ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

DOES OIL CORRUPT? EVIDENCE FROM A NATURAL EXPERIEMNT IN WEST AFRICA

Pedro C. Vicente

Number 317

April 2007

Manor Road Building, Oxford OX1 3UQ

Does Oil Corrupt?
Evidence from a Natural Experiment in West Africa*

Pedro C. Vicente[1]

First Draft: November 2004; This Draft: November 2006

Abstract

This paper explores an oil discovery natural experiment to assess the role of natural resources in determining corruption. We argue that an anticipated oil boom may increase corruption by boosting the value attributed by an elite to being in power when the actual oil exploration begins. We test this proposition by analyzing the impact of the oil discovery announcements that took place in 1997-99 in São Tomé and Príncipe (West Africa). For this objective we conducted purposely-designed household surveys on perceived corruption in the public services/sector. These were carried out in São Tomé and Príncipe and in Cape Verde, a control West African country sharing strong cultural ties and important contemporary economic/political shocks. The unique survey instrument was retrospective and used personal histories to elicit memories from the respondents. Urban subjects, public officials, and respondents with higher reported experience with the services/issues at stake are used as internal treatment groups. Comparisons are also made with corresponding groups in Cape Verde. In addition, the regressions control for well-known 'good old times' bias: this is done by using data from direct questions on optimism and from the inclusion of a 'placebo' period (when no major occurrence had arisen). We conclude that a clear increase in perceived corruption has occurred in São Tomé and Príncipe in recent years, ranging from 21 to 38% of the subjective scale. Consistently with our theoretical mechanism, which underlines the importance of being in power when the oil boom occurs, these effects are most robust in vote buying, education, and state jobs.

JEL Codes: D73, O13, O55, P16.
Keywords: Corruption, Influence, Political Economy, Natural Resources, Oil, West Africa.

---

* I wish to thank primarily Cátia Batista, Marianne Bertrand, Casey Mulligan, and Roger Myerson for detailed comments and encouragement. I am also grateful for suggestions to Abigail Barr, Marcel Fafchamps, Nicola Gennaioli, Steven Levitt, Kenneth Rasinski, Allen Sanderson, Gerhard Seibert, Susan Stokes, Robert Townsend, and Christopher Udry. I thank participants in seminars at Stockholm School of Economics, Cornell, Oxford, CSAE Conference, GPRG Conference on Corruption, ISEG-UTL, EEA Congress, Econometric Society World Congress, CESifo Summer Institute, Yale, Navarra, Aarhus, UQAM, and especially Chicago for comments. I am grateful for institutional support to Lúcio Pinto at Instituto Superior Politécnico of São Tomé e Príncipe and Deolinda Reis at the National Statistics Office of Cape Verde, to my co-interviewers in the surveys conducted, C. Castro, A. Coelho, E. Lima, A. Pinto, H. Santos, D. Solé, E. Solé, D. Vicente (São Tomé and Príncipe), and C. Andrade, E. Fernandes, A. Gonçalves, V. Lenine, H. Mendes, R. Moniz, L. Pina, A. Semedo, N. Varela (Cape Verde), and to the 1907 respondents who answered our questions with generosity. I gratefully acknowledge financial support from Fundação para a Ciência e a Tecnologia (BD1215/2000), Henry Morgenthau Jr. Memorial Fund, Department of Economics at the University of Chicago, and ESRC-funded Global Poverty Research Group. All remaining errors are my responsibility alone.

[1] Email: pedro.vicente@economics.ox.ac.uk. Post: Department of Economics/CSAE, University of Oxford, Manor Road, Oxford, OX1 3UQ, UK.

Contents

"*It's because of oil that they want to take over power.*" - Fradique de Menezes, President of the Democratic Republic of São Tomé and Príncipe, after July 2003 coup attempt
"*The international community should investigate the activities of successive governments in São Tomé.*" - Fernando Pereira "Cobo", July 2003 coup attempt leader

1 Introduction

Natural resources are a key topic in the development literature: this perspective has been grounded on the widespread idea that they represent a "curse" in developing countries. Indeed, anecdotal evidence has supported the association of resource abundance with lack of transparency and rent-seeking economies. From Gabon, Nigeria, Angola, Congo throughout the last decades, to the new oil-producing Caucasian countries (Azerbaijan, Kazakhstan) or even the extreme case of Equatorial Guinea in the 90s, a clear pattern seems to arise: these are among the most corrupt countries in the world[2].

Simultaneously, corruption has been approached as one of the main direct impediments to economic growth, with dramatic consequences in the poorest areas of the globe - see for instance Mauro (1995), Gupta, Davoodi, and Alonso-Terme (1998), Kaufmann and Kraay (2002). A better understanding of the relation between natural resources and corruption may represent an important leap forward in the opening of the 'resource curse' black box. This paper is an attempt at deepening that knowledge, specifically by addressing the research questions: is there a causal relationship from natural resource abundance to corruption? And how does this relation begin – i.e. which are the sectors and allocations where corruption starts appearing?

To answer these questions we analyze the case of São Tomé and Príncipe, a two-island country, situated in the Gulf of Guinea, with three well-known resource-cursed neighbors (Gabon, Nigeria, and Equatorial Guinea). We argue that a natural experiment occurred in this country stemming from a simple exogenous event: the announcements of a significant oil discovery in the period 1997-1999.

We argue, using a simple rent-seeking adapted model that, through competition within the elite, a higher rent (from the new oil sector) increases corruption efforts. If that higher rent is anticipated, as happened in São Tomé and Príncipe, we should see increased competition for an elite seat, which translates into increased corruption efforts. This is the theoretical proposition we take to the data in this paper.

[2] As given by cross-country indices like Transparency International's Corruption Perceptions Index (http://www.transparency.org/) or World Bank's Governance Indicators (www.worldbank.org/wbi/governance).

Our measurement is the result of fully and purposely-designed data collection efforts. These took place not only in São Tomé and Príncipe but also in Cape Verde, a control West African country. Cape Verde shared a Portuguese colonial past with São Tomé and Príncipe, which included an important migrant contingent during the second half of the 20th century, meaning strong cultural links between the two countries. In addition, after independence granted to both countries in 1975, these nations had a remarkably similar political and institutional timeline: their first regime was socialist-like, in 1991 the first free elections took place in both countries with both incumbents leaving power, aid levels and timing were comparable. We argue this comparison will allow us to control for the most important macro-level shock occurred closest to the oil discovery: democratization/aid-induced economic reform.

Our workhorse method was composed by representative household surveys on perceived corruption, including 841 interviews in São Tomé and Príncipe, and 1066 interviews in Cape Verde. Perceived corruption questions were asked regarding Health Care, Education (in general, and concerning the allocation of Scholarships[3]), Courts, Police, State Jobs, Customs, Subsidies / Supplier Positions for the State, Public Infrastructures, Licenses, Vote Buying. These questions were centered in exchange of favors or personal influence[4].

A unique retrospective survey instrument was used following two basic principles: eliciting memories from respondents by making use of their personal histories (in the perceived corruption questions, information from factual questions in the beginning of the questionnaire was used to refer to the different periods of interest - before and after the oil discovery announcements); asking specific questions on corruption only to respondents showing experience with the specific service/allocation. In addition, controls for a "Good Old Times" bias were included: direct questions on assessing optimism of respondents and a Placebo period introduced in a question on perceived changes in corruption (where no change was supposed to be perceived provided the lack of significant modifications in the macro characterization of the countries). Note that the fact that the same instrument, with corresponding techniques, was used in both countries may also help controlling for the retrospective nature of the design.

In terms of econometric strategy, we measure the effect of the oil discovery announcements on corruption by estimating (using difference-in-differences regressions): the change perceived by highly exposed respondents in excess of the other subjects' perceived changes (internal control group); the change perceived in São Tomé and Príncipe in excess of the Cape Verdean survey

[3] Note that since STP does not have any university, for higher education, São-Tomeans have to apply to study abroad. Provided the fact that most of the people cannot afford that kind of possibility, the average citizen sees scholarships as central for a good future. The higher education institutional context in Cape Verde is very similar.
[4] We found empirically that bribery has been seen as less a serious problem in these countries than personal influence.

respondents' perceived changes (external control group). Internal comparisons can be interpreted in the view of the model: we take those connected agents as having better information on the allocation of the "elite seats". These agents are mainly taken to be empirically the competitors/auctioneers in the allocations we asked about in the surveys. As empirical proxies for these connected groups we use urban and public administration respondents; we also use the information on the individual experience with the services/allocations to compose a "Connectivity" index.

We conclude that a clear increase in perceived corruption arose after the period 1997-1999. This change was highest (and most significant) for Vote Buying (as perceived in all treated/control comparisons), Education (as perceived in all treated/control comparisons, with special prominence in the Scholarships variable), State Jobs (specially when considered by the connected respondents), and Health Care (especially when assessed by the urban and public administration groups). Note that, as the model suggests, most of these sectors/allocations embed a clear notion of competition for being in power when oil production begins.

Important cross-country empirical work on the effects of natural resources has already been performed. The influential work by Sachs and Warner (1995) opened the growth literature on natural resources by noticing that natural resource abundant economies (as measured by natural resource exports to GDP) have tended to grow slower than economies without substantial resources[5]. Later the study by Mehlum, Moene, and Torvik (2006) - MMT, claimed the main reason for diverging experiences by resource-rich countries (in terms of growth) is differences in the quality of institutions. This is the general growth "resource curse" picture that we take as given in our work.

Theoretically, different versions of the MMT empirical result were established: Baland and Francois (2000) find that, depending on initial conditions, resource booms bias gains in favor of rent-seeking over entrepreneurship; Robinson, Torvik, and Verdier (2006) argue that, under weak institutions, natural resource booms increase resource misallocation by raising the value of having the political power and by providing politicians with more resources which they can use to influence the outcome of elections.[6] In fact, one of our main findings, regarding a clear increase in vote buying, fits closely this proposed channel.

More closely to our paper, the cross-country work of Ades and Di Tella (1999) and Leite and Weidmann (1999) shows explicitly that natural resource abundance is an important factor in determining corruption. While the first argues that this is due to lack of competition in product

[5] Other interesting associations with the presence of natural resources were explored by Bannon and Collier (2003), and Lam and Wantchekon (2002), with the analysis of civil wars and dictatorships, respectively.
[6] Other theoretical foundations for the resource curse are presented by Tornell and Lane (1999), and Torvik (2002).

market, the second emphasizes the link to the growth debate. Our empirical treatment reaches compatible results by analyzing a natural experiment, suited to assess causality, using microdata, therefore not depending on the shortcomings of aggregate data and cross-country regressions. We hope that this methodology, while being complementary, may also shed light on specific mechanisms/sectors by which the corruption resource curse may happen.

Our work also links with the disperse but growing literature on microeconomic measurement of corruption. Duggan and Levitt (2000) pioneered this literature by providing evidence of match rigging in Japanese sumo wrestling. Naturally-occurring experimental settings are explored by Fisman (2001) and Fisman and Miguel (2006): the first aims at measuring the extent of corrupted political links in Indonesia by using news on the health of former dictator Suharto to explore the evolution of local stock prices; the second presents evidence on the effect of culture on corrupt behavior by using evidence on diplomatic (not-sanctioned) parking tickets in New York. Explicit experimental designs are used recently in Olken (2005), who analyzes the effect of monitoring on corruption in road construction in Indonesia, and in Bertrand et al (2006), who study efficiency of corruption from individual and social perspectives in the allocation of driving licenses in India. Our paper, while addressing a fundamental development question, contributes to this micro-measurement literature by using a clear natural experiment design, and by adding to that natural framework explicit experimental components in data gathering (through tailored household surveys on corruption).

In section 2 we present a simple theoretical framework to analyze an anticipated oil shock. Section 3 provides an historical overview on the natural experiment we explore in São Tomé and Príncipe and Cape Verde. Following that section, we describe the basic ingredients of our experimental design, which precedes a representation of the data collection methods in section 5. In section 6 we explain the estimation approach, and in section 7 we offer the econometric results. Section 8 concludes.

2 Theoretical Mechanism

Corruption, generally defined as "abuse of public office for private gain" [7], has commonly been seen as a form of rent-seeking. As originally phrased by Krueger (1974), this concept corresponds to "resource wasting activities of individuals in seeking transfers of wealth through the aegis of the state" (in the words of Buchanan, Tollison and Tullock, 1980). Clearly, these are

[7] Bardhan (1997) presents a comprehensive look at the economic literature on corruption, including its usual definition. See Becker and Stigler (1974), Rose-Ackerman (1978), Cadot (1987), Klitgaard (1988), Myerson (1993) for early approaches to the economics of corruption.

linked and similar concepts, both related with activities undertaken for private gain in the context of state intervention. Nevertheless, relative to corruption, rent-seeking focuses on waste and relaxes on illegality (as embedded in the word 'abuse').

Throughout this paper we focus on measuring corruption. That means we are not specifically involved in assessing efficiency, but are extremely interested in measuring the importance/frequency of exchange of favors or influence, and of bribery in the public sector, two prominent forms of the above corruption definition. We do not take strictly the illegality dimension of the classic notion[8].

As a proper illustration of the mechanisms at stake in the empirical part of this paper, we propose a version of the seminal rent-seeking model by Tullock (1980). Since in his theoretical treatment inefficiency considerations associated with the concept of rent-seeking are a matter of interpretation, we are able to connect the concept of corruption with that of rent-seeking in this framework. Note also that this rent-seeking model has implicit an allocation game, essential in characterizing corruption[9].

We therefore consider a two-period game, with $m \times n$ players (with both $m, n > 1$). In the first period each player competes in a group of *m* players for a seat in the elite of the country[10]. In the second date the members of the elite, assumed to be *n*, compete for the public sector rent *R*.

The crucial assumption is that at each stage players may undertake corrupt activities that are costly. These costs may be interpreted in terms of time (e.g. exchange of favors), risks (e.g. legal punishment, social unrest), or money (e.g. bribery). Corruption is assumed to increase chances of winning the elite's seat or the public rent at each period. For player *i*, we denote the corrupt activities (and their costs) as x_i^k, where $k = 1,2$ is the time period. The probability of being successful at both periods for player *i* is postulated to be

$$p_i\,(x_1^k,...,x_l^k) = \frac{x_i^k}{\sum_{j=1}^{l} x_j^k},$$

with $l = m, n$ for periods 1 and 2 respectively.

Solving this simple game by backward induction, at period 2 each elite player maximizes

$$\max_{x_i^2} p_i(x_1^2,...,x_m^2)R - x_i^2 .$$

[8] Kaufmann and Vicente (2005) explore the cross-country difference between legal and illegal forms of corruption.
[9] Corruption-like notions in the context of allocation games have been explored by other authors: Bernheim and Whinston (1986) introduce influence in the context of first-price auctions; Banerjee (1997) analyzes mechanisms with red tape and asymmetric information.
[10] Note that this is a very simplified form of allocation of political power – see Vicente (2006) for a general theory with a specifically modeled corruption definition, explicit hierarchy of political power, and the derivation of effects of windfalls.

The unique pure strategies Nash Equilibrium, which is symmetric, of this sub-game is:

$$x^{2*} = \frac{(m-1)R}{m^2}. \quad (1)$$

In the first period, each agent in the economy will solve the problem

$$\max_{x_i^1} p_i(x_1^1,...,x_m^1)E - x_i^1,$$

where $E \equiv p_i(x^{2*},...,x^{2*})R - x^{2*}$ is an elite member expected rent.

The unique pure strategies sub-game perfect equilibrium is embedded in

$$x^{1*} = \frac{(n-1)E}{n^2}. \quad (2)$$

As in the classic model, under competition (here, $m, n \to \infty$), the total value of corruption costs will be approaching the total rent at stake in the model[11].

We now consider which would be the effects of an increase in *R* coming from a new oil sector under the control of the state. As can easily be seen from (1) and (2), corruption levels at both periods 1 and 2 increase as a result.

Corruption will increase at date 2 because competition drives up the amount bidden by elite members for the state rent:

$$\frac{\partial x^{2*}}{\partial R} = \frac{m-1}{m^2} > 0$$

If the oil announcement happens in this period only, this change in corruption may be thought of as the unanticipated effect of an oil shock, where an elite fights for newly arrived resources.

At date 1, we also predict an increase in the level of corruption. This is due to competition for the increased-value elite seat. This comes from:

$$\frac{\partial E}{\partial R} = \frac{1}{m^2} > 0, \quad \frac{\partial x^{1*}}{\partial R} = \frac{n-1}{n^2 m^2} > 0$$

This would add to the date 2 increase in composing the effect of an anticipated oil shock. This is, we argue, what we face in São Tomé and Príncipe – an oil shock announcement has

[11] Note that the stage model presented corresponds to a specific version from Tullock's analysis. Namely, we consider rent-seeking sensitivity as measured by the parameter *q* in the probability function

$$p_i(x_1^k,...,x_l^k) = \frac{(x_i^k)^q}{\sum_{j=1}^{l}(x_j^k)^q},$$

to be equal to 1. This rules out the extreme cases of no rent-seeking ($q = 0$) or of no pure strategies equilibrium (that can arise with $q > 1$). The low rent-seeking result may be recovered by considering open-ended repetitions of the static game (Leininger and Yang, 1994). These variations, although providing a rationale for realistic levels of rent-seeking, are mostly neutral to our variation of interest: an increase in the value of the rent).

happened but most related increases in the size of the public sector rent are still to come. Therefore, the increase in corruption at date 1 is what we attempt to measure in the remaining of the paper.

In the next sections we look at an empirical counterpart of the natural resource boom we have just analyzed. We begin with some historical context for the series of announcements regarding the existence of vast oil reserves made in the period 1997-1999 in São Tomé and Príncipe.

3 Historical Background

São Tomé and Príncipe (STP), the second smallest country in Sub-Saharan Africa, traditionally a cocoa producer (80% of exports[12]), and an aid-dependent country (aid accounted for approximately 49% of the state revenues in 2003[13]), is composed of two main islands in West Africa (Gulf of Guinea) and has 148 thousand inhabitants[14]. A Portuguese colony for close to 500 years, STP gained independence in 1975. Following the common trend of the decolonization process in Lusophone Africa, the first political regime after independence followed the soviet socialist model. Like in many other countries, in 1989 the democratization process was announced, leading to the first free elections in 1991[15]. Aid per capita rose sharply in the democratization period as can be seen below in Chart 1 - this was accompanied by IMF and World Bank-sponsored economic reforms (macroeconomic stabilization/price liberalization, privatization plans - namely with respect to land ownership). Despite increasing aid and stable cocoa prices (see Chart 2[16]), STP still ranked 115 out of 136 countries in terms of GDP per capita PPP in 2000[17], after a stagnating decade (GDP per capita growth rates averaged -0.77% in the period 1989-2000) and no major changes in sectoral labor activity[18].

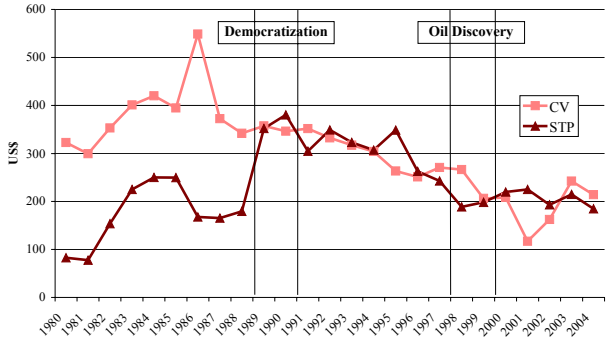
The events that constitute our focus in this paper occurred late in the nineties. During the period 1997-1999, a series of announcements regarding the existence of offshore oil took place in STP. For all main details of the oil chronology, see Appendix A – its key ingredients are Nigeria, with whom a joint exploration was agreed, and international oil companies such as Chevron and

[12] CIA World Factbook.
[13] IMF São Tomé and Príncipe Country Report 04-107.
[14] This is for 2000 (World Development Indicators, 2002).
[15] Seibert (1999), who presents a comprehensive study on the history of STP, underlines that the democracy path was the result of the recognition that the support of western countries and institutions would mean better offers of aid, in a context of a decrease in world prices of cocoa.
[16] Note that the cocoa price is expressed in constant dollars - this is argued to be a good measure of the impact of cocoa price in STP provided tradable goods represent a high proportion of the expenditure of the economy (imports of goods and services as a percentage of GDP averaged 86% in the period 1989/2000 - World Development Indicators).
[17] Penn World Tables 2002.
[18] World Development Indicators, World Bank.

ExxonMobil, who added credibility to the findings (e.g. in 2003 bidding for a first set of joint blocks amounted to 237% of STP GDP). We argue these facts created the exogenous variation in our natural experiment[19].

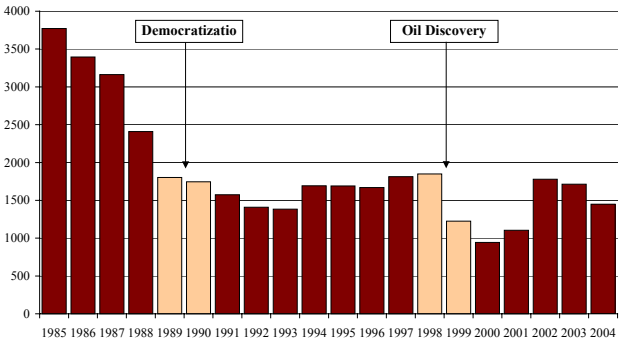
In this context, we defend that the West African Cape Verde (CV) may serve as a suitable control country for this experiment. CV is a nine-island country with 441 thousand inhabitants[20], which became independent from Portugal in 1975 after a similar (to STP) long period of colonization by that European country. As STP, CV faced a socialist regime until 1991, when free elections took place[21]. This democratization meant in CV that similar (to STP) aid inflows happened and similar reforms were sought, namely those sponsored by international agencies. It is worth mentioning that election cycles and changes of parties in power were also very close in both countries[22]. A note is also due to the strong cultural links these two countries share: important waves of emigration from land-impoverished CV to fertile STP took place in the twentieth century, the last of which in its second half. Carreira (1982) reports that in the context of the colonial strategy almost 35 thousand Cape Verdeans left to STP in the period 1950-1970[23]. Crucially, oil was never found in CV to date and is said to be very unlikely to exist or to have viable exploration in its territory[24].

Chart 1: Net ODA Grants per capita in STP and CV (2003 prices, US$)



Source: Development Assistance Committee, OECD, 2005; World Development Indicators, World Bank, 2005.

Chart 2: World Cocoa Annual Average Price (2002 US$ per Tonne)



Source: IMF Primary Commodity Prices; CPI used to deflate from US Department of Labor.

[19] In Shadish, Cook, and Campbell (2002), a natural experiment is defined to stem from a "naturally occurring event, one that cannot be manipulated".
[20] This is for 2000 (World Development Indicators, 2002).
[21] Note that these elections happened in the same month as the first STP elections – interestingly and against the African norm, both incumbents were ousted in these elections.
[22] Elections in STP: 1991 (Presidential and Parliamentary); 1994 (Parliamentary), 1996 (Presidential), 1998 (Parliamentary), 2001 (Presidential), 2002 (Parliamentary). Elections in CV: 1991 (Presidential and Parliamentary); 1995 (Parliamentary), 1996 (Presidential), 2001 (Presidential and Parliamentary).
[23] While many Cape Verdeans are accounted to have returned, many stayed after independence and are regarded as the main constituent of the so-called "contratados" STP ethnic group.
[24] See Prime-Minister Neves side-declarations to Radio Comercial of CV in June 2003.

These historical and cultural similarities with STP make CV the best possible comparison country for a late nineties shock in STP[25]. We hope, specifically, that by using CV, we are able to control for any persistent effects of the democratization/reform process of late 80s, early 90s (arguably, the main post-independence shock in STP) - see Chart 1 for the evolution of aid in CV, alike and contemporaneous to the one in STP from 1989.

Finally, we motivate the importance of the STP oil discovery and its consequences by referring to some direct questions asked in our STP survey (at the end of the interviews). Namely, the population was clearly convinced in 2004 that significant oil revenues were to reach the country (more than 50% of the sample is expecting "high" revenues, as opposed to "medium" or "low"). In addition, agents competing for political power are clearly seen (in public opinion) as more "anxious" to grab power after the oil discovery announcements (more than 80% stated "more anxious" as opposed to equally or less), and the competitiveness for public sector jobs before-after the oil discovery in STP is seen to have increased considerably more in STP (26%) than in CV (9%). These general preliminary findings as well as reports of lack of transparency in early oil-related contracts signed by the STP government - see Frynas et al (2003) for a detailed description – point in the direction of our model. We now describe our experiment in detail.

4 Basic Experimental Design

We base the answer to our empirical questions - on the effect of the oil discovery announcements on corruption in STP - on comparing changes in the corruption perceived by the population before and after the announcements.

For this purpose, we designed and conducted household surveys on corruption where we asked questions on perceived corruption in a broad range of public services/allocations.

The way we derive the effect of the oil discovery announcements is by comparing responses of subjects close to the allocation processes with answers of respondents without that close contact (internally), and responses of Sao-Tomean residents with those of subjects living in CV (externally). As proxies for informed respondents (internally) we take urban, public administration, and experienced-with-allocations (as given by service-specific information gathered in the survey) subjects. Accordingly, the estimation method is based on difference-in-differences regressions.

[25] Nevertheless, CV was already, in the end of the nineties, a more successful economy. CV combined aid and remittances to achieve comfortable growth rates in the end of the 80s and 90s. This is a clear difference between the two countries that makes us aware that a comparison with STP is not free from imperfections.

Importantly, as our data will be based on one-time surveys, the survey instrument design was specifically oriented to elicit earlier memories accurately - using information gathered on personal milestones in the beginning of the interviews, before questions on corruption. In addition, we control for the effect of the well-known "Good Old Times" bias, which makes people systematically negative about their views regarding the passage of time. For a measure of this variable to be available, we include several questions in the survey and a fictitious period of interest (which we call placebo period). This is in addition to the use of the same instrument and techniques both in treatment and control groups, which already helps controlling for homogeneous "Good Old Times" bias. This is particularly suitable when we use the Cape Verdian control group, since the same type of respondents are compared, and no a priori reason for differential bias exists.

5 Data Collection: Tailored Household Surveys on Corruption

Data analyzed in this paper come from surveys conducted by a team recruited and trained by the author, which included the author, in STP (Apr./May 2004), and in CV (Dec. 2005/Feb. 2006)[26]. In STP, the survey was submitted to 841 households in 30 of the 149 (20%) census areas of the country[27] (on average approximately 28 interviews per area were conducted). In CV, the survey was submitted to 1066 households (though only 997 respondents completed the full length of the interview) in 30 of the 561 (6%) census areas of the country[28] (on average 36 interviews per census area).

5.1 Sampling Process

The sampled census areas were chosen randomly weighting by the number of households[29]. Households chosen were distributed as evenly as possible by the census area[30]. In each household an individual with 30 years or older, with residence or direct relatives' residence in the respective country in the relevant periods of the survey instrument (these included our periods of interest,

[26] For details of fieldwork activities in both surveys, visit http://users.ox.ac.uk/~econ0192/fieldwork.htm
[27] Although, the smaller established administrative division of the country is the "district" (STP is composed of 6 districts, Água Grande, Cantagalo, Lembá, Lobata, Mé-Zochi, and Príncipe), the STP National Statistics Office has available a census-aimed map-based division of the country (which may be better described as a division by neighborhoods - urban areas - or sets of villages - rural areas). This was the sampling basis used in our STP survey.
[28] Although administrative divisions of CV are finer than in STP, for sampling purposes, we ended up using the CV National Statistics Office census-based grid, which differed from the administrative map.
[29] These data - referring to the 2000/2001 censuses - were provided by the National Statistics Offices of both countries.
[30] Standard techniques as seeking the nth (depending on the number of households at each area) house in a census area were used.

before and after the oil discovery announcements, 1991-1997 and 2000-present[31]), was invited to answer the questionnaire.

We can then infer that a two-stage sample design, first selecting clusters and then households, was the basic method pursued, with the probability that a given household was chosen being a priori the same across all households[32] (we aimed at national representative samples of households). Note also that, since we aimed at assessing perceived corruption by the household, any representative of the household with the above stated conditions would be adequate - this was the reason for the adoption of the described simple method of sampling within the household, which naturally implies a non-random sample of individuals in the survey.

Two sources of imperfection in the sampling of households should be referred: non-respondents[33] and differences in the number of identified respondents in the different census areas[34]. The first problem was mitigated in the process of collecting the data because when potential respondents refused to answer, interviewers would take note of their gender, approximate age, schooling, and (only in CV) income. These data were used to weight observations appropriately (with respect to gender, age, schooling, and income). Tables with information on the sample geographical distribution are provided in Appendix B (Table A1 and A2), as well as detailed maps on the sampled census areas (Maps A1-A6).

To have a grasp of the differences between responding and non-responding sub-samples, we present the pyramid graphs (gender and age) for the respective groups: middle-aged men (same for women in CV) refused to respond disproportionately more in both STP and CV (these are in Charts A1-A2 and A4-A5 in Appendix B). Concerning education, as could be anticipated, more schooled subjects answer more frequently in both countries (this is in Charts A3 and A6 in Appendix B).

All statistics and econometric work we will present use weights to account for the two referred departures from perfect household random sampling[35]. This is for consistency with the sampling approach. However, differences to unweighted resulting statistics are negligible throughout the analysis.

[31] Note that the survey was also intended to assess the impact of the changes occurred in the country in the period 1989-1990, which included the democratization process – therefore questions were also asked about the pre-democratization period 1985-1988. However, these data do not concern the research questions of this paper, and therefore are not used here.
[32] Note that in CV a further cluster stage was used: the choice of 4 out of the 9 areas weighting by the population.
[33] In this survey this problem was not significant by common standards in the survey literature: identified non-respondents amounted to 8.5% (STP) and 16% (CV) of the identified sample.
[34] Namely, an over-sampling in the two census areas of the island of Príncipe took place.
[35] Standard techniques were followed - see Deaton, 1997.

5.2 Survey Instrument Design

The questionnaire[36] was constructed following two basic principles[37]: asking questions about specific public services to the individuals who had experience with those particular services (this was assessed by asking directly or indirectly about that experience); and eliciting respondents' memories by making use of their personal history[38].

In terms of general structure, the questionnaire began with a question regarding the current general quality of different public services[39] and continued with non-threatening demographic questions (e.g. children's ages). This demographic information enabled the interviewer to "draw a precise picture" of the respondent, which was useful in identifying legitimate respondents to specific questions - e.g. a sufficient condition to answer the questions on the health care services was to have had at least one child - and in eliciting memories / probing (the interviewers referred to the periods of interest in a personal way - e.g. for the subject that had her first child born in 1994, in the health care services question, the interviewer referred to the period before-oil announcement, 1991-97, as the period when the respondent's first child was born[40]).

Corruption-related questions were then posed in the following sector order: health care public services, public education, justice (courts of law and police), private sector related public activity (choice of state suppliers, subsidy beneficiaries, and public infrastructures; license emission services), public sector jobs, customs, political system (existence of campaigning gifts, influence of gifts in voting). These questions were always asked in a general, non-personal way, making sure respondents felt comfortable and safe in answering. Whenever experience with a particular service/issue was not assessed during the first part of the interview, specific questions regarding that experience were posed just before asking the corruption-perception questions - see Table A3.1 in Appendix C for a precise phrasing of these questions.

Note that the questions on corruption were centered on personal influence[41] (these were questions repeated for the relevant periods, and the basic/direct source for our econometric results

[36] The final version of the survey instruments is available in the original Portuguese and as an English translation at: http://users.ox.ac.uk/~econ0192/fieldwork.htm

[37] See Bertrand and Mullainathan (2001) for a review of the main difficulties of using/designing subjective data.

[38] This technique is well known in the survey design literature (see for instance Rasinski, Rips, and Tourangeau, 2000).

[39] This was basically for attachment to the interview since the survey was presented to subjects as one on the quality of public services in their countries of residence.

[40] The question regarding children's ages implied, from an age input, and from the use of tables of ages, the identification for that child of the relevant birth period. Although interviewers in STP did this procedure manually, it was done automatically in CV provided the use of handheld computers and survey conduction software. In both countries the use of tables was aimed at simplifying the questions asked (taking out complexity from the respondent's side) and at increasing accurateness.

[41] In fact, this prior, dominance of influence types of corruption, had considerable support on pre-survey anecdotal evidence and on the testing phase (piloting) of the survey questions.

- see Table A3.2 in Appendix C for a precise phrasing of these questions) - some were posed on the explicit/negative side and some were posed on the positive side[42]. In addition a question regarding bribery was added for each clearly decentralized public service or each allocation implying infrequent contact between allocator and recipient (where crude bribery would be more likely to arise), which worked as a qualifier in terms of monetarization of corruption - see Table A3.3 in Appendix for a precise phrasing of these questions. A summary of the constraints (on respondents), elicitors of memory, and formats of the corruption-perception questions is presented in the next table.

Table 1: Questions on Corruption - Constraints, Memory Elicitors, and Format

Questions on:	Accessing Categories (constraints)	Elicitors of Memory (in the periods of interest)	Way Question Asked	Asked about Bribery ?
Health Care	occup: public administration OR occup: health OR at least one child OR at least one visit to health care w/ illness	occupation, children birth, visits to health care/illnesses	-	Yes
Education	occup: public administration OR occup: education OR at least one child in primary school	respondent's schooling, occupation, children schooling	+	Yes
Scholarships	occup: public administration OR occup: education OR at least one child in secondary school	respondent's schooling, occupation, children schooling	-	No
Courts	occup: public administration OR contact with courts of law (qabs)	contact with courts of law (qabs)	-	Yes
Police	occup: public administration OR contact with police (qabs)	contact with police (qabs)	-	Yes
State Jobs	got public sector job or applied (qabs)	got public sector job or applied (qabs)	-	No
Customs	occup: public administration OR went through customs (qabs)	went through customs (qabs)	-	Yes
Subsidies / Supplier Positions	supplier/beneficiary or applicant to these (qabs)	supplier/beneficiary or applicant to these (qabs)	+	Yes
Public Infrastructures	supplier/beneficiary or applicant to these (qabs) AND occup: construction	supplier/beneficiary or applicant to these (qabs)	+	No
Licenses	got license or applied (qabs)	got license or applied (qabs)	-	Yes
Vote Buying	NA	elections history	-	NA

Notes: qabs - question asked at the beginning of the section.

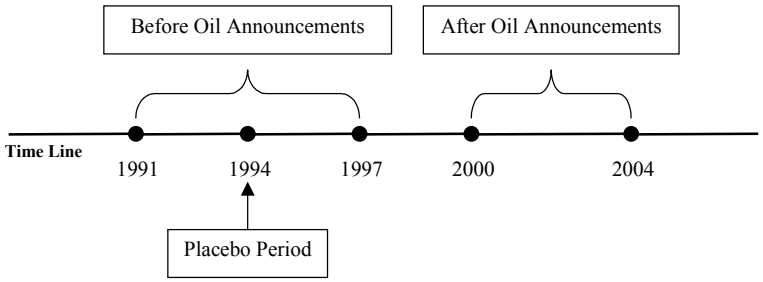
General questions regarding the expectations about the impact of the oil reserves were posed near the end of the STP interviews (not to suggest any link between the questions on corruption and the oil news, that could raise 'public opinion'-type of answers)[43]. Also at the end of the questionnaire, questions regarding more sensitive demography were presented (e.g. political preferences, income).

[42] This worked as a balancer of tone in the interviews - we did not want to sound negative, so that we could bias respondents to "always worse with time" kind of answers.

[43] For an exact phrasing regarding the two oil related questions we will use later in econometric work see Table A3.4 in Appendix C.

Finally, a remark is due regarding the kind of tools we used to control the credibility of answers. Concerning the possibility of a "good old times" bias (tendency of respondents to report worsening conditions over time), two questions were included in the beginning of the questionnaire concerning general optimism (for an exact phrasing see Table A3.4 in Appendix C), and a placebo period (the relevant period 1991-1997 was divided into 1991-1994 and 1994-1997) was included in the main question regarding corruption in the health care services[44]. An illustration of the relevant periods regarding the oil discovery announcements and the placebo period is provided in Figure 1.[45]

Figure 1: An Illustration of the Periods Used to Assess The Impact of the Oil Discovery Announcements



It should be noted that all questions in the survey instrument were close-ended, with the questions on subjective assessments being given in the 1-7 scale. Importantly, all these questions were approached in a two-stage iterative process: when asking a question, the interviewer referred first to the three basic scale options 1-3, 4, and 5-7; when the subject positioned herself in one of these, the interviewer, in the first and last cases, asked about a precise number within the class already identified by the respondent. Notice that, in these questions, all numbers had a precise scaling language counterpart (using adjectives), which was the only way used by the interviewers to refer to scales. An example is provided in the next table.

[44] This question was chosen to be the one where a Placebo period would be included provided it is the most widely answered question on standard perceived corruption we could expect.

[45] In addition, two questions were presented at the end of the questionnaire regarding overall reliability (degree of comfort while answering the questionnaire - for an exact phrasing see Table A3.4 in Appendix C -, and degree of understanding, assessed by the interviewer, about the questionnaire. Although the language of the questionnaire was Portuguese, the official language of both STP and CV, other local languages are used in these countries (in STP: Forro or São-tomense, and Angolar; in CV: Creole or Capeverdian). Occasionally in STP (particularly with elderly respondents), but frequently in CV, translations to these languages were performed by the local interviewers.

Table 2: Scaling Approach in Subjective Questions

1st oral approach	Not Necessary				Necessary		
2nd oral approach	Not at All	Not a Lot	Not Much	More or Less	Somewhat	Very	Extremely
	(Necessary)						
Coding	1	2	3	4	5	6	7

In Appendix D we show descriptive statistics of the sample using the demographic variables included in our instrument. In some cases we compare with census data. Our main purpose there is to provide a deeper characterization of STP and CV, while comparing both samples (we find reassuringly similar statistics in most indicators[46]).

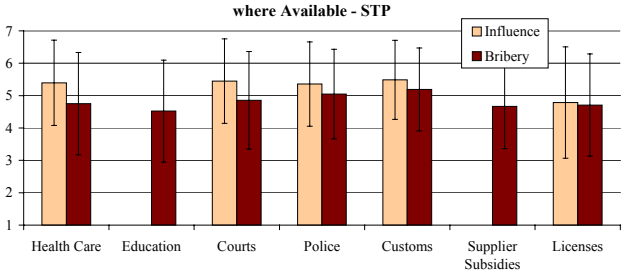
6 Perceived Corruption Data and Estimation Approach

As the basic econometric question in this paper, we want to identify/quantify the effect on perceived corruption of the announcement of the existence of oil reserves in STP.

We base the answer to this question on comparing data from the questions on influence (these are the ones referenced "Main" in Table A3.2 in Appendix relative to Health Care, Education, Scholarships, Courts, Police, Subsidies / Supplier Positions Recipients, Public Infrastructures, Licenses, State Jobs, Customs) and Vote Buying. In fact, from the observation of the average ratings of bribery (highest for Customs, Police and Courts) for the period 2000-present in our surveys - shown in Charts 3 - we verify our prior (in the design of the questionnaire) that influence is generally seen as the most prevalent form of corruption in both STP and CV (a comparison with influence averages, where possible[47], is provided in the same Charts)[48]. Note however that bribery in CV is generally perceived as not so frequent as in STP – this is another reason to base our comparisons on influence.

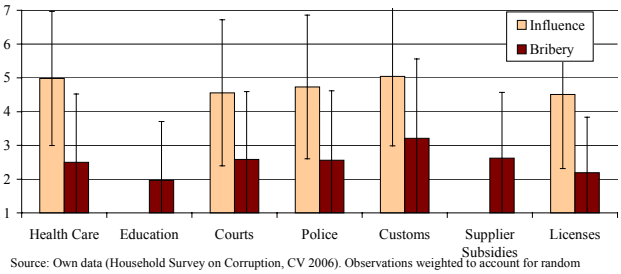
[46] This is despite the fact that we will control for these demographic variables in the regressions ahead.
[47] This is for the issues where influence and bribery questions are comparable. Note that since the questions on influence for Education and Subsidies / Supplier Positions Recipients are posed in a positive tone, their answers are not comparable with those regarding the bribery questions (posed on the negative manner). Bribery questions are described in Table A3.3 in Appendix.
[48] It can also be argued that influence is just a lagged (and possibly repeated) transaction equivalent to the usual bribery notion - therefore more general.

Chart 3a: Bribery Perceptions 2000-present (Averages and Std. Devs.), with Comparison with Influence (Averages and Std. Devs.) where Available - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Lines represent +- Standard Deviations. Sample Questions (Health Care): "In the Sao-Tomean reality, with respect to being assisted in the public health care services, what has been the need to: know someone who works there? (Influence) Offering bribes? (Bribery)".

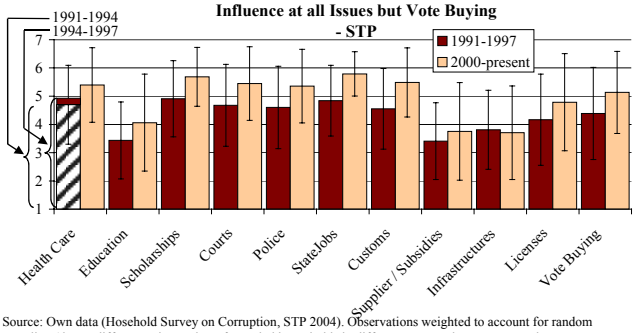
Chart 3b: Bribery Perceptions 2000-present (Averages and Std. Devs.), with Comparison with Influence (Averages and Std. Devs.) where Available - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and income). Sample Questions (Health Care): "In the Sao-Tomean reality, with respect to being assisted in the public health care services, what has been the need to: know someone who works there? (Influence) Offering bribes? (Bribery)".

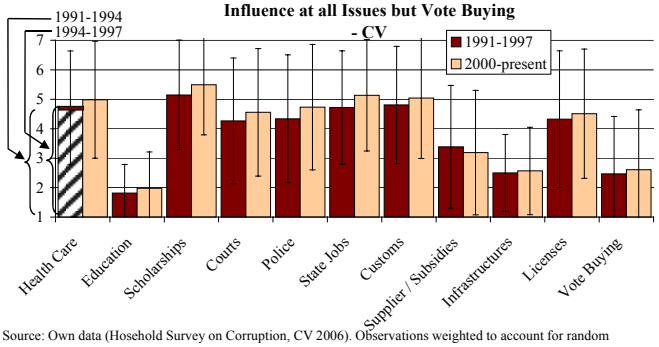
A crude observation of the data available on perceived corruption for the periods 1991-1997 and 2000-present (in Charts 4 below) leads us to conclude for a less clear increase in this variable over the referred periods in CV than in STP (despite the slight decrease in the perception of corruption in public infrastructures[49] for STP and in Subsidies/Supplier Positions for CV) - the difference before/after placebo period also seems to be clearly lower than the difference before/after oil discovery announcements in STP.

Chart 4a: Corruption Perceptions (Averages and Std. Devs.) - as Influence at all Issues but Vote Buying - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Lines represent +- Standard Deviations (1994-1997 in the first column). Questions on Education, Subsidies / Supplier Positions Recipients, and Public Infrastructures are not comparable with the other questions provided, contrary to these, the first are posed in a positive manner.

Chart 4b: Corruption Perceptions (Averages and Std. Devs.) - as Influence at all Issues but Vote Buying - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and income). Lines represent +- Standard Deviations (1994-1997 in the first column). Questions on Education, Subsidies / Supplier Positions Recipients, and Public Infrastructures are not comparable with the other questions provided, contrary to these, the first are posed in a positive manner.

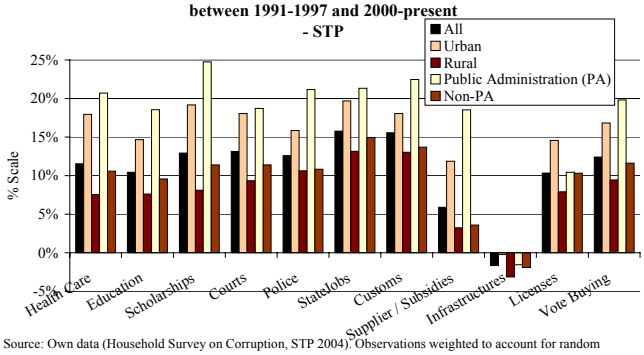
However, in light of our theoretical model, as the oil announcements may imply a change in behavior of the elite-seat candidates (in terms of corruption), a better/more precise answer to the above question should be given by the comparison between answers from better and worse informed (on the actions of those candidates to becoming the country's elite) respondents

[49] The data concerning this question should be seen with some reserve provided it was only responded by 94 and 93 subjects for the periods 1991-1997 and 2000-2004 (respectively) in STP, and for 62 and 60 respondents in CV.

(internally) or by the comparison of the STP with the CV data (externally). For the first we pose an important hypothesis: we assume that urban subjects, public administration-employed subjects, and (more generally) respondents who show (in the questions of the survey) familiarity with the public services/allocations are better informed on the referred actions.[50]

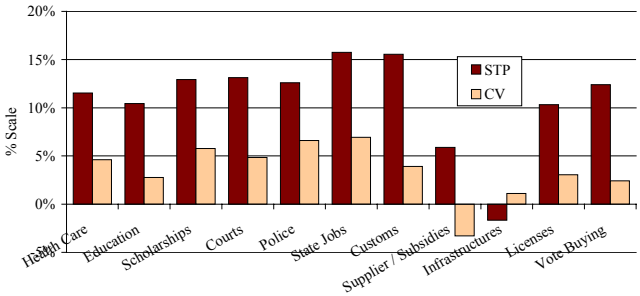
In this view we review the crude STP data by focusing on perceived changes by urban/rural[51] and public administration/other respondents[52]: clear higher changes seem to be perceived by urban and public administration groups relative to their respective counterparts. When comparing STP with CV we see a consistent/compatible pattern[53]. These are shown in Charts 5 below.

Chart 5a: Differences (% Scale) in Perceived Corruption between 1991-1997 and 2000-present - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).Questions on Education, Subsidies / Supplier Positions Recipients, and Public Infrastructures are not comparable with the other questions provided, contrary to these, the first are posed in a positive manner.

Chart 5b: Differences (% Scale) in Perceived Corruption between 1991-1997 and 2000-2004 - STP vs CV



Source: Own data (Household Surveys on Corruption, STP 2004, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - income). Questions on Education, Subsidies / Supplier Positions Recipients, and Public Infrastructures are not comparable with the other questions provided, contrary to these, the first are posed in a positive manner.

We have also composed a variable that makes use of the questions in Table A3.1 in Appendix C relative to specific experience with services/allocations by the public sector. This variable is intended to capture the degree of exposure of the respondents relative to information regarding the behavior of the elite in a most precise manner - we denominate it "Connectivity Indicator". Note that it also includes information on occupation and children's birth/schooling - in what can

[50] This familiarity may be derived from a higher experience in using the services or applying for the allocations (this is the main idea behind using urban respondents) or may be derived from inside information (which is more the case of the public administration employees, who are themselves allocators of the state services).

[51] This distinction was constructed by considering:

- for STP: all census areas in Água Grande (the district composed of the capital city and its suburbs) and the census area of Trindade-Cruzeiro (corresponding to the center of the second city of the country) to be urban; all the other census areas visited in the survey were considered as rural.
- for CV: all census areas in Praia (with the exception of São Martinho Grande) and in Mindelo, the two major cities of the country, are considered to be urban; all others are taken as rural.

[52] To be classified as part of the public administration, a respondent had to answer at least once (relative to the two periods: 1991-1997 and 2000-2004) that one of her professional occupations was in the public administration.

[53] Regarding Vote Buying, additional evidence from plainly asking if parties have distributed non-trivial gifts in the respondent's town - question referenced VB.1 in Table A2.2 in Appendix - indicates that the percentage seeing no vote buying decreased from 26% (in 1994-1998 elections) to 19% (in 2001-2002 elections), whereas in CV it decreased from 76% (in 1995-1996) to 72% (in 2001).

be seen as an index of the information in the accessing information column in Table 1 above (since the questions of the survey were posed only to respondents showing a minimal amount of experience with the issues at stake). The weights given to each piece of information (when available) regarding each respondent are presented in Table A4 in Appendix E. They assume that the different issues and respective number of questions (concerning perceived corruption) included in the survey instrument form a representative set of implied relevant actions by the elite-seat candidates, so that accessing questions are weighted according to their use in the actual completed interview[54]. Correlations between the different internal treatment variables are shown in Tables A5 in Appendix E: as expected the highest correlation is between the public administration dummy and the Connectivity Indicator in both countries.

The above assumptions on higher exposure of specific groups allow us to use difference-in-differences type regressions (DID) to estimate the relevant impact of the announcements; this econometric specification is obvious for the external (CV comparison) exercise. The general regression equation (leading to the DID estimator), we will use, can be presented as:

$$K = \beta_0 + \beta_1 X + \delta_0 dT + \delta_1 dL + \delta_2 dTdL + \varepsilon ,$$

where K is a variable of perceived corruption, X is a vector of individual controls,

$$dT = \begin{cases} 1, \textit{ if observation is taken after the treatment } (2000 - \textit{present}) \\ 0, \textit{ if otherwise } (1991 - 1997) \end{cases},$$

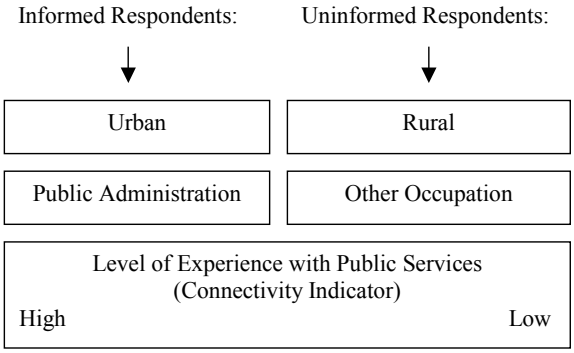
ε is an error term. Note that

$$dL = \begin{cases} 1, \textit{ if observation is taken within the treatment group} \\ 0, \textit{ if otherwise} \end{cases}$$

(if the treatment/control group is binary[55]) or dL equals the Connectivity Indicator (otherwise)[56]. An illustration of the internal comparison groups used is showed next.

[54] For each entry, a value 0, 0.5, or 1 was given for no experience at any period, experience at 1991-1997 or 2000-2004, and experience at both periods (respectively). This indicator yields values between 0 and 1.
[55] Here we use the dummy variables already presented in the graphs above.
[56] Note that, as Bertrand, Duflo, and Mullainathan (2004) point out, the fact that we collapse time series variation into before/after periods, minimizes standard error biases from serial correlation.

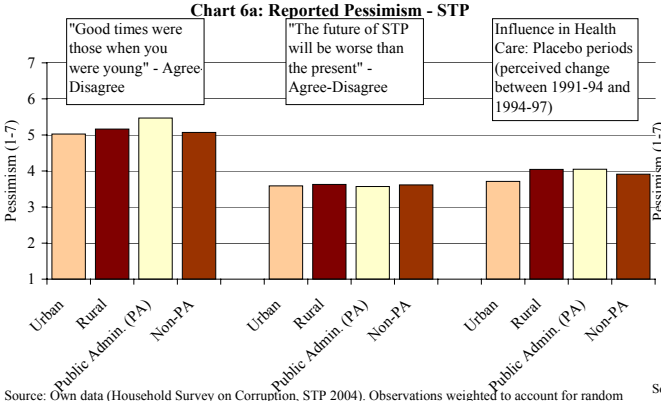
Figure 2: An Illustration of the Internal Comparison Groups



6.1 Psychological Controls: the "Good Old Times" Problem

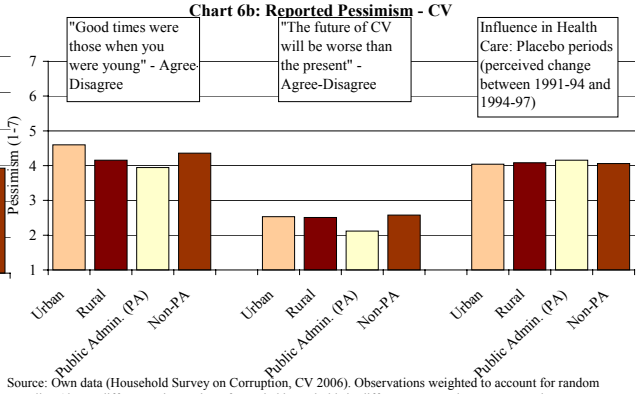
Finally a note is due on pessimism and the usually encountered "Good Old Times" bias. We control for this bias by using data from the Optimism questions in Table A3.4 in Appendix C and from the inclusion of a "Placebo" period in the question on perceived corruption in health care. Data from these four questions are showed in Charts 6 for our two main internal treatment/control groups in the two country samples: rural respondents seem to be generally the most pessimistic in STP; respondents in CV seem to be generally less pessimistic than in STP. However overall differences are small across these groups, which is a symptom that the use of the same survey techniques in all treatment and control groups could already be close to taking care of this kind of time bias.

Chart 6a: Reported Pessimism - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).

Chart 6b: Reported Pessimism - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures).

Data from these three sources will be used ahead for controlling the effects of interest not only in levels but also in differences. Questions on the impact of oil and the degree of comfort

with the interview (Other Opinion / Psychological Questions in Table A3.4 in Appendix C) will also have a role in our regressions, since they assess other psychological dimensions of the respondents mainly related with trust.

7 Econometric Results

7.1 Internal Controls

Our baseline internal control results are founded on the above described comparisons of STP respondents. Several distinct sets of internal controls and a wide range of individual characteristics are considered. The detailed results for all sectors/allocations are presented in Tables A6-A15 in Appendix F.

All the estimated coefficients of interest show the expected positive sign across the various specifications and the different sectors of public service, although with differing degrees of statistical significance. These are shown in Charts 7 and 8 below. In terms of magnitude of the effects: when urban/rural or public administration/non-public administration are used, we observe values ranging from 5% to 12%; once we employ the Connectivity Indicator, the increase in perceived corruption ranges from 21% to 38% of the subjective scale. Note that while the first set of results concerns group-derived effects, the second refers to changes perceived by the highest informed respondent as compared with the lowest knowledgeable subject – it is therefore expectable that they are generally higher.

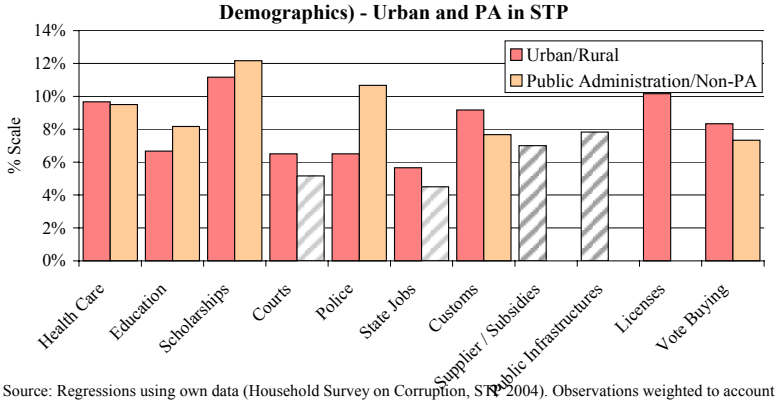
Regarding the urban-derived effects, these are highest for Scholarships (11% of the subjective scale), Licenses, Health Care, Customs, and Vote Buying (not lower than 8%). On the Public Administration regressions, we see highest interaction terms in Scholarships (12%), Police, Health Care, Education, Customs (not lower than 7%). Effects on State Jobs, though relatively low (close to 6%), are seen to be solid in the urban/rural comparisons (with statistical significance at the 1% level); the lack of significance in the Public Administration group may be due to underreporting by self-concerned public officials.

When using the respondents showing higher connectedness with the services and allocations, we find that Scholarships (37%), Vote Buying, Police, Education, and State Jobs (27%) have the highest perceived effects on corruption after 1997-1999. These are most significant (at the 1% level) for Scholarships, Vote Buying and State Jobs. [57]

[57] Note that generally samples sizes may be a problem in not finding significant results in Public Infrastructures and, less clearly, in Licenses.

Contrary to what one could expect, psychological and demographic controls[58] do not seem to substantially or systematically alter results. Indeed, differences in the estimated cross coefficients are generally minimal when the psychological controls are introduced (slightly more towards decreasing the estimates). Note that this is despite the fact that clear patterns of explanatory power arise with respect to these variables: usually pessimism variables are positive on differences and are negative on levels: this is consistent with the idea that indeed more pessimistic respondents perceive a higher difference in corruption over time, but report less corruption in the period before the oil discovery (the good old times). Usually Public Administration interaction effects seem to be decreasing more than urban-derived effects (which is consistent with Chart 6a above). The changes introduced by demographic results are generally slightly larger, but clearly both positive (Scholarships, Police, Vote Buying) and negative (Health Care, Education), depending on the sector.

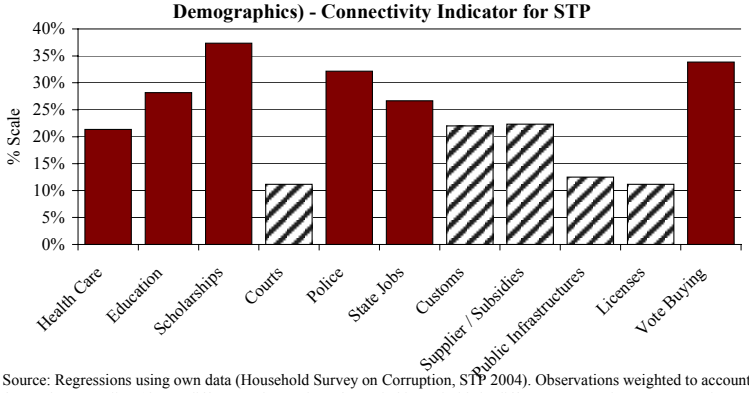
Chart 7: Perceived Additional Change in Corruption (% Scale) as Seen by "Connected" Groups (Controlled by Pessimism and Demographics) - Urban and PA in STP



Source: Regressions using own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Listed columns correspond to statistically insignificant estimates (at the 10% level).

[58] Interesting patterns arise here. For instance, regarding vote buying, we find that respondents with higher expenditure report lower vote buying influence in voting, which is consistent with Brusco, Nazareno, and Stokes (2004). We also obtain that house ownership increases perceived corruption - the apparent contradiction with the first finding may be explained by considering house owners - who are not necessarily richer, provided many houses in the country are constructed by their occupants with relatively cheap materials - have higher geographical stability and social connectedness and are therefore more attractive to vote buyers.

Chart 8: Perceived Additional Change in Corruption (% Scale) as Seen by "Connected" Groups (Controlled by Pessimism and Demographics) - Connectivity Indicator for STP



Source: Regressions using own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Listed columns correspond to statistically insignificant estimates (at the 10% level).

7.2 External Controls

We now focus our attention on comparisons of the Sao Tomean and Cape Verdean samples within the DID framework. Generally, the results observed in the last section pointing to significant increases in perceived corruption in STP are confirmed. However, a somewhat different (with some clear validations) alignment of the sectors/allocations emerges.

Note that, consistently with the hypothesis used for the internal comparisons, in addition to using full samples for both countries, we also compare urban subsamples, public administration subsamples, and connected subsamples (as given by a value of the Connectivity Indicator above the country median) in both countries.

These results can be seen in Tables A17-A27 in Appendix G, as well as in Chart 9, where interaction effects are plotted for the different questions.

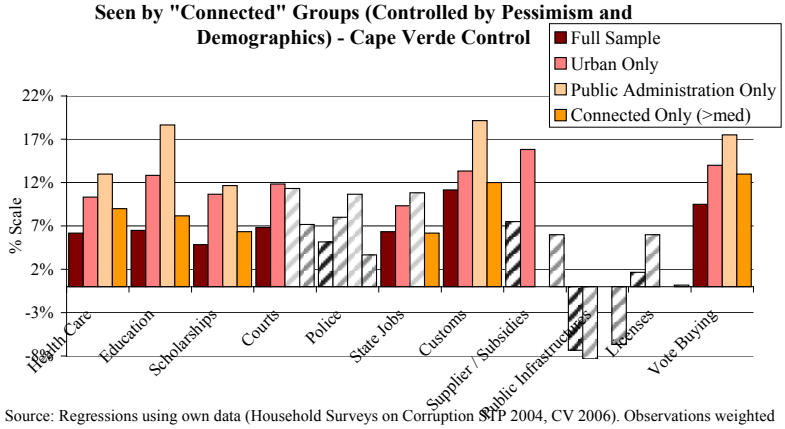
Two major conclusions are worth stressing from the analysis of these regressions.

First, we observe that, overall, full samples provide lower estimates than those obtained by using informed group samples. This is consistent with the idea that, when using full samples, effects are diluted, since a smaller proportion of the sample is able to recognize the effect of interest - i.e. does not have a close contact with the elite-seat competition from our model. Generally, public administration-based comparisons across these countries yield the highest effects, followed by urban and connected subsamples.

Second, in terms of sectoral differences in the changes of perceived corruption, significant effects are highest for Customs and Vote Buying (respectively, for full samples 11% and 10% of the subjective scale). Health Care, Education, Scholarships, and State Jobs Allocations maintain

solid effects (in terms of significance) but decrease their magnitude, when comparing with the internal control-based results of the last sub-section.

Chart 9: Perceived Additional Change in Corruption (% Scale) as Seen by "Connected" Groups (Controlled by Pessimism and Demographics) - Cape Verde Control



Source: Regressions using own data (Household Surveys on Corruption SIP 2004, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures). Listed columns correspond to statistically insignificant estimates (at the 10% level).

7.3 Robustness of the Corruption Scale Hypothesis: Ordered Probit

We now analyze the consequences (in terms of resulting estimates) of our hypothesis on the corruption scale use.

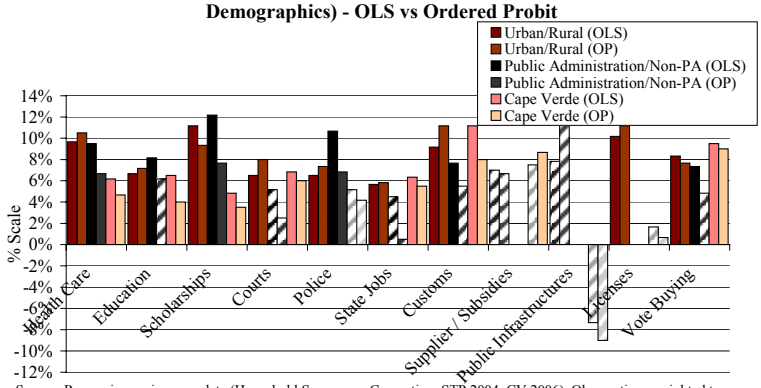
In the OLS estimations above, we have implicitly assumed an equidistant 7-point scale for perceived corruption. This assumption was based on the fact that, as described in section 5 above, the survey has implied a symmetric and stepwise verbal approach on a scale with subjective categories that were themselves close to being equidistant. We challenge this hypothesis by estimating an Ordered Probit model.

Our findings are very much supportive of the equidistant perceived corruption scale hypothesis, as we find very small differences between Ordered Probit and OLS estimates. This can be attested from the observation of Table A28 in Appendix H, where Ordered Probit results are displayed, and where partial effects correspond to the second column for each regression. Charts 10 and 11 below provide a comparison of interaction effects of interest. Note that all regressions are run for the same sectors as in internal/external comparisons in the last sections, and use psychological and demographic controls – in fact the exact same variables are used as in the corresponding OLS regressions. Nevertheless, only full sample results are displayed in the external control regressions, which (as seen in the last sub-section) yield lower bounds on the effects of interest.

The most substantial differences relative to OLS internal control estimation are that: urban interaction terms in Ordered Probit are higher than in OLS (with the exception of Scholarships, Public Infrastructures, and Vote Buying), whereas the opposite happens with the Connectivity Indicator (with decreased statistical significance in the Police sector); Public Administration estimates with Ordered Probit are also generally lower than those obtained using OLS, with decreased statistical significance beyond the 10% threshold in Education, Customs, and Vote Buying.

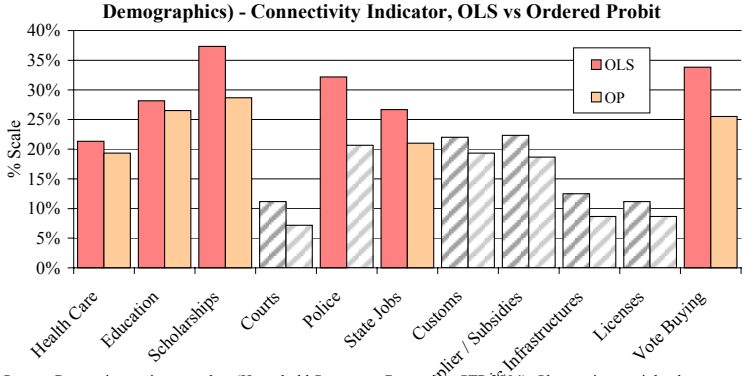
A final note is deserved regarding the Cape Verdean external control: coefficients are generally lower with OLS than with Ordered Probit; in addition, statistical significance is maintained or slightly increased (Subsidies/Supplier Positions).

Chart 10: Perceived Additional Change in Corruption (% Scale) as Seen by "Treated" Groups (Controlled by Pessimism and Demographics) - OLS vs Ordered Probit



Source: Regressions using own data (Household Surveys on Corruption, STP 2004, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures). Listed columns correspond to statistically insignificant estimates (10% level).

Chart 11: Perceived Additional Change in Corruption (% Scale) as Seen by "Treated" Groups (Controlled by Pessimism and Demographics) - Connectivity Indicator, OLS vs Ordered Probit



Source: Regressions using own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Listed columns correspond to statistically insignificant estimates (10% level).

7.4 Summary of Results and Interpretation

Given the evidence presented in this section, we can generally conclude that corruption has increased after the period 1997-1999, as we would expect. This change was perceived as higher within respondents close to the public allocations of the country[59], controlling for pessimism, other psychological factors, and demography. This pattern may be interpreted as the oil announcements causing agents with proximity to the elite of the country (or competing for a position that will enable them to access the oil rent in the future) to increase the frequency of their corrupt acts (and therefore the corruption effort or cost we explicitly have modeled in section 2).

The second research question we proposed to address regarded the channels by which corruption starts increasing after the expectation of increased public rents. Overall, summarizing all methods and robustness tests performed, we can conclude that clear changes appeared in several sectors where the mechanism of our model is patent: Vote Buying (not only but specially when using the external control group), which may be interpreted strictly in light of our model (that *is* competition for the *political* elite of the country); Education/Scholarships (over all treatment/group evaluations, in particular using internal comparisons), which has an obvious relation with future status recognition in the STP society (specially with respect to higher education, and therefore the allocation of scholarships to study abroad we analyzed); State Jobs (using all comparison groups but the possibly underreporting Public Administration), corresponding to the competition for the public administration elite. The effect perceived on Health Care (through all comparison groups) is less clear in light of our model. We can always interpret it under the same lens, which in this case turns out to be extreme: health care allocations represent the competition for being in good health to witness the oil rent. But we prefer to interpret this effect in an indirect fashion: it may be the result of decreased effort from the leaders of the country in improving the population's welfare from the re-orientation of efforts towards rent-seeking.

Overall, the effects observed are sizeable: they indicate an increase of 21% to 38% of the subjective scale, when using the broadly encompassing Connectivity Indicator (which offers the cleanest interpretation). We believe these findings may provide us with a better idea of the mechanisms by which the natural resource curse may work.

[59] Note that generally the coefficient on oil is positive and significant, which leads us to conclude that the control groups also see more corruption after the period 1997-1999.

8 Concluding Remarks

This paper analyzes increasingly important questions in the development literature: what is the impact of natural resources in the level of corruption of a country? And which are the channels by which corruption starts increasing after a natural resource discovery?

We postulate a theoretical mechanism whereby an anticipated increase in a country's public sector rent (through the appearance of a new oil sector) increases competition for a place in the elite of that country (which will be able to grab the public sector rent). This implies competitors for those seats will be willing to spend more resources in improving their perspectives of success (meaning higher corruption) – i.e. they will want to win more present-time allocations that have implications in their future status.

We have found, using a tailored design of collection of household survey data, that not only does corruption seem to have indeed increased in STP as a consequence of the oil discovery announcements of 1997-1999, but also most sectors with higher and more significant estimated effects were those directly implying future power to grab rents (Vote Buying, Education, State Jobs).

These facts consolidate the idea that natural resources may be harmful for the quality of governance in developing countries, and that the political arena can see a fierce increase in competition in the event of a natural resource shock. Monitoring of new resource-rich countries in the areas connected with political power, although sensitive, seems to be particularly adequate if these effects are to be prevented and counteracted. This is therefore evidence that top-based efforts towards accountability and transparency of oil revenues may be crucial and to-the-point in weakening the expectation of rents-for-grab – i.e. improving governance today.

On the methodological side, we have shown that we can make use of the population as a witness to evaluate the impact of mainly top-based/political/macro-level exogenous shocks. Household surveys of the kind used in this paper may prove to be, if suitably designed, a powerful and effective instrument in such endeavors, not only to tackle mainly-research questions (as the ones addressed here) but also to undertake policy evaluation.

In our view, two interesting areas regarding the issues at stake in this paper remain in need of further research. Namely, a deeper study of inefficiency of corruption in empirical terms would provide an important qualifier of the problems created by the reported increases in corruption, i.e. accomplish a closer link with growth considerations. In addition, a parallel and complementary effort to gather sector-specific measures of corruption, drawing upon direct quantification (as opposed to perceptions), can be a way forward in studies of corruption like ours – the design of

direct measures of public sector corruption based on potentially systematic data-gathering may be an important path for the improvement of development policies.

References

Ades, Alberto and Rafael Di Tella (1999), *Rents, Competition, and Corruption*, American Economic Review, 89(4), pp. 982-993;

Baland, Jean-Marie and Patrick Francois (2000), *Rent-Seeking and Resource Booms*, Journal of Development Economics, 61, pp. 527-542;

Banerjee, Abhijit V. (1997), *A Theory of Misgovernance*, Quarterly Journal of Economics, 62, pp.1289-1332;

Bannon, Ian and Paul Collier (editors) (2003), *Natural Resources and Violent Conflict: Options and Actions*, The World Bank;

Bardhan, Pranab (1997), *Corruption and Development: A Review of Issues*, Journal of Economic Literature, 35, pp. 1320-1346;

Becker, Gary S. and George J. Stigler (1974), *Law Enforcement, Malfeasance, and Compensation of Enforcers*, Journal of Legal Studies, 3(1), pp. 1-18;

Bernheim, B. Douglas and Michael D. Whinston (1986), *Menu Auctions, Resource Allocation, and Economic Influence*, Quarterly Journal of Economics, 101(1), pp. 1-32;

Bertrand, Marianne, Simeon Djankov, Rema Hanna, and Sendhil Mullainathan (2006), *Does Corruption Produce Unsafe Drivers?*, NBER, Working Paper 12274;

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan (2004), *How Much Should We Trust Difference-in-Differences Estimates?*, Quarterly Journal of Economics, 119(1), pp. 249-275;

Bertrand, Marianne and Sendhil Mullainathan (2001), *Do People Mean What They Say? Implications for Subjective Survey Data*, American Economic Review, 91(2), pp. 67-72;

Brusco, Valeria, Marcelo Nazareno, and Susan Stokes (2004), *Vote Buying in Argentina*, Latin American Research Review, 39(20, pp. 66-88;

Buchanan, James M., Robert D. Tollison, and Gordon Tullock (Eds) (1980), *Toward a Theory of the Rent-Seeking Society*, Texas A&M University Press;

Cadot, Olivier (1987), *Corruption as a Gamble*, Journal of Public Economics, 33, pp. 223-244;

Carreira, António (1982), *The People of the Cape Verde Islands - Exploitation and Emigration*, C. Hurst & Company;

Deaton, Angus (1997), *The Analysis of Household Surveys: a Microeconometric Approach to Development Policy*, The John Hopkins University Press;

Duggan, Mark and Steven D. Levitt (2002), *Winning Isn't Everything: Corruption in Sumo Wrestling*, American Economic Review, 92(5), pp. 1594-1605;

Fisman, Raymond (2001), *Estimating the Value of Political Connections*, American Economic Review, 91(4), pp. 1095-1102;

Fisman, Raymond and Edward Miguel (2006), *Cultures of Corruption: Evidence From Diplomatic Parking Tickets*, NBER, Working Paper 12312;

Fowler, Floyd J.(1993), *Survey Research Methods*, 2nd Edition, Sage Publications;

Frynas, Jedrzej G., Geoffrey Wood, and Ricardo S. Oliveira (2003), *Business and Politics in São Tomé and Príncipe: From Cocoa Monoculture to Petro-State*, African Affairs, 102(1), pp. 51-80;

Gupta, Sanjeev, Hamid Davoodi, and Rosa Alonso-Terme (1998), *Does Corruption Affect Income Inequality and Poverty?*, IMF, Working Paper 76;

Hodges, Tony and Malyn Newitt (1988), *São Tomé and Príncipe: From Plantation Colony to Microstate*, Westview Press;

Kaufmann, Daniel, and Aart Kraay (2002), *Growth without Governance*, Economia, 3, pp.169-229;

Kaufmann, Daniel and Pedro C. Vicente (2005), *Legal Corruption*, World Bank Institute, Governance and Anti-Corruption, Working Paper;

Klitgaard, Robert (1988), *Controlling Corruption*, University of California Press;

Krueger, Anne O. (1974), *The Political Economy of the Rent-Seeking Society*, American Economic Review 64, pp. 291-303;

Lam, Ricky and Leonard Wantchekon (2002), *Political Dutch Disease*, New York University, mimeograph;

Leininger, Wolfgand and Chun-Lei Yang (1994), *Dynamic Rent-Seeking Games*, Games and Economic Behavior 7, pp. 406–427;

Leite, Carlos and Jens Weidmann (1999), *Does Mother Nature Corrupt? Natural Resources, Corruption, and Economic Growth*, IMF, Working Paper 85;

Mauro, Paolo (1995), *Corruption and Growth*, Quarterly Journal of Economics, 110(3), pp. 681-712;

Mehlum, Halvor, Karl Moene and Ragnar Torvik (2006), *Institutions and the Resource Curse*, The Economic Journal, 116(508), pp. 1-20;

Myerson, Roger B. (1993), *Effectiveness of Electoral Systems for Reducing Government Corruption: A Game-Theoretic Analysis*, Games and Economics Behavior, 5, pp. 118-132;

Olken, Benjamin (2005), *Monitoring Corruption: Evidence from a Field Experiment in Indonesia*, NBER, Working Paper 11753;

Rasinski, Kenneth, Lance J. Rips, and Roger Tourangeau (2000), *The Psychology of Survey Response*, Cambridge University Press;

Robinson, James A., Ragnar Torvik and Thierry Verdier (2006), *Political Foundations of the Resource Curse*, Journal of Development Economics, forthcoming;

Rose-Ackerman, Susan (1978), *Corruption: A Study in Political Economy*, Academic Press;

Sachs, Jeffrey D. and Andrew M. Warner (1995), *Natural Resource Abundance and Economic Growth*, NBER, Working Paper 5398;

Seibert, Gerhard (1999), *Comrades, Clients and Cousins: Colonialism, Socialism and Democratization in São Tomé and Príncipe*, Leiden University, CNWS Publications 73;

Shadish, William R., Thomas D. Cook, and Donald T. Campbell (2002), *Experimental and Quasi-Experimental Designs for Generalized Causal Infere*nce, Houghton Mifflin Company;

Tornell, Aaron and Philip R. Lane (1999), *The Voracity Effect*, American Economic Review, 89, pp. 22-46;

Torvik, Ragnar (2002), *Natural Resources, Rent Seeking and Welfare*, Journal of Development Economics, 67, pp. 455-470;

Tullock, Gordon (1980), *Efficient Rent-Seeking*, in J. M. Buchanan, R. D. Tollison, and G. Tullock (Eds.), *Toward a Theory of the Rent-Seeking Society*, Texas A&M University Press, pp. 97-112;

Vicente, Pedro C. (2006), *A Theory of Corruption, Political Exclusion, and Windfalls*, in M. Gradstein and K. Konrad (Eds.), *Institutions and Norms in Economic Development*, MIT Press, forthcoming.

Appendix A – Oil Chronology in São Tomé and Príncipe

All started in 1997, when for USD 5m the government signed a controversial exclusive deal with a US-based oil company (ERHC - Environmental Remediation Holding Corporation) for oil soundings and implied exploration to be carried out (this company announced the existence of oil in the maritime waters of STP but did not advance too credible estimates of reserves). In 1998 a joint-venture named STPETRO was constituted between the government and ERHC, and ExxonMobil was added to the deal, with preferential rights, later in that year (adding credibility to the announcements, given further realized soundings). This series of events is what we consider to be the source of our natural experiment.

Nevertheless, in 1999 ERHC started having financial problems and the STP government cancelled the contract with ERHC alleging that company had failed to pay the contract-signing bonus - this was also the result of the pressure of international institutions, which viewed the initial contract as clearly damaging for STP. It was then that Nigerian interests close to President Obasanjo took over ERHC. Sources connect this happening with the following steps: still in 1999 an important maritime demarcation accord with Nigeria was signed, leading in 2001 to an agreement on the joint São-Tomean-Nigerian exploration of several off-shore oil blocks ("Joint Development Zone", implying a 40%/60% split between STP and Nigeria).

First round auctions were opened in 2003 regarding the allocation of 9 out of 25 of these blocks (a total USD 506m was bided in 7 blocks, which represented 237% of the GDP of the country in 2003[60]). Following those happenings, in 2004, only one block was allocated, namely to a joint venture between ExxonMobil (which exercised its preferential rights) and Chevron (the highest bidder), with a small participation by a Nigerian consortium, which was inserted at the political level. Recent news (as of March 2006) on the results of the first well drilled in this block were optimistic (sources point to the possibility that it can hold more than 1 billion barrels). The second round of bidding, related to 5 of the first round blocks, happened in the end of 2004, with decisions in the first half of 2005, attributing shares to a number of small Nigerian companies – ERHC was among them, exercising its preferential rights. However, much controversy happened thereafter, with STP attorney general publishing an unsolicited report saying that round was "subject to serious procedural deficiencies and political manipulation"[61].

[60] GDP estimate for 2003 from CIA World Factbook.
[61] See the Finantial Times article, "Sao Tome - where the champagne swills in before the oil gushes out", March 25th, 2006.

Production is expected to start in the next years (recent estimates by the IMF point to 30 thousand barrels per day in 2013), but contract-signing revenues have already happened (USD 49m in 2005 relative to the first allocated block) and are expected in the short term.

Appendix B – Survey Sampling: Tables, Maps, and Charts

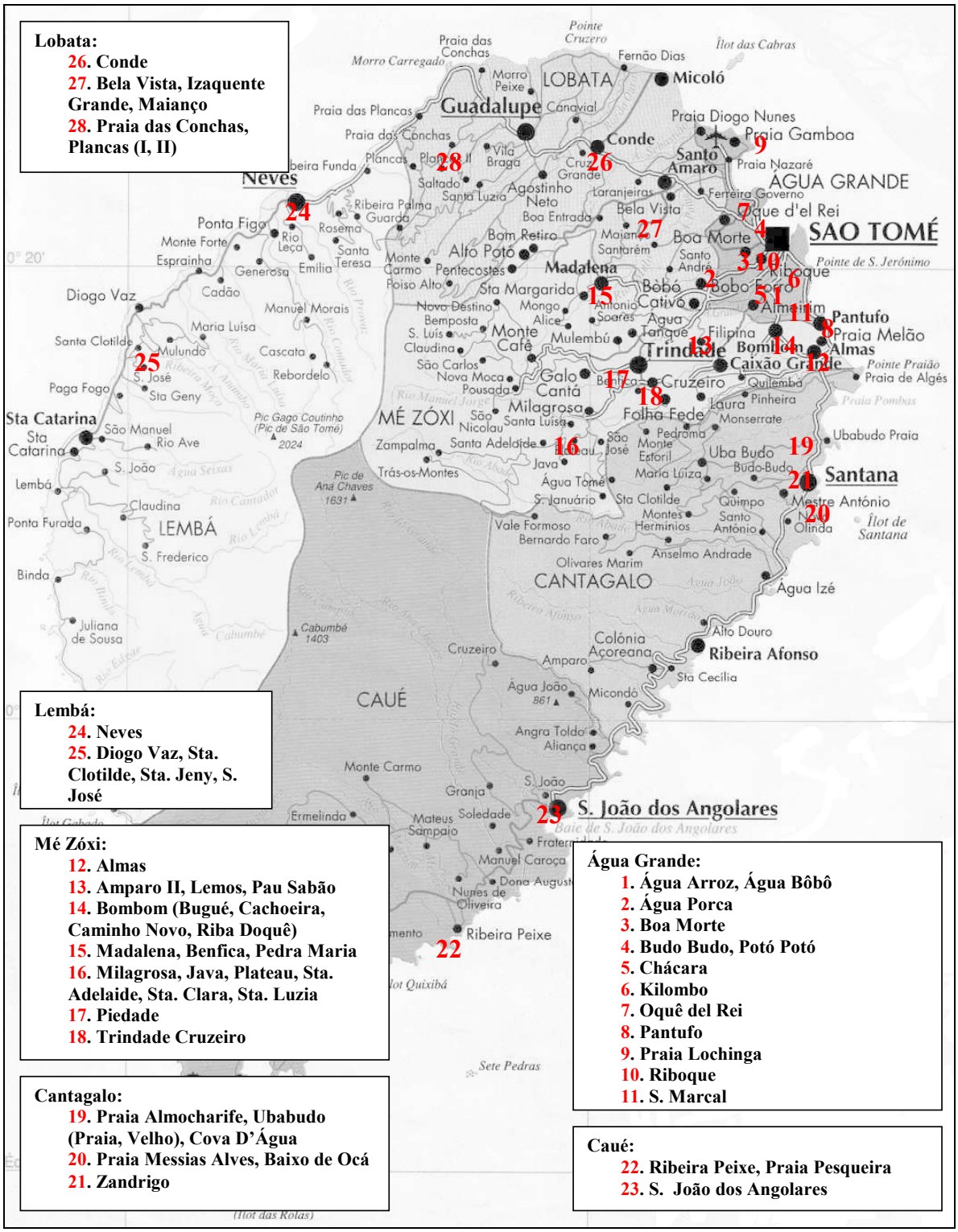
Table A1: Geographical Distribution of the Sample (STP)

	Sampled Households	Identified Non-respondents	Total
Água Arroz, Água Bôbô	25	2	27
Água Porca	25	2	27
Boa Morte	27	2	29
Budo Budo, Potó Potó	31	3	34
Chácara	25	2	27
Kilombo	29	1	30
Oquê del Rei	25	1	26
Pantufo	25	0	25
Praia Lochinga	26	0	26
Riboque	29	4	33
S. Marçal	24	4	28
District of Água Grande	*291*	*21*	*312*
Almas	23	2	25
Amparo II, Lemos, Pau Sabão	27	1	28
Bombom (Bugué, Cachoeira, Caminho Novo, Riba Doquê)	22	4	26
Madalena, Benfica, Pedra Maria	25	0	25
Milagrosa, Java, Plateau, Sta. Adelaide, Sta. Clara, Sta. Luzia	29	4	33
Piedade	19	2	21
Trindade Cruzeiro	26	7	33
District of Mé-Zóxi	*171*	*20*	*191*
Praia Almocharife, Ubabudo (Praia, Velho), Cova D'Água	27	3	30
Praia Messias Alves, Baixo de Ocá	28	5	33
Zandrigo	32	8	40
District of Cantagalo	*87*	*16*	*103*
Ribeira Peixe, Praia Pesqueira	32	1	33
S. João dos Angolares	29	3	32
District of Caué	*61*	*4*	*65*
Neves	24	0	24
Diogo Vaz, Sta. Clotilde, Sta. Jeny, S. José	33	2	35
District of Lembá	*57*	*2*	*59*
Conde	25	5	30
Bela Vista, Izaquente Grande, Maianço	25	3	28
Praia das Conchas, Plancas (I, II)	25	1	26
District of Lobata	*75*	*9*	*84*
Sto. António	46	5	51
Porto Real (Recta, Town), Bela Vista, S. Joaquim, Sto. António II	53	1	54
District of Príncipe	*99*	*6*	*105*
Total	***841***	***78***	***919***

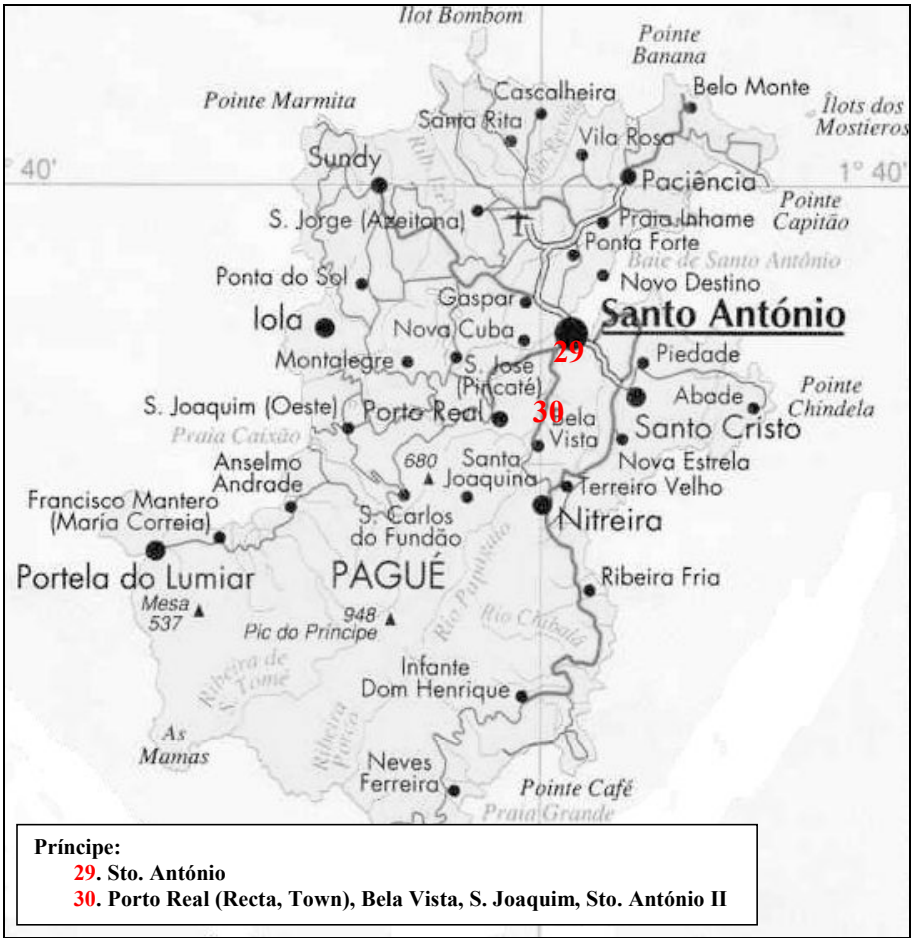
Table A2: Geographical Distribution of the Sample (CV)

	Sampled Households	Complete Interviews	Identified Non-Respondents	Total
(Praia) Plateau, Chã de Areia, Achada Santo António	55	46	20	75
(Praia) Achada Santo António	34	32	6	40
(Praia) Tira Chapéu	26	26	4	30
(Praia) Achada Eugénio Lima	48	40	6	54
(Praia) Achadinha	34	30	13	47
(Praia) Vila Nova	37	34	7	44
(Praia) Safende	33	29	4	37
(Praia) São Martinho Grande	54	52	3	57
(Tarrafal) Figueira Muita	28	27	2	30
(Tarrafal) Tarrafal – Codje Bicho	30	28	7	37
(São Miguel) Achada Monte	32	28	6	38
(Santa Catarina) Boa Entrada, Poilão da Boa Entrada	31	27	6	37
(Santa Catarina) Assomada – Trás de empa	35	35	4	39
(Santa Catarina) Chão Tavares, Cruz Grande	28	27	0	28
(Santa Cruz) Pedra Badejo – Achada Fátima	31	29	0	31
(Santa Cruz) Achada Fazenda	35	34	3	38
(Santa Cruz) São Jorge, Ribeirão Galinha	27	26	2	29
(São Domingos) Cutelo Gomes, Mato Afonso, Mato Martins	31	29	2	33
Santiago Island	*629*	*579*	*95*	*724*
(São Vicente) Mindelo - Morada	30	25	42	72
(São Vicente) Mindelo – Madeiralzinho, Chã de Alecrim	37	37	22	59
(São Vicente) Mindelo – Monte Sossego	42	39	14	56
(São Vicente) Mindelo – Bela Vista	38	36	3	41
(São Vicente) Mindelo – Fonte Inês, Espia	33	33	4	37
São Vicente Island	*180*	*170*	*85*	*265*
(Ribeira Grande) Fajã Domingas Benta, Chã de Enrique, Lombo	36	34	1	37
(Ribeira Grande) Chã de Pedras – Pia de Cima, Várzea, Pia de	38	37	1	39
(Porto Novo) Porto Novo	36	34	5	41
(Paúl) Ribeira de Janela, Pontinha de Janela	38	37	0	38
Santo Antão Island	*148*	*142*	*7*	*155*
(Mosteiros) Queimada Guincho	38	37	1	39
(São Filipe) São Filipe - Fonte Aleixo, Cobon	37	37	6	43
(São Filipe) Curral Grande	34	32	3	37
Fogo Island	*109*	*106*	*10*	*119*
Total	***1066***	***997***	***197***	***1263***

Map A1: São Tomé Island (STP) Sampled Census Areas

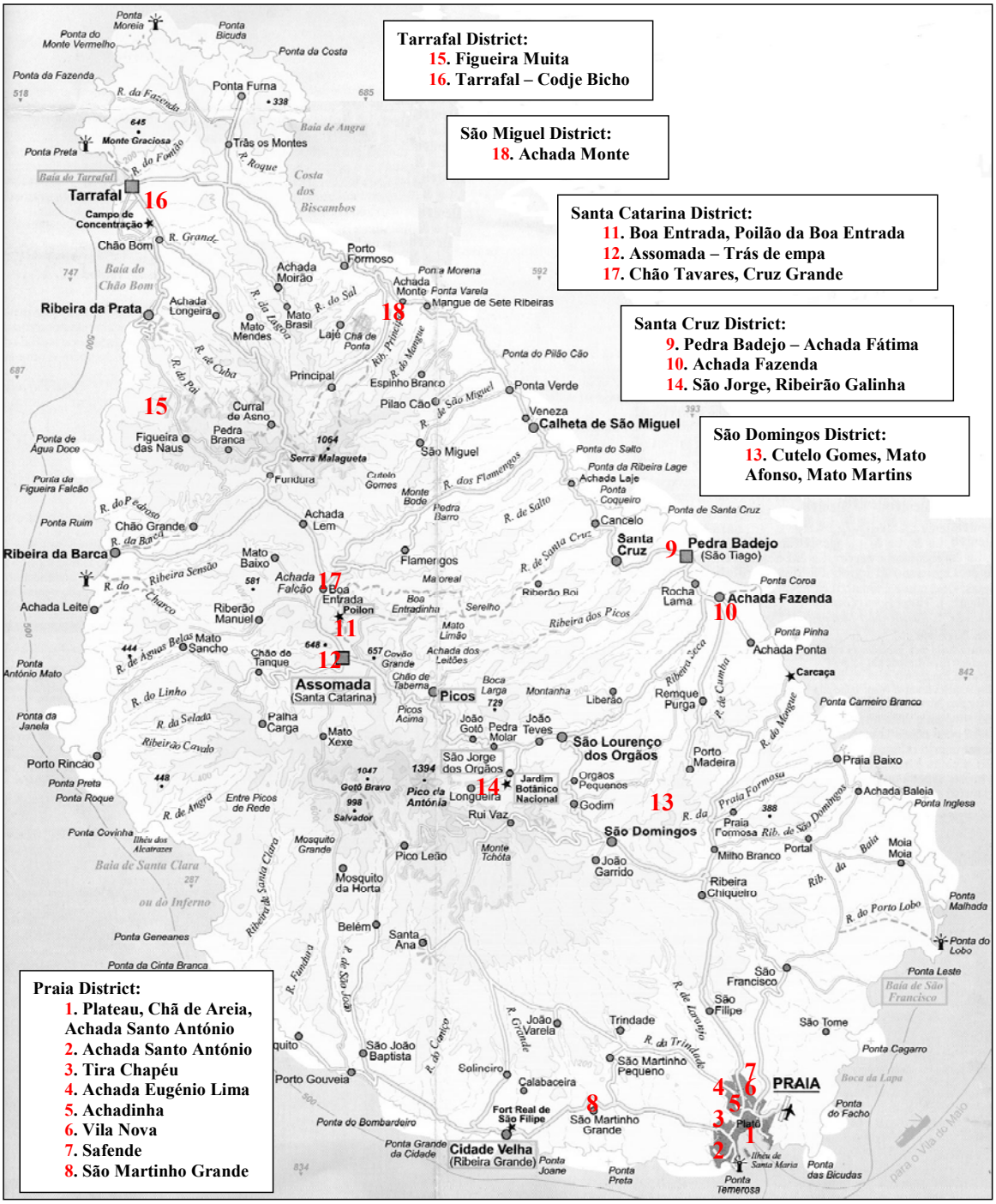


Map A2: Príncipe Island (STP) Sampled Census Areas

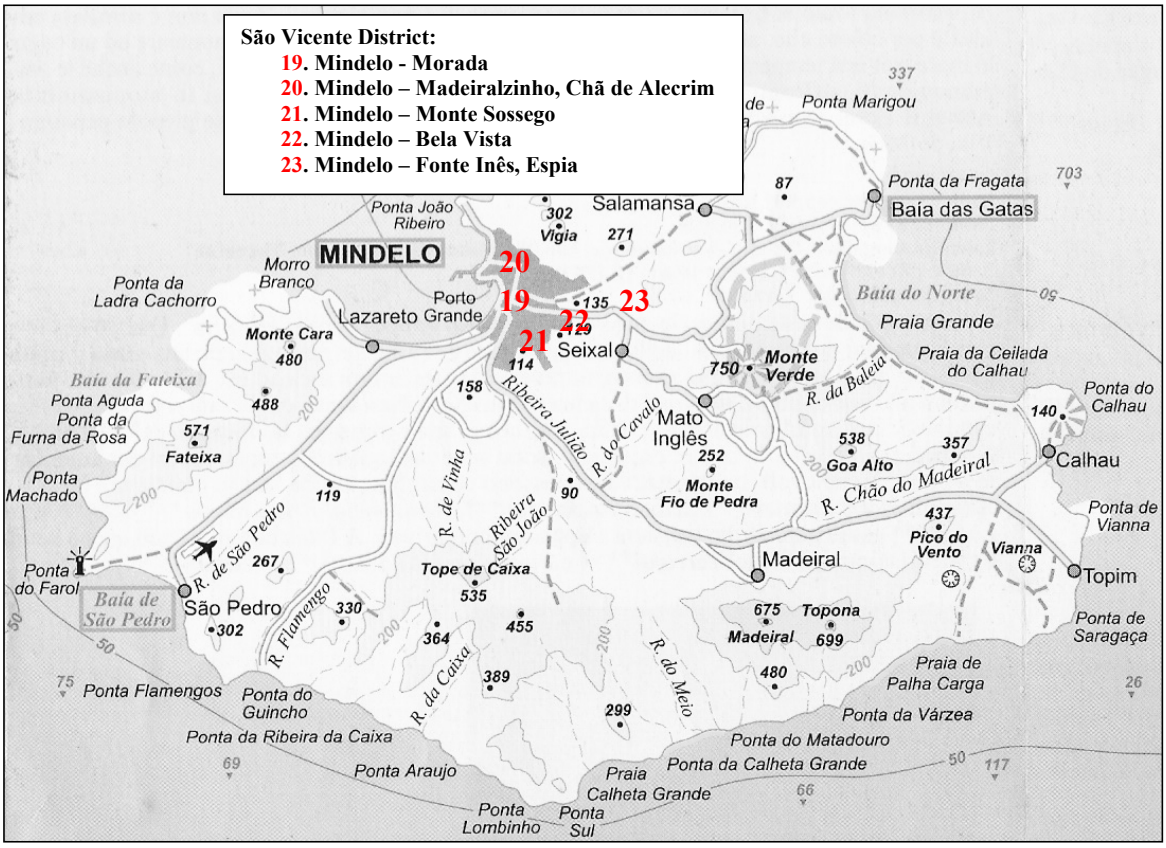


Príncipe:
29. Sto. António
30. Porto Real (Recta, Town), Bela Vista, S. Joaquim, Sto. António II

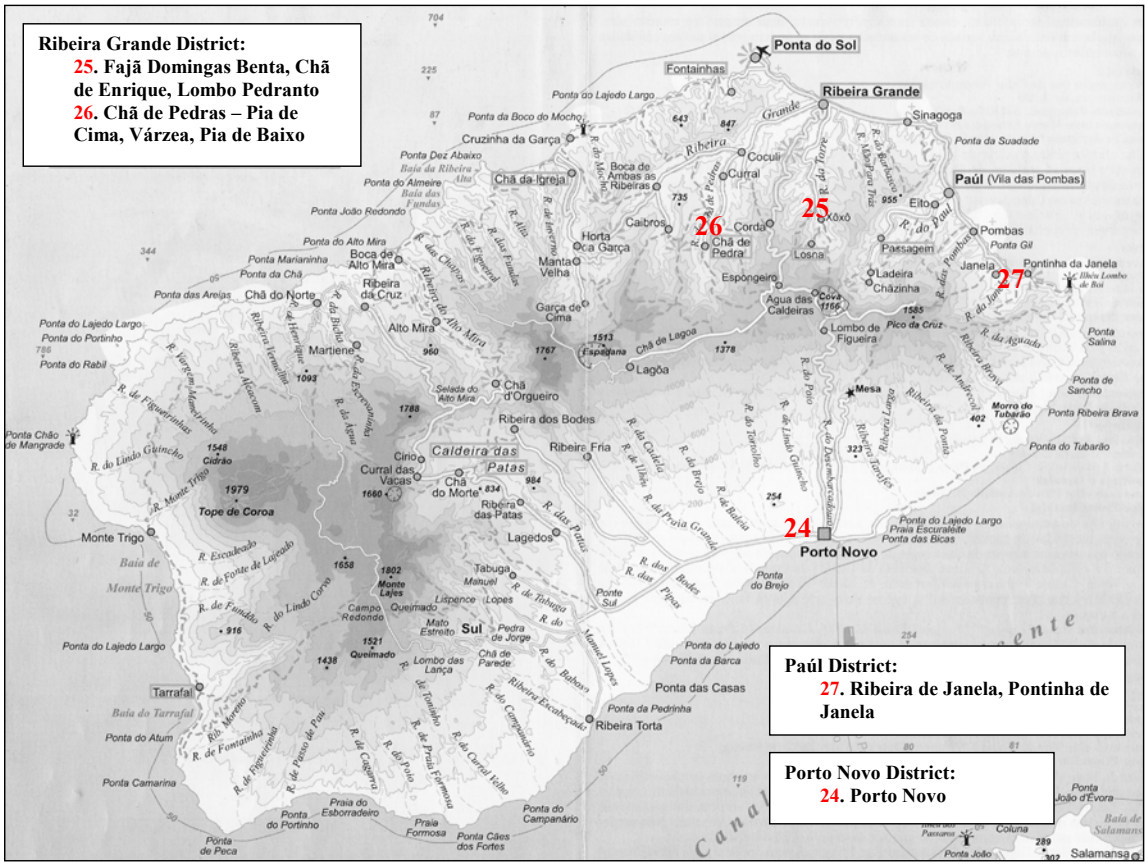
Map A3: Santiago Island (CV) Sampled Census Areas



Map A4: São Vicente Island (CV) Sampled Census Areas



Map A5: Santo Antão Island (CV) Sampled Census Areas



Map A6: Fogo Island (CV) Sampled Census Areas

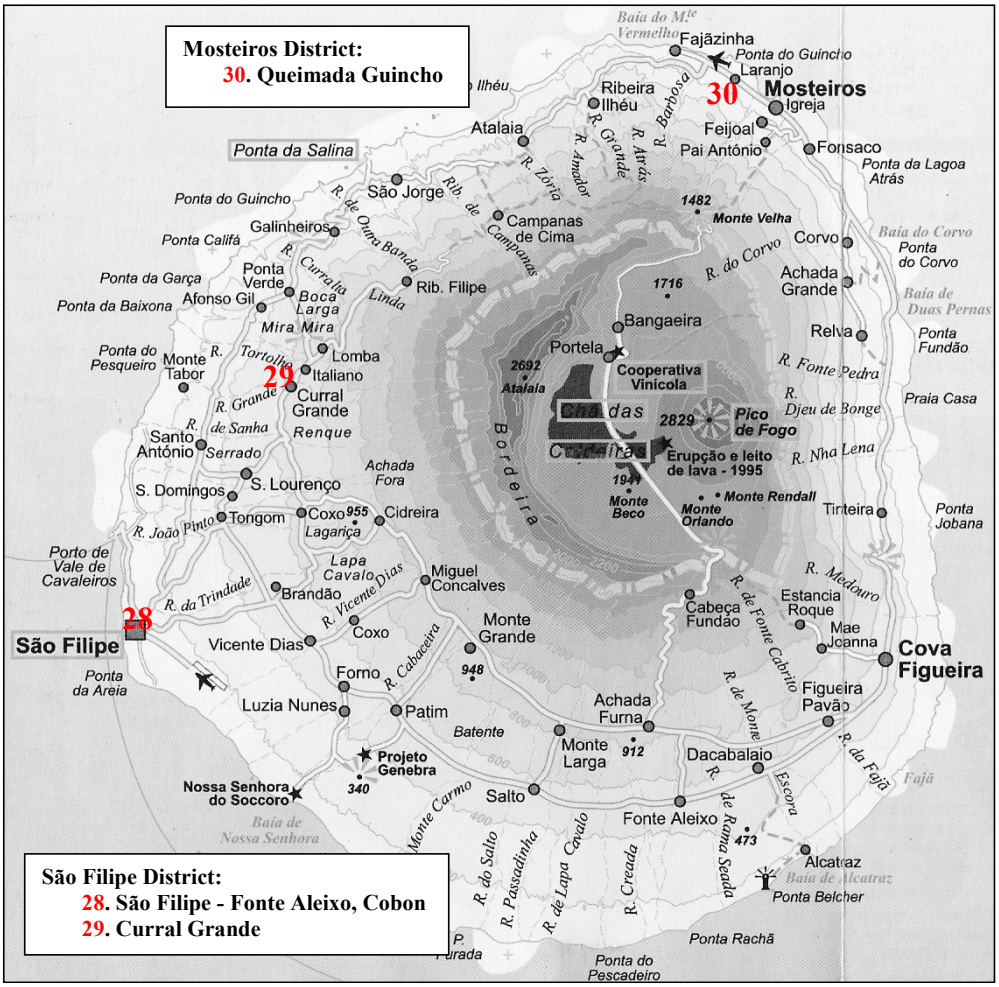
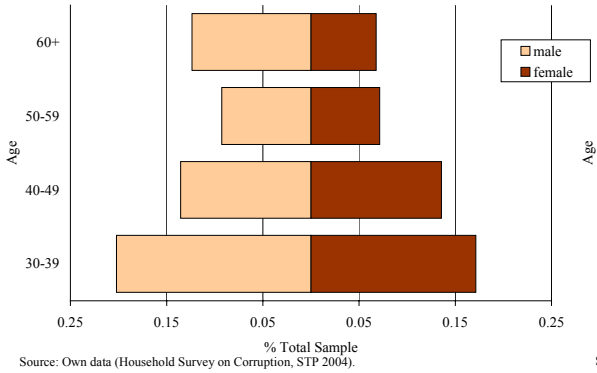
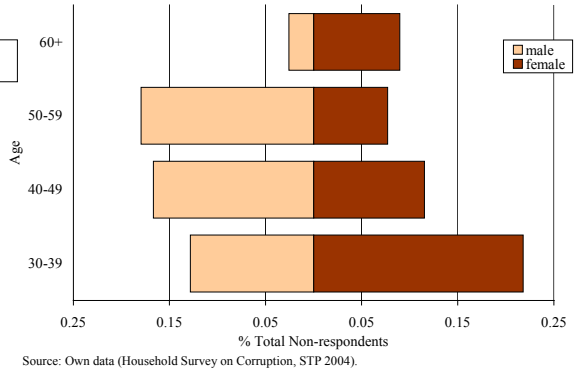


Chart A1: Age Pyramid - Sample - STP



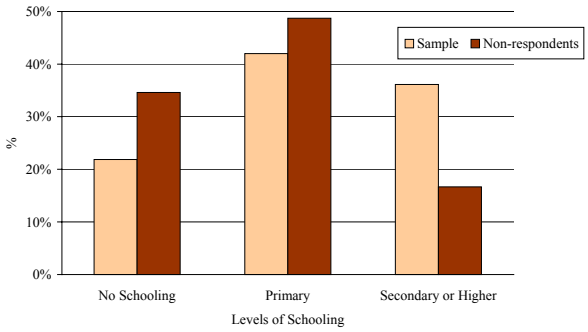
Source: Own data (Household Survey on Corruption, STP 2004).

Chart A2: Age Pyramid - Non-respondents - STP



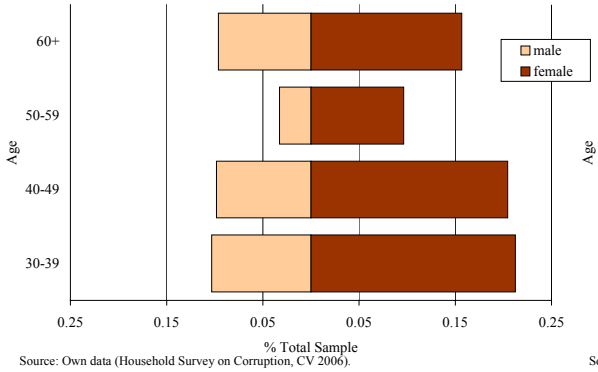
Source: Own data (Household Survey on Corruption, STP 2004).

Chart A3: Education (Sample vs Non-respondents) - STP



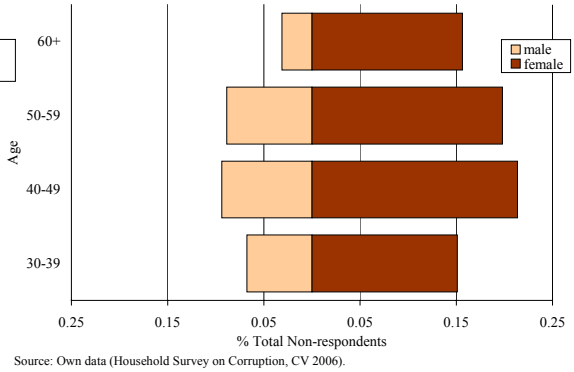
Source: Own data (Household Survey on Corruption, STP 2004).

Chart A4: Age Pyramid - Sample - CV



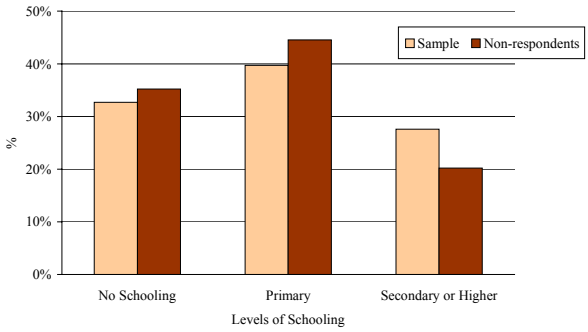
Source: Own data (Household Survey on Corruption, CV 2006).

Chart A5: Age Pyramid - Non-respondents - CV



Source: Own data (Household Survey on Corruption, CV 2006).

Chart A6: Education (Sample vs Non-respondents) - CV



Source: Own data (Household Survey on Corruption, CV 2006).

Appendix C - Survey Instrument Questions[62]

Table A3.1: Phrasing of the Main Questions - Specific Accessing Questions

Areas	Ref.	Question	Concerning Time	Scale
Courts	CL.0	Have you or a member of your household had contact with the courts of law (e.g.: as a witness, as the complainer or the subject of a process)?	All Periods of Interest	Yes - No
Police	P.0	Have you or a member of your household had contact with the police (ex: were/was fined, complainer or subject of the complaint)?	All Periods of Interest	Yes - No
Subsidies / Supplier Positions Recipients	SS0a	Have you or your firm supplied or applied to supply your products or services to entities of the state?	All Periods of Interest	Yes - No
	SS0b	Have you or your firm received or applied to receive any state subsidy?	All Periods of Interest	Yes - No
Licenses	L.0	Have you or the firm where you worked obtained or tried to obtain from the state any type of license for the exercise of your professional activity?	All Periods of Interest	Yes - No
State Jobs	SJ.0	Have you or a member of your household tried to get a job in the state?	All Periods of Interest	Yes - No
Customs	C.0	Have you or a member of your household had any contact with the customs of STP?	All Periods of Interest	Yes - No

Table A3.2: Phrasing of the Main Questions: Perceived Corruption (Influence)

Areas	Ref.	Question	Concerning Time	Scale
Health Care	HC.Main	In the Sao Tomean reality of public health services, what has been the need to: Know someone who works there?	All Periods of Interest and a PLACEBO Period	Not Necessary - Necessary (7-grade scale)
Education	E.Main	In the Sao Tomean reality, when passing annual exams in primary and secondary schools, what has been the need for: The student to have quality/merit?	All Periods of Interest	Not Necessary - Necessary (7-grade scale)
Scholarships	S.Main	In the Sao Tomean reality, when allocating scholarships for higher education abroad, what has been the need to: Know someone important?	All Periods of Interest	Not Necessary - Necessary (7-grade scale)
Courts	CL.Main	In the Sao Tomean reality of the processes in the courts of law, what has been the need to: Know someone important?	All Periods of Interest	Not Necessary - Necessary (7-grade scale)
Police	P.Main	In the Sao Tomean reality of the treatment offered by the police, what has been the need to: Knowing someone important?	All Periods of Interest	Not Necessary - Necessary (7-grade scale)
Subsidies / Supplier Positions Recipients	SS.Main	In the Sao Tomean reality, in the choice of state suppliers and/or subsidy recipients, what has been the need for: The candidates to be competent?	All Periods of Interest	Not Necessary - Necessary (7-grade scale)
Public Infrastructures	PI.Main	In the Sao Tomean reality, in the choice of public infrastructures construction in STP (public buildings, roads, schools, health centers), what has been the importance given by politicians to the needs of the population in general?	All Periods of Interest	Not Important - Important (7-grade scale)
Licenses	L.Main	In the Sao Tomean reality of the public services of licensing and registration, what has been the need to: Know someone who works there?	All Periods of Interest	Not Necessary - Necessary (7-grade scale)
State Jobs	SJ.Main	In the Sao Tomean reality of allocation of jobs in the state, what has been the need to: Know someone important?	All Periods of Interest	Not Necessary - Necessary (7-grade scale)
Customs	C.Main	In the Sao Tomean reality of functioning of the customs service, what has been the need to: Know someone important?	All Periods of Interest	Not Necessary - Necessary (7-grade scale)
Vote Buying	VB.1	In the electoral campaigns, which parties have distributed gifts in your town (do not consider campaign material as posters, t-shirts, stickers,...)?	*L1991/P1991; L1994/P1996/L1998; P2001/L2002	"Large" Parties/Candidates, "Small" Parties/Candidates, None
	VB.Main	In the Sao Tomean reality, for the voting decision, what has been the importance of: Gifts/favors offered by politicians?	1991-1997 and 2000-present	Not Important - Important (7-grade scale)

Notes: *L stands for Parliamentary Elections, and P stands for Presidential Elections.

[62] Note that all questions in this section concern specifically the STP survey. However all were asked in CV (with the exact same phrasing except for references to the country and different election years), with the exception of oil-related

Table A3.3: Phrasing of the Main Questions: Perceived Corruption (Bribery)

Areas	Ref.	Question	Concerning Time	Scale
Health Care	HC.Bribe	In the Sao Tomean reality of public health services, what has been the need to: Offering bribes?	2000-present only	Not Necessary - Necessary (7-grade scale)
Education	E.Bribe	In the Sao Tomean reality, when passing annual exams in primary and secondary schools, what has been the need for: Offering bribes?	2000-present only	Not Necessary - Necessary (7-grade scale)
Courts	CL.Bribe	In the Sao Tomean reality of the processes in the courts of law, what has been the need to: Offering bribes?	2000-present only	Not Necessary - Necessary (7-grade scale)
Police	P.Bribe	In the Sao Tomean reality of the treatment offered by the police, what has been the need to: Offering bribes?	2000-present only	Not Necessary - Necessary (7-grade scale)
Subsidies / Supplier Positions Recipients	SS.Bribe	In the Sao Tomean reality, in the choice of state suppliers and/or subsidy recipients, what has been the need for: Offering bribes?	2000-present only	Not Necessary - Necessary (7-grade scale)
Licenses	L.Bribe	In the Sao Tomean reality of the public services of licensing and registration, what has been the need to: Offering bribes?	2000-present only	Not Necessary - Necessary (7-grade scale)
Customs	C.Bribe	In the Sao Tomean reality of functioning of the customs service, what has been the need to: Offering bribes?	2000-present only	Not Necessary - Necessary (7-grade scale)

Table A3.4: Phrasing of the Main Questions - Optimism and Other Opinion/Psychological Questions

Areas	Ref.	Question	Concerning Time	Scale
Optimism	Pessim.1	"Good times were those when you were young".	NA	Agree - Disagree (7-grade scale)
	Pessim.2	The future of STP will be better than the present.	NA	Agree - Disagree (7-grade scale)
Other Opinion / Psychological Questions	Oil.Rev.Expec	The revenues for STP from oil exploration are going to be, following your expectation:	NA	Small - Large (7-grade scale)
	Oil.Population	You think the announcement of the existence of oil was, for the population of STP as a whole:	NA	Negative - Positive (7-grade scale)
	Rely.1	How much comfortable have you felt during the interview?	NA	Not Comfortable - Comfortable (7-grade scale)

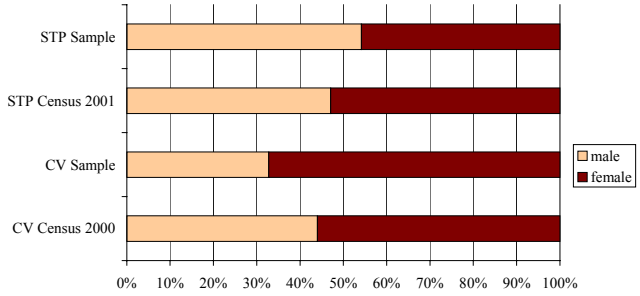
questions.

Appendix D - Descriptive Statistics of the Sample

In this section we present the samples in terms of their characterizing features, which yield a comparison between the two countries under analysis. In some key demographic indicators we also provide a comparison with 2001 (STP)/2000 (CV) census data[63].

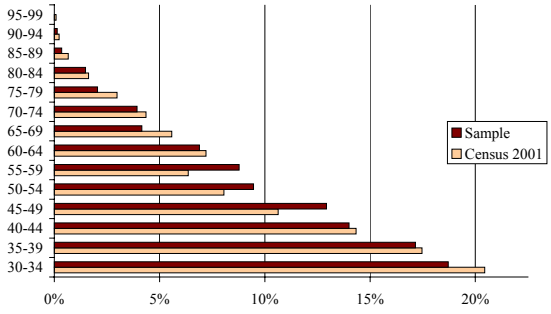
We can therefore conclude the following regarding the various essentially demographic indicators we collected that will be used ahead as controlling variables:

Chart A7: Gender (Aged 30+) - Sample vs Census, STP vs CV



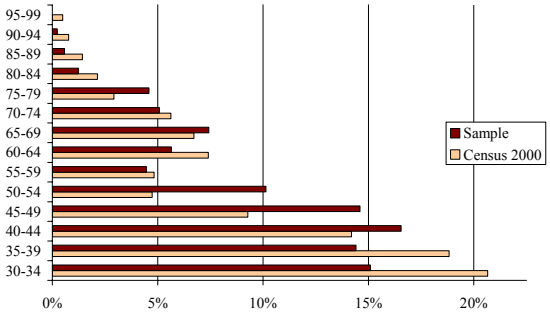
Source: Own data (Household Surveys on Corruption, STP 2004, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures). STP Census 2001 from National Statistics Office of STP and CV Census 2000 from National Institute of Statistics of CV.

Chart A8a: Age Distribution (30+) - Sample vs Census - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). STP Census 2001 from National Institute of Statistics of STP.

Chart A8b: Age Distribution (30+) - Sample vs Census - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures). CV Census 2001 from National Institute of Statistics of CV.

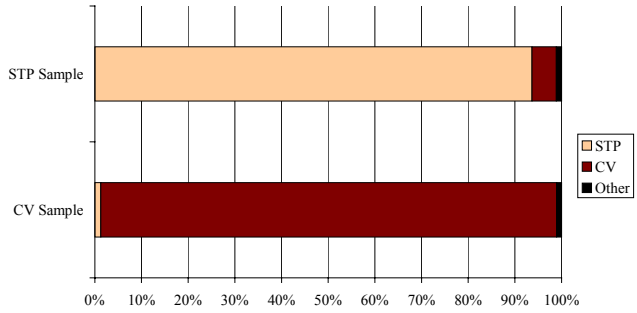
Gender and Age (respectively Charts A7 and A8): 46% female and 54% male respondents in STP denote a clear propensity for men to be called to respond the questionnaire (as heads of the household) in STP. However, in CV we see the opposite tendency with 67% women and 33% men, which may be due to an important emigrant tradition, which led women to take representation responsibilities in the household. Census data point to a slight propensity towards women in both countries (higher in CV); regarding age, the average respondents had 47 (STP)

[63] These were made available by the National Institutes of Statistics of both countries. A note should be stressed however that in many indicators our data is not representative of the population, since our sampling method was aimed at being representative of households including at least one member aged 30 years or more with significant experience in the country - some differences to census data are therefore anticipated and realized.

and 49 (CV) years old, with the youngest members of both samples being apparently undersampled (relative to older categories, according to the census data)[64].

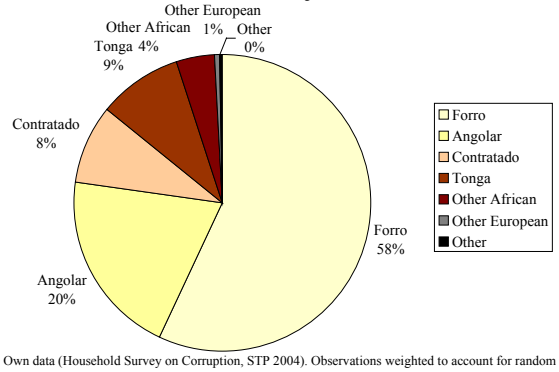
Nationality and Ethnicity/Origin (respectively Charts A9 and A10): 94% of our STP sample was said to be Sao-Tomean, 5% was identified as Cape Verdean; we found a similar but more uniform pattern in CV, with 98% of respondents of Cape Verdean nationality; regarding Ethnicity in STP, a majority is considered to be "Forro" (58%), with the second larger group being that of "Angolar" denomination (20%)[65]. CV does not have clear ethnic groups, with the closest distinctions being based on island of birth[66] – the majority of our sample was born in Santiago island (54%).

Chart A9: Nationality - STP vs CV



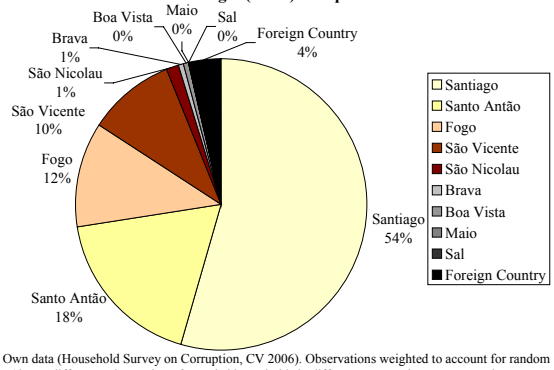
Source: Own data (Household Surveys on Corruption, STP 2004, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures).

Chart A10a: Ethnic Composition - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).

Chart A10b: Origin (Birth) Composition - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures).

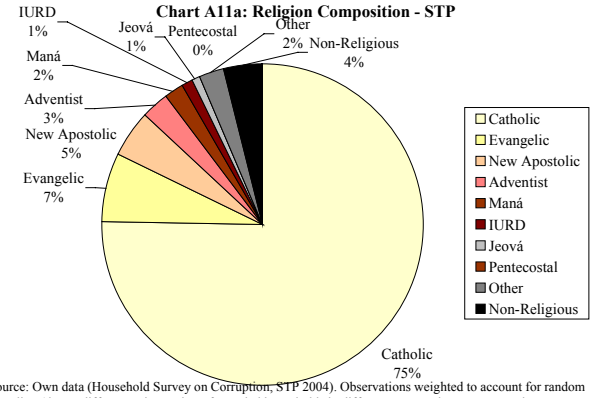
[64] This may be due to a measured (in our surveys) higher employment rate among the younger.

[65] Following Seibert (1999), "Forro" are the descendents of the first free slaves in STP (who began to arise in the XVIth century), while "Angolar" are descendents of runaway slaves - said, in a legend, to be from Angola, that cultivated a sense of autonomy/isolation - this group based its subsistence on fishing (contrary to the majority of the population, who has always been living on agriculture). The other ethnic groups are connected to more recent (XXth century) waves of African (namely Cape Verdean) immigrants, who were contracted to work in agriculture ("Contratado"): the "Tonga" group refers to their descendents.

[66] The ethnic denominations used in CV are mainly "Sampadjudo" and "Badiu", and are associated with origins from the northern and southern sets of islands of the country, respectively.

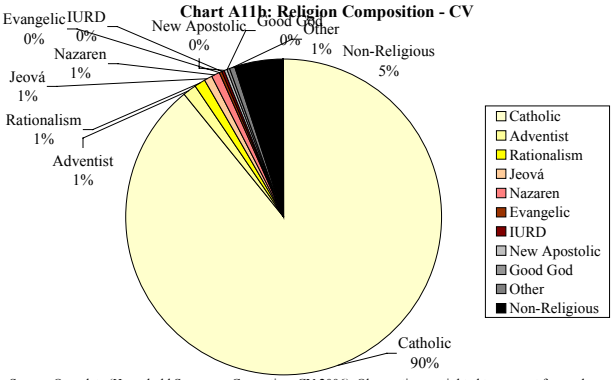
Religion (Charts A11): clear majorities of 75% (STP) and 90% (CV) were said to be Catholic; however, some differences are identified across the two countries in terms of prevalence of other Christian religions (more popular in STP).

Chart A11a: Religion Composition - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).

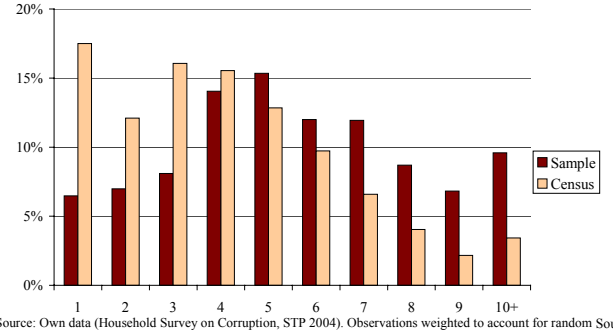
Chart A11b: Religion Composition - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures).

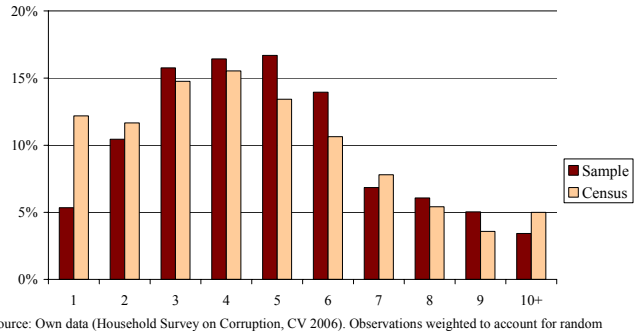
Household Size, Number of Children, and Children Schooling (Charts A12, A13, and A14): the average households had 5.8 (STP) and 4.9 (CV) elements; Cape Verdean households generally have a lower number of members; differences to census data arise in both STP and CV (where small households are more frequent)[67]; the average number of children was 5.6 in STP while in CV it was 4.4, with only close to 5% of the respondents in both countries not having children; average children school attendance reported was 91% (STP) and 95% (CV) in primary school and 63% (STP) and 59% (CV) in secondary school - these are as percentage of the children with 18 or lower age, and older than 5 (for primary school) or 9 (for secondary school) - with significant geographic variation (highest values in STP are for the capital district, Água Grande; interestingly, in CV the capital does not rank top in these indicators).

Chart A12a: Household Size - Sample (Respondents Aged 30+) vs Census (All) - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Census 2001 - National Institute of Statistics of STP.

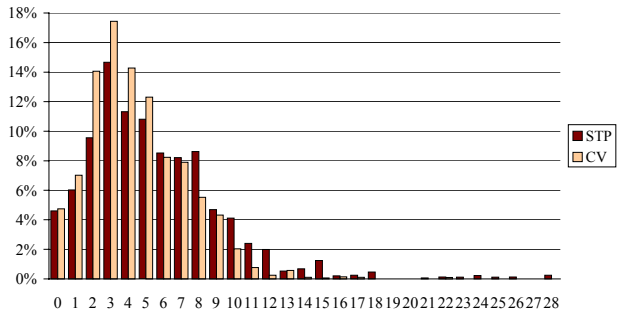
Chart A12b: Household Size - Sample (Respondents Aged 30+) vs Census (All) - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures). Census 2000 - National Institute of Statistics of CV.

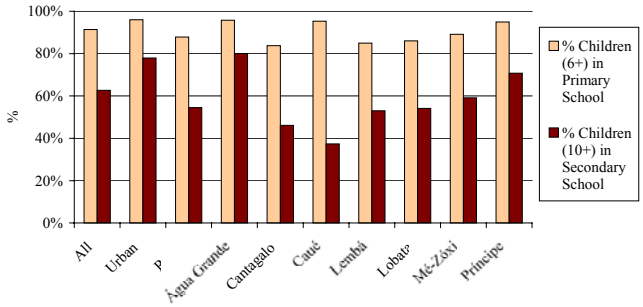
[67] This may be due to the age profile of interviewed families (where we had to find someone with more than 29 years of age).

Chart A13: Number of Children - STP vs CV



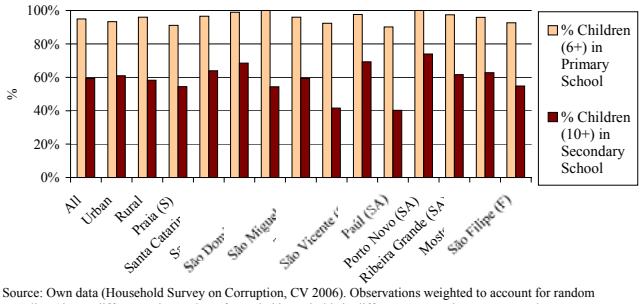
Source: Own data (Household Survey on Corruption, STP 2004, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures).

Chart A14a: Children School Attendance (as % of Children =<18 Years of Age) - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Urban area is composed of all census areas visited in the district of Água Grande and of the census area of Trindade-Cruzeiro.

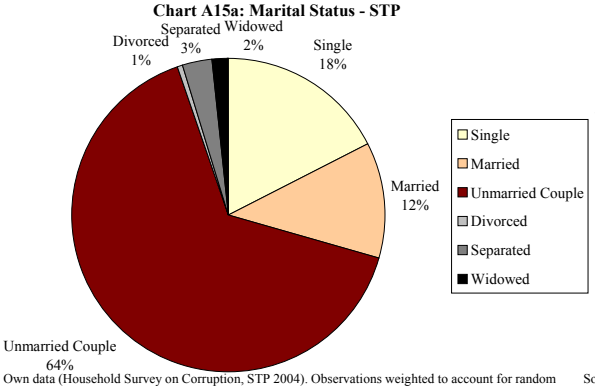
Chart A14b: Children School Attendance (as % of Children =<20 Years of Age) - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures). Urban area is composed of all census areas visited in the districts of Praia (with the exception of São Martinho Grande) and São Vicente (city of Mindelo). S - Santiago; SV - São Vicente; SA - Santo Antão; F - Fogo.

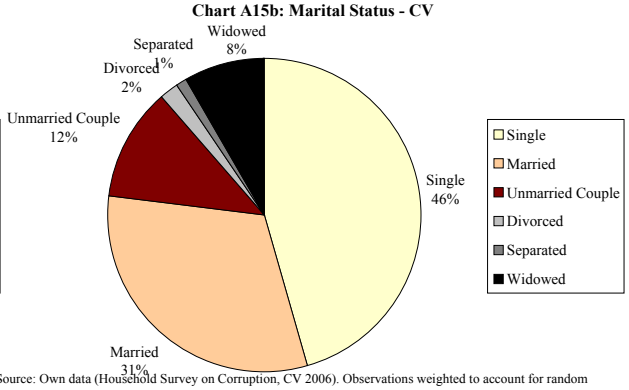
Marital Status (Charts A15): despite a strong Catholic influence, STP is also characterized by an informal regime in marriage with 65% of the respondents of the survey saying they are part of unmarried couples; in CV formal marriages are more popular, with single respondents being most prominent (which may be taken as a way to refer to unmarried relationships).

Chart A15a: Marital Status - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).

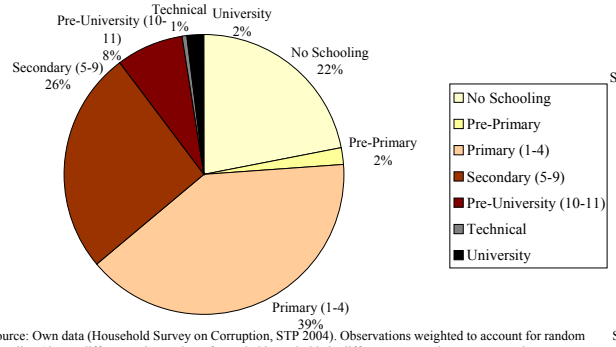
Chart A15b: Marital Status - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures).

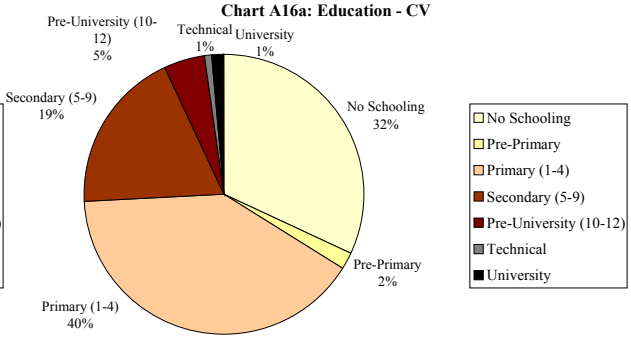
Respondents' Education (Charts A16): our data concern attendance of the given schooling levels (not necessarily its completion); this fact may explain the relatively low number of respondents with no schooling (22% in STP and 32% in CV); only 11% (STP) and 7% (CV) are reported to have more than 9 years of formal education[68].

Chart A16a: Education - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Each schooling category represents attendance of that level, not necessarily its completion.

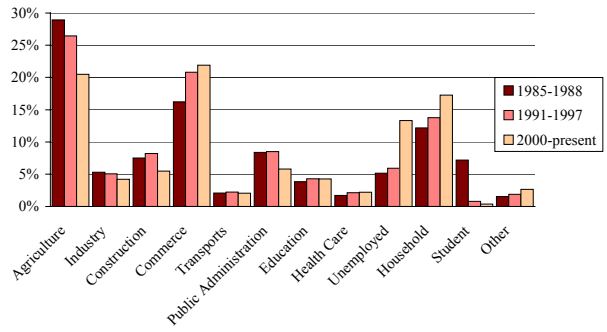
Chart A16a: Education - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures). Each schooling category represents attendance of that level, not necessarily its completion.

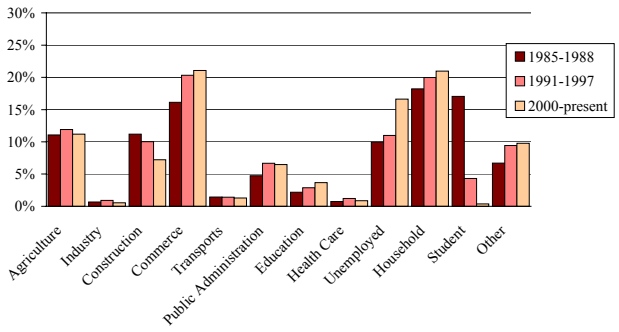
Occupation and Job Stability (Charts A17 and A18): the importance of agriculture and of stable jobs (respectively 21% and 65% in the period 2000-present) is clearly decreasing over time in STP, with that of commerce and unemployment increasing over the same 20 years (respectively 22% and 13% in the period 2000-present); in CV we can observe a similar trend in commerce and unemployment (however, the importance of agriculture has been smaller in this country than it has been in STP).

Chart A17a: Occupation - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).

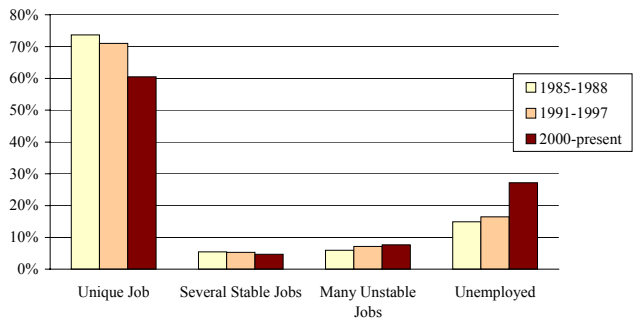
Chart A17b: Occupation - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures).

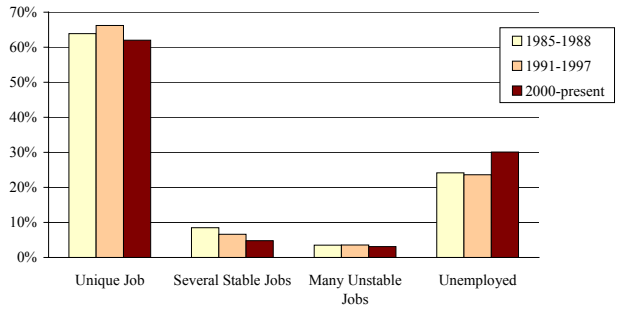
[68] Note that this is despite the fact that Cape Verdeans are widely seen as being highly educated for African standards; however, most of these educated Cape Verdeans reside abroad.

Chart A18a: Job Stability - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).

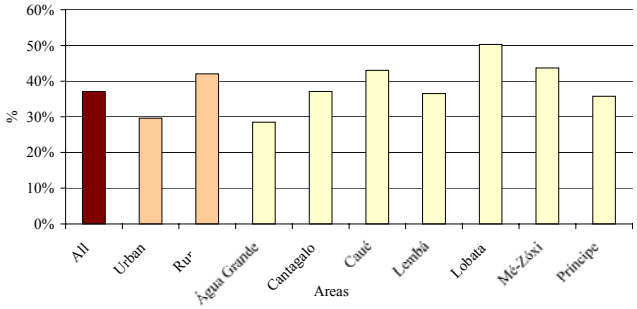
Chart A18b: Job Stability - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures).

Malaria Incidence (Chart A19): a very high number of 37% of the STP households reported to have had some form of malaria in the period 2000-present; however a clear difference appears between rural and urban areas (less affected by the disease)[69]. CV has a small prevalence of this disease.

Chart A19: % Households with Malaria 2000-2004 - STP



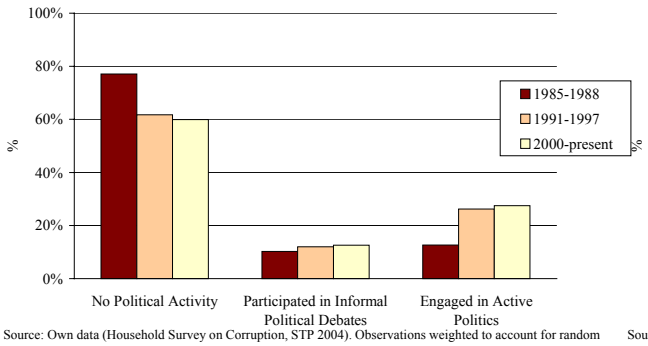
Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Urban area is composed of all census areas visited in the district of Água Grande and of the census area of Trindade-Cruzeiro.

Reported Political Activity and Preferences (Charts A20 and A21): reported political activity in STP clearly increased (by 17%) after democratization, with only a slight increase of 2% reported to have occurred after 1999; CV sees a more homogenous increase in reported political activity from democratization, without reaching the STP level of active participation (which was 28%, against the Cape Verdean 23%[70]); regarding political preferences[71] we look at support for the main political "families"; in STP these are Pinto da Costa/MLSTP - the dictatorship president and independence party, ADI/Trovoada - the second two-term president of

[69] This is despite the fact that, after a major health campaign, this disease was under control in the 70s - as reported in Hodges and Newitt (1988).
[70] This may be simply due to the fact that the absolute number of politically connected individuals, that run the state, may be similar in these countries, while population size is clearly different.
[71] It should be noted that our samples are not perfectly representative of the electorate (provided the sampling process was aimed at households with representatives aged 30+ and not at individuals aged 18+).

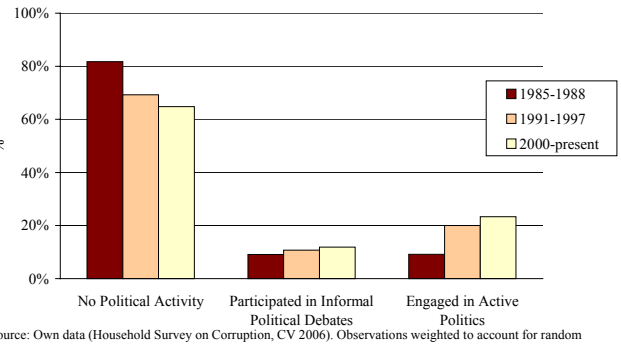
the country, from 1991 to 2001, democratically elected, and his supporting party created in 1992 - , and Fradique de Menezes/MDFM - the present president (2001-) and his supporting party created before the 2002 parliamentary elections: we can observe that the first and the third groups are the main forces in the 2004 political scene of the country; in CV, bi-partidarism is more rooted, with PAICV (the unique party after independence) and MPD (formed with democratization, ruling in the 90s) dividing the electorate and embedding the emergence of the most important politicians.

Chart A20a: Reported Political Activity - STP



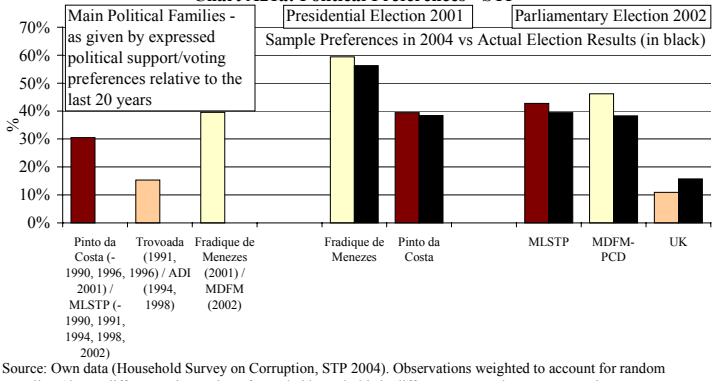
Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).

Chart A20b: Reported Political Activity - CV



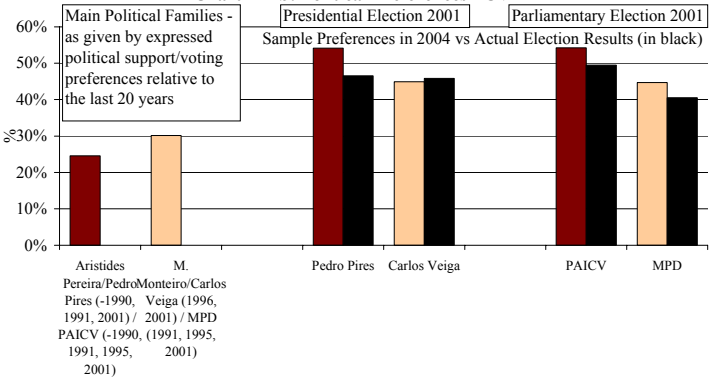
Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures).

Chart A21a: Political Preferences - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling). Note that regarding the main political families, while the first is exclusive of the second and third, the latter two may not be exclusive of each other - e.g. in 2001, Trovoada endorsed Fradique. Political Families: % of sample; Presidential, Parliament Elections: % of votes. UK coalition included ADI in 2002.

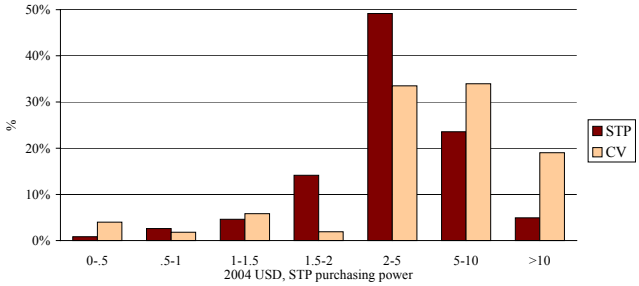
Chart A21b: Political Preferences - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures). Political Families: % of sample; Presidential, Parliament Elections: % of votes.

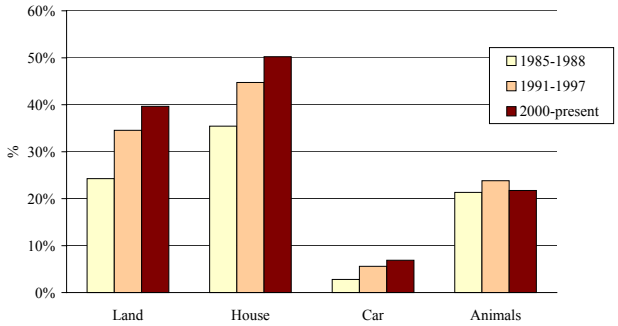
Household Expenditure and Property (Charts A22 and A23): almost 50% of respondents reported household expenditure per day to be between USD 2 and USD 5 in STP; using middle points at all classes of expenditure (USD125 was used for the USD100+ class), we have an average expenditure per capita of USD 0.98 in STP; in CV we can observe a distribution of expenditures which is shifted to the right (with USD 1.6 average expenditures per capita); ownership of property has been increasing from the late eighties (for land, houses, and cars in STP, and for land, houses, cars, and animals in CV) – note that CV has a higher rate of respondents owning houses, cars, and animals, but lower rate with regard to land.

Chart A22: Distribution of Household Expenditure per Day (2004 USD, STP purchasing power) - STP vs CV



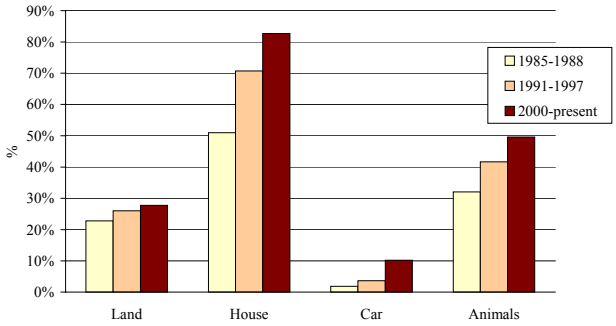
Source: Own data (Household Survey on Corruption, STP 2004, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and - for CV only - expenditures). Inflation Rate from Cape Verde Central Bank (May 04 to Jan 06), and Penn World Tables PPP CV-STP factor (1998-2000) were used to adjust CV raw data.

Chart A23a: Household Property - STP



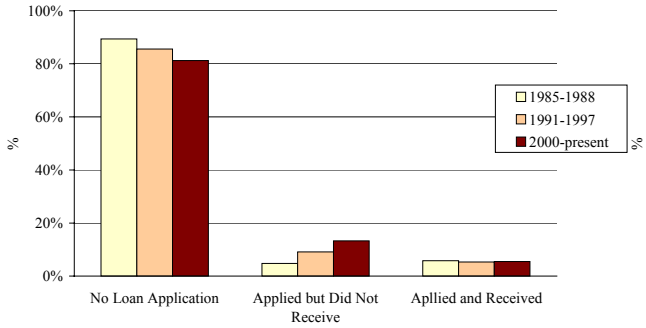
Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).

Chart A23b: Household Property - CV



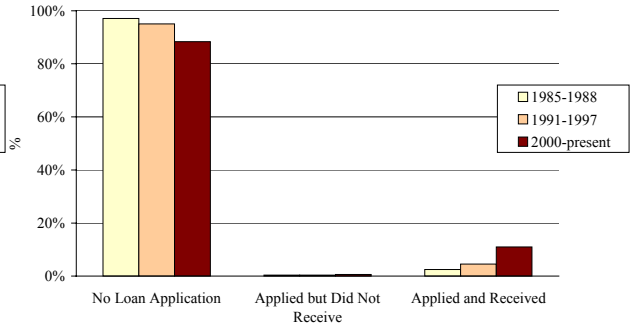
Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures).

Chart A24a: Credit Constraints - STP



Source: Own data (Household Survey on Corruption, STP 2004). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling).

Chart A24b: Credit Constraints - CV



Source: Own data (Household Survey on Corruption, CV 2006). Observations weighted to account for random sampling (due to differences in number of sampled households in different areas and to non-respondents characteristics - sex, age, schooling, and expenditures).

Credit Constraints (Charts A24): although an average of 85% of respondents to the STP survey reported no loan application in the last 20 years, the number of applications without success has been increasing (with a clear jump in the period 1997-99), pointing to an increase in credit constraints[72] (13% credit-applying credit-constrained households in 2000-present) - the

[72] The first jump in credit applications (1988-1991) may be explained with the wave of privatizations and market-oriented reforms that accompanied democratization; the second (1997-2000) may possibly be already associated with the oil discovery.

number of successful applications has been constant; in CV, it is the number of successful applications that has been increasing (with 11% in the latest period) with failed applications remaining constant and close to zero.

Appendix E – Treatment Variables

Table A4: Weights for Connectivity Indicator

Experience Categories	Number of Questions where Accessing Category Used	Weight
At least one child born	1	4.55%
At least one child in primary school	1	4.55%
At least one child in secondary school	2	9.09%
Occupation: Public Administration	5	22.73%
Occupation: Education	2	9.09%
Occupation: Health Care	1	4.55%
Main Occupation: Student in secondary school or higher education	2	9.09%
At least one member of the household ill	1	4.55%
At least one member of the household in court	1	4.55%
At least one member of the household in police	1	4.55%
Applied to supply state	1	4.55%
Applied to state subsidy	1	4.55%
Applied to license for professional activity	1	4.55%
At least one member of the household tried to get a job in the state	1	4.55%
At least one member of the household with experience with customs	1	4.55%

Note: The question on public infrastructures was not included provided the very few observations regarding that question.

Table A5: Correlations (Treatment Variables)

STP	dummyurban	dummypa	indexconnect
dummyurban	1	0.03	0.08
dummypa		1	0.58
indexconnect			1
CV	dummyurban	dummypa	indexconnect
dummyurban	1	0.12	0.07
dummypa		1	0.58
indexconnect			1

Appendix F – Baseline Internal Control Econometric Results

In Tables A6-A12 and A16, the first and second specifications regard the simple DID regressions where (respectively) urban and public administration respondents are the treatment groups; in the third specification we add the two treatment groups in the same regression; in the fourth specification we add psychological controls (analyzed in the end of section 6)[73]; in the fifth specification we put in demographic controls[74]; in the sixth to eighth specifications we use the Connectivity Indicator in a simple DID equation, add psychological controls and the demographic variables (respectively);

In Tables A13-A15, we present the same sequence as before apart from the fact that we do not use the public administration dummy variable - this is provided the respective dependent variables were only asked to respondents with occupation in the private sector at any period of interest of the questionnaire[75];

[73] We always use the pessimism variables in levels and differences (the placebo variable always enters the regressions; the most highly significant variable from Pessim.1, Pessim.2 and the index of the two with equal weight for each is also used); these range from 1-7. Regarding the other psychological controls, they are only used in levels and when significant; these range from 1-7 (increasing trust).

[74] The demographic controls are formed by at least one variable from each of the following sets:

- Basic Demographics: sex (dummy - 1 male), age, household size, number of children, STP nationality (dummy), catholic (dummy), non-religious (dummy);
- Ethnic Groups: forro, angolar, contratado, tonga (dummies);
- Marital Status: single, married, unmarried couple, widow (dummies);
- Health: malaria (0-2; 0: no; 1: yes for one period - 1991-1997 or 2000-2004; 2: yes for the two periods);
- Household Schooling: no schooling (dummy), schooling (1-7), children in primary school (% of children 6+ who went to primary school), children in secondary school (% of children 10+ who went to secondary school), fluency of the respondent (1-7);
- Occupation: agriculture, industry, construction, commerce, transports, public administration, education, health, housewife, unemployed - all 0-2 variables (0: no; 1: yes for one period; 2: yes for the two periods); job insecurity (0-6; sum of the ratings, 0-3, for each of the periods of interest);
- Political Preferences: political indifference (2-10; sum of the ratings, 1-5, for each period of interest); MLSTP/Pinto da Costa, ADI/Trovoada, MDFM/Fradique (dummies);
- Financial Variables: expenditure (1-7) (inflation rate from Cape Verde Central Bank - May 2004 to January 2006 – and Penn World Tables PPP CV-STP factor 1998-2000 were used to adjust CV data), expenditure per capita (1-11), loans (0-4; sum of the ratings, 0-2, for each period of interest), land, house, car, animal (0-2; 0: no; 1: yes for one period; 2: yes for the two periods), property (dummy).

Statistical significance is the criterion for the referred choice at the margin; when significant, several variables may appear for a certain set.

[75] However, some respondents, who have moved into the private sector (from public administration) or out from the public administration (into the private sector) were asked these questions. These cases are not numerous.

Table A6: Health Care Regressions

	Dependent Variable ------>		Perceived Corruption in Health Care							
Main Explanatory Variables	oil	coef	**0.45**	**0.64**	**0.39**	**0.31**	**0.29**	**0.39**	**0.4**	**0.41**
		std err	0.08***	0.07***	0.09***	0.6	0.61	0.14***	0.61	0.63
	urban	coef	**0.1**		**0.1**	**0.09**	**0.07**			
		std err	0.1		0.1	0.09	0.1			
	oil*urban	coef	**0.62**		**0.61**	**0.59**	**0.58**			
		std err	0.14***		0.14***	0.13***	0.13***			
	pub_adm_dummy	coef		**-0.3**	**-0.3**	**-0.35**	**-0.43**			
		std err		0.16*	0.16*	0.16**	0.16***			
	oil*pub_adm_dummy	coef		**0.62**	**0.6**	**0.57**	**0.57**			
		std err		0.23***	0.23***	0.22***	0.22***			
	index_connect	coef						**-0.38**	**-0.6**	**-0.71**
		std err						0.44	0.43	0.47
	oil*index_connect	coef						**1.49**	**1.5**	**1.28**
		std err						0.62**	0.61**	0.62**
Psychological Controls	placebo	coef				**4.52**	**4.43**		**4.5**	**4.55**
		std err				0.79***	0.80***		0.80***	0.83***
	oil*placebo	coef				**-0.62**	**-0.56**		**-0.83**	**-0.8**
		std err				1.11	1.12		1.13	1.17
	pessim1	coef				**-0.07**	**-0.08**		**-0.08**	**-0.08**
		std err				0.03**	0.03**		0.03**	0.03***
	oil*pessim1	coef				**0.08**	**0.08**		**0.08**	**0.08**
		std err				0.04*	0.04*		0.04*	0.04*
	oil_rev_expec	coef				**-0.07**	**-0.06**		**-0.08**	**-0.07**
		std err				0.03***	0.03**		0.03***	0.03**
Demographic Controls			**No**	**No**	**No**	**No**	**Yes**	**No**	**No**	**Yes**
	Constant		4.66	4.73	4.69	2.99	3.78	4.77	3.19	3.32
			0.06***	0.05***	0.06***	0.43***	0.51***	0.10***	0.44***	0.48***
	Number of Observations		1634	1632	1632	1532	1462	1634	1532	1434
	Adjusted R-squared		0.09	0.06	0.1	0.13	0.17	0.06	0.1	0.14

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A7: Education Regressions

	Dependent Variable ------>		Perceived Corruption in Education							
Main Explanatory Variables	oil	coef	**0.46**	**0.58**	**0.41**	**-1.11**	**-1.1**	**0.25**	**-1.17**	**-1.17**
		std err	0.10***	0.08***	0.10***	0.69	0.71	0.17	0.72	0.73
	urban	coef	**0.56**		**0.57**	**0.62**	**0.45**			
		std err	0.11***		0.11***	0.11***	0.12***			
	oil*urban	coef	**0.43**		**0.43**	**0.41**	**0.4**			
		std err	0.15***		0.15***	0.16***	0.16**			
	pub_adm_dummy	coef		**-0.11**	**-0.12**	**-0.11**	**-0.26**			
		std err		0.19	0.18	0.18	0.19			
	oil*pub_adm_dummy	coef		**0.54**	**0.53**	**0.51**	**0.49**			
		std err		0.26**	0.25**	0.25**	0.26*			
	index_connect	coef						**-0.15**	**-0.07**	**-0.44**
		std err						0.52	0.52	0.58
	oil*index_connect	coef						**1.8**	**1.78**	**1.69**
		std err						0.74**	0.73**	0.73**
Psychological Controls	placebo	coef				**-0.41**	**-0.72**		**-0.64**	**-0.77**
		std err				0.91	0.94		0.94	0.96
	oil*placebo	coef				**1.63**	**1.3**		**1.42**	**1.09**
		std err				1.28	1.33		1.33	1.35
	pessim1	coef				**0.01**	**-0.05**		**0**	**-0.07**
		std err				0.04	0.04		0.04	0.04*
	oil*pessim1	coef				**0.13**	**0.16**		**0.13**	**0.17**
		std err				0.05**	0.05***		0.05**	0.05***
	oil_rev_expec	coef				**0.07**	**0.08**		**0.06**	**0.08**
		std err				0.03**	0.03**		0.03**	0.03**
	oil_population	coef							**-0.07**	**-0.05**
		std err							0.04*	0.04
	rely1	coef							**-0.09**	**-0.02**
		std err							0.04**	0.04
Demographic Controls			**No**	**No**	**No**	**No**	**Yes**	**No**	**No**	**Yes**
	Constant		3.21	3.45	3.22	3.15	4.04	3.47	4.25	4.5
			0.07***	0.06***	0.07***	0.50***	0.54***	0.12***	0.58***	0.60***
	Number of Observations		1566	1563	1563	1447	1275	1566	1447	1275
	Adjusted R-squared		0.1	0.04	0.1	0.12	0.19	0.04	0.06	0.16

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A8: Scholarships Regressions

	Dependent Variable ------>		Perceived Corruption in Scholarships							
Main Explanatory Variables	oil	coef	0.49	0.69	0.39	-1.64	-2.09	0.29	-1.62	-2.24
		std err	0.09***	0.07***	0.09***	0.58***	0.58***	0.15*	0.58***	0.61***
	urban	coef	-0.27		-0.28	-0.3	-0.34			
		std err	0.09***		0.09***	0.09***	0.10***			
	oil*urban	coef	0.67		0.67	0.67	0.67			
		std err	0.13***		0.13***	0.13***	0.13***			
	pub_adm_dummy	coef		-0.73	-0.73	-0.6	-0.68			
		std err		0.14***	0.14***	0.14***	0.15***			
	oil*pub_adm_dummy	coef		0.8	0.81	0.66	0.73			
		std err		0.20***	0.20***	0.20***	0.20***			
	index_connect	coef						-1.88	-1.79	-2.19
		std err						0.42***	0.42***	0.46***
	oil*index_connect	coef						2.23	2.2	2.24
		std err						0.60***	0.60***	0.62***
Psychological Controls	placebo	coef				-1.53	-2.03		-1.36	-1.97
		std err				0.75**	0.76***		0.75*	0.79**
	oil*placebo	coef				2.83	3.74		2.48	3.67
		std err				1.05***	1.06***		1.06**	1.11***
	pessim1	coef				-0.09	-0.08		-0.1	-0.09
		std err				0.03***	0.03**		0.03***	0.03***
	oil*pessim1	coef				0.11	0.1		0.12	0.11
		std err				0.04**	0.04**		0.04***	0.05**
Demographic Controls			No	No	No	No	Yes	No	No	Yes
	Constant		5.03	4.99	5.11	6.35	6.72	5.32	6.51	6.83
			0.06***	0.05***	0.06***	0.41***	0.48***	0.10***	0.41***	0.46***
	Number of Observations		1343	1341	1341	1274	1226	1343	1274	1194
	Adjusted R-squared		0.11	0.11	0.13	0.13	0.16	0.11	0.12	0.14

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A9: Courts Regressions

	Dependent Variable ------>		Perceived Corruption in Courts							
Main Explanatory Variables	oil	coef	0.56	0.66	0.46	-0.94	-0.83	0.54	-0.86	-0.82
		std err	0.15***	0.13***	0.16***	0.93	0.99	0.26**	0.93	1.07
	urban	coef	-0.02		-0.02	0.04	-0.03			
		std err	0.16		0.16	0.16	0.17			
	oil*urban	coef	0.48		0.47	0.42	0.39			
		std err	0.22**		0.22**	0.23*	0.23*			
	pub_adm_dummy	coef		-0.29	-0.29	-0.19	-0.19			
		std err		0.19	0.19	0.19	0.2			
	oil*pub_adm_dummy	coef		0.46	0.46	0.33	0.31			
		std err		0.26*	0.26*	0.26	0.27			
	index_connect	coef						-1.28	-1.06	-1.9
		std err						0.65*	0.65	0.79**
	oil*index_connect	coef						0.93	0.99	0.67
		std err						0.93	0.92	1.02
Psychological Controls	placebo	coef				-3.11	-3.34		-3.07	-3.14
		std err				1.19***	1.28***		1.20**	1.35**
	oil*placebo	coef				1.39	1.2		1.28	1.14
		std err				1.67	1.79		1.68	1.89
	pessim1	coef				-0.08	-0.09		-0.09	-0.14
		std err				0.05*	0.05*		0.05*	0.06**
	oil*pessim1	coef				0.14	0.14		0.14	0.15
		std err				0.07**	0.07*		0.07*	0.08*
	rely1	coef				0.12	0.1		0.1	0.12
		std err				0.06**	0.06*		0.06*	0.07*
Demographic Controls			No	No	No	No	Yes	No	No	Yes
	Constant		4.69	4.74	4.75	6.15	6.66	5	6.48	7.57
			0.10***	0.09***	0.11***	0.76***	0.82***	0.18***	0.76***	0.89***
	Number of Observations		619	617	617	578	564	619	578	476
	Adjusted R-squared		0.08	0.07	0.08	0.11	0.13	0.07	0.1	0.15

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A10: Police Regressions

	Dependent Variable ------>		Perceived Corruption in Police							
Main Explanatory Variables	oil	coef	**0.64**	**0.65**	**0.54**	**-0.56**	**-0.54**	**0.29**	**-0.75**	**-0.74**
		std err	0.12***	0.11***	0.13***	0.79	0.79	0.23	0.8	0.81
	urban	coef	**0.08**		**0.09**	**0.15**	**-0.06**			
		std err	0.14		0.14	0.15	0.15			
	oil*urban	coef	**0.31**		**0.28**	**0.34**	**0.39**			
		std err	0.2		0.2	0.21*	0.20*			
	pub_adm_dummy	coef		**-0.44**	**-0.44**	**-0.35**	**-0.41**			
		std err		0.18**	0.18**	0.18*	0.19**			
	oil*pub_adm_dummy	coef		**0.62**	**0.61**	**0.57**	**0.64**			
		std err		0.26**	0.26**	0.26**	0.25**			
	index_connect	coef						**-2.17**	**-1.9**	**-2.87**
		std err						0.60***	0.61***	0.62***
	oil*index_connect	coef						**1.9**	**1.84**	**1.93**
		std err						0.85**	0.86**	0.84**
Psychological Controls	placebo	coef				**-1.22**	**-0.23**		**-1.22**	**0.25**
		std err				1.02	1.04		1.03	1.04
	oil*placebo	coef				**1.04**	**1.14**		**0.92**	**0.98**
		std err				1.44	1.45		1.45	1.45
	pessim1	coef				**-0.08**	**-0.07**		**-0.08**	**-0.07**
		std err				0.04*	0.04*		0.04*	0.04
	oil*pessim1	coef				**0.11**	**0.09**		**0.11**	**0.1**
		std err				0.06*	0.06		0.06*	0.06*
	rely1	coef				**0.25**	**0.22**		**0.22**	**0.22**
		std err				0.05***	0.05***		0.05***	0.05***
Demographic Controls			**No**	**No**	**No**	**No**	**Yes**	**No**	**No**	**Yes**
	Constant		4.57	4.68	4.64	4.29	2.53	5.13	4.92	2.95
			0.09***	0.07***	0.09***	0.65***	0.78***	0.16***	0.66***	0.79***
	Number of Observations		821	818	818	770	714	821	770	714
	Adjusted R-squared		0.07	0.07	0.08	0.12	0.17	0.08	0.11	0.17

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A11: State Jobs Regressions

	Dependent Variable ------>		Perceived Corruption in State Jobs Allocations							
Main Explanatory Variables	oil	coef	**0.79**	**0.9**	**0.75**	**-0.59**	**-0.72**	**0.61**	**-0.58**	**-0.78**
		std err	0.07***	0.06***	0.08***	0.5	0.52	0.13***	0.5	0.52
	urban	coef	**-0.15**		**-0.13**	**-0.15**	**-0.12**			
		std err	0.08*		0.08	0.08*	0.09			
	oil*urban	coef	**0.39**		**0.37**	**0.37**	**0.34**			
		std err	0.12***		0.12***	0.12***	0.13***			
	pub_adm_dummy	coef		**-0.43**	**-0.43**	**-0.36**	**-0.3**			
		std err		0.13***	0.13***	0.13***	0.14**			
	oil*pub_adm_dummy	coef		**0.38**	**0.37**	**0.31**	**0.27**			
		std err		0.18**	0.18**	0.18*	0.19			
	index_connect	coef						**-1.49**	**-1.46**	**-1.48**
		std err						0.37***	0.37***	0.43***
	oil*index_connect	coef						**1.58**	**1.61**	**1.6**
		std err						0.52***	0.52***	0.55***
Psychological Controls	placebo	coef				**-1.24**	**-1.38**		**-1.06**	**-1.23**
		std err				0.65*	0.67**		0.65	0.67*
	oil*placebo	coef				**1.08**	**1.37**		**0.74**	**1.06**
		std err				0.92	0.95		0.92	0.94
	pessim1	coef				**-0.08**	**-0.07**		**-0.08**	**-0.06**
		std err				0.03***	0.03**		0.03***	0.03**
	oil*pessim1	coef				**0.15**	**0.14**		**0.15**	**0.15**
		std err				0.04***	0.04***		0.04***	0.04***
	rely1	coef				**0.11**	**0.06**		**0.12**	**0.06**
		std err				0.03***	0.03*		0.03***	0.03*
Demographic Controls			**No**	**No**	**No**	**No**	**Yes**	**No**	**No**	**Yes**
	Constant		4.9	4.9	4.95	5.42	5.4	5.16	5.53	5.24
			0.05***	0.04***	0.05***	0.39***	0.45***	0.09***	0.39***	0.45***
	Number of Observations		1331	1327	1327	1262	1035	1331	1262	1035
	Adjusted R-squared		0.18	0.17	0.18	0.2	0.24	0.18	0.2	0.25

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A12: Customs Regressions

	Dependent Variable ------>		Perceived Corruption in Customs							
Main Explanatory Variables	oil	coef	0.78	0.82	0.64	-2	-2.27	0.52	-1.8	-1.88
		std err	0.16***	0.13***	0.17***	1.03*	1.11**	0.28*	1.02*	1.09*
	urban	coef	0.2		0.15	-0.01	-0.07			
		std err	0.16		0.16	0.16	0.16			
	oil*urban	coef	0.3		0.34	0.51	0.55			
		std err	0.23		0.22	0.23**	0.23**			
	pub_adm_dummy	coef		-0.74	-0.73	-0.55	-0.63			
		std err		0.19***	0.19***	0.20***	0.20***			
	oil*pub_adm_dummy	coef		0.53	0.56	0.46	0.46			
		std err		0.28*	0.27**	0.28*	0.27*			
	index_connect	coef						-2.44	-1.88	-1.78
		std err						0.68***	0.68***	0.72**
	oil*index_connect	coef						1.52	1.38	1.32
		std err						0.95	0.96	0.96
Psychological Controls	placebo	coef				-5.35	-5.39		-5.18	-5.14
		std err				1.39***	1.50***		1.40***	1.52***
	oil*placebo	coef				4.02	4.52		3.62	3.82
		std err				1.96**	2.10**		1.98*	2.12*
	pessim1	coef				-0.03	-0.07		-0.05	-0.07
		std err				0.06	0.05		0.05	0.06
	oil*pessim1	coef				0.11	0.1		0.1	0.1
		std err				0.08	0.08		0.08	0.08
	oil_population	coef				-0.1	-0.12		-0.09	-0.13
		std err				0.05*	0.05**		0.05*	0.05***
Demographic Controls			No	No	No	No	Yes	No	No	Yes
	Constant		4.45	4.71	4.63	7.79	8.67	5.22	8.14	8.85
			0.11***	0.09***	0.12***	0.74***	0.82***	0.20***	0.74***	0.82***
	Number of Observations		543	543	543	497	480	543	497	480
	Adjusted R-squared		0.12	0.13	0.14	0.17	0.24	0.13	0.16	0.22

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A13: Subsidies / Supplier Positions Regressions

	Dependent Variable ------>		Perceived Corruption in Subsidies / Supplier Positions Allocations					
Main Explanatory Variables	oil	coef	0.2	-1.1	-1.09	-0.03	-1.18	-1.1
		std err	0.16	1.23	1.22	0.3	1.25	1.29
	urban	coef	0.27	0.16	0.04			
		std err	0.2	0.2	0.21			
	oil*urban	coef	0.5	0.41	0.42			
		std err	0.29*	0.29	0.29			
	index_connect	coef				-0.77	-0.57	-1.07
		std err				0.79	0.76	0.83
	oil*index_connect	coef				1.55	1.7	1.34
		std err				1.11	1.08	1.09
Psychological Controls	placebo	coef		-2.72	-2.67		-2.87	-2.42
		std err		1.66	1.65		1.65*	1.72
	oil*placebo	coef		2.41	2.43		1.92	1.77
		std err		2.34	2.32		2.34	2.41
	pessim2	coef		-0.13	-0.12		-0.13	-0.1
		std err		0.06**	0.06*		0.06**	0.06
	oil*pessim2	coef		0.01	0.01		0.02	0.04
		std err		0.09	0.09		0.09	0.09
	rely1	coef		-0.3	-0.2		-0.32	-0.19
		std err		0.06***	0.07***		0.06***	0.07***
Demographic Controls			No	No	Yes	No	No	Yes
	Constant		3.32	6.85	6.48	3.59	7.19	7.17
			0.11***	0.97***	1.06***	0.21***	0.97***	1.06***
	Number of Observations		535	518	506	535	518	499
	Adjusted R-squared		0.04	0.07	0.11	0.01	0.06	0.1

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A14: Public Infrastructures Regressions

	Dependent Variable ------>		Perceived Corruption in Public Infrastructures					
Main Explanatory Variables	oil	coef	-0.2	-4.31	-5.02	-0.25	-3.99	-3.26
		std err	0.32	2.46*	2.66*	0.55	2.43	3.05
	urban	coef	0.51	0.28	0.59			
		std err	0.32	0.33	0.31*			
	oil*urban	coef	0.19	0.33	0.47			
		std err	0.45	0.47	0.41			
	index_connect	coef				-2.51	-2.53	-3.37
		std err				1.71	1.73	1.70**
	oil*index_connect	coef				0.71	0.25	0.75
		std err				2.43	2.45	2.24
Psychological Controls	placebo	coef		-6.94	-3.02		-6.55	-1.37
		std err		3.40**	3.78		3.41*	4.11
	oil*placebo	coef		6.03	7.03		5.87	4.72
		std err		4.69	5.1		4.71	5.55
	pessim1	coef		-0.02	0.04		-0.05	-0.03
		std err		0.14	0.12		0.14	0.14
	oil*pessim1	coef		0.19	0.21		0.16	0.12
		std err		0.2	0.17		0.19	0.19
	oil_rev_expec	coef		-0.18	-0.2		-0.2	-0.23
		std err		0.10*	0.10**		0.10**	0.11**
	rely1	coef		-0.34	-0.55		-0.36	-0.56
		std err		0.12***	0.12***		0.12***	0.13***
Demographic Controls			No	No	Yes	No	No	Yes
	Constant		3.55	9.63	9.9	4.33	10.42	7.89
			0.22***	2.09***	2.17***	0.39***	2.05***	2.53***
	Number of Observations		181	164	157	181	164	143
	Adjusted R-squared		0.03	0.06	0.36	0	0.06	0.34

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A15: Licenses Regressions

	Dependent Variable ------>		Perceived Corruption in Licenses					
Main Explanatory Variables	oil	coef	0.47	-1.07	-0.79	0.47	-0.88	-0.56
		std err	0.21**	1.26	1.2	0.41	1.25	1.21
	urban	coef	-0.25	-0.35	-0.42			
		std err	0.25	0.26	0.27			
	oil*urban	coef	0.4	0.5	0.61			
		std err	0.35	0.37	0.36*			
	index_connect	coef				-3.17	-3.13	-3.79
		std err				1.08***	1.11***	1.15***
	oil*index_connect	coef				0.63	0.42	0.67
		std err				1.53	1.58	1.51
Psychological Controls	placebo	coef		-3.7	-3.74		-3.35	-3.81
		std err		1.67**	1.64**		1.64**	1.62**
	oil*placebo	coef		1.55	0.65		1.37	0.43
		std err		2.36	2.29		2.32	2.26
	index_pessim	coef		-0.09	-0.7		0.1	-0.52
		std err		0.58	0.58		0.57	0.57
	oil*index_pessim	coef		1.24	1.41		1.16	1.27
		std err		0.82	0.80*		0.81	0.79
Demographic Controls			No	No	Yes	No	No	Yes
	Constant		4.26	6.24	8.25	4.94	6.6	8.25
			0.15***	0.89***	0.96***	0.29***	0.88***	0.97***
	Number of Observations		388	362	287	388	362	287
	Adjusted R-squared		0.03	0.04	0.22	0.06	0.07	0.24

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A16: Vote Buying Regressions

	Dependent Variable ------>		Vote Buying							
Main Explanatory Variables	oil	coef	**0.57**	**0.7**	**0.52**	**-0.33**	**-0.61**	**0.36**	**-0.35**	**-0.61**
		std err	0.10***	0.08***	0.10***	0.69	0.69	0.16**	0.69	0.7
	urban	coef	**0.03**		**0.03**	**0.02**	**-0.21**			
		std err	0.11		0.11	0.11	0.11*			
	oil*urban	coef	**0.44**		**0.44**	**0.46**	**0.5**			
		std err	0.15***		0.15***	0.16***	0.15***			
	pub_adm_dummy	coef		**-0.27**	**-0.27**	**-0.23**	**-0.24**			
		std err		0.18	0.18	0.18	0.19			
	oil*pub_adm_dummy	coef		**0.49**	**0.49**	**0.44**	**0.44**			
		std err		0.26*	0.26*	0.26*	0.25*			
	index_connect	coef						**-1.33**	**-1.23**	**-1.56**
		std err						0.50***	0.50**	0.53***
	oil*index_connect	coef						**1.89**	**1.91**	**2.03**
		std err						0.70***	0.70***	0.70***
Psychological Controls	placebo	coef				**-1.68**	**-2.88**		**-1.58**	**-2.46**
		std err				0.90*	0.90***		0.90*	0.91***
	oil*placebo	coef				**1.12**	**1.66**		**0.82**	**1.31**
		std err				1.27	1.27		1.28	1.29
	pessim1	coef				**-0.02**	**-0.01**		**-0.03**	**0.01**
		std err				0.04	0.04		0.04	0.04
	oil*pessim1	coef				**0.05**	**0.05**		**0.05**	**0.05**
		std err				0.05	0.05		0.05	0.05
	pessim2	coef								
		std err								
	oil*pessim2	coef								
		std err								
	oil_rev_expec	coef								
		std err								
	rely1	coef							**-0.07**	**-0.02**
		std err							0.04*	0.04
	Demographic Controls		**No**	**No**	**No**	**No**	**Yes**	**No**	**No**	**Yes**
	Constant		4.38	4.41	4.4	5.38	6.46	4.66	5.95	6.78
			0.07***	0.06***	0.07***	0.49***	0.60***	0.11***	0.54***	0.64***
	Number of Observations		1664	1660	1660	1584	1544	1664	1584	1498
	Adjusted R-squared		0.06	0.06	0.06	0.07	0.12	0.06	0.06	0.12

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Appendix G – External Control Econometric Results

In this Appendix, we display we display only regressions both with and without psychological/demographic controls[76].

Table A17: Health Care Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in Health Care							
			full sample	urban	public admin	connect>med	full sample	urban	public admin	connect>med
Main Explanatory Variables	oil	coef	0.28	0.42	0.4	0.27	-0.14	1.01	1.23	0.18
		std err	0.08***	0.12***	0.24*	0.12**	0.49	0.8	1.63	0.63
	stpcv	coef	-0.01	0.26	-0.53	-0.18	1.5	1.66	1.25	0.94
		std err	0.08	0.13**	0.25**	0.11	0.17***	0.26***	0.45***	0.24***
	oil*stpcv	coef	0.42	0.65	0.86	0.54	0.37	0.62	0.78	0.54
		std err	0.11***	0.18***	0.35**	0.16***	0.12***	0.19***	0.41*	0.17***
Psychological Controls			No	No	No	No	Yes	Yes	Yes	Yes
Demographic Controls			No	No	No	No	Yes	Yes	Yes	Yes
	Number of Observations		3626	1442	327	1528	3001	1075	243	1251
	Adjusted R-squared		0.02	0.08	0.07	0.04	0.13	0.22	0.25	0.12

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A18: Education Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in Education							
			full sample	urban	public admin	connect>med	full sample	urban	public admin	connect>med
Main Explanatory Variables	oil	coef	0.17	0.11	0.04	0.13	-0.63	-0.2	-0.46	-0.51
		std err	0.06***	0.1	0.19	0.1	0.44	0.7	1.14	0.59
	stpcv	coef	1.62	1.76	1.43	1.58	1.18	2.05	2.54	1.51
		std err	0.06***	0.11***	0.20***	0.10***	0.16***	0.62***	0.41***	0.21***
	oil*stpcv	coef	0.46	0.78	1.07	0.64	0.39	0.77	1.12	0.49
		std err	0.09***	0.15***	0.28***	0.14***	0.11***	0.16***	0.32***	0.16***
Psychological Controls			No	No	No	No	Yes	Yes	Yes	Yes
Demographic Controls			No	No	No	No	Yes	Yes	Yes	Yes
	Number of Observations		3473	1379	329	1543	2473	1149	292	1244
	Adjusted R-squared		0.34	0.38	0.4	0.34	0.38	0.47	0.52	0.41

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A19: Scholarships Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in Scholarships							
			full sample	urban	public admin	connect>med	full sample	urban	public admin	connect>med
Main Explanatory Variables	oil	coef	0.35	0.4	0.48	0.31	-2.01	-1.94	-2.06	-1.96
		std err	0.08***	0.13***	0.26*	0.12**	0.50***	0.79**	1.43	0.68***
	stpcv	coef	-0.24	-0.19	-0.73	-0.34	0.01	0.24	-0.67	-0.14
		std err	0.08***	0.13	0.26***	0.12***	0.11	0.17	0.33**	0.16
	oil*stpcv	coef	0.43	0.76	1	0.62	0.29	0.64	0.7	0.38
		std err	0.12***	0.19***	0.37***	0.17***	0.12**	0.19***	0.41*	0.18**
Psychological Controls			No	No	No	No	Yes	Yes	Yes	Yes
Demographic Controls			No	No	No	No	Yes	Yes	Yes	Yes
	Number of Observations		2698	1185	314	1404	2327	1014	274	1203
	Adjusted R-squared		0.04	0.07	0.1	0.05	0.09	0.15	0.3	0.12

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

[76] Note that the demographic controls used in these regressions follow the same criteria described in the last appendix corresponding footnote, and are drawn from the same group of variables with the following additions relating to CV:

- Basic Demographics: CV nationality (dummy);
- Ethnic/Origin Groups: santiago, são vicente, santo antão, fogo (dummies);
- Political Preferences: PAICV, MPD (dummies).

Table A20: Courts Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in Courts							
			full sample	urban	public admin	connect>med	full sample	urban	public admin	connect>med
Main Explanatory Variables	oil	coef	0.29	0.19	0.23	0.3	-1.72	-0.83	-2.66	-1.15
		std err	0.14**	0.21	0.29	0.19	0.83**	1.26	1.72	1.02
	stpcv	coef	0.41	0.37	0.22	0.26	1.11	0.96	0.34	1.67
		std err	0.15***	0.22*	0.3	0.18	0.22***	0.29***	0.4	0.40***
	oil*stpcv	coef	0.48	0.85	0.89	0.5	0.41	0.71	0.68	0.43
		std err	0.20**	0.30***	0.42**	0.26*	0.22*	0.34**	0.46	0.28
Psychological Controls			No	No	No	No	Yes	Yes	Yes	Yes
Demographic Controls			No	No	No	No	Yes	Yes	Yes	Yes
	Number of Observations		1257	552	293	799	1078	430	262	674
	Adjusted R-squared		0.05	0.08	0.07	0.04	0.13	0.15	0.16	0.15

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A21: Police Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in Police							
			full sample	urban	public admin	connect>med	full sample	urban	public admin	connect>med
Main Explanatory Variables	oil	coef	0.4	0.49	0.29	0.34	-1.54	-1.92	-3.19	-1.74
		std err	0.14***	0.20**	0.29	0.17**	0.73**	1.11*	1.64*	0.87**
	stpcv	coef	0.27	0.24	0.22	0.16	1.45	1.09	3.09	0.73
		std err	0.13**	0.2	0.31	0.16	0.22***	0.31***	0.53***	0.31**
	oil*stpcv	coef	0.36	0.46	0.98	0.52	0.31	0.48	0.64	0.22
		std err	0.18*	0.28	0.43**	0.23**	0.2	0.29	0.47	0.26
Psychological Controls			No	No	No	No	Yes	Yes	Yes	Yes
Demographic Controls			No	No	No	No	Yes	Yes	Yes	Yes
	Number of Observations		1486	608	302	943	1245	503	232	715
	Adjusted R-squared		0.04	0.06	0.08	0.05	0.14	0.21	0.35	0.2

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A22: State Jobs Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in State Jobs Allocations							
			full sample	urban	public admin	connect>med	full sample	urban	public admin	connect>med
Main Explanatory Variables	oil	coef	0.42	0.49	0.26	0.43	-1.09	-0.64	-1.16	-0.8
		std err	0.09***	0.15***	0.27	0.13***	0.54**	0.85	1.67	0.7
	stpcv	coef	0.12	-0.12	-0.13	0	1.59	1.82	1.8	1.44
		std err	0.09	0.14	0.27	0.12	0.25***	0.25***	0.40***	0.24***
	oil*stpcv	coef	0.53	0.69	1.02	0.68	0.38	0.56	0.65	0.37
		std err	0.12***	0.19***	0.38***	0.17***	0.14***	0.20***	0.49	0.20*
Psychological Controls			No	No	No	No	Yes	Yes	Yes	Yes
Demographic Controls			No	No	No	No	Yes	Yes	Yes	Yes
	Number of Observations		2392	937	289	1218	1786	739	195	958
	Adjusted R-squared		0.07	0.1	0.08	0.09	0.15	0.22	0.35	0.2

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A23: Customs Jobs Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in Customs							
			full sample	urban	public admin	connect>med	full sample	urban	public admin	connect>med
Main Explanatory Variables	oil	coef	0.23	0.17	0.24	0.14	-1.56	-1.45	-2.78	-2.34
		std err	0.13*	0.2	0.37	0.18	0.92*	1.33	1.54*	1.07**
	stpcv	coef	-0.25	0.03	-1.23	-0.63	0.96	0.74	-0.95	0.16
		std err	0.14*	0.21	0.34***	0.18***	0.24***	0.31**	0.35***	0.28
	oil*stpcv	coef	0.7	0.91	1.11	0.85	0.67	0.8	1.15	0.72
		std err	0.20***	0.29***	0.48**	0.25***	0.24***	0.34**	0.39***	0.29**
Psychological Controls			No	No	No	No	Yes	Yes	Yes	Yes
Demographic Controls			No	No	No	No	Yes	Yes	Yes	Yes
	Number of Observations		1220	605	190	730	865	414	139	538
	Adjusted R-squared		0.03	0.05	0.12	0.05	0.13	0.18	0.66	0.2

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A24: Subsidies / Supplier Positions Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in Subsidies / Supplier Positions Allocations					
			full sample	urban	connect>med	full sample	urban	connect>med
Main Explanatory Variables	oil	coef	-0.2	-0.29	-0.09	-0.97	-1.21	-1.11
		std err	0.23	0.4	0.32	1.04	1.81	1.3
	stpcv	coef	0.02	0.96	0.03	0.56	0.28	0.37
		std err	0.19	0.34***	0.26	0.24**	0.48	0.49
	oil*stpcv	coef	0.54	0.99	0.47	0.45	0.95	0.36
		std err	0.27**	0.48**	0.37	0.29	0.42**	0.4
Psychological Controls			No	No	No	Yes	Yes	Yes
Demographic Controls			No	No	No	Yes	Yes	Yes
	Number of Observations		763	233	454	648	199	405
	Adjusted R-squared		0.01	0.15	0.01	0.14	0.4	0.17

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A25: Public Infrastructures Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in Public Infrastructures					
			full sample	urban	connect>med	full sample	urban	connect>med
Main Explanatory Variables	oil	coef	0.07	0.21	0.08	-0.11	0.4	0.37
		std err	0.26	0.57	0.35	1.57	1.69	1.46
	stpcv	coef	1.31	1.07	1.72	2.45	3.37	4.78
		std err	0.24***	0.46**	0.31***	0.38***	0.67***	0.50***
	oil*stpcv	coef	-0.17	-0.23	-0.18	-0.44	-0.5	-0.4
		std err	0.34	0.66	0.44	0.4	0.56	0.41
Psychological Controls			No	No	No	Yes	Yes	Yes
Demographic Controls			No	No	No	Yes	Yes	Yes
	Number of Observations		306	120	150	216	93	119
	Adjusted R-squared		0.14	0.05	0.26	0.34	0.53	0.59

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A26: Licenses Regressions (Cape Verde Control)

	Dependent Variable ------>		Perceived Corruption in Licenses					
			full sample	urban	connect>med	full sample	urban	connect>med
Main Explanatory Variables	oil	coef	0.18	0.29	0.2	-1.08	-2.37	-1.34
		std err	0.32	0.46	0.45	1.16	1.65	1.62
	stpcv	coef	-0.16	-0.23	-0.27	-0.21	-2.29	0.1
		std err	0.26	0.39	0.36	0.38	0.67***	0.66
	oil*stpcv	coef	0.44	0.58	0.46	0.1	0.36	0.01
		std err	0.37	0.55	0.51	0.4	0.52	0.5
Psychological Controls			No	No	No	Yes	Yes	Yes
Demographic Controls			No	No	No	Yes	Yes	Yes
	Number of Observations		515	207	314	433	146	258
	Adjusted R-squared		0.02	0.02	0.02	0.19	0.48	0.25

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Table A27: Vote Buying Regressions (Cape Verde Control)

	Dependent Variable ------>		Vote Buying							
			full sample	urban	public admin	connect>med	full sample	urban	public admin	connect>med
Main Explanatory Variables	oil	coef	0.15	0.19	-0.06	0.11	-0.15	-0.01	0.86	-0.68
		std err	0.08*	0.13	0.24	0.13	0.57	0.89	1.47	0.75
	stpcv	coef	1.92	1.93	1.83	1.81	2.32	2.83	2.03	1.96
		std err	0.08***	0.13***	0.25***	0.13***	0.17***	0.22***	0.40***	0.19***
	oil*stpcv	coef	0.6	0.82	1.24	0.8	0.57	0.84	1.05	0.78
		std err	0.12***	0.19***	0.35***	0.18***	0.15***	0.19***	0.40***	0.20***
Psychological Controls			No	No	No	No	Yes	Yes	Yes	Yes
Demographic Controls			No	No	No	No	Yes	Yes	Yes	Yes
	Number of Observations		3656	1449	330	1532	2491	1059	261	1274
	Adjusted R-squared		0.28	0.32	0.39	0.3	0.32	0.45	0.51	0.34

Note: * significant at 10%; ** significant at 5%; *** significant at 1%

Appendix H – Robustness of Econometric Results

Table A28: Ordered Probit Regressions (with Psychological and Demographic Controls)

	Dependent Variable ------>		Health Care						Education					
			OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff
Main Explanatory Variables	oil	coef	-0.21	-0.23	0	0	-0.13	-0.19	-0.79	-1.08	-0.8	-1.11	-0.29	-0.34
		std err	0.53		0.54		0.34		0.54		0.54		0.35	
	urban	coef	0.03	0.03					0.3	0.42				
		std err	0.08						0.09***					
	oil*urban	coef	0.65	0.63					0.3	0.43				
		std err	0.12***						0.12**					
	pub_adm_dummy	coef	-0.29	-0.34					-0.13	-0.18				
		std err	0.13**						0.14					
	oil*pub_adm_dummy	coef	0.41	0.4					0.26	0.37				
		std err	0.19**						0.19					
	index_connect	coef			-0.58	-0.64					-0.29	-0.4		
		std err			0.39						0.42			
	oil*index_connect	coef			1.04	1.16					1.13	1.59		
		std err			0.53**						0.54**			
	stpcv	coef					0.18	0.26					1.2	1.45
		std err					0.18						0.13***	
	oil*stpcv	coef					0.19	0.27					0.17	0.21
		std err					0.08**						0.08**	
	Number of Observations		1462		1434		3023		1275		1275		2734	
	Likelihood Ratio (Chi2)		286.58		227.93		366.47		255		216.27		1319.73	

Note: * significant at 10%; ** significant at 5%; *** significant at 1%;

Table A28 (continued): Ordered Probit Regressions (with Psychological and Demographic Controls)

	Dependent Variable ------>		Scholarships						Courts					
			OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff
Main Explanatory Variables	oil	coef	-1.69	-1.7	-1.75	-1.78	-1.14	-1.44	-0.52	-0.62	-0.39	-0.48	-0.8	-1.29
		std err	0.56***		0.57***		0.38***		0.79		0.85		0.52	
	urban	coef	-0.26	-0.26					-0.08	-0.1				
		std err	0.09***						0.13					
	oil*urban	coef	0.62	0.56					0.43	0.48				
		std err	0.13***						0.18**					
	pub_adm_dummy	coef	-0.52	-0.57					-0.12	-0.14				
		std err	0.14***						0.16					
	oil*pub_adm_dummy	coef	0.53	0.46					0.13	0.15				
		std err	0.19***						0.21					
	index_connect	coef			-1.69	-1.71					-1.21	-1.48		
		std err			0.43***						0.62*			
	oil*index_connect	coef			1.7	1.72					0.35	0.43		
		std err			0.58***						0.8			
	stpcv	coef					-0.19	-0.24					0.33	0.53
		std err					0.18						0.17**	
	oil*stpcv	coef					0.16	0.2					0.19	0.3
		std err					0.09*						0.13	
	Number of Observations		1227		1194		2285		565		476		1087	
	Likelihood Ratio (Chi2)		208.15		186.03		221.48		93.54		83.58		101.65	

Note: * significant at 10%; ** significant at 5%; *** significant at 1%;

Table A28 (continued): Ordered Probit Regressions (with Psychological and Demographic Controls)

	Dependent Variable ------>		Police						State Jobs Allocations					
			OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff
Main Explanatory Variables	oil	coef	-0.39	-0.44	-0.43	-0.49	-0.7	-1.07	-0.62	-0.52	-0.59	-0.49	-0.73	-0.89
		std err	0.66		0.68		0.48		0.59		0.59		0.43*	
	urban	coef	-0.06	-0.07					-0.11	-0.09				
		std err	0.13						0.1					
	oil*urban	coef	0.41	0.44					0.45	0.35				
		std err	0.17**						0.14***					
	pub_adm_dummy	coef	-0.22	-0.25					-0.13	-0.11				
		std err	0.15						0.15					
	oil*pub_adm_dummy	coef	0.39	0.41					0.04	0.03				
		std err	0.21*						0.21					
	index_connect	coef			-1.98	-2.25					-1.22	-1.01		
		std err			0.51***						0.48**			
	oil*index_connect	coef			1.09	1.24					1.53	1.26		
		std err			0.7						0.62**			
	stpcv	coef					0.07	0.11					0.39	0.49
		std err					0.25						0.14***	
	oil*stpcv	coef					0.17	0.26					0.25	0.29
		std err					0.13						0.11**	
	Number of Observations		714		714		1189		1036		1036		1698	
	Likelihood Ratio (Chi2)		152.49		151.57		145.17		292.96		295.74		241.58	

Table A28 (continued): Ordered Probit Regressions (with Psychological and Demographic Controls)

	Dependent Variable ------>		Customs						Subsidies / Supplier Positions Allocations					
			OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff
Main Explanatory Variables	oil	coef	-2.47	-2.59	-1.87	-2.04	-0.8	-1.2	-0.54	-0.74	-0.55	-0.75	-0.53	-0.8
		std err	0.99**		0.96*		0.59		0.88		0.93		0.69	
	urban	coef	-0.14	-0.15					0.07	0.09				
		std err	0.14						0.15					
	oil*urban	coef	0.7	0.67					0.28	0.4				
		std err	0.21***						0.21					
	pub_adm_dummy	coef	-0.53	-0.61										
		std err	0.17***											
	oil*pub_adm_dummy	coef	0.33	0.33										
		std err	0.24											
	index_connect	coef			-1.37	-1.51					-0.61	-0.84		
		std err			0.60**						0.59			
	oil*index_connect	coef			1.05	1.16					0.81	1.12		
		std err			0.84						0.78			
	stpcv	coef					-0.23	-0.35					0.18	0.27
		std err					0.18						0.22	
	oil*stpcv	coef					0.36	0.51					0.33	0.5
		std err					0.14***						0.19*	
	Number of Observations		480		480		1043		507		499		679	
	Likelihood Ratio (Chi2)		157.42		132.93		132.01		75.37		66.71		84.96	

Note: * significant at 10%; ** significant at 5%; *** significant at 1%;

Table A28 (continued): Ordered Probit Regressions (with Psychological and Demographic Controls)

	Dependent Variable ------>		Public Infrastructures						Licenses					
			OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff	OProbit	PartEff
Main Explanatory Variables	oil	coef	-5.27	-4.38	-3.38	-3.43	-0.03	-0.04	-0.69	-0.96	-0.46	-0.63	-0.69	-1.04
		std err	2.36**		2.61		1.33		0.93		0.96		0.79	
	urban	coef	0.59	0.71					-0.28	-0.4				
		std err	0.28**						0.19					
	oil*urban	coef	0.56	0.68					0.52	0.67				
		std err	0.37						0.27*					
	index_connect	coef			-2.92	-3.65					-2.53	-3.5		
		std err			1.49**						0.85***			
	oil*index_connect	coef			0.42	0.52					0.38	0.52		
		std err			1.97						1.12			
	stpcv	coef					2.23	2.3					-0.22	-0.33
		std err					0.33***						0.26	
	oil*stpcv	coef					-0.47	-0.54					0.02	0.04
		std err					0.33						0.28	
	Number of Observations		157		143		216		287		287		434	
	Likelihood Ratio (Chi2)		89.67		78.57		122.38		87.68		93.94		111.94	

Note: * significant at 10%; ** significant at 5%; *** significant at 1%;

Table A28 (continued): Ordered Probit Regressions (with Psychological and Demographic Controls)

	Dependent Variable ------>		Vote Buying					
			OProbit	PartEff	OProbit	PartEff	OProbit	PartEff
Main Explanatory Variables	oil	coef	-0.2	-0.26	-0.14	-0.19	-0.12	-0.22
		std err	0.51		0.52		0.34	
	urban	coef	-0.14	-0.18				
		std err	0.08*					
	oil*urban	coef	0.37	0.46				
		std err	0.11***					
	pub_adm_dummy	coef	-0.14	-0.19				
		std err	0.13					
	oil*pub_adm_dummy	coef	0.23	0.29				
		std err	0.19					
	index_connect	coef			-0.91	-1.21		
		std err			0.38**			
	oil*index_connect	coef			1.15	1.53		
		std err			0.52**			
	stpcv	coef					1.16	2.06
		std err					0.11***	
	oil*stpcv	coef					0.31	0.57
		std err					0.08***	
	Number of Observations		1545		1499		3121	
	Likelihood Ratio (Chi2)		211.94		195.05		918.44	

Note: * significant at 10%; ** significant at 5%; *** significant at 1%;