Modeling the Role of Social Interactions in Antisocial Behavior

Technical Appendix to:

Thinking about Corruption in Greece[1]

Costas Azariadis and Yannis M. Ioannides[2]

February 20, 2015

1 Introduction

In this appendix we approach corrupt practices as outcomes of decisions in a social context. One can distinguish conceptually among pure individual propensity towards corruption, other things being equal, and the externality effect that corrupt action by others has on each individual. Greek citizens normally observe corrupt practices at varying degrees, and such observations may feed perceptions that "corruption pays." In fact, the widely used Corruption Index of Transparency International[3] is based on reported perceptions. So, the important question to put to a model is: When does the widespread perception of corruption becomes corruption?

This question can be modeled by the social interactions literature. In a standard formulation, [see Brock and Durlauf (2001), Durlauf and Ioannides (2010), and Ioannides (2013), Ch. 2], aggregate behaviors such as corrupt practices and tax evasion emerge through interactions in a population. It is possible that different occupational groups have different exposure to practices. In such a model imposing self-consistency (in effect the counterpart here of rational expectations) in a manner which connects the expected behavior of the typical agent in a nonlinear fashion with the perceptions of corruption, leads in general to multiple equilibria in actual corruption. This means that different subgroups of the population may cluster into different practices, some at low levels of corruption and others at

---

[1] Submitted to: Meghir, C., C. A. Pissarides, D. Vayanos and N. Vettas, eds. *Crisis in the Eurozone Periphery: Policy Options for Greece*, MIT Press, 2015.

[2] Correspondence: `yannis.ioannides@tufts.edu`

[3] `http://www.transparency.org/cpi2012/results`

higher ones, which are typically stable, and the middle level ones are typically unstable.

For the perception of corruption not to translate into actual corruption individuals must feel incentives not to conform to social perceptions of corruption and of other practices and thus in effect not to "coordinate" on the worst outcomes (which is loosely speaking here, because a Bayes-Nash equilibrium is invoked). Appropriate incentives can be studied on the basis of the model introduced in this appendix. Some of them are discussed informally in the main part of the paper. These incentives must include individual enforcement mechanisms, like legal and administrative sanctions, as well as socially integrated ones, like stigma.

This appendix offers a model which underlies the narrative exposition in the main text. A key result of the model identifies actual practices as emerging as social equilibria. One can base policy design on such foundations, where policy does not just bring about a marginal reduction of corruption but moves the economy to a completely different equilibrium. This is a feature of non-uniqueness, which is particularly attractive in the context of Greece, where it is highly desirable to overturn the impact of expectations that things will go unchanged.

2 Model

Corruption practices can be very different and can adapt to specific cultural features of given societies. So, a very general model is called for. We borrow from Durlauf and Ioannides (2010) a basic model of social interactions that involves decisions over a discrete set of choices. It aims at capturing many of the interesting implications of integrating feedback from social factors and conditions into individual behavior.

Consider a population of I individuals each of whom chooses between L different alternatives; individual choices are denoted by ω_i, and the choice set is: $S = \{0, 1, \ldots, L-1\}$. Each agent i is associated with a group $g(i)$, which is defined as those members of the population whose behaviors and characteristics enter as direct arguments in i's decision problem. For example, different groups of the population may have different options in engaging in corrupt practices. Whereas everyone deals with providers of public services of different kinds, depending on their nature of the incomes, individuals have different options in engaging in tax evasion or even (legal) tax avoidance. Those on salary, wage or pension incomes have income taxes withheld at source. Self-employed individuals can be small shopkeepers with more or less visible business practices, or lawyers with access to information regarding the risk associated with different modes of tax compliance. We start by assuming that each actor is a member of a single group. We describe how social interactions affect individual and aggregate outcomes regarding specific practices.

Each of the possible choices ℓ produces utility $V_{i,\ell}$ for individual i. We conceptualize choice-specific utility as having three distinct components. The first, $h_{i,\ell}$, is *private deterministic utility.* It is private in that it does not exhibit direct dependence on the choices of others and is deterministic as it is treated as known to the modeler; in econometric work this is operationalized by assuming that it is a known function of observables and estimable parameters. The second component is *deterministic social utility* and captures the dependence of individual i's utility on specific choice by others. If individual i chooses $\omega_i = \ell$ and j chooses $\omega_j = s$, then individual i receives $J_{i,j,\ell,s}$. For example, a young aspiring lawyer cannot be seen driving other than a sparkling BMW. This is quite general as each pair of individuals and pair of choices is assigned a separate payoff. The payoffs will be restricted in order to produce tractable results. We assume that the payoffs from the choices of others are additive and so are both deterministic private and social utility. A third component is a *random utility* term, $\epsilon_{i,\ell}$. These random utility components are assumed to be independent across choices and individuals; this assumption can be relaxed in a straightforward way (e.g., by the nested logit model, as we see further below).

Together, these three components are summed so that

$$V_{i,\ell} = \mathcal{E}\{V_{i,\ell}\} + \epsilon_{i,\ell},$$

with the expected utility taking the form:

$$\mathcal{E}\{V_{i,\ell}\} = h_{i,\ell} + \sum_{s=0}^{L-1} \sum_{j \neq i} J_{i,j,\ell,s} p^e_{j,s|i}, \tag{1}$$

where $p^e_{j,s|i}$ denotes the probability i assigns to choice s on the part of j. For example, individual i is not sure whether individual j tax evades, say chooses $\omega_j = s$, but thinks that is so with probability $p^e_{j,s|i}$. In the language of Transparency International, it is i's perception that individual j tax evades. For later use, we define $\mathbf{J}$ as the array of interaction coefficients, an $I \times I \times L \times L$ array with element $J_{i,j,\ell,s}$, and $\mathbf{I}$ as the $I \times I$ identity matrix.

This expected utility function allows for an explicit characterization of the equilibrium choice probabilities once the probability distributions for the random utility terms is specified. We assume that the $\epsilon_{i,\ell}$'s are distributed according to the multinomial logit model with mean zero and dispersion parameter ς. The variance is given by $\frac{\pi^2}{\varsigma^2}$. Individual i's choice probabilities are given by:

$$\text{Prob}(\omega_i = \ell) = p_{i,\ell} = \frac{\exp\left[\varsigma\left(h_{i,\ell} + \sum_{s=0}^{L-1}\sum_{j\neq i} J_{i,j,\ell,s} p^e_{j,s|i}\right)\right]}{\sum_{\ell'=0}^{L-1}\exp\left[\varsigma\left(h_{i,\ell'} + \sum_{s=0}^{L-1}\sum_{j\neq i} J_{i,j,\ell',s} p^e_{j,s|i}\right)\right]}, \ell \in S, i = 1, \ldots, I. \quad (2)$$

Higher ς implies lower variance. The case of $\varsigma = 0$ implies purely random choice, where all outcomes are equally likely because the private random utility density is so diffused that the maximum of the random utility shocks will control the choice. In contrast $\varsigma = \infty$ means that choices are deterministic in the sense that the private random utility terms are all equal to 0 with probability 1.

Self-consistency of beliefs requires that beliefs are validated at equilibrium, that is, perception of corruption is confirmed. For this model we require that

$$\text{Prob}(\omega_i = \ell) = p_{i,\ell} = \frac{\exp\left[\varsigma\left(h_{i,\ell} + \sum_{s=0}^{L-1}\sum_{j\neq i} J_{i,j,\ell,s} p_{j,s}\right)\right]}{\sum_{\ell'=0}^{L-1}\exp\left[\varsigma\left(h_{i,\ell'} + \sum_{s=0}^{L-1}\sum_{j\neq i} J_{i,j,\ell',s} p_{j,s}\right)\right]}, \quad \ell \in S, \ i = 1, \ldots, I. \quad (3)$$

It is straightforward to verify that under the Brouwer fixed point theorem, at least one such fixed point exists for each of individual i's choice probabilities. Equation (3) defines $L-1$ independent equations for each individual's *equilibrium* beliefs, and an additional one follows by the need for probabilities to sum up to one. Thus, the matrix $p_{i,\ell}, \ i = 1, \ldots, I, \ell = 0, L-1$, of all individuals' *equilibrium* beliefs is determined, although multiple equilibria are possible. We can think of different individuals as representative of different groups of the population. For example, the different tax evading behaviors of the groups identified in the study by Artavanis, Morse and Tsoutsoura (2012) can be analyzed fully.

A simple though still interesting case, that is commonly explored in the social interactions literature, is to simplify the interaction structure by restricting social utility so that each individual only cares about the fraction of the entire population making the same choice he does. This renders the agent indifferent as to who makes the choices within the economy — individuals are anonymous. What particular choices others make is irrelevant to an individual who believes that shameless tax evasion is practiced by all. Under this simplification, the object of interest is not the matrix with elements $p_{i,\ell}$, the individual choice probabilities, but rather the aggregate choice probabilities, $p_\ell = I^{-1}\sum_i p_{i,\ell}$. This leads to:

$$p_\ell^e = p_\ell = \frac{1}{I}\sum_i \frac{\exp\left[\varsigma\left(h_{i,\ell} + J_{i,\ell}p_\ell\right)\right]}{\sum_{s=0}^{L-1}\exp\left[\varsigma\left(h_{i,s} + J_{i,s}p_s\right)\right]}, \quad \ell \in S, \tag{4}$$

where $J_{i,\ell}$ is the social utility weight i assigns to the share among the population of others making the choice ℓ.

2.1 Binary Choice

Simplifying further, let us assume that the choice set is binary, say `tax evade` or `not tax evade`. In that case, we can without loss of generality define the choice set as $\{1, -1\}$. The convenience of this definition will be shown shortly. We proceed under the assumption that social utility, $\sum_{j\neq i} J_{i,j,\ell,s}p_{j,s|i}^e$ in (1), is a function of the *expected average choice* of others, i.e. $\sum_{j\neq i} J_{i,j}m_{i,j}^e$,[4] where

$$m_{i,j}^e = 1 \times p_{j,1|i}^e + (-1) \times p_{j,-1|i}^e, \tag{5}$$

and make the additional assumption of *self-consistency of beliefs*, that is, individual i's perception of individual j's tax evading is validated at equilibrium:

$$m_j = m_{i,j}^e.$$

Let $\mathbf{m}$ denote the $I-$vector with the m_j's as elements. In this case, it is convenient to

---

[4] It is easy to see this as a specification of $\mathbf{J}$ in (1) above. That is, by going back to the original notation of equation (1) and setting $J_{i,j,\ell,1} = J_{ij}$, $J_{i,j,\ell,-1} = -J_{ij}$, then

$$\sum_{j\neq i} J_{i,j,\ell,s}p_{j,s|i} = \sum_j J_{ij}m_j.$$

use the hyperbolic tangent function,

$$\tanh(x) :\equiv \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}, \ -\infty < x < \infty, \ \tanh(x) \in (-1, 1).$$

We rewrite (3) as:

$$m_i = \tanh\left[\varsigma h_i + \varsigma \mathbf{J}_i \mathbf{m}\right], \ i = 1, \ldots, I, \tag{6}$$

where $h_i = h_{i,1} - h_{i,-1}$, and $\mathbf{J}_i$ denotes the ith row of the array, now a matrix, of interaction coefficients J_{ij}. Brouwer's fixed point theorem guarantees that the system of social interactions with an interactions matrix $\mathbf{J}$ admits an equilibrium that satisfies (6).

To understand the properties of the binary choice model, we consider the case where all heterogeneity across agents is due to random utility, i.e. we assume that $h_{i,\ell}$, and $J_{i,\ell}$, are constant across agents. This implies $\mathcal{E}_i\{\omega_j\} = m$, for all individuals, so that the Nash equilibria associated with (6) simplify to

$$m = \tanh(\varsigma h + \varsigma J m). \tag{7}$$

The properties of this special case are straightforward to describe. Referring to Fig. 3, the model of this section suggests three cases, in general. First, if $\varsigma J > 1$, and $h = 0$, then the function $\tanh(\varsigma h + \varsigma J m)$ is centered at $m = 0$, its slope at that point exceeds one, and equation (7) has three roots: a positive one ("upper"), (m^*_+), zero ("middle"), and a negative one ("lower"), (m^*_-), where $m^*_+ = |m^*_-|$. In this case, private deterministic utility does not favor either choice. A situation where individuals believe either choice is chosen with equal probability is confirmed purely because of the statistical dispersion of the unobservable component of utility. However, the condition $\varsigma J > 1$, equivalently written as $J > \frac{1}{\varsigma}$, can be interpreted as saying that the social interactions coefficient is strong enough relative to the dispersion of the unobservable component of utility to induce an outcome where the belief that in the average many individuals tax evade, $m^*_+ \geq 0$, (alternatively, few individuals tax evade, $m^*_+ < 0$) becomes self-fulfilling. The symmetry, $m^*_+ = |m^*_-|$, is of course an outcome of the analytics of the special case of equation (7) $h = 0$.

Second, if $h \neq 0$, and $J > 0$, then the inherent attractiveness, if $h > 0$, (or lack thereof, if $h < 0$) of either choice as expressed by the private utility component, if individuals are conformist (positive social interactions coefficient), then there exists a threshold H^*, which depends on ς and J, such that if $\varsigma h < H^*$, equation (7) has a unique root, which agrees with h in sign. This is depicted by curves A and C, Fig. 3. In other words, given a private utility difference h, say in favor of tax evading (because it leaves everyone with higher after tax income), given fundamentals, and a sufficiently large dispersion of the random utility component (that is, there are many individuals with large enough unobserved propensity to disobey the tax laws $h < H^*\frac{1}{\varsigma}$,), the random component dominates choice in the direction indicated by the sign of inherent attractiveness. These are depicted by points $m = \tilde{m}$ and $m = m^{**}$. If, on the other hand, $h > H^*\frac{1}{\varsigma}$, then equation (7) has three roots: one with the same sign as h, and the others of the opposite sign. See Fig. 3, Curve B. That is, given a private utility difference, if the dispersion of the random utility component is small, $h > H^*\frac{1}{\varsigma}$, then the social component dominates choice and is capable of producing multiplicity in conformist behavior.

Third, if $h \neq 0$ and anti-conformist behavior, $J < 0$, then there is a unique equilibrium that agrees with h in sign. That is, for any given level of deterministic utility, beliefs that others engage in tax evasion will induce behavior in the opposite direction. So, the stronger is the inherent attractiveness of tax evasion, the lower the probability that others would engage in it. Very high such attractiveness will elicit self-confirming behavior in spite of non-conformism. Uniqueness follows because as the expectation of engaging in tax evasion increases, the likelihood that a particular individual engages in it decreases. For the purpose of simplicity, this possibility is not depicted on Fig. 3, but it would be represented by a sigmoid but downwards sloping curve.

The intuition of relationship between the number of equilibria and the parameters J, h, and ς can be summarized as follows. Holding ς constant, it is not surprising that multiple equilibria emerge when the strength of the externality (here taking the form of conformity effects), measured by J, is large relative to the strength of inherent attractiveness as privately evaluated by the individual, measured by h. The role of ς is more subtle. The parameter ς measures the degree of heterogeneity in payoffs across individuals in the population. Higher ς means smaller heterogeneity. The degree of heterogeneity, in turn, determines how private and social incentives interact to produce equilibria. When ς is small, which means that dispersion of unobservable propensity is large, then relatively large fractions of the population will experience draws such that either $\epsilon_{i,\ell} - \epsilon_{i,\ell'}$ or $\epsilon_{i,\ell'} - \epsilon_{i,\ell}$ is large. This means in turn that a relatively high fraction will have their decisions overwhelmingly influenced by their idiosyncratic payoffs in the sense that the realization of the idiosyncratic part of the payoffs is large enough that it dominates the common private and social incentives. By symmetry of the density for $\epsilon_{i,\ell} - \epsilon_{i,\ell'}$, equal percentages of the population, in expectation, will make choices 1 and -1 because their payoffs are dominated by the idiosyncratic terms. But this means that a relatively small percentage of the population remains that can engage in self-consistent conformism because of social utility effects. Put differently, when enough agents make choices driven by symmetrically distributed payoff differences, the magnitude of the social utility terms is reduced, since it restricts the m term in Jm.

3 Uses of the Model

The model we study makes several policy-related suggestions. One is to assess whether multiplicity in the observed outcomes of tax evasion or corrupt behavior is more likely in certain settings and try to relate that to the structure of the model. For example, in its most general setting, as expressed by (2) and (3), the model may be used to study tax evasion behavior by different groups of the population. The groups may be defined as different income groups or as different occupations, where it is presumed that the members of the groups are in social contact with one another. Estimating the model with data from different cities or regions can be used to guide allocation of resources for enforcement. The simple binary choice model may be used to study tax evasion behavior across countries. The empirical tax evasion literature may be used to guide selection of regressors, either with individual or aggregate data.

The presence of multiple equilibria also poses interesting questions about enforcement. That is, can enforcement be designed so that individuals are induced to move away from low-compliance to high-compliance equilibria. Note that the macroeconomic effects of such a change in compliance can be orders of magnitude bigger than changes in marginal tax rates or other features of tax policy. The good news here is that the better of the two good equilibria, say the one associated with high compliance, is stable; the bad news is that the bad equilibrium, too, is stable. Thus, the policy must be substantial enough to move the economy from the bad to the good equilibrium, but not to drastic so as to blunt individual incentives. The model can deliver this basic intuition as follows.

Let us define as $h_{i,+1}$ ($h_{i,-1}$), individual's i private deterministic utility when he complies (does not comply) with tax laws. To simplify the problem (at the risk of abusing the model), we assume that the h_i's are given by an expected subutility under the respective regime. Following the classical formulation of Allingham and Sandmo (1972), let Y_i be income, t_i the corresponding tax rate, and β_i the share of income that i declares to the tax authorities, which we will refer to as the level of compliance. The individual is not audited with probability $1 - p_i$, in which case his after-tax income is $(1 - t_i\beta_i)Y_i$; he is audited with probability p_i, in which case, the actual income is determined and he pays the tax on the declared income, plus the tax along with a penalty, θ, on the evaded tax, so that his after tax income is: $Y_i - t_i\beta_iY_i - (1+\theta)t_i(1-\beta_i)Y_i$, where $\theta > t_i$. To fix ideas, let's limit consideration to either full compliance, $\beta_i = 1$, or total non-compliance, $\beta_i = 0$, and assume that the individual's subutility associated with the tax system is the log of after-tax income. Thus compliance versus non-compliance depends on the comparison of the following quantities, treated for simplicity as deterministic,utility from compliance, $h_{i,+1} = \ell n[(1-t_i)Y_i]$, and utility from non-compliance, $h_{i,-1} = (1-p_i)\ell nY_i + p_i\ell n[(1-(1+\theta)t_i)Y_i]$. Therefore, individual i chooses compliance if:

$$\left[\frac{\ell n(1-t_i)}{\ell n(1-(1+\theta)t_i)}\right] \leq p_i.$$

So, the higher p_i, the audit probability, or the higher θ, the penalty, the higher the probability of compliance.

In general, the audit probability depends on income and perhaps its composition as well, which introduces another layer of richness in the above comparison. But most significantly, individuals might not know their actual audit probability, and may infer through their social connections, perhaps on the basis of experiences of others, such as friends or professional acquaintances. This introduces a social element into individuals' decisions, which in the terminology of the social interactions literature would be a contextual effect.

More consequential in terms of outcomes, however, is if individuals are sensitive to the compliance behavior of their social contacts. Suppose individuals are conformist and value their expectation of the compliance behavior of others. This brings us back to the model of section 2.1 above. If individual i complies, he enjoys income $(1 - t_i)Y_i$, regardless of whether or not he is audited. If he does not comply, he may be audited with probability p_i, and enjoys income $(1 - (1 + \theta)t_i)Y_i$, or not audited with probability $1 - p_i$, in which case he enjoys income Y_i. So, conditional on non-compliance, the expected income is equal to $p_i[(1 - (1 + \theta)t_i)]Y_i + (1 - p_i)Y_i$. Defining h_i as expected difference in incomes under compliance relative to non-compliance, we have:

$$h_i = h_{i,+1} - h_{i,-1} = t_i Y_i[p_i\theta - (1 - p_i)].$$

Rewriting (7) we have:

$$m = \tanh(\varsigma t_i Y_i[p_i\theta - (1 - p_i)] + \varsigma J m). \tag{8}$$

Conformism is expressed by the presence of a positive effect of m, expected compliance, in the utility comparison. Recall that in this case, we may have, in general, a unique equilibrium or three equilibria. See Figure 3. If h_i is small relative to the threshold $\frac{H^*}{\varsigma}$, and expresses that non-compliance is inherently more attractive, then there will be a unique equilibrium agreeing in sign with h_i, m_-. That is, the inherent attractiveness of non-compliance prevails. This is mapped by curve A, Figure 3. In such a case, improvement in enforcement, either by means of more likely audits, that is higher p_i, and/or larger penalties, higher θ, can increase the inherent attractiveness of compliance. However, they need to be drastic enough to overcome the threshold, so that h_i becomes greater than $\frac{H^*}{\varsigma}$. In that case, the map of $\tanh(\varsigma h_i + \varsigma J m)$ shifts up to curve B, Figure 3. If the increase in the attractiveness of compliance is large enough, then the map becomes like curve C, Figure 3, and the economy shifts to m^{**}. This line of reasoning illustrates that timid enforcement can lessen non-compliance, like moving the curve A insufficiently higher to allow for three equilibria, and still affecting equilibrium non-compliance However, it is sufficiently vigorous enforcement that allows the economy to avail itself of a unique good equilibrium with compliance, m^{**}.

Finally, we note that in the context of the previous example (see (8 above), if $p_i > (<)(1+\theta)^{-1}$, then individuals with higher (lower) incomes and/or those facing higher (lower) tax rates are more likely to comply.

4 Empirical Application with Eurobarometer Data

The main part of the paper discusses and the Box presents in more detail results from an empirical study, based the Eurobarometer survey data. The Eurobarometer has conducted starting in 2005 biennial surveys (for years 2005, 2007, 2009, 2011, 2013) of the attitudes of the publics of the EU countries by means of a large number of questions. These surveys are an official activity of the European Commission; see Eurobarometer Data, Various Years. The latest available micro data are for 2013 and the samples are roughly 1000 observations from each EU country, amounting to a maximum total of 26856. We interpret as perception of corruption the following question (2011, QC4): "In (our country), do you think that the giving and the taking of bribes, and the abuse of positions of power for personal gain, are widespread among any of the following?" The categories listed range from among people working in the police services, customs, the judiciary, politicians at various levels, official awarding tenders, or permits, people working in the public education or the public health sectors to inspectors in health, construction, food quality, sanitary control and licensing. We recode the micro data as a the sum of affirmative responses under these categories into a categorical variable ranging from 0 to 13. We interpret as experience of corruption the following question (2011, QC5): "Over the last 12 months, has anyone (in our country) asked you, or expected you, to pay a bribe for his or her services?" We recoded the micro data as a categorical variable, that is equal to 0, if the answer is "no, nobody did," or equal to 1, if the answer is at least one in the above categories.

After much experimentation with different empirical models, we performed an ordered logit regression with the experience of corruption as a dependent variable, and perception of corruption as an explanatory variable, along with a country dummy for each of the EU countries, while allowing for 13 discrete thresholds, to be estimated, plus a large number of explanatory variables based on individual demographics. The thresholds allow naturally for the nonlinearity associated with the increasing severity of corruption experienced by respondents. Moreover, the ordered logit model, which depicts the likelihood for different outcomes through the cumulative of the logistic function as a function of respondents' perception, matches the intuition of the model introduced in section 2.1 for binary choice. It differs only in allowing for more categories. The main part of the paper and Ioannides and Murthy (2014) provide more details.