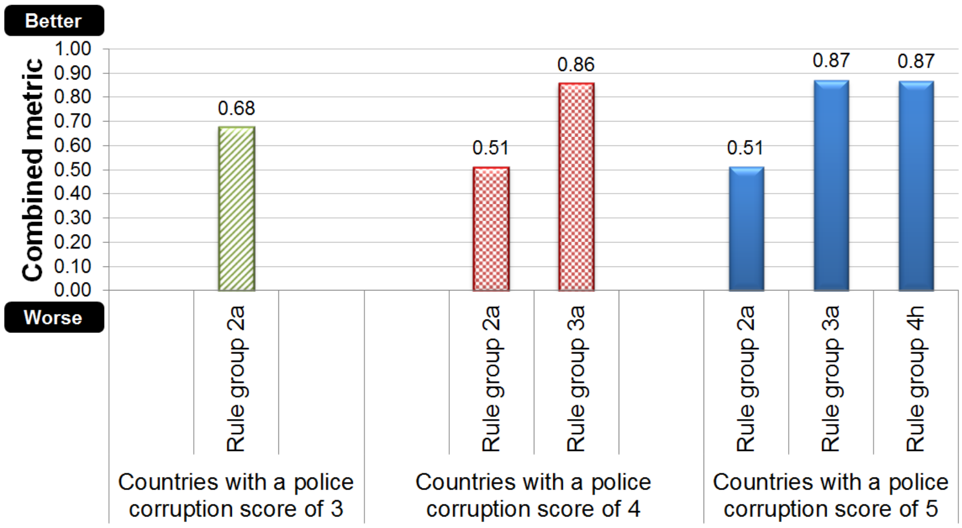

Fig. 5. Evaluation of rule groups based on the combined metric.

Observation 8. In general, the values of the social indicators of countries with a PCS of 4 are (i) closer to the values of the decision nodes of rule group 3a and (ii) farther from the values of the decision nodes of rule group 2a.

Analysis. As can be observed in Fig. 5, countries with a PCS of 4 are closer to be classified as countries with a PCS of 3 by using rule group 3a than they (i.e., countries with a PCS of 4) are to be classified as countries with a PCS of 2 by using rule group 2a. The separation distance (Fig. 3) of countries with a PCS of 4 from rule group 3a is relatively small, only 0.04, compared to the separation distance of 0.2 of rule group 2a. In fact, on average, countries with a PCS of 4 comply with 75% of the rules of rule group 3a, in contrast to complying with only an average of 22% of the rules of rule group 2a, see Fig. 4. This result suggests that the more suitable and reachable strategies to improve the PCS of countries with a PCS of 4 are those involving social indicators included in the rule group 3a rather than social indicators included in the rule group 2a. As stated by North (1993), the majority of changes in institutions (e.g., governments) is progressive and gradual. Then, adopting a reachable strategy that involves a gradual

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

change in the public policies of a country may be more adequate than adopting a strategy that involves abrupt, large changes.

In summary, this result suggests that countries with a PCS of 4 should adopt strategies supported by rule group 3a, and once they have improved their PCS of 4 to 3, the countries should adopt strategies supported by the decision nodes of rule group 2a.

Observation 9. In general, the values of the social indicators of countries with a PCS of 5 are (i) equally closer to the values of the decision nodes of rule groups 3a and 4h and (ii) farther from the values of the decision nodes of rule group 2a.

Analysis. As shown in Fig. 5, when the social indicators of countries with a PCS of 5 were evaluated against rule groups 3a and 4h, both their separation distance and rule compliance percentage were similar, see Figs. 3 and 4. These results suggest that the effort required from countries with a PCS of 5 to improve its PCS by adopting a strategy supported either by rule group 3a or 4h would be the same. As a consequence, countries with a PCS of 5 should adopt a strategy supported by the decision nodes of rule group 3a to improve their PCS from 5 to 3 directly. Then, to establish a progressive change (North 1993), once they have reached a PCS of 3, they can adopt a strategy supported by the decision nodes of rule group 2a.

Observation 10. The social indicators of countries with a PCS of 3 were the best evaluated against the rule group 2a.

Analysis. As shown in Fig. 5 and analyzed in *Observations 8* and *9*, in general, the closer the PCS of a country to the level of the rule group to which its social indicators are being evaluated, the better the value of the combined metric for the country. This result suggests that the metrics, namely rule compliance, separation distance and combined metric, appropriately measure the distances among the countries with different PCSs. In addition, this result (supported by *Observations 8* and *9*) also

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

suggests that rule groups 2a, 3a, and 4h effectively characterized the social indicators that correspond to the different PCSs. Moreover, from *Observations 8*, *9*, and *10*, we can conclude that rule groups 2a and 3a are gradual and reachable steps that countries should follow to improve their PCSs.

Nevertheless, it should be noted that rule groups must not be interpreted as explicit strategies to be followed. For instance, in order for a country to be classified as a country with a PCS of 3 using rule group 3a, the *infant mortality rate* should be greater than a normalized value of 0.2, which is a negative aspect. Rule group 3a classifies countries with a relatively high level of police corruption, and as consequence, it also includes negative aspects. For this reason, rule groups must not be used blindly without any interpretation and/or contextualization.

6. Conclusion

The novelty of this research effort lies in its pioneering effort to analyze social determinants of police corruption by creating country-level socio-economic profiles that were accurately classified according to their PCSs using a decision tree. The branches of the decision tree provide a set of guiding principles for assisting policymakers in designing societal level strategies to control police corruption by improving social and economic conditions. By adopting a societal perspective, this work can be used to complement anti-corruption strategies at the level of police departments (Bayley 2011; Prenzler 2001; Punch 2000).

The added value of this paper is twofold: (i) to identify cross-national social indicators that are relevant to police corruption, which give insights into the causes of police corruption at a societal level; (ii) to determine the most representative rule groups that characterize countries based on their PCSs, which provide guiding principles for

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

tackling police corruption according to the socio-economic situation and the current level of police corruption of a given country.

The significance of our work is that by analyzing the relationship between the social indicators and the level of police corruption of a country (see Section 5.1), we can conclude that:

- The most relevant social indicator to police corruption is *male school life expectancy*;
- Countries with a high *male school life expectancy* tend to have a low level of police corruption;
- In addition to having a well-educated population, in order for a country to have a low level of police corruption, it should also reduce its informal labor force;
- The active involvement of women in society as self-employed workers, employees, employers, and students is relevant to police corruption;
- Economic development and work-related social indicators are relevant to police corruption.

Moreover, from the evaluation and analysis of the rule groups extracted from the decision tree (see Section 5.2), we can conclude that countries should gradually improve social indicators with the objective of reducing police corruption.

We hope that our analysis of social determinants of police corruption will shed new light in designing and adopting strategies aimed at undertaking a reform of society in order to undertake a reform of the police.

It should be noted that more social indicators could be added to the present study to explore other potential determinants of police corruption. Nonetheless, we only included the minimum list of social indicators for monitoring social development of countries (proposed by the United Nations) to guarantee a baseline level of social indicators.

We acknowledge that the present work is limited by the incompleteness of the dataset of countries' social indicators provided by the United Nations Statistics Division (2014), e.g., only a few countries have information about the number of people per room. We prepared the data to deal with the missing information (see Section 3), however, complete information of all the social indicators for all the countries may detect other relevant features to police corruption. We also acknowledge that the PCSs of the countries were determined by surveying citizens, a valid mechanism for measuring police corruption (Ivković 2003). However, a highly accurate quantitative level of police corruption is extremely difficult to determine due to the absence of completely reliable data (Klockars 2000), which is not reported by citizens or recorded by police institutions. More accurate PCSs can be obtained by triangulating different data sources of police corruption (Ivković 2003), but unfortunately, as far as we know, there are no other data sources containing data on police corruption of the 103 countries analyzed in the present research effort.

Acknowledgements. This work has been supported by Asociación Mexicana de Cultura, A.C. and by the Mexican Council for Science and Technology (CONACYT) through grant number 216101.

References

Arrigo, B.A. and Claussen N. (2003). Police corruption and psychological testing: A strategy for preemployment screening. *International Journal of Offender Therapy and Comparative Criminology*, 47(3), 272-290.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Bachors B. (2009). *Rwanda gets tough on corruption*. The Telegraph. Retrieved from http://www.telegraph.co.uk/expat/expatnews/6752425/Rwanda-gets-tough-on-corruption.html.

Bardhan P. (1997). Corruption and development: a review of issues. *Journal of Economic Literature*, 1320-1346.

Bashkirov A.G. and Vityazev A.V. (2000). Information entropy and power-law distributions for chaotic systems. *Physica A: Statistical Mechanics and its Applications* 277(1), 136-145.

Bayley, D.H. and Perito, R. (2011). *Police corruption: What past scandals teach about current challenges*. US Institute of Peace.

Botero, J., Ponce, A., and Shleifer, A. (2012). *Education and the quality of government* (No. w18119). National Bureau of Economic Research.

Brown, B. and Benedict, W.R. (2002). Perceptions of the police: Past findings, methodological issues, conceptual issues and policy implications. *Policing: An International Journal of Police Strategies & Management*, 25(3), 543-580.

Brown, B., Benedict, W.R., and Wilkinson, W.V. (2006). Public perceptions of the police in Mexico: a case study. *Policing: An International Journal of Police Strategies & Management*, 29(1), 158-175.

Bruce, D. (2008). Benign neglect? The politics and practice of controlling police corruption in South Africa. In *Crime, Poverty and Police Corruption Annual Bank Conference on Development Economics*, 1-21.

Carvajal, R. (1999). Large-scale corruption: definition, causes, and cures. *Systemic Practice and Action Research*, 12(4), 335-353.

Dantzker, M.L. (2011). Psychological preemployment screening for police candidates: Seeking consistency if not standardization. *Professional Psychology: Research and Practice*, 42(3), 276

De Graaf, G. (2007). Causes of corruption: Towards a contextual theory of corruption. *Public Administration Quarterly*, 31(1), 39-86.

Dimant, E. (2013). The nature of corruption: An interdisciplinary perspective. *Economics Discussion Papers* No. 2013-59, 1-57.

Di Puppo, L. (2010). Police reform in Georgia: Cracks in an anti-corruption success story. *U4 Practice Insight*, 2010(2).

Dreher, A. and Herzfeld, T. (2008). The economic costs of corruption: a survey of the empirical evidence. In F. Columbus (Ed.), *Economic Corruption and its Impact*. New York: Nova Science.

Dutta, N., Kar, S., and Roy, S. (2011). Informal sector and corruption: An empirical investigation for India. *IZA Discussion Paper* No. 5579, 1-35.

Gingerich, T.E., Chu, D.C., and Chang, C.K. (2011). Public perceptions of police performance in Taiwan: A historical review (1945–2009). *Asian Politics & Policy*, 3(1), 49-68.

Girodo, M. (1998). Undercover probes of police corruption: Risk factors in proactive internal affairs investigations. *Behavioral Sciences & the Law*, 16(4), 479-496.

Hall M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I.H. (2009). The WEKA data mining software: An update. *SIGKDD Explorations*, 11(1), 10-18.

Husted, B.W. and Instituto Tecnológico y de Estudios Superiores de Monterrey (1999). Wealth, culture, and corruption. *Journal of International Business Studies*, 30(2), 339-359.

International Association of Chiefs of Police (2014). *Model policy on standards of conduct*. Retrieved from http://www.theiacp.org/Model-Policy-on-Standards-of-Conduct.

Ivković, S.K. (2003). To serve and collect: Measuring police corruption. *Journal of Criminal Law and Criminology*, 93(2), 593-650.

Jain, A.K. (2001). Corruption: a review. *Journal of economic surveys*, 15(1), 71-121.

Johnson, R.A. and Sharma, S. (2004). About corruption. In Johnson, R.A. (Ed.), *The struggle against corruption: A comparative study* (1st ed., pp. 1-19). Hampshire, England: Palgrave Macmillan.

Kies, M.F. (2011). Policing the police: Freedom of the press, the right to privacy, and civilian recordings of police activity. *The George Washington Law Review*, 80(1), 274-310.

Klaver, J.J. (2013). Law enforcement ethics and misconduct. In Fitch, B.D. (Ed.), *Law enforcement ethics: Classic and contemporary issues* (1st ed., pp. 3-28). Thousand Oaks, California: SAGE Publications, Inc.

Klockars, C.B., Ivkovich, S.K., Harver, W.E., and Haberfeld, M.R. (2000). *The measurement of police integrity*. Washington D.C.: National Institute of Justice.

The Version of Record of this manuscript has been published and is available in Policy Studies (ISSN 0144-2872 (Print), 1470-1006 (Online)) 12-February-2016 http://dx.doi.org/10.1080/01442872.2016.1144735

Lamboo, T., Lasthuizen, K., and Huberts, L.W. (2008). 10. How to encourage ethical behavior: the impact of police leadership on police officers taking gratuities. In Leo, W.J.C., Maesschalck, J., Jurkiewicz, L.C. (Eds.), *Ethics and integrity of governance: Perspectives across frontiers* (1st ed., pp. 159-180). Cheltenham, UK: Edward Elgar Publishing Limited.

Macintyre, S., Prenzler, T., and Chapman, J. (2008). Early intervention to reduce complaints: an Australian Victoria police initiative. *International Journal of Police Science and Management*, 10(2), 238-250.

McCusker, R. (2006). Review of anti-corruption strategies. Technical and background paper, No. 23, *Australian Institute of Criminology*, Canberra, Australia.

Newburn, T. (1999). Understanding and preventing police corruption: lessons from the literature (No. 110). In Webb B. (Ed.), *Home Office, Policing and Reducing Crime Unit, Research, Development and Statistics Directorate*, 1-56.

Newham, G. (2000). Towards understanding and combating police corruption. *Crime and Conflict*, 19, 21-25.

Newham, G. (2002). Tackling Police Corruption in South Africa. *Africa do Sul: Centre for the Study of Violence and reconciliation*.

North, D.C. (1993). Five propositions about institutional change. Economic History, EconWPA.

Polinsky, A.M., and Shavell, S. (2001). Corruption and optimal law enforcement. *Journal of Public Economics*, 81(1), 1-24.

Prenzler, T., and Ronken, C. (2001). Police integrity testing in Australia. *Criminology and Criminal Justice*, 1(3), 319-342.

Prenzler, T., and Ronken, C. (2003). A survey of innovations in the development and maintenance of ethical standards by Australian police departments. *Police Practice and Research*, 4(2), 149-161.

Prenzler, T. (2009). *Police corruption: Preventing misconduct and maintaining integrity*. Boca Raton, FL: CRC Press.

Punch, M. (1994). Rotten barrels: systemic origins of corruption. In Kolthoff E.W. (Ed.) *Strategieen voor corruptie-beheersing bij de politie* (1st ed.), Gouda Quint: Arnhem.

Punch, M. (2000). Police corruption and its prevention. *European Journal on Criminal Policy and Research*, 8(3), 301-324.

Quinlan, J.R. (1993). *C4.5: Programs for machine learning*. San Francisco, CA: Morgan Kaufmann Publishers Inc.

Rokach, L. and Maimon O. (2008). *Data mining with decision trees: theory and applications*. Danvers, MA: World Scientific.

Rose-Ackerman, S. (2007). *International handbook on the economics of corruption* (1st ed.) Cheltenham, UK: Edward Elgar Publishing Limited.

Rugege, S. (2006). Judicial independence in Rwanda. *Global Business and Development Law Journal*, 19, 411-425.

Saar-Tsechansky M. and Provost F. (2007). Handling missing values when applying classification models. *Journal of Machine Learning Research*, 8(7), 1625-1657.

Saunders, R. (2003). *Defining vulnerability in the labour market*. Ottawa: Canadian Policy Research Networks.

Savage, S.L. (2003). *Decision Making with Insight*. Cole Belmont, CA: Thomson Brooks.

Shannon C.E. and Weaver W. (1946). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Sherman, L.W. (1978a). *Controlling Police Corruption: The Effects of Reform Policies - Summary Report*. National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, US Department of Justice.

Sherman, L.W. (1978b). *Scandal and reform: Controlling police corruption*. Berkeley, CA: University of California Press.

Small, M.W. and Dickie, L. (2003). Conjoining ethical theory and practice: An Australian study of business, accounting, and police service organizations. *Teaching Business Ethics*, 7(4), 379-393.

Stevenson, J. C., Mahoney, E. R., Walker, P. L., and Everson, P. M. (2009). Technical note: prediction of sex based on five skull traits using decision analysis (CHAID). *American Journal of Physical Anthropology*, 139(3), 434-441.

Swamy, A., Knack, S., Lee, Y., and Azfar, O. (2001). Gender and corruption. *Journal of Development Economics*, 64(1), 25-55.

Transparency International (2014). *Global corruption barometer 2013*. Retrieved from http://www.transparency.org/gcb2013/in_detail.

Tonoyan, V. (2004). The bright and dark sides of trust: Corruption and entrepreneurship. In Hönmann H.-H. and Welter F. (Eds.) *Trust and*

Entrepreneurship: A West-East-Perspective (1st ed., pp. 1-25), Cheltenham, UK: Edward Elgar Publishing Limited.

United Nations Statistics Division (2014). *Social indicators*. Retrieved from http://unstats.un.org/unsd/demographic/products/socind/.

United Nations working group on international statistical programmes and coordination (2014). *Report of the expert group on the statistical implications of recent major United Nations conferences*. Retrieved from http://unstats.un.org/unsd/demographic/products/socind/xgrp2.htm.

Weaver W., (1949) Some recent contributions to the mathematical theory of communication. In Shannon C.E. and Weaver W., *The Mathematical Theory of Communication* (1st ed., p. 1-28), Urbana, IL: University of Illinois Press.

Weitzer, R. and Tuch, S.A. (2005). Determinants of public satisfaction with the police. *Police Quarterly*, 8(3), 279-297.

Welsh, B. C., Rocque, M., and Greenwood, P.W. (2014). Translating research into evidence-based practice in juvenile justice: brand-name programs, meta-analysis, and key issues. *Journal of Experimental Criminology*, 10(2), 207-225.

Williams, H. (2002). Core factors of police corruption across the world. In *Forum on Crime and Society*, 2(1), 85-99, UCL Press.

Wu, Y., Lambert, E.G., Smith, B.W., Pasupuleti, S., Jaishankar, K., and Bhimarasetty, J.V. (2012). An exploratory comparison of policing views between Indian and US College Students. *International Criminal Justice Review*, 22(1), 68-82.

Zurada, J., and Lonial, S. (2011). Comparison of the performance of several data mining methods for bad debt recovery in the healthcare industry. *Journal of Applied Business Research*, 21(2), 37-54.