

Corrupted Scholarships*

Pedro C. Vicente¹

First Draft: June 2007; This Draft: August 2007

Abstract:

This paper studies the mechanics of the prominent African natural resource curse. We ask: how does corruption change after an oil discovery? We closely followed the example of São Tomé and Príncipe, a West African island-country where a significant discovery announcement took place in 1997-1999. We postulate that an anticipated oil shock increases the value of maintaining power by the political elite, implying higher corruption after the discovery. We focus on the case of scholarships for higher education abroad – as in many other developing countries, these have been a key source of social status for elite families. We apply a rich set of measurement methods that includes both direct and perceived proxies of corruption. The first are based on a unique dataset including the characteristics of all possible applicants (e.g. grades) and all winners of scholarships in the period 1995-2005, and the second are based on three specifically-tailored household surveys, in different years, both in São Tomé and Príncipe and a suitable control West African island-country. The comparison of these different measures sheds light on the long-standing debate on the validity of corruption perceptions in the empirical measurement of corruption. We find that corruption increased after the announcements (by close to 10%) but decreased slightly after 2004. Corruption perceptions perform well since they capture the signs of these changes – they seem however to exaggerate magnitudes.

JEL Codes: D73, O13, O55, P16.

Keywords: Corruption, Influence, Scholarships, Natural Experiment, Political Economy, Natural Resources, Oil, São Tomé and Príncipe, West Africa.

* I wish to thank primarily Cátia Batista for comments and encouragement throughout all stages of this project. Marianne Bertrand, Paul Collier, Marcel Fafchamps, Casey Mulligan, Roger Myerson, and Gerhard Seibert provided helpful suggestions. I am grateful to Artur Cota and Christian Helmers for excellent research assistance. I thank numerous seminar participants for useful comments. This paper builds on an unusually-comprehensive set of fieldwork activities - I therefore can only name a few contributors: most importantly, I want to thank my co-interviewers in the three household surveys conducted in São Tomé and Príncipe and Cape Verde – their efforts were crucial for the success of this project; I am also grateful for kind cooperation to the Ministry of Education and all secondary schools of São Tomé and Príncipe, the Portuguese Development Agency, the French Government, and the Calouste Gulbenkian Foundation. I wish to acknowledge financial support from the Department of Economics at the University of Chicago, the ESRC-funded Global Poverty Research Group, and the Open Society Institute. All errors are my responsibility.

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

Table of Contents

1 Introduction.....	3
2 Theoretical Mechanism.....	7
3 Historical and Institutional Background.....	8
4 Real Corruption.....	10
4.1 Data for Scholarship Contests and Estimation Approach.....	10
4.2 Econometric Results.....	14
5 Perceived Corruption.....	16
5.1 Experimental Design.....	16
5.2 Data Collection: Tailored Household Surveys on Corruption.....	18
5.3 Descriptive Data and Estimation Approach.....	20
5.4 Econometric Results.....	23
6 Summary of Results and Interpretation.....	25
7 Concluding Remarks.....	27
References.....	29
Appendix.....	31

“[...] Because there is oil. It is the salvation of the island, small insular African state. Is it?”

- in ‘Lenin Oil’, by Pedro Rosa Mendes, fiction

1 Introduction

Corruption² has been at the center of the debate on institutions. And over the past decade it completely changed the face of development policy. Today, no discussion on how to tackle poverty lacks the mention of the ‘corruption’ factor. In parallel, natural resources were increasingly flagged as a curse for developing countries (Sachs and Warner, 1995). The power of this idea made it to the popular discourse on African malaises in a few years. Today, oil and diamond multinationals reason differently, in face of informed western consumers.

Although causality was traced from natural resources to lower growth, we still lack general knowledge of what the mechanism is - as identifying it through standard, cross-country work is patently challenging. Namely, is corruption really determined by natural resources?

Most interestingly, in 1997, a tiny, insular West African country, São Tomé and Príncipe, neighbored by well-known, resource-cursed Nigeria, Equatorial Guinea, and Gabon, found significant amounts of oil reserves off its coast. That represented a unique opportunity to study the workings of a resource curse (if it existed) at the micro level of a country, using a natural experiment. This is the context of this paper. Our main research question is therefore how corruption changed and evolved after the oil discovery announcements in São Tomé and Príncipe.

Our hypothesis is that the perspective of a future oil boom increases the will to hold political power in the future; corruption as way to improve one’s chances of holding power becomes more attractive. This theoretical idea makes stellar sense in the case of the allocations of scholarships for higher education abroad, the focus of this paper. Not too differently from many other African countries, São Tomé and Príncipe relies exclusively on scholarships to

² See Becker and Stigler (1974), Rose-Ackerman (1978) for early economics of corruption; see Bardhan (1997) for an early review of literature on corruption.

study abroad, offered by foreign donors, for higher education. Interestingly, that has been the primary way to educate the political elite in the post-colonial times³.

In pursuing our exercise, we however stumble on a methodological problem that has been challenging empirical economists for at least the last decade: the measurement of corruption. Initially tackled through cross-country, macro, perception indices, that quickly gained popularity (Transparency International, World Bank Governance)⁴, measurement of corruption started to be approached through more direct (though operationally difficult), micro-founded ways. ‘Missing money’ techniques (Reinikka and Svensson, 2004), which compare reported money (by the public official) with actually spent money (reaching the target), acquired recognition.

This paper uses a unique and comprehensive set of data, which encompasses both direct and perception-based measures of corruption over the pre and post-oil discovery periods in São Tomé and Príncipe. This enables answering the second main research question of the paper: assessing the performance of corruption perceptions over time in exposing the actual changes in the corruption reality those perceptions target. That is a fundamental question that is currently faced by international donor institutions, when trying to quantify governance improvements in development countries, based on perception indicators.

We collected data on the complete set of students, 4392 individuals, at the end of high school, from all schools in São Tomé and Príncipe, over the period 1995-2005, as well as on the complete set of scholarships that were allocated in the same period (from a wide range of disperse donor agencies). Notably we have all grades for all students in the last year of studies in the secondary school system of São Tomé and Príncipe. That enables us to assess which were the grades of the winners of scholarships, when compared with those of the population of applicants – i.e. measuring the distortion (corruption) of allocations relative to grades⁵. Provided the existence of long full names in the Lusophone tradition of the country, we are able to match the names of students to those of all politicians holding office in the post-

³ A further remark goes for the fact that despite being seen as an essential channel of foreign aid (even for foreign aid skeptics), not much research attention (from the perspective of aid assessment) has been devoted to scholarships for higher education.

⁴ See Mauro (1995) for an influential article on the effects of corruption on growth, which used early macro indices of perceived corruption.

⁵ This corruption measurement procedure is in fact close to the idea of ‘missing money’ accounting referred above. In our case, we can refer to it as measuring the ‘missing grades’ in the scholarship-winners’ group.

colonial period of the country – we are then able to compute a measure of family favoritism that is tested as a determinant of scholarship allocations.

On the perceptions side, we purposely-conducted three nationwide, representative, tailored household surveys. Two of these surveys were completed in São Tomé and Príncipe (2004, 2006), while one of them was carried out in a control West African, insular country, Cape Verde (2006), sharing much of its colonial history and post-colonial shocks (not oil) with São Tomé and Príncipe. These datasets (841, 1275, and 1066 interviews, respectively) form the basis of an experimental design, where we identify differences over time (between before oil, after oil until 2004, and after 2004) by contrasting perceptions among different treatment and control groups. Internally (to São Tomé and Príncipe), we compare more and less informed groups – which mainly controls for public opinion, though at the expense of different demographic profiles; externally we compare the Sao Tomean with the Cape Verdean perceptions – which takes care of common macro trends while keeping the same demographic profiles, but (as in most natural experiments) still relies on the comparison of two different countries. Perceptions about the pre-oil discovery period (only) are retrospective, but are gathered using a set of techniques to deal with any ‘Good Old Times’ bias: these include the use of personal and specifically-related (to the allocations at stake) memory markings during survey interviews, and the gathering of several measures of psychological time pessimism.

We find a clear increase in corruption of scholarship allocations of 7% after the oil discovery announcements, in the period until 2004, as measured by ‘grade distortions’. This is followed by a small adjustment after 2004 (decrease of 1%). We also found that family connectedness is an important determinant of the allocation outcomes. Using perceptions we achieved comparable numbers for the effect of oil, and indeed a smaller downward adjustment after 2004. However, both are larger (in absolute terms) than the directly-measured effects. That leads us to conclude for evidence in support of an increase in corruption after the oil announcement in São Tomé and Príncipe, and of the general validity of household perceptions in portraying time changes in corruption – although some caution should be exercised when targeting magnitudes.

This paper mainly relates to the literatures on the natural resource curse and on the microeconomic measurement of corruption.

Several important and related contributions to the understanding of the natural resource curse followed the work by Sachs and Warner. In the context of cross-country empirical work, Ales and Di Tella (1999) and Leite and Weidmann (1999) showed explicitly that natural resource abundance is associated with corruption. More recently, Mehlum et al (2006) argue that the main reason for diverging experiences by resource-rich countries (in terms of growth) is differences in the quality of institutions. Theoretically, different contributions underlined the importance of initial conditions in determining the effects of resource booms on growth: Baland and Francois (2000) emphasized the structure of the economy in terms of rent-seeking vs. entrepreneurship, Robinson et al (2006) focused on the quality of political institutions and incentives. Our paper extends the debate to the micro-level of a resource boom in a specific developing country, while focusing on corruption effects.

Although the growing literature on microeconomic measurement of corruption seems to be scattered along a wide range of themes and approaches, our paper relates to a large number of these contributions. We infer corruption in the allocation of scholarships by looking into the actual characteristics of the allocation outcomes (although we also use perceptions about these); and we use a naturally-occurring experimental setting (the oil discovery) to infer variation in corruption. Most of the following relevant literature contributions embed both these components. Like Reinikka and Svensson, Olken (2007) focuses on measuring missing expenditures in local public goods provision; differently however, this author applies a specific field experimental setting (partially created by the researcher). Bertrand et al (2006) also implement a field experiment, this time targeting corruption efficiency. In terms of strictly naturally-occurring frameworks: Duggan and Levitt (2000) used a non-linearity in sumo wrestling incentives, Fisman (2001) looked at news on the health of former dictator Suharto of Indonesia, and Fisman and Miguel (2006) applied the specifics of the diplomatic legal setting in New York to investigate various effects on corruption. Finally a mention should be made of Olken (2006), who contrasts local perceptions with direct measurements of corruption in road-building projects across Indonesian villages: perceptions are found to contain true information, but some biases are encountered when contrasting absolute levels across villages. In a complementary way, our exercise in this paper concerns the validity of differences in corruption perceptions over time.

In the remaining of the paper we begin by describing the main theoretical mechanism we propose for the impact of the oil discovery (Section 2). Then we introduce the reader to the historical and institutional context of São Tomé and Príncipe (Section 3). The analysis of direct data on the allocation of scholarships is presented in Section 4. Section 5 is devoted to the exploration of our perception-based experimental design. Finally, Sections 6 and 7 provide summary of results and concluding remarks.

2 Theoretical Mechanism

The theoretical mechanism we propose for the effect of a natural resource discovery on corruption is based on the simple idea that: if rents from natural resources can be appropriated by politicians, holding political power becomes more valuable after an oil shock; corruption, as an instrument to increase the likelihood of holding that power in the future, becomes more attractive, once the oil shock is anticipated.

To convey this idea formally, we present a very simple two-period version of Tullock's (1980), classical, rent-seeking model. We consider a two-period game, with n players ($n > 1$). In the first period each player competes with all other players for a 'seat' in the elite of the country. In the second date the members of the elite share the public sector rent; for simplicity we assume the winner of the seat in the first period gets R in the second period as a member of the elite.

The crucial assumption is that, in period 1, 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 the chances of winning the elite's seat. For player i , we denote the corrupt activities (and their costs) as x_i . The probability of getting the elite's seat for player i is postulated to be

$$p_i(x_1, \dots, x_n) = \frac{x_i}{\sum_{j=1}^n x_j}.$$

Solving this simple game by backward induction (assuming no discount factor), at period 2 each elite player maximizes

$$\max_{x_i} p_i(x_1, \dots, x_n)R - x_i.$$

The unique pure strategies Nash Equilibrium, which is symmetric, is:

$$x^* = \frac{(n-1)R}{n^2}.$$
⁶

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, known in period 1, i.e. anticipated (what we have in São Tomé and Príncipe). As can easily be seen from x^* , as a result, corruption increases in period 1.

This is due to competition for the increased-value elite seat.

We argue the case of the allocation of scholarships for higher education abroad is a suitable application of this model: increased corruption in the choice of scholarship recipients results in increased likelihood of ‘getting an elite seat’ in the future. In the remaining of the paper we focus on measuring this increase in corruption.

3 Historical and Institutional Background

São Tomé and Príncipe (STP)⁷ is a two-island West African country with 193,000 inhabitants⁸. It is also one of the poorest countries in the world (USD 1200 in 2003⁹, ranking 207 out of 233 countries). It achieved independence from Portugal in 1975, after five centuries of colonization. As a common trend in Lusophone Africa, STP began as a socialist regime. Democratization was initiated 1989, with the first multiparty elections being run in 1991. This process was accompanied by a significant increase in Western foreign aid, which

⁶ As in the classic model, under competition ($n \rightarrow \infty$), the total value of corruption costs will be approaching the total rent at stake in the model.

⁷ See Hodges and Newitt (1988), Seibert (2006) for a thorough historical coverage of this country.

⁸ 2006 estimate at CIA World Factbook.

⁹ CIA Factbook, 2006.

was associated with economic reform (e.g. macroeconomic stabilization, privatization of land).

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 happened in STP. After those, joint exploration with Nigeria was agreed, and several concession auctions took place (e.g. in 2003 bidding for a first set of joint blocks amounted to 237% of STP GDP). IMF estimates production will begin soon, with 30,000 barrels per day in 2013. We will argue ahead these facts can be seen as creating the exogenous variation in a natural experiment. Recent work (Vicente, 2006) showed these happenings seem to have led to a substantial increase in a variety of sector-perceptions of corruption after 1999, with specific prominence in education and vote buying¹⁰ (sectors where future political power is most clearly at stake – which renders support to our theoretical proposition).

Also in the middle of this post-oil discovery process, a coup attempt took place in 2003, with both President and plotters linking it to the increased struggle for political power, the fundamental thesis of this paper. This coup was halted through the intervention of the international community. It is worth noting that several positive steps were taken in its aftermath, during 2004: a National Forum (deliberative democracy) on oil was held¹¹, a ‘state of the art’ law of oil revenues management was passed, and international standards for good oil governance were adopted (Extractive Industries Transparency Initiative).

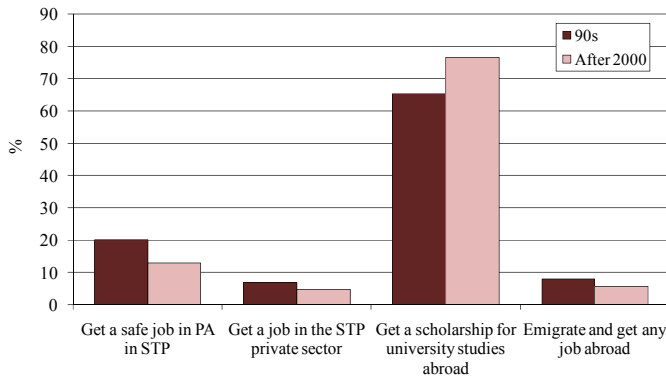
Specifically on education, STP does not have any university¹² and only 0.7% of the population is reported to have attended higher education (see Table A1 in Appendix for census data). Most interestingly, beginning in colonial times and similarly to most African countries, there has been a long tradition of scholarships for higher education abroad (in donor countries, mostly former colonies). These scholarships remain extremely popular to this date in STP: when asked about which would be the best option after secondary school for a high school graduate, 75% of households in our 2006 survey chose ‘getting a scholarship for university studies abroad’ (see Chart 1).

¹⁰ Frynas et al (2003) corroborate an increase in funds available to finance politicians in the post-discovery years. Vicente (2007), through a field experiment, shows that vote buying is effective in driving the allocation of political power in STP.

¹¹ See Humphreys et al (2007) for description of this countrywide exercise of deliberative democracy.

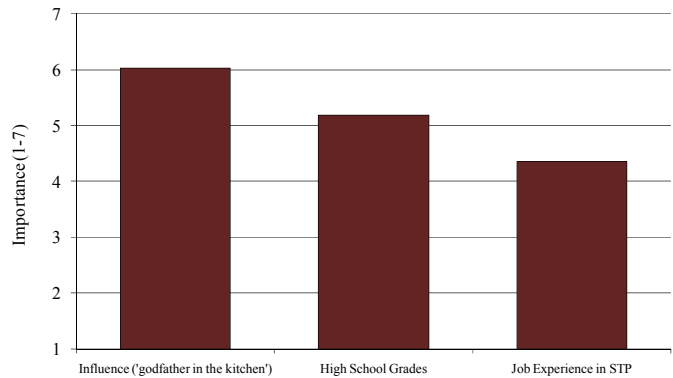
¹² This is despite the existence, from the end of the 90s, of two small professional schools.

Chart 1: When a student finishes high school in STP, what is the best next step for him or for her?



Source: Own data (STP 2006).

Chart 2: Criteria for Allocation of a Scholarship - Applicant Households



Source: Own data (STP 2006).

Crucially, when asked about the importance of different criteria in deciding the allocation of the scholarships for studying abroad, households were resolute in stating ‘personal connections’ as the most important item (Chart 2). We explore in this paper measures of corruption associated with the use of these ‘personal connections’ or influence (or ‘godfather in the kitchen’, to use a popular and specifically Sao Tomean expression).

4 Real Corruption

4.1 Data for Scholarship Contests and Estimation Approach

We gathered data on the characteristics of all students in the 11th year of schooling (last one being offered in the STP high-school system) in the period 1995-2005, and on the characteristics of all scholarships offered in the period 1994-2006. These will enable an analysis of the quality of the winners of scholarship contests (relative to their pools of applicants) over time.

For the characteristics of students we collected data from all schools offering the 11th year in STP. These schools are: the Liceu Nacional of São Tomé city (the main school of the country, located in its capital¹³), the Liceu of Príncipe (in Santo António, Príncipe Island), and the Instituto Diocesano de Formacao (in São Tomé city, private school, funded by Portuguese

¹³ A small set of data, for some evening classes, was gathered from a preparatory school (Patrice Lumumba).

agencies). The variables we assembled for each student were: full name, year of birth, country of birth, and individual final grades for all courses. Table A2 shows descriptive statistics for these data, 4392 students in total. We can see that Liceu Nacional accounts for 94% of the students and that the number of students finishing studies generally increases from 1995-1996 to 2004-2005. Note that we will use ahead simple averages of course grades; however, provided the fact that many students do not have grades for all courses (i.e. did not finish the 11th year) and that (despite that) they may have been awarded scholarships as well, we adjust the average grade by the number of courses (the overall average of this adjusted grade is 9.1 on the original grade scale of 0 to 20).

Concerning the characteristics of the scholarships, we collected data from the Ministry of Education in STP, the Portuguese Development Agency (IPAD), the Calouste Gulbenkian Foundation (based in Portugal), and the French Foreign Affairs Ministry. The latter is the only source of scholarships that also has been retaining allocation decisions (through its local embassy); all others rely on the first (Ministry of Education of STP) for those decisions. In the period 1994-2006, the Ministry of Education of STP decided on scholarships to study in Brazil, Cuba, Macao, Morocco, Mozambique, Portugal (funded through IPAD and Gulbenkian), and Taiwan. We were therefore able to cross-check information for both Portuguese sources of scholarships. The variables we collected for each scholarship allocated were: full name of student, year of allocation, and institution/degree attended. Tables A3 and A4 show the destination and year distributions of scholarships (respectively). Note that these numbers regard those students who were found in our student records, i.e. matched students¹⁴, 700 in total. From the observation of these tables we conclude Portugal (mainly through its Aid Agency), Cuba, and Brazil took on the largest contingents of scholarship recipients, and that the number of scholarships allocated increased substantially after 2001. In Appendix, in Table A5 and Chart A1, we also show grade percent ranks per destination and destination preferences (respectively).

In Charts A2-11 in Appendix, we show for each of the 10 years of data we have, from 1995 to 2005, the number of students by grade category, i.e. the absolute distributions of grades per

¹⁴ Most likely, winners of scholarships not found in our records of students correspond to students finishing high school studies before 1995. Since all sources point to a very small number of students in the 11th year in early 90s, and given we will control ahead for time trends, we are hopeful no important bias for our exercise will arise from not having data from earlier (than 1995) potential applicants.

year, for both winners of scholarships and potential applicants. There we make an important assumption in defining the pools of potential applicants: any student after her last year in high school remains as applicant to getting a scholarship until she is allocated one.

These comparisons between grades of potential applicants and winners are the basis of our first proxy of corruption, the average lag of scholarship winners (displayed in Table 1). This is the average difference between applicants and winners, when both are ordered from highest to lowest. If all highest grade applicants are allocated scholarships, this average difference would be 0.

Table 1: Average Grade Lag of Scholarship Winners

1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
1.4	1.4	1.5	2.6	3.8	4.1	3.6	2.6	3.6	2.4
1.4					3.5			2.4	

Sources: STP Liceu Nacional, Liceu Principe, Instituto Diocesano de Formacao;
Ministry of Education STP, IPAD (Portugal), FCG (Portugal), French Government.

The main assumption here is considering that potential applicants correspond to actual applicants – in defending it we make use of the 75% number we presented in Chart 1: three quarters of the population seem to favor getting a scholarship to study abroad over any other career choice¹⁵. A better assumption would need information on the actual pools of applicants: this is made impossible in face of the frequent use of informal application processes in STP, themselves based on personal knowledge of interested students on the part of the officials in charge of the allocations.

The other implicit assumptions in interpreting these measures as corruption are that: average grades should be the unique criterion for the allocation of scholarships; and not following the grades is corruption. The first is considered a reasonable procedure when dealing with high school graduates all over the world; in addition and in practice, there is no other formal means of assessment in STP (e.g. interviews). We will however consider the other possible resumé-based items in the econometric analysis below, aimed at explaining scholarship allocations:

¹⁵ In addition, it is worth noting that the 75% figure does not rule out that, in the remaining 25% households, getting a scholarship is just a less frequent real possibility. Within the households that stated ‘getting a scholarship’ at the top of the list, we may also see that several applicants correspond to one household. We may therefore interpret our 75% number as a lower bound on the percentage of potential applicants who are in fact interested/applying to scholarships.

age (as proxy for experience) and school attended (to control for differing quality of education for the same grade). The second assumption just rules out that the officials who allocate the scholarships follow unexpected criteria such as randomness.

Note that in Table 1 we see a clear increase in corruption from the period 1996-1997 to the period 2000-2004, corresponding to 11% of the grade scale, with a downward adjustment after 2004 of 6%.

In the econometric analysis below we will attempt to identify corruption over time by exploring specific family connections that can be captured through surnames. In the Portuguese-speaking tradition, full names in STP are unusually long, with average 4.52 names (standard deviation 1.08) in our student database. This enables a good probability that names in common between a specific student and a specific politician mean that student has a family link to that politician. That is the matching exercise we conduct: we construct a measure of family connectedness to the political elite by matching names. We will test ahead whether this variable explains scholarship allocations.

For that purpose, we assembled a database of politicians, containing the full names of all presidents, members of government, and members of national assembly in the post-colonial period (1975-2006), with information on specific mandates. 438 individual politicians were identified – Table A6 provides summary statistics on these politicians and the corresponding distribution of mandates.

We use ahead two measures of family connectedness to the political elite: one based uniquely on the number of names in common between the student and the set of politicians (minimum 2 names in common with a given politician for a non-zero degree of connectedness to be attributed), and one based both on the names but also on the number/type of mandates by each connected politician (i.e. weighting for the importance of connected politicians).

The latent specification we consider in running our Probit regressions of the likelihood of being attributed a scholarship is as follows:

$$R_{i,t} = a + bX_i + cI_i + dt + eG_i + ftG_i + \varepsilon_{i,t},$$

where subscript i is used for individual i , t is the time dummy variable (before 0; after 1), R is the rating of the applicant (if $R \geq 0$, the applicant gets a scholarship), X is a vector of

individual controls (e.g. year born, school attended), I is the measure of connectedness with the political elite, and G is the grade.

Note that we continue interested in differences in the importance of the grades over time (coefficient f). In the specification above, contrary to our simple exercise of Table 1, we control for other characteristics of applicants. Importantly, this approach also controls for differences over time that are not associated with the cross-section of grades, such as differences in the number of allocated scholarships (embedded in the time coefficient).

4.2 Econometric Results

Tables A7-A10 in Appendix show the simple year-regressions (versions of the above specification with no time dimension included). We begin by displaying regression results using average grades and the private school dummy as the explanatory variables; then add connectedness to politicians - version using names only; then substitute the latter by the connectedness version using politician ratings as well; finally, we keep the more complete measure of connectedness and substitute the private school dummy for the year-born variable (since this variable is only available for public schools, we lose the private school observations in this regression).

We can easily conclude that grades are important determinants of the likelihood of being allocated a scholarship: apart from 1996 and 1997, where the number of observations is lowest, the average grade is a solid (at the 1% statistical significance) determinant of these allocations – an additional grade point seems to increase the probability of getting a scholarship on the 1-3% range. In addition, attending the private school seems to increase a student's probability of being funded to study abroad (only significant for years where the number of observations is highest). On the other hand, experience (age) appears to damage those chances (statistically significant at the 1 % level). Interestingly, family connectedness to politicians seems to increase the likelihood of receiving a scholarship - although effects are relatively low, statistical significance ranges from 1 to 5% in half of the years (see Table A9). It is worth pointing out that the more complete measure of family connectedness to the political elite (embedding information on the rating of connected politicians) features higher

statistical significance than the version based on matching names only. We employ this name and mandate version in the remaining regressions.

In Tables 2 and 3 below, we display the results of regressions applying the full time-varying specification introduced above. The first table concerns the pre (before 1998) and post-oil discovery (2000-2004) periods; the second table regards the follow-up comparison between the 2000-2004 and the post-2004 periods. Apart from featuring the explanatory variables already used in the last set of regressions, these regressions add a dummy variable on allocation of scholarships by the French embassy, the only agent deciding on who gets the scholarships apart from the STP Ministry of Education – it controls for possible unobservable differing characteristics targeted by this allocating body.

Table 2: Scholarship Allocations - Pre- and Post-Oil Discovery

		Dependent Variable ----->		Allocation of Scholarships			
				Probit			
Main Explanatory Variables	time	m.e.	0.205	0.202	0.2	0.146	
		std err	0.046***	0.045***	0.043***	0.034***	
	average grade	m.e.	0.088	0.086	0.085	0.042	
		std err	0.029***	0.029***	0.028***	0.029	
	time*average grade	m.e.	-0.07	-0.068	-0.068	-0.025	
		std err	0.030**	0.030**	0.029**	0.03	
	private school	m.e.	0.014	0.015	0.013		
		std err	0.022	0.023	0.022		
	politicians (name & mandate)	m.e.		0.001	0.001	0	
		std err		0.000*	0.000**	0	
	french allocation	m.e.			0.204	0.228	
		std err			0.035***	0.042***	
	year born (public schools only)	m.e.				0.008	
		std err				0.002***	
Number of Observations			3092	3092	3092	2586	
Lik. Ratio (Chi2)			208.53	212.17	255.17	210.92	

Note: Standard errors reported. * significant at 10%; ** significant at 5%; *** significant at 1%

The results of Tables 2 and 3 mainly re-state what we have seen in Table 1 above: a decrease in the importance of grades in determining scholarship allocations after the oil discovery, and a small increase on that effect after 2004. The first corresponds to a -7% change in the probability of getting a scholarship with an additional grade point (after oil, over the baseline effect given by the coefficient of ‘average grade’); the second is a small adjustment of +1%. Note that private school and year born retain the signs already discussed, and that family connectedness to the political elite yields significant and positive effects on the probability of

being allocated a scholarship. This is clear evidence that family favoritism is an important part of the choice of recipients of scholarships¹⁶.

Table 3: Scholarship Allocations - Pre- and Post-2004

		Dependent Variable ----->		Allocation of Scholarships			
				Probit			
Main Explanatory Variables	time	m.e.	-0.134	-0.135	-0.135	-0.137	
		std err	0.029***	0.029***	0.029***	0.032***	
	average grade	m.e.	0.014	0.013	0.013	0.012	
		std err	0.001***	0.001***	0.001***	0.001***	
	time*average grade	m.e.	0.005	0.005	0.005	0.005	
		std err	0.002**	0.002**	0.002**	0.002**	
	private school	m.e.	0.035	0.036	0.033		
		std err	0.015**	0.015**	0.015**		
	politicians (name & mandate)	m.e.		0.001	0.001	0.001	
		std err		0.000***	0.000***	0.000**	
	french allocation	m.e.			0.168	0.171	
		std err			0.021***	0.025***	
	year born (public schools only)	m.e.				0.005	
		std err				0.001***	
Number of Observations			6139	6139	6139	5165	
Lik. Ratio (Chi2)			448	458.63	532.04	440.27	

Note: Standard errors reported. * significant at 10%; ** significant at 5%; *** significant at 1%

5 Perceived Corruption

5.1 Experimental Design

In this section, we pursue an alternative to the exercise of corruption measurement pursued above, which was based on actual data of the allocations of scholarships. Note that in the last section we could only identify differences in the importance of the grades in determining allocations. Although those differences can be directly interpreted as differences in corruption, we aim at finding more specific evidence of differences in ‘abuse of public office

¹⁶ As an extension of the specification explored in the regressions tables shown in the paper, we also attempted to fit time differences associated with the connectedness variables. However, these turned out to be statistically insignificant at standard levels. Although differences over time on the effect of grades are patent symptoms of varying corruption (as argued about the raw measures of Table 1), we could not directly identify these differences for the specific ‘family-favouritism’ type of corruption that is captured by our connectedness variables. That may be due to measurement noise (still possible, even after proving these variables are capable of explaining allocations) and/or simply to the limits of families in providing ‘connections’ (i.e. they may be stretched to their limits during the whole time frame of this exercise, leaving the bulk of the interim time variation to other non-family -based favouritism).

for private gain'. For that purpose we designed an experimental framework focusing on perceptions of corruption by households.

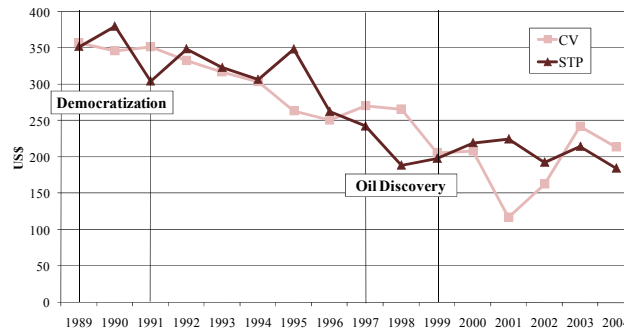
We will derive corruption differences over time (pre and post-oil discovery, i.e. 1991-1997 vs. 2000-2004; and pre and post-2004, i.e. the follow-up comparison, 2000-2004 vs. 2004-2006) by making use of the comparison of perceptions of more vs. less informed households (internally to STP), and of STP versus control-country households (externally). These measures stem from three different purposely-tailored, nationwide, representative household surveys: two conducted in STP and one conducted in the control-country, Cape Verde (CV).

The internal comparison uses a set of alternative options for defining an informed group: namely, we explore urban, public administration, and experienced (with public services) respondents as proxies for informed groups. Note that identification is achieved through cleaning the differences captured by informed groups from 'public opinion' differences portrayed by uninformed groups.

The external exercise is secured by a comprehensive set of similarities between STP and CV, and the fact that CV has not found significant natural resources in its territory to date. Both countries are island countries in West Africa¹⁷ and former Portuguese colonies gaining independence in 1975. Their post-independence political and economic shocks are extremely alike: both countries started as socialist regimes until steps towards democratization were taken in 1989, leading to their first multiparty elections of 1991 in both countries; economic reform associated with foreign aid took on as the main driving force of the early 90s in STP and CV (Chart 3 below depicts the remarkable similarity of aid per capita in both countries after 1988); post-democratization political cycles/changes in ruling parties were mostly contemporaneous in both countries. A note is also due to the strong direct cultural links STP and CV share: important waves of emigration from CV to STP took place in the twentieth century, the last of which in its second half. Carreira (1982) reports that, embedded in the colonial strategy, almost 35 thousand Cape Verdeans left to STP in the period 1950-1970.

¹⁷ Cape Verde is a nine-island country with 423 thousand inhabitants (CIA World Factbook estimate, 2007).

Chart 3: Net ODA Grants per capita in STP and CV (2003 prices, US\$)



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

This natural setting enables us to explore an alternative (to the internal) identification strategy. This relies on contrasting the opinions of similar households, as given by their demographics, while at the same time helping to isolate the effect of the oil discovery. On the latter, a finer estimate will be at stake through controlling for competing trend effects of the potentially-relevant democratization process of the early 90s (arguably the main post-independence shock in both countries).

Importantly, as some of our data will be based on retrospective questions (with regard to the pre-oil discovery period, 1991-1997), the survey instrument design was specifically oriented to elicit 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 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 denominate ‘placebo year’). 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” bias¹⁸.

5.2 Data Collection: Tailored Household Surveys on Corruption

¹⁸ This is particularly suitable when we use the Cape Verdean control group, since the same type of respondents are compared, and no a priori reason for differential bias exists.

Data analyzed in this section come from three purposely-conducted household surveys on perceived corruption: STP 2004 (April/May) including 841 respondents in 30 out of the 149 census areas (20%) of the country; CV 2006 (January/February) counting 1066 observations in 30 out of the 561 census areas (6%) of the country; and STP 2006 (July) featuring 1275 households in 50 out of the 149 census areas (33%) of the country¹⁹. These surveys were conducted by teams of local interviewers; importantly, all teams were recruited, trained, and accompanied in-field (at all locations) by the author. Maps for the overall 80 sampled census areas are displayed in Appendix²⁰ (Maps A1 and A2).

All three surveys followed the same basic household sampling procedures: random choice of census areas, weighting by the number of households (using last available census data from both countries); at each census area, households chosen were distributed as evenly as possible (the n^{th} house was systematically sought). We therefore can conclude that all three surveys have representative samples of households. However, since the first two surveys asked questions retrospectively, they additionally required that specific respondents had at least 30 years of age; the STP 2006 survey, which only asked about the period 2004-2006, required that subjects had at least 18 years. Note also that since we obviously faced non-respondents (on whom we gathered some proxy data during the fieldwork – on gender and approximate age, schooling, and income), and since there were small differences in the number of attempted interviews in the different census areas, we opted for using weights in our regressions ahead – this is done for consistency with the sampling methodology, with differences for un-weighted results being negligible.

The questionnaire on perceived corruption was the same in all three surveys. It was constructed following two basic principles: asking questions about specific public services/allocations to the individuals who had experience with those particular services/allocations; and eliciting respondents' memories by making use of their personal history (this was not needed in STP 2006, since no retrospective questions were asked)²¹. On

¹⁹ The census areas in STP 2006 included all census areas visited in 2004, plus an additional set of 20 locations.

²⁰ Details on the survey fieldwork conducted, including questionnaires used, are available at <http://users.ox.ac.uk/~econ0192/fieldwork.htm>.

²¹ The interviewers referred to the periods of interest in a personal way - e.g. for the subject that had her first child in secondary school in 1994, in the scholarship corruption question, the interviewer referred to the period before-oil announcement, 1991-97, as the period when the respondent's first child was in secondary school.

the former, experience was assessed by asking directly or indirectly²² about proximity to public services. This information is compiled into an index of connectivity/usage to/of public allocations/services that we use ahead to define one of our proxies of informed households²³.

The crucial question asked concerning corruption in scholarships was ‘In the Sao-Tomean (Cape Verdean) reality, when allocating scholarships for higher education abroad, what has been the need to know someone important?’, with the 1-7 response scale going from ‘Not at all necessary’ to ‘Extremely necessary’. As for all other corruption scales in the questionnaire, these answers were approached qualitatively, in a stepwise fashion, starting with ‘Not necessary’ (1-3), ‘More or less’ (4), and ‘Necessary’ (5-7), and continuing with having the interviewer deepen the options of one of the two opposite sides in case that corresponded to the first answer by the respondent. This way, we are confident the scale was approached in a linear way²⁴.

Finally, a description is due on how we proxied the ‘Good Old Times’ bias (tendency of respondents to report worsening conditions over time), relevant in controlling perceptions about the past (1991-1997): two direct questions were included in the beginning of the questionnaire concerning general pessimism associated with time²⁵, and a placebo shock (the relevant period 1991-1997 was divided into 1991-1994 and 1994-1997) was included in one of the perceived corruption questions. On the second measure, any difference perceived between the two sub-periods will be taken as an alternative measure of pessimism over time (provided no prominent shocks happened in the placebo period, 1994, in either STP or CV).

5.3 Descriptive Data and Estimation Approach

We now provide a first look into the perceptions data, as well as the outline of the specific econometric strategy we follow regarding these data.

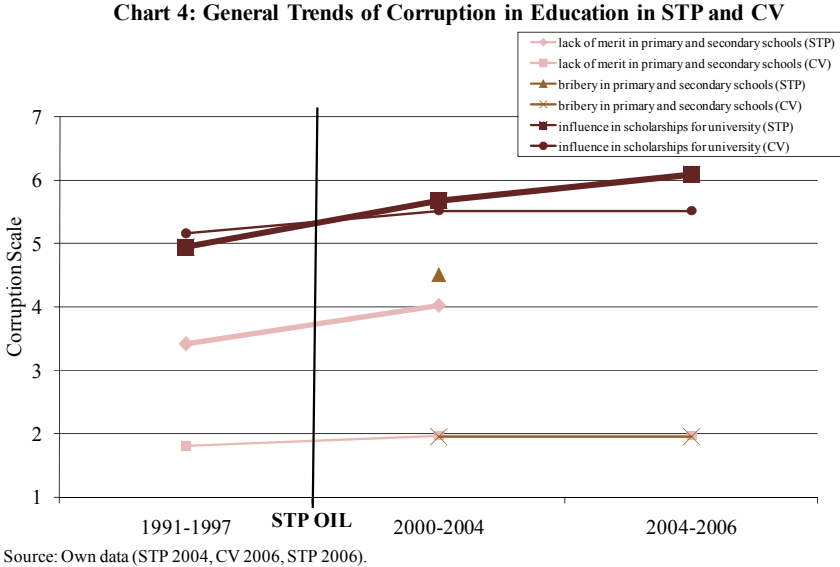
²² Direct questions on experience with public questions were posed immediately before the relevant corruption questions. Indirect information was elicited through standard demographic questions posed in the beginning of the questionnaire.

²³ The specifics of the construction of this index, including weights and information used, can be consulted in Vicente (2006).

²⁴ In Vicente (2006) ordered probit regressions of perceived corruption variables (where their latent scales are estimated) are contrasted with OLS regressions (assuming the referred linearity in the scale): no relevant differences arise in these results.

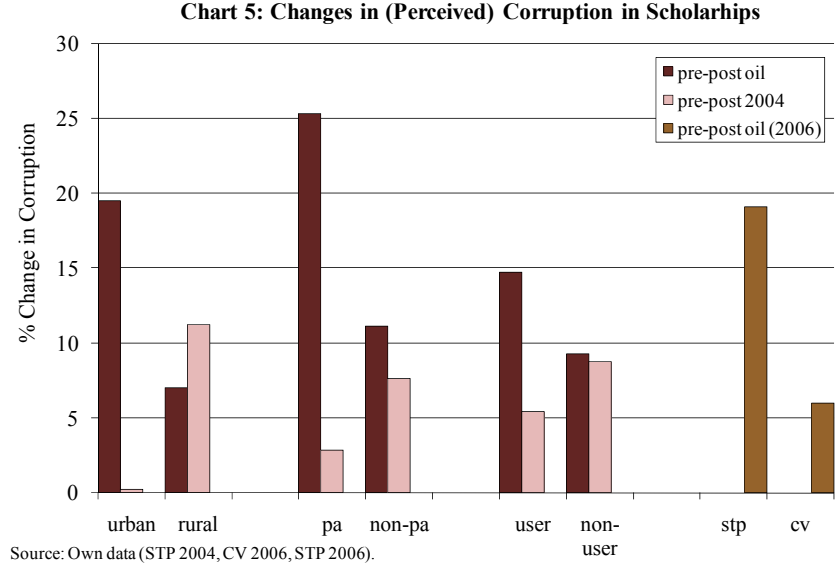
²⁵ Agree/disagree questions on a stepwise 1-7 scale: ‘Good times were those when you were young’; ‘The future of STP (CV) will be better than the present’.

In Chart 4 we present the general trends of average perceived corruption in education in both STP and CV, as given in our household surveys. These include data on the corrupt nature of the allocation of scholarships as well as perceptions of corruption earlier in the education system, in primary and secondary schools. We observe that although the level differences between STP and CV are low on scholarships but high on primary and secondary schools (STP has much higher perceived corruption, seen either as lack of merit or crude bribery), it is clear that the difference before-after oil discovery years is much steeper for STP. We also note that on STP scholarship favoritism we seem to see a slow-down of the increasing trend.



In Chart 5 we deepen our analysis through the comparison of our treatment and control groups, both internally and externally. We provide percentage differences in corruption over time, pre-post oil (with post oil being 2000-2004), pre-post 2004, and pre-post oil (with post oil being 2000-2006). We detect clear internal effects pre-post oil, as urban, public administration, and experienced (with public services/allocations) individuals report higher increases in scholarship corruption during those periods. The opposite is true for the pre-post 2004 move: informed respondents seem to see a lower increase than their less-informed counterparts. However, these differences are smaller than the first ones in absolute terms, which is consistent with time changes uncovered in the real corruption section. Indeed, the

external comparison STP vs. CV confirms that corruption seems to have increased from 1991-1997 to the whole period of 2000-2006.



As the econometric strategy we follow to derive the time differences of interest in perceived corruption on scholarship allocations, we employ simple difference-in-difference estimators. These can be described in the context of the following specification:

$$K_{i,t,l} = a + bX_i + cY_l + dt + eT_{i,l} + ftT_{i,l} + \varepsilon_{i,t,l},$$

where subscripts i and l stand for individual i and location l , t is the time dummy variable (before 0; after 1), $K \in \{1, \dots, 7\}$ is the variable of perceived corruption, X is a vector of individual controls²⁶, Y is a vector of district fixed effects, and T is the treatment variable. Note that when using the connectivity indicator (measuring the respondents' proximity to public services/allocations), we face a 'continuous' treatment – that index can take any value on the 0-1 range -, which has implications for estimate interpretation. Most importantly, our

²⁶ These controls are chosen among a wide and rich set of individual and household measures including basic demographics (gender, age, household size, number of children), religion, ethnicity and nationality, marital condition, health, schooling, occupation, political preferences, financial condition, and psychological measurements (including the ones relating to the 'Good Old Times' bias, described above). Refer to Vicente (2006) for a more thorough characterization of these control variables.

coefficient of interest is f , which can be interpreted as the corruption time change when differencing treatment and control perceptions.

5.4 Econometric Results

We now analyze the results of our regressions of perceived corruption in the allocation of scholarships. Table 4 is dedicated to the assessment of the 1991-1997/2000-2004 time change, based on internal-to-STP comparisons only. Table 5 completes the sequence by being devoted to 2000-2004/2004-2006. And Table 6 focuses on the overall pre-post oil discovery announcement periods, 1991-1997/2000-2006, with the STP sample being contrasted with its CV counterpart (external comparison only). Each basic regression comparison appears in the following specification order: the uncontrolled difference-in-difference regression is displayed first; then fixed effects are added; then demographic controls are added.

In Table 4, concerning the pre-post oil discovery (until 2004) time change, we can observe clear increases in perceived corruption, as given by differences between more informed and less informed respondents. Urban-rural and public administration-other occupations yield time-difference estimates ranging from 11 to 13% of the subjective scale; the estimate provided by using the connectivity indicator, interpreted differently (as the difference between highest and lowest possible individual proximity to public services/allocations, yields a 33% difference). The group-based estimates (first two numbers) are close to the ones advanced for the direct corruption measures of last section. Note that the coefficients of interest are highly statistically significant (at the 1% level) and extremely robust to the addition of controls (including ‘Good Old Times’, GOT, controls) to the regression equations.

Table 5 displays less statistically significant coefficients for the 2000-2004 vs. 2004-2006 time passage (still 1% level for urban, but 10% level for public administration, and none for the connectivity indicator -based comparison). Estimates are still fairly robust to the different specifications (least robust for when demographics are added). Most importantly, we depict a clear decrease in corruption, which is estimated at 7.8 to 8.2% of the subjective scale (group-based comparisons), i.e. clearly lower than the pre-post oil discovery increase.

Table 4: Regressions for 1991-1997/2000-2004, STP Data Only

		Dependent Variable ----->	Perceived Influence in Scholarship Allocations								
			OLS								
Main Explanatory Variables	time1	coef	0.49	0.49	-1.65	0.71	0.71	-1.28	0.41	0.41	-1.71
		std err	0.09***	0.09***	0.59***	0.07***	0.07***	0.58**	0.13***	0.13***	0.59***
	urban	coef	-0.27	-0.05	-0.14						
		std err	0.09***	0.2	0.2						
	time1*urban	coef	0.67	0.67	0.65						
		std err	0.13***	0.13***	0.14***						
	pa	coef				-0.7	-0.71	-0.71			
		std err				0.18***	0.18***	0.18***			
	time1*pa	coef				0.88	0.88	0.8			
		std err				0.25***	0.25***	0.25***			
index_connect	coef							-1.23	-1.33	-1.84	
	std err							0.36***	0.36***	0.41***	
time1*index_connect	coef							1.68	1.68	1.96	
	std err							0.50***	0.50***	0.52***	
Constant	coef	5.03	4.79	5.66	4.96	4.98	5.61	5.18	5.23	6.11	
	std err	0.06***	0.20***	0.56***	0.05***	0.06***	0.52***	0.09***	0.10***	0.53***	
District Fixed Effects			No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Demographic Controls (inc. GOT)			No	No	Yes	No	No	Yes	No	No	Yes
Number of Observations			1343	1343	1225	1341	1341	1225	1343	1343	1225
Adjusted R-squared			0.11	0.12	0.12	0.1	0.11	0.12	0.1	0.11	0.12

Note: Standard errors reported. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Regressions for 2000-2004/2004-2006, STP Data Only

		Dependent Variable ----->	Perceived Influence in Scholarship Allocations								
			OLS								
Main Explanatory Variables	time2	coef	0.67	0.66	0.66	0.48	0.48	0.53	0.56	0.51	0.53
		std err	0.09***	0.09***	0.10***	0.08***	0.07***	0.08***	0.15***	0.15***	0.16***
	urban	coef	0.39	0.29	0.22						
		std err	0.09***	0.22	0.23						
	time2*urban	coef	-0.63	-0.63	-0.47						
		std err	0.15***	0.15***	0.15***						
	pa	coef				0.17	0.15	0.08			
		std err				0.17	0.17	0.18			
	time2*pa	coef				-0.57	-0.55	-0.49			
		std err				0.26**	0.26**	0.26*			
index_connect	coef							0.45	0.33	0.11	
	std err							0.34	0.34	0.4	
time2*index_connect	coef							-0.6	-0.38	-0.19	
	std err							0.53	0.53	0.56	
Constant	coef	5.52	5.62	5.87	5.67	5.77	6.01	5.59	5.71	5.95	
	std err	0.06***	0.23***	0.41***	0.05***	0.06***	0.34***	0.09***	0.10***	0.36***	
District Fixed Effects			No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Demographic Controls (inc. GOT)			No	No	Yes	No	No	Yes	No	No	Yes
Number of Observations			1081	1081	1023	1079	1079	1023	1081	1081	1023
Adjusted R-squared			0.05	0.06	0.07	0.03	0.05	0.06	0.03	0.04	0.06

Note: Standard errors reported. * significant at 10%; ** significant at 5%; *** significant at 1%

Finally, Table 6 features the overall comparison between 1991-1997 and the full post-oil discovery period, 2000-2006, using the STP/CV treatment/control groups. Interestingly, we find quite high effects at par to the highest numbers encountered internally for the 1991-1997/2000-2004 comparison: 13% of the subjective scale. These estimates are highly significant (at the 1% level) and robust to the inclusion of controls (although the estimates decrease slightly in the process of adding controls).

Table 6: Regressions for 1991-1997/2000-2006, STP and CV Data

Dependent Variable ----->		Perceived Influence in Scholarship Allocations			
		OLS			
Main Explanatory Variables	time	coef	0.35	0.35	0.37
		std err	0.09***	0.09***	0.09***
	stpcv	coef	-0.24	-1.57	-1.11
		std err	0.09***	0.27***	0.29***
	time*stpcv	coef	0.85	0.82	0.78
		std err	0.13***	0.13***	0.14***
	Constant	coef	5.15	6.31	6.05
		std err	0.06***	0.26***	0.41***
District Fixed Effects		No	Yes	Yes	
Demographic Controls (inc. GOT)		No	No	Yes	
Number of Observations		2434	2434	2174	
Adjusted R-squared		0.06	0.09	0.1	

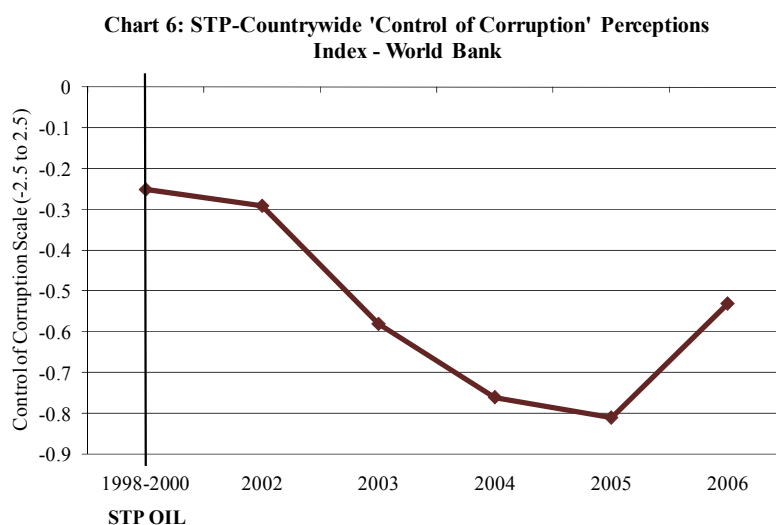
Note: Standard errors reported. * significant at 10%; ** significant at 5%; *** significant at 1%

6 Summary of Results and Interpretation

We are now in position to summarize our findings. This paper targeted two main research questions: the effect of the oil discovery of 1997-1999 in STP on corruption, as measured in the allocation of scholarships for studies abroad, a prominent and highly-valued state allocation in STP; and the comparison of real, direct corruption measurement with perceived corruption measurement over time. For the first question, we theoretically proposed that after an anticipated oil boom shock, the political elite will want to increase corruption in order to raise the probability of being in power (which is now more attractive, even if the boom has not happened yet, as is the case in STP): getting a scholarship to study abroad indeed constitutes a strong instrument to perpetuate membership of the elite of the country.

In Section 4 we explored data on the characteristics of the population of students (crucially their grades), and of the scholarships allocated, in the period 1995-2005. We posed the identifying assumptions that the population of students was a good approximation for the pool of applicants, and that we could define corruption as departures from the objective individual grades. We reached the conclusion that, defined this way, corruption increased by 7% across the periods before 1998, 2000-2004, and decreased by 1% across the periods 2000-2004, after 2004. Importantly we could prove directly that family connectedness is an important determinant of allocations.

In Section 5 we embraced perceptions of corruption but lifted the implicit notion of corruption based on the departure from grades in the allocation process. We pursued identification of time changes in perceived corruption in two ways: through the comparison of internal-to-STP groups (which enabled controlling for public opinion, but relied on comparison of different demographic groups), and through the comparison of a STP sample with a CV sample (which enabled contrasting same demographic groups and controlling for any democratization trends, but relied on the comparison of two different countries, with corresponding limitations). We found the following internally-computed effects: 11 to 13% increase from the period 1991-1997 to the period 2000-2004, and 8% decrease from the period 2000-2004 to the period 2004-2006. Externally we could only compute one pre-post oil effect (to 2006): 13 to 14%. Interestingly, this pattern of perceptions is very similar to the one reported by the World Bank for countrywide corruption. In Chart 6 we show the evolution of the ‘Control of Corruption’ index by Kaufmann et al (2007) – it displays a 10% increase in corruption after 2000 (until 2004) and a 5% overall decrease after that.



Source: World Bank - WBI, Kaufmann et al (2007).

If we rely more heavily on our estimates of real, direct corruption differences, we conclude that oil seems to have induced an increase in corruption on scholarship allocations that was on average 7% in 2004. We also have evidence that a small downward adjustment happened after 2004. This finding lends support to our theoretical proposition. We could also check that perceptions back up the signs of these changes; but, if one unconditionally believes in direct

measurement, perceptions overestimate the magnitudes of these changes – by at least 4% in the first, and by 7% in the second periods. Note however that the direct measurement of corruption we pursued in this paper, although solid, relies on a definition of corruption that targets ‘personal influence’ as ‘missing grades’; our perception measures, on conceptual terms, directly target that influence. This means we do not want to put the whole bottom-line emphasis on the direct-measurement results and their induced perception overestimation.

7 Concluding Remarks

This paper embeds two main policy implications. First, natural resources indeed seem to provide damaging incentives in developing countries like São Tomé and Príncipe, in the sense of leading to increases in corruption. Second, household corruption perceptions, when used to derive changes over time, and using suitable comparison groups, seem to be informative but prone to magnitude overestimation.

The first implication, while highlighting the need to provide careful aid incentives in newly resource-rich countries, provides specific indication that all initiatives aimed at diminishing the ability to capture resource rents (by politicians) are worth considering. In this context and on the specifics of São Tomé and Príncipe, some good news may be at stake for development policy in that the downward adjustment on corruption identified after 2004 is contemporaneous to a set of externally-induced initiatives aimed at improving oil governance (described in Section 3).

The second implication is encouraging for the use of perceptions to draw development policy over time, as turning points seem to be finely captured by these measures, and perceptions are more easily available for a wide range of countries and sectors. However, that does not preclude us from underlining the importance of running full-scale surveys and identifying appropriate comparison groups. Prudence should also be exercised when dealing with magnitudes.

A side note is also due on assessing scholarships for studying abroad as an aid channel. The above-explored data is definitive in conveying the idea that scholarships may be a means to consolidate an elite’s power, while embedding a clearly-inefficient allocative process. Although our prior was that having donors directly run application processes could decrease

corruption, our evidence (although limited in variety) is not consistent with that hypothesis. As we could observe in Table A5, students given scholarships to study in France, the only in our sample to be subject to application processes run by foreign-to-STP bodies, rank among the lowest grade-students (on average) offered scholarships²⁷.

A final remark goes to efficiency. Although we have focused on a corruption setting where the definition of corruption corresponds to that of inefficiency (lower-ability individuals getting scholarships), that is not the case in most settings/allocations/sectors. We believe that, in future research, it is fundamental that ways of identifying the inefficiency associated to corruption are developed empirically.

²⁷ This is despite a good ranking in the preferences of applicants, as seen in Chart A11.

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;
- 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;
- Bertrand, Marianne, Simeon Djankov, Rema Hanna, and Sendhil Mullainathan (2006), *Obtaining a Driving License in India: An Experimental Approach to Studying Corruption*, Quarterly Journal of Economics, forthcoming;
- Carreira, António (1982), *The People of the Cape Verde Islands - Exploitation and Emigration*, C. Hurst & Company;
- 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;
- Frynas, Jędrzej 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;
- Hodges, Tony, and Malyn Newitt (1988), *São Tomé and Príncipe: From Plantation Colony to Microstate*, Westview Press;
- Humphreys, Macartan, William A. Masters, and Martin E. Sandbu (2007), *The Role of Leaders in Democratic Deliberations: Results from a Field Experiment in São Tomé and Príncipe*, World Politics, 58(4), pp. 583-622;
- Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi (2007), *Governance Matters VI: Governance Indicators for 1996-2006*, World Bank Policy Research Working Paper 4280;
- 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*, Economic Journal, 116(508), pp. 1-20;

- Olken, Benjamin (2006), *Corruption Perceptions vs. Corruption Reality*, NBER, Working Paper 12428;
- Olken, Benjamin (2007), *Monitoring Corruption: Evidence from a Field Experiment in Indonesia*, *Journal of Political Economy*, 115(2), pp. 200-249;
- Reinikka, Ritva, and Jakob Svensson (2004), *Local Capture: Evidence from a Central Government Transfer Program in Uganda*, *Quarterly Journal of Economics*, 119(2), pp. 679-706;
- Robinson, James A., Ragnar Torvik and Thierry Verdier (2006), *Political Foundations of the Resource Curse*, *Journal of Development Economics*, 79(2), pp. 447-468;
- 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 (2006), *Comrades, Clients and Cousins: Colonialism, Socialism and Democratization in São Tomé and Príncipe*, 2nd Edition, Brill Academic Publishers;
- 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), *Does Oil Corrupt? Evidence from a Natural Experiment in West Africa*, Oxford University, Working Paper;
- Vicente, Pedro C. (2007), *Is Vote Buying Effective? Evidence from a Field Experiment in West Africa*, Oxford University, Working Paper;

Appendix

Table A1: Schooling in STP

Pre-School	1.8%
Primary (1-4)	38.3%
Basic Secondary (5-9)	23.8%
Pre-University (10-11)	5.6%
Technical	0.5%
Superior	0.7%
No Schooling	29.3%

Source: Census 2001

Table A2: Student Population - Descriptive Statistics

	Number	%	Average	St. Dev.
Total Students	4392			
Liceu Nacional	4126	94%		
Liceu Principe	26	1%		
Instituto Diocesano de Formacao	227	5%		
1996	25	1%		
1997	214	5%		
1998	254	6%		
1999	160	4%		
2000	343	8%		
2001	440	10%		
2002	350	8%		
2003	449	10%		
2004	649	15%		
2005	705	16%		
Born in STP (Public Schools Only)	3737	97%		
Year of Birth (Public Schools Only)			1981	4
Average Grade (raw)			12.4	1.1
Number of Courses Completed			5.9	2.5
Average Grade (adjusted)			9.1	4.0

Sources: STP Liceu Nacional, Liceu Principe, Instituto Diocesano de Formacao.

Table A3: Scholarship Destinations (matched)

Destination	Number
Brazil	149
Taiwan	3
Cuba	172
France	50
Morocco	33
Mozambique	38
Portugal	226
Gulbenkian (Por)	29
Total	700

Sources: Ministry of Education STP, IPAD (Portugal), FCG (Portugal), French Government.

Table A4: Scholarship Allocation Year (matched)

Year	Number
1996	2
1997	7
1998	11
1999	11
2000	15
2001	12
2002	48
2003	208
2004	177
2005	201

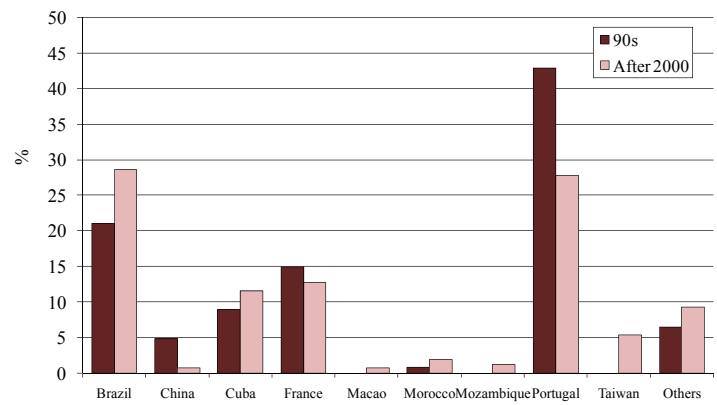
Sources: Ministry of Education STP, IPAD (Portugal), FCG (Portugal), French Government.

Table A5: Average Winner Quality

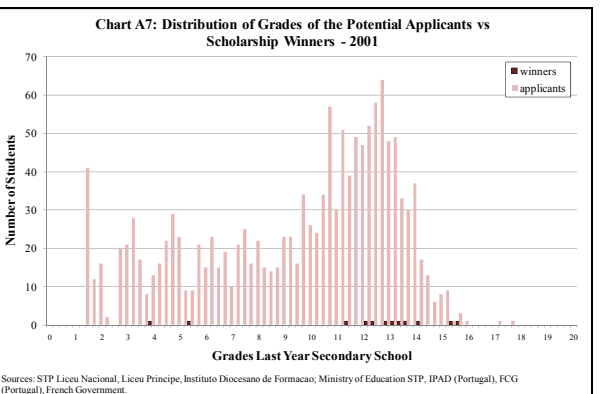
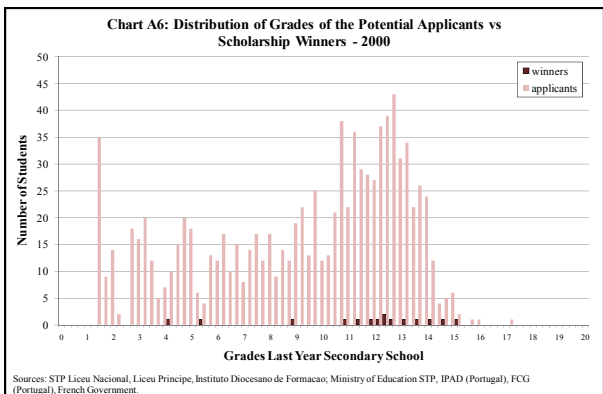
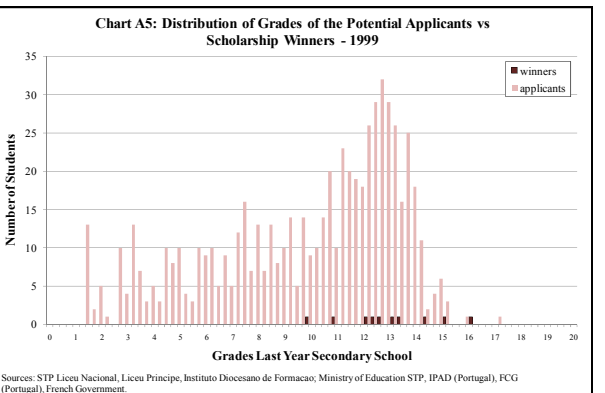
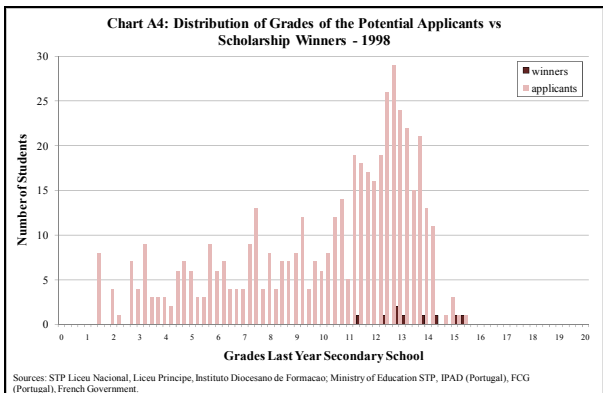
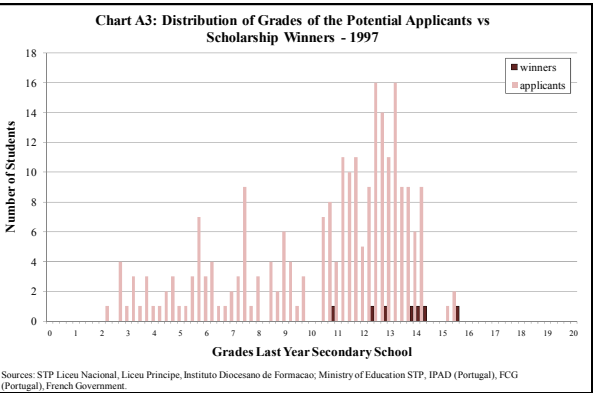
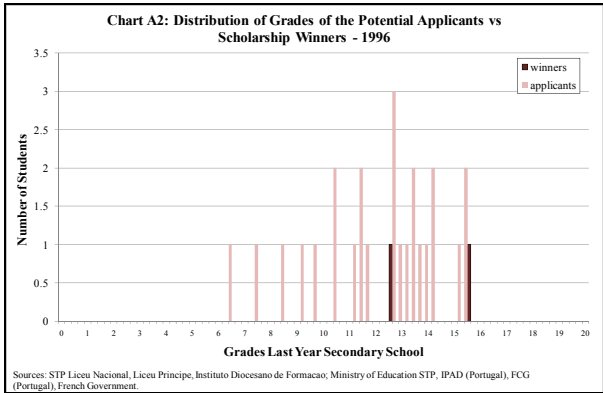
Destination	Grade Percent Rank
Brazil	0.76
Taiwan	0.90
Cuba	0.71
France	0.66
Morocco	0.90
Mozambique	0.51
Portugal	0.73
Gulbenkian (Por)	0.78

Sources: STP Liceu Nacional, Liceu Principe, Instituto Diocesano de Formacao; Ministry of Education STP, IPAD (Portugal), FCG (Portugal), French Government.

Chart A1: Most Preferred Destination for a Scholarship - Applicant Households



Source: Own data (STP 2006).



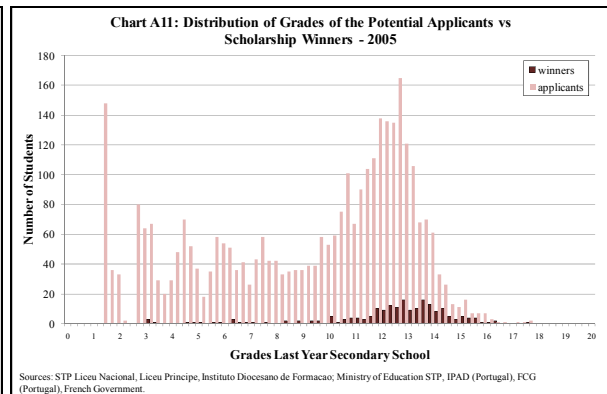
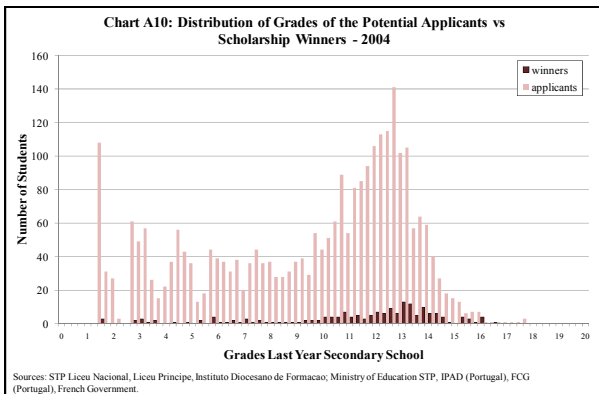
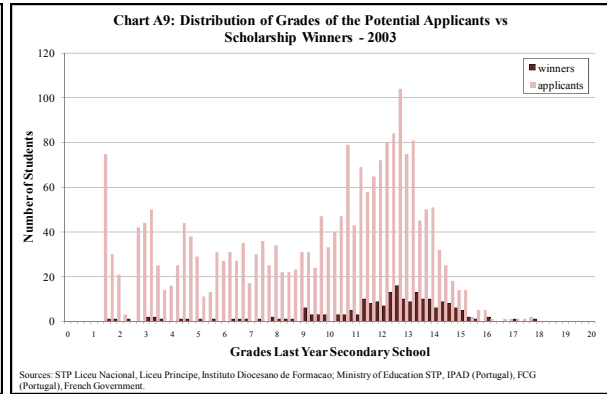
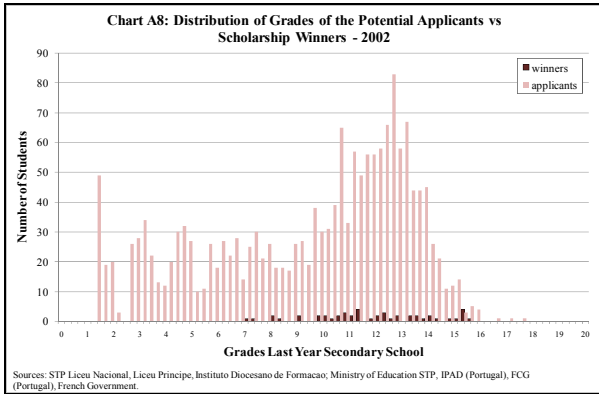


Table A6: Politicians Data - Summary

	President	Government	National Assembly	National Assembly (Subs)	Total
Number of Politicians	3	127	316	67	513
Number of Mandates	5	306	559	235	1105

Sources: STP National Assembly records, Seibert (2006)

Table A7: Scholarship Allocations (1996-2005) - Grades and School

		Dependent Variable ----->	Allocation of Scholarships									
			Probit									
		Year ----->	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Main Explanatory Variables	average grade	m.e.	0.028	0.008	0.014	0.016	0.02	0.026	0.032	0.033	0.031	0.019
		std err	0.046	0.007	0.005***	0.004***	0.003***	0.003***	0.003***	0.002***	0.002***	0.001***
	private school	m.e.		-0.003	-0.039	0.016	0.051	0.015	0.03	0.037	0.069	0.076
		std err		0.075	0.059	0.061	0.058	0.052	0.045	0.039	0.034**	0.027***
Number of Observations			25	237	487	640	974	1401	1736	2171	2773	3284
Lik. Ratio (Chi2)			9.29	1.24	9.09	15.49	40.15	76.14	154.09	214.63	270.1	225.98

Note: Standard errors reported. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A8: Scholarship Allocations (1996-2005) - Politician's Influence (proxy based on names only)

		Dependent Variable ----->	Allocation of Scholarships									
			Probit									
		Year ----->	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Main Explanatory Variables	average grade	m.e.	0.039	0.007	0.014	0.016	0.02	0.025	0.032	0.033	0.031	0.019
		std err	0.049	0.007	0.005***	0.004***	0.003***	0.003***	0.003***	0.002***	0.002***	0.001***
	private school	m.e.		-0.001	-0.036	0.024	0.054	0.02	0.029	0.038	0.069	0.076
		std err		0.076	0.061	0.063	0.058	0.053	0.046	0.04	0.034**	0.027***
	politicians (name)	m.e.	0.004	0.008	0.004	0.005	0.002	0.004	0.005	0.003	0.004	0.003
		std err	0.011	0.005	0.003	0.003	0.003	0.002*	0.002**	0.002*	0.001***	0.001***
Number of Observations			25	237	487	640	974	1401	1736	2171	2773	3284
Lik. Ratio (Chi2)			9.98	3.69	10.86	18.1	40.78	79.31	159.6	217.98	278.14	233.68

Note: Standard errors reported. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A9: Scholarship Allocations (1996-2005) - Politician's Influence (proxy based on names and mandates)

		Dependent Variable ----->	Allocation of Scholarships									
			Probit									
		Year ----->	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Main Explanatory Variables	average grade	m.e.	0.029	0.007	0.014	0.015	0.02	0.025	0.032	0.033	0.031	0.019
		std err	0.048	0.007	0.005***	0.004***	0.003***	0.003***	0.003***	0.002***	0.002***	0.001***
	private school	m.e.		0	-0.034	0.023	0.056	0.021	0.032	0.04	0.072	0.079
		std err		0.076	0.061	0.063	0.058	0.053	0.046	0.04	0.035**	0.027***
	politicians (name & mandate)	m.e.	0	0.001	0.002	0.002	0.001	0.002	0.002	0.001	0.002	0.001
		std err	0.002	0.002	0.001	0.001*	0.001	0.001**	0.001***	0.001**	0.001***	0.000***
Number of Observations			25	237	487	640	974	1401	1736	2171	2773	3284
Lik. Ratio (Chi2)			9.32	1.83	11.09	18.83	41.83	80.96	161.03	218.65	280.07	236.4

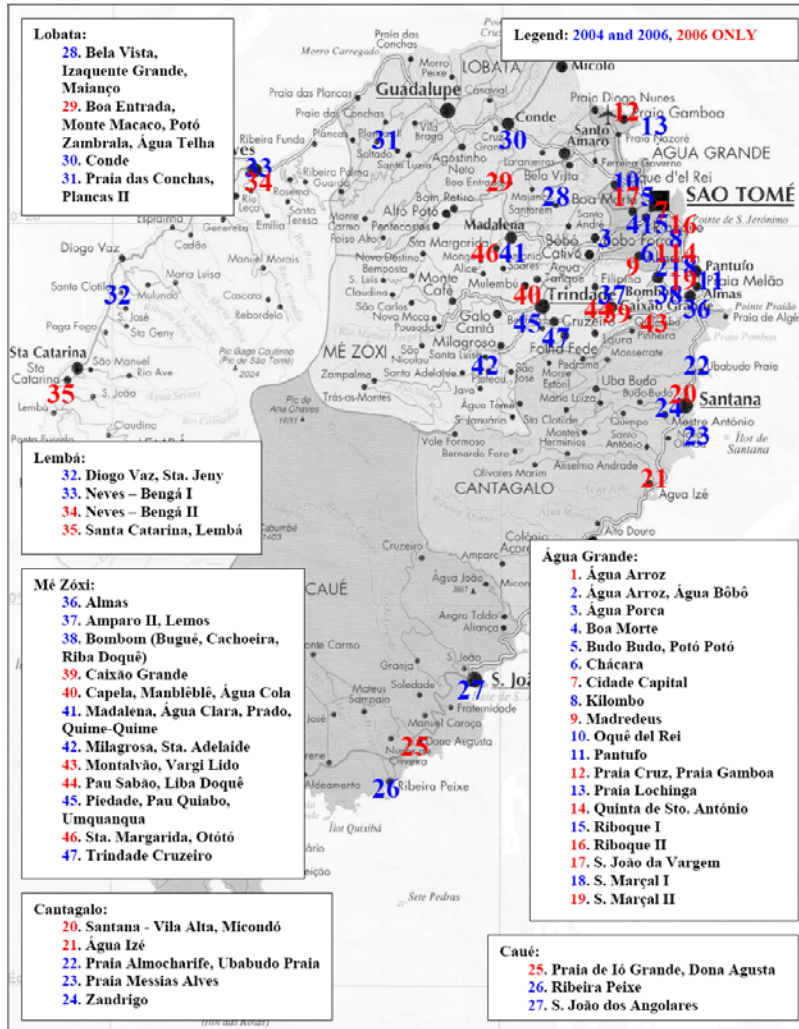
Note: Standard errors reported. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A10: Scholarship Allocations (1997-2005) - Politician's Influence (proxy based on names and mandates)

		Dependent Variable ----->	Allocation of Scholarships								
			Probit								
		Year ----->	1997	1998	1999	2000	2001	2002	2003	2004	2005
Main Explanatory Variables	average grade	m.e.	0.003	0.01	0.012	0.018	0.022	0.028	0.03	0.028	0.017
		std err	0.007	0.004**	0.004***	0.004***	0.004***	0.003***	0.003***	0.002***	0.001***
	year born (public schools only)	m.e.	0.057	0.042	0.044	0.046	0.046	0.038	0.029	0.02	0.008
		std err	0.018***	0.008***	0.008***	0.006***	0.006***	0.005***	0.004***	0.003***	0.002***
	politicians (name & mandate)	m.e.	0.001	0.001	0.002	0.001	0.002	0.002	0.001	0.001	0.001
		std err	0.002	0.001	0.001*	0.001	0.001*	0.001**	0.001*	0.001**	0.000**
Number of Observations			203	434	560	802	1121	1411	1788	2313	2782
Lik. Ratio (Chi2)			10.38	34.6	49.1	83.75	129.25	197.6	243.48	271.17	191.24

Note: Standard errors reported. * significant at 10%; ** significant at 5%; *** significant at 1%

Maps A1: São Tomé and Príncipe Islands – Sampled Census Areas



Maps A2: Cape Verde Islands (Santiago, Sao Vicente, Santo Antao, and Fogo): Sampled Census Areas

