TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

Operations Research and Control Systems

A Dynamic Analysis of Schelling's Binary Corruption Model: A Competitive Equilibrium Approach

J.P. Caulkins, G. Feichtinger, D. Grass, R.F. Hartl, P.M. Kort, A.J. Novak, A. Seidl, F. Wirl

Research Report 2011-10

October, 2011

Operations Research and Control Systems
Institute of Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Argentinierstraße 8/E105-4,
1040 Vienna, Austria
E-mail: orcos@eos.tuwien.ac.at

A Dynamic Analysis of Schelling's Binary Corruption Model: A Competitive Equilibrium Approach

Jonathan P. Caulkins[a], Gustav Feichtinger[b], Dieter Grass[b], Richard F. Hartl[c], Peter M. Kort[d,e], Andreas J. Novak[c], Andrea Seidl[b,*], and Franz Wirl[c]

[a]Carnegie Mellon University, H. John Heinz III School of Public Policy & Management, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA,*e-mail:* caulkins@andrew.cmu.edu
[b]Department for Operations Research and Control Systems, Institute for Mathematical Methods in Economics, Vienna University of Technology, Argentinierstr. 8, 1040 Vienna, Austria, *e-mail:* gustav@eos.tuwien.ac.at (Gustav Feichtinger), dieter@eos.tuwien.ac.at (Dieter Grass), aseidl@eos.tuwien.ac.at (Andrea Seidl)
[c]Department of Business Administration, University of Vienna, Bruennerstr. 72, 1210 Vienna, Austria, *e-mail:* richard.hartl@univie.ac.at (Richard F. Hartl), andreas.novak@univie.ac.at (Andreas J. Novak), franz.wirl@univie.ac.at (Franz Wirl)
[d]Department of Econometrics and Operations Research & CentER, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, *e-mail:* kort@uvt.nl
[e]Department of Economics, University of Antwerp, Prinsstraat 13, 2000 Antwerp, Belgium
[*]Corresponding author. *e-mail:* aseidl@eos.tuwien.ac.at ,*tel./fax:* +43 1 58801 105471

October 7, 2011

Abstract

Schelling (1978) suggested a simple binary choice model to explain the variation of corruption levels across societies. His basic idea was that expected profitability of engaging in corruption depends on its prevalence. The key result of the so-called Schelling diagram is the existence of multiple equilibria and a tipping point. The present paper puts Schelling's essentially static approach into an intertemporal setting. We show how the existence of an unstable interior steady state leads to thresholds such that history alone or in addition expectations (or coordination) are necessary to determine the longrun outcome. In contrast to the related literature, which classifies these two cases according to whether the unstable equilibrium is a node or focus, the actual differentiation is more subtle because even a node can lead to an overlap of solution paths such that the initial conditions alone are insufficient to uniquely determine the competitive equilibrium. Another insight is that a (transiently) cycling competitive equilibrium can dominate the direct and monotonic route to a steady state, even if the direct route is feasible.

JEL: D73, C61, C62

Keywords: corruption, Schelling diagram, intertemporal competitive equilibria, thresholds, history versus expectations.

1 Introduction

This paper embeds Schelling's static description of corruption within a dynamic framework along the lines suggested in Krugman (1991) for labor markets. In this way we correct the crucial and important distinction between history and/or expectation dependent outcomes and find an additional and interesting point concerning the potential Pareto ranking among the intertemporal equilibria (in short, the short cut to the steady state may be dominated by taking first a few rounds).

Corruption is an old phenomenon that has a substantial impact on economic growth (mostly negative) and affects all societies, albeit to different degrees. For example, the organisation *Transparency International* compares corruption internationally and documents a substantial variation across states. Several papers try to explain why similar socio-economic structures can give rise to such different levels of corruption (see, e.g., Andvig and Moene, 1990). Given the long history it is no surprise that corruption is analyzed already in the ancients (e.g. in Plato) and in the political science literature. Friedrich (1972) documents historical examples of corruption (Roman empire, monarchical England, Prussia, Russia, France, United States). These examples as well as the recent revolutions in the Arab world reveal a link between corruption and dictatorships confirming the famous quotation from Lord Acton (taken from Popper, 1966, p. 137) that all power corrupts, and absolute power corrupts absolutely. Although democracy itself seems not a sufficient guaranty against bribery, some political scientists, e.g., Friedrich (1972), claim at least an inverse relation between corruption and popular support. However, corruption need not always harm economic development if it provides the grease to oil the bureaucracy. Some even argue that 'politics needs all these dubious practices, it cannot be managed without corruption', Friedrich (1972), and these practices may help to mitigate the abuse of state authorities. Lui (1986) goes even further and argues that corruption may improve social welfare because bureaucrats speed up the administration process to obtain more bribes.

Early economic investigations of corruption are Rose-Ackerman (1975, 1978) (with a recent follow up, Rose-Ackerman, 2010) and Schelling (1973, 1978). Shleifer and Vishny (1993) differentiate between beneficial and harmful corruption. Concerning the latter, think for instance about the disastrous effect of multiple marginalizations if different bureaucracies sell licenses as is practiced e.g. in post-communist Russia. These papers as well as the bulk of this literature is static. Therefore, one major objective here is to investigate corruption within a dynamic framework accounting for rational actors (e.g., bureaucrats) and (social) interdependence among them. Starting point is the model of Schelling (1973, 1978) that illustrates how the *same* bureaucrat within the *same* political and economic system can be corrupt or not depending on the aggregate level of corruption.

More precisely, this investigation focuses on both dynamics and social interactions. Dynamics is a crucial characteristic of corruption that evolves, spreads and turns out to be hard to eradicate. Examples of dynamic treatment of the evolution of corruption are investigations of (optimal) deterrence of corruption. Lui (1986) reports on corruption in Communist China that is characterized by episodes of pervasive corruption followed by anti-corruption campaigns. Feichtinger and Wirl (1994) attempt to endogenize such episodes of crusades against corruption alternating with little or no deterrence. Recently, Dong and Torgler (2011b) presents an overlapping generations model in order to assess also empirically the link between democracy, inequality, property rights and corruption.

The incidence of corruption varies strongly across nations, regions, societies and also time. E.g., Singapore, which occupies today always a top position among the least corrupt countries in the rankings of *Transparency International*, was considered to be very corrupt in the 1950-ies. Russia and other ex-Soviet republics moved from corrupt (a characteristic of the Socialist economies that was overlooked in the West, see Friedrich, 1972) to worse. Given these experiences, an important question is why are not more people corrupt? An important observation is that the profitability of a corrupt transaction compared to rejecting a bribe depends inter alia on the number of other people accepting bribes (compare Andvig, 1991, p.69). A number of papers recognize this importance of social pressure for understanding corruption. Social interactions are the subject in Glaeser et al. (2003), and this (static) concept is applied to crime in Glaeser et al. (1996). More recently, Dong and Torgler (2011a) find social interaction significant for corruption in a cross-province panel in China 1998-2007. However, papers that account for both, dynamics and social interdependence, are rare. E.g., the rather comprehensive survey by Aidt (2003) does not include dynamic approaches to corruption in spite of referring to multiple equilibria. Wirl (1998) and Epstein (2002) investigate dynamics and social interactions within the context of corruption via cellular automata. This approach adds spatial elements but ignores intertemporal optimization. Feichtinger et al. (1998) try to explain corruption as a cyclical phenomenon within a dynamic but only descriptive model.

The objective of this paper is to account for rational agents who optimize intertemporally and have rational expectations about the other agents' behavior. A corresponding starting point is Krugman (1991) who considers competitive agents who can choose between two activities: working either in agriculture, which yields a constant wage, or in manufacturing where the individual wage depends positively on the aggregate size of the industry (Krugman stresses the importance of these increasing returns, although this property is not necessary, see Wirl and Feichtinger, 2006). In our application, bureaucrats must choose to accept (or even ask for) or to reject a bribe. Facing costs for moving between these two activities, agents must solve an intertemporal optimzation problem where future payoffs depend on the other agents' actions. More precisely, payoffs depend on some *aggregate* level X, which is a given parameter that is not subject to *individual* influence although individual and aggregate coincide along any symmetric competitive equilibrium.

The contribution of this paper is twofold. Firstly, for the corruption literature we provide a dynamic formulation that by and large corroborates Schelling's insight produced 30 years ago within a dynamic setting. Secondly, the entire literature on rational expectation intertemporal competitive equilibria (starting with Krugman, 1991) applies the following criterion. If the interior and unstable steady state is a node, then history (i.e., the initial state(s)) provides all the information needed to determine the future evolution. If the unstable steady state is a focus, then history alone is insufficient and additional information about the all other agents' expectations is needed. However, this straightforward distinction is wrong. The reason is, and this is demonstrated within our application to corruption, that even unstable node can lead to an overlap, which requires, as in the case of a focus, additional information to determine future outcomes. As a consequence, the differentiation between the two cases, history versus history+expectations, is unfortunately more subtle and complicated. Another and related point is that starting from the interior spiral of the saddle point path can Pareto-dominate the (usual) alternative of starting at the 'envelope' and moving directly and monotonically to the steady state.

The paper is organized as follows. In section 2 we sketch the basic idea based on the so-called Schelling diagram. The purpose of our analysis is to extend Schelling's model to an intertemporal setting. The model presented and analyzed in section 3 shows how interior and boundary equilibria may occur: In particular, it is illustrated how the existence of an unstable interior steady state leads to a tipping point. In section 4 another (quadratic) utility function is considered to obtain additional insights. Finally, section 5 concludes. That is, the (political) decision makers are able to control corruption the obtained results are put in a more general social interaction context.

2 The Schelling diagram

Thomas Schelling sketched almost four decades ago how individual agents find it rational to accept or to reject bribes (compare Schelling, 1973, p. 388 and Schelling, 1978, Chap. 7). The following exposition follows Andvig (1991, p. 70-75). Denote by $X \in [0, 1]$ the level of corruption within a society; if $X = 0$, then everybody is 'clean', while $X = 1$ refers to a totally corrupt society.

Schelling assumes that the individual has only a binary choice $x \in \{0, 1\}$ to be either corrupt ($x = 1$) or not and is informed beforehand how corrupt the others are. Schelling's core idea is that the profitability of individual corruption depends on the level of corruption in the society, X. Denoting the agent's payoff by $U(x, X)$, the decision to be corrupt or not depends on $U(0, X)$ and $U(1, X)$, see Fig. 1, and the analysis can be restricted to the following according to the survey in Andvig (1991). Firstly, in a clean society the profitability of being corrupt is below of those of being honest[1]. Thus, it is reasonable to assume that

$$U(0, 0) > U(1, 0). \tag{1}$$

Conversely, overall corruption renders individual corruption profitable (e.g., there is no loss of individual reputation anymore). Thus, we have

$$U(0, 1) < U(1, 1). \tag{2}$$

For our purposes, the assumptions (1) and (2) are sufficient. As a consequence there is at least one intersection of the two utility functions in the interior of the utility interval; the precise shapes of the two utility functions [2] are not crucial. Let us assume that there is a unique intersection of $U(0, X)$ and $U(1, X)$, denoted as B, as illustrated in Fig. 1. At B, the agent is *indifferent* between a non-corrupt and a corrupt action. However, this point B is an *unstable* equilibrium in the following sense: If only one more member of the reference group is corrupt, it will pay to become corrupt. And since the higher profitability of corruption prevails right of B, the reference group will move to the high equilibrium C, i.e. the society becomes 'dirty'. If one starts slightly below B, the agents have a private incentive to be honest, and the system converges to the stable low ('clean') boundary equilibrium denoted as A.

Thus, small changes in initial conditions around the unstable indifference point B will have a large impact on the long-run behavior. This 'history-dependence' has important economic and political consequences. A strong, but short-lived anti-corruption campaign may move the society below the tipping point B. And since non-corrupt behavior is there more profitable, the whole system then will progress by its own movements to the clean stable equilibrium. The campaign may be lifted provided that no exogenous shocks drive the society again above point B.

Schelling's approach provides a microfoundation of the observed widely different levels of corruption by linking macro-level variables to individual (micro-) profitability using the *same* set of economic assumptions. Although the Schelling diagram is vague concerning the dynamics off an equilibrium, it is a useful tool to explain some basic facts of economic corruption such as multiple equilibria and tipping points.

We extend Schelling's analysis (at least) in two directions. First, we replace the binary corrupt/non-corrupt behavior with a continuous spectrum of various levels of individual consumption. Second, and more important, we generalize Schelling's static approach to rational agents with rational expectations solving a dynamic optimization model.

[1] E.g., if (almost) everybody is non-corrupt, then breaking the rules may lead to higher feelings of guilt than in a 'dirty' society. Furthermore, a black sheep in a herd of white ones can be more easily caught for a given capacity of public interventions.

[2] In particular, we omit the discussion of the shape of the marginal utilities $U_X(i, X)$ for $i = 0, 1$; see, however, Andvig (1991).

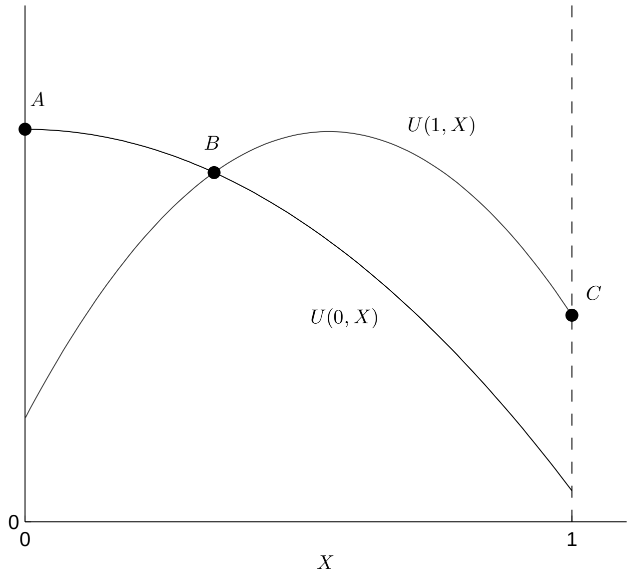

Figure 1: **The Schelling diagram.** X ... level of corruption in a reference group; $U(0, X)$ ... individual utility for being non-corrupt; $U(1, X)$ ... individual utility for a corrupt action; A ... stable boundary equilibrium: 'clean' society; B ... stable boundary equilibrium: 'dirty' society; C ... unstable interior equilibrium

3 Dynamic models of corruption

Let $x \in [0, 1]$ denote the individual decision maker's own degree of private corruption, i.e. the (continuous) share of corrupt acts of an agent and $X \in [0, 1]$ be the aggregate or average level of corruption. An agent's individual payoff is given by $U(x, X)$ since it depends on the average degree of corruption in the population, X. The dynamics results from sluggish individual behavior, more precisely,

$$\dot{x}(t) = u(t), x(0) = x_0 \in (0, 1), \tag{3}$$

where u denotes the individual's decision on which level to increase or decrease his degree of corruption. Any substantial change in individual behavior is costly, where the costs are given by $k(u, x, X)$. Costs may be due to moral (bad conscience) and due to economic reasons because expanding corruption requires an expansion of business; a reduction requires avoiding some otherwise fruitful contacts, alienation of former clients, etc. Marginal adjustment costs may increase or decrease in private or aggregate corruption; moral costs for getting more corrupt decline as aggregate corruption increases, while the cost for expanding become higher as everyone fights for the last niche to collect bribes. Therefore, the (competitive) agent's individual dynamic optimization problem is to

$$\max_{\{u(t)\in\mathbb{R}\}} \int_0^\infty e^{-rt} \left[U(x(t), X(t)) - k(u(t), x(t), X(t))\right] \mathrm{d}t, \tag{4}$$

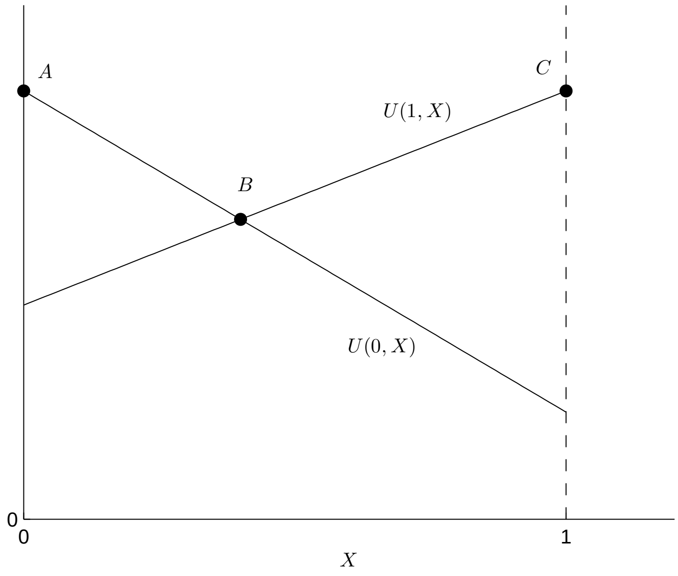

Figure 2: Linear utility function $U(0, X)$ and $U(1, X)$ in the Schelling diagram

subject to (3). Although, the agents take the aggregate X as given, i.e., X cannot be influenced by individual actions, they know (rational expectations, i.e., perfect foresight) that

$$X(t) = x(t)\,\forall t \geq 0, \tag{5}$$

for identical agents and a symmetric equilibrium. Summarizing, we have transformed Schelling's static approach to an infinite time horizon dynamic optimization problem, where x is the state, u acts as the control variable, and X is the external factor affecting the decisions of individual agents. Appendix 6.1 gives a clarification of the theoretical setting. In spite of identity (5), X is not part of the optimization, i.e., it is only a parameter within the optimality conditions.

3.1 Linear payoff U

The most simple case is that of a linear payoff: individual bribes increase linearly in aggregate corruption, while the payoff from honesty $(1 - x)$ decreases with overall corruption. Therefore,

$$U(x, X) = x(a + bX) + (1 - x)(c - dX) - \delta X^2, a > 0, b > 0, c > 0, d > 0, \delta > 0, \tag{6}$$

and Fig. 2 illustrates the above utility function. This chart ignores the last term, which is the public good externality of living in a corrupt society ($\delta > 0$), because this (additive) term is irrelevant for individual decisions.

Emphasizing simplicity following the above linear payoff, a quadratic cost function k depending only on u and some parameter γ, describing the adjustment costs, is assumed.

The resulting model is very close to Krugman's labour market model (however, Krugman, 1991 does not cite Schelling, 1978). The only difference is that the payoff of rejecting bribes is declining with respect to X while the wage in agriculture is constant. Hence, the following analysis can be kept short. The advantage of (6)'s quadratic structure is that closed form solutions of the steady states and precise characterizations of their stability properties are possible. The Hamiltonian of the associated agent's optimization problem (expressed below in current-values), equals

$$\mathcal{H} = x(a + bX) + (1 - x)(c - dX) - \frac{\gamma}{2}u^2 + \lambda u.$$

It is strictly jointly concave in the for the optimization relevant variables (u, x) such that all solution paths, that satisfy the first order conditions below, are optimal since they satisfy also a limiting transversality condition.

By using Pontryagin's maximum principle we find that

$$u^* = \frac{\lambda}{\gamma}. \tag{7}$$

The co-state equation is

$$\dot{\lambda} = r\lambda - (b + d)X + (c - a) \tag{8}$$

Differentiating (7) with respect to time, and using $X = x$ and (7), we get

$$\dot{u} = ru + \frac{(c - a) - (b + d)x}{\gamma}. \tag{9}$$

Therefore, the dynamic system (9) and (3) has one steady state at the intersection shown in Fig. 2 [3],

$$\hat{x}_B = \frac{c - a}{b + d}, \qquad \hat{u}_B = 0, \qquad \hat{\lambda}_B = 0. \tag{10}$$

This steady state is admissible if $c \geq a$ and $b + d \geq c - a$, and is unstable as the determinant of the Jacobian is $\det(J) = \frac{b+d}{\gamma} > 0$. Considering the eigenvalues of the Jacobian $\xi_{1,2} = \frac{1}{2}(r \pm \sqrt{r^2 - \frac{4(b+d)}{\gamma}})$, the point is a focus if $\frac{4(b+d)}{\gamma} > r^2$. Using the Lagrangian,[4]

$$\mathcal{L} = \mathcal{H} + \nu_1 x + \nu_2(1 - x)$$

we can find following steady states at the boundary of the admissible region

$$\begin{array}{lllll} \hat{x}_A = 0, & \hat{u}_A = 0, & \hat{\lambda}_A = 0, & \hat{\nu}_{A1} = c - a, & \hat{\nu}_{A2} = 0, \\ \hat{x}_C = 1, & \hat{u}_C = 0, & \hat{\lambda}_C = 0, & \hat{\nu}_{C1} = 0, & \hat{\nu}_{C2} = (b + d) - (c - a). \end{array}$$

Thus, assuming $\gamma = 1$, the following classification is possible, which has crucial economic implications.

[3] Subscript B refers to the corresponding point in the Schelling diagram.

[4] For a detailed presentation of the Lagrangian technique taking into consideration inequality constraints see Grass et al. (2008, Sect. 3.6).

- $c - a < 0 < b + d$: one admissible steady state with maximal corruption,
- $0 < b + d < c - a$: one admissible steady state with no corruption,
- $0 < c - a < b + d$: three admissible steady states, where the interior steady state is
  - an unstable focus iff $\frac{b+d}{\gamma} > \frac{r^2}{4}$,
  - an unstable node iff $\frac{b+d}{\gamma} \leq \frac{r^2}{4}$.

The characterization as a node or focus seems superficial given the instability of the steady state, but it has crucial economic implications. More precisely, the following distinction has been used in the literature since Krugman (1991):

1. Unstable node. Since the saddlepoint paths heading to the two boundary equilibria, $x \to 0$ or 1, do not overlap, history is sufficient to determine the future. More precisely, starting to the left of the unstable steady state implies $x \to 0$, while starting to the right implies an entirely corrupt society, $x \to 1$. Later we will show that this equivalence between node and history, i.e. starting to the left (right) of the node implies convergence to the steady state being situated on the left (right), does not hold general, although it is assumed in all related papers. Flat linear functions (b and d are small), high adjustment costs (γ large) and high discounting favor this scenario.
2. Unstable focus. Here, the two saddlepoint paths overlap, see the discussion of the example in Fig. 3 later on. This implies that history alone is insufficient to determine the future. Of course, the opposite properties, steep linear functions (b and d are large), small adjustment costs (γ small) and low discounting lead to fast adjustments that require additional information to determine the individual course of action.
3. Stable (focus or node) implies (local) indeterminacy, which is impossible in the above case of linear utilities and additive costs. However, it cannot be excluded in general. Examples that allow for indeterminacy are Karp and Thierry (2007), who consider an interaction between agriculture and polluting manufacturing, and Wirl (2011), who assumes the adjustment costs to increase with respect to X (aggregate employment in manufacturing).

If the unstable steady state is a focus, then initial conditions (i.e., history) are insufficient to determine the competitive outcome transiently and in the long run. The reason is that both saddlepoint strategies can be reached for a set of initial conditions; in the example of Fig. 3 this domain where history is insufficient is extremely large, because it contains the whole feasible region. While in the case of unilateral optimization the maximization objective provides an additional criterion to choose among different saddlepoint paths, a competitive equilibrium can only rely on the Nash property, i.e., individual behavior must be the best response given all others' actions[5]. Therefore, even if an individual knew that the way to $x = X = 0$, and thus toward no corruption, Pareto-dominates the other direction to the long run outcome $x = X = 1$, only a fool (in the context of the model) or a saint[6] would choose to go to $x = 0$ when all others 'agree' to move to the right. Krugman (1991) calls this expectation dependence, because agents must hold the same expectations within a symmetric competitive equilibrium. An alternative interpretation is that the agents must somehow coordinate their actions to choose either one of the two stable saddlepoint branches (e.g., via cheap talk). The particular feature of Fig. 3 is that, in contrast to Krugman's model, a very large overlap[7] results. More precisely, an initial overall corruption between $\bar{x}_1 = 5$ and $\bar{x}_2 = 99\%$ allows to reach both (extreme) steady states of all or no corruption. Expectations and/or coordination is crucial to choose between the two. Therefore, an improvement in coordination such as the one facilitated by modern media (facebook, twitter, mobile phones, sms) may explain the drastic switch from one equilibrium to the other. Kuran (1989) suggests that revolutions come often as a surprise since people hide their true support for the opposition until it pays off to come forward; in fact this static analysis bears some resemblance to the Schelling diagram with respect to individual payoffs from openly supporting the opposition depending (expected) aggregate support of the opposition and unstable interior steady states. In our framework, fast changes occur only if the adjustment costs are very small, which they are presumably getting as the improved communications lower uncertainties about individual exposures. This can explain why the 'liberal' dictators are usually removed fastest and those who play hardball much latter if at all. More precisely, let us apply the insights from our corruption model to revolutions given the topical events in Arab countries. For this purpose let the optimal control u describe the changing support of the government that depends on the adjustment costs γ. If a dictator leads a very strict regime and is able to bloc communications old (by prohibiting public meeting and of course demonstrations) and new (internet, twitter, facebook, mobile phones, etc.) then it is of course riskier and costlier to change towards supporting the opposition. As a consequence, such a dictator can at least delay his removal as Bashar al Assad is proving.

4 Quadratic payoff U

Given that the above simple, essentially linear payoff with respect to corruption, produces already interesting outcomes, the question arises how generalizations of payoffs and adjustment costs affect and in particular complicate the picture. It makes sense to keep the relation linear with respect to private corruption (since individual acts of corruption are negligible

[5]For this reason the payoff comparisons are irrelevant. To stress this point, note that the pure externality terms, here $-\delta X^2$, affects the payoffs but NOT the dynamics. Hence, by choosing a proper δ, I can choose the Pareto-dominant outcome at will.

[6]Well there are some. The recent self sacrifice by a Tunesian triggered the Tunesian revolution and Jan Palach's sacrifice was also crucial for the Czechs in 1968.

[7]More precisely, Krugman (1991) shows also a large overlap but only due to an error while the true overlap is small. Fukao and Benabou (1993) and Krugman (1991): both longrun outcomes are attainable for almost all admissible initial conditions, $x_0 \in (0.05, 0.99)$.

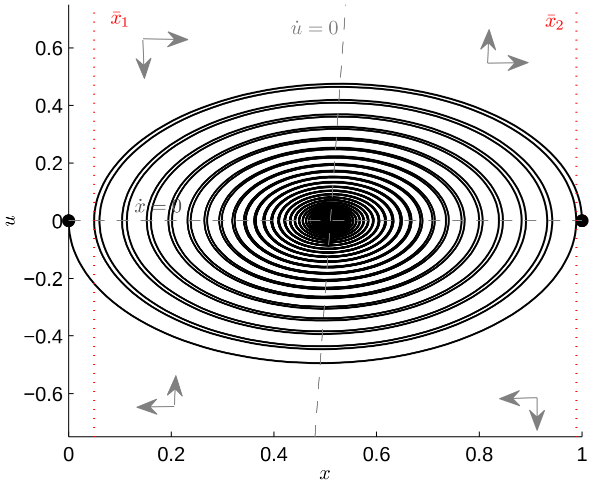

Figure 3: Phase portrait for $0 < c - a < b + d$ with $r = 0.04$, $a = b = d = 0.5$ and $c = 1.01$

in the total). On the other hand we impose that the payoff from each act of corruption depends on the aggregate in a non-monotonic way: increasing at low levels, because few offers of corruption will be made even if the bureaucrat in question is corrupt, and decreasing at high levels due to many other bureaucrats 'selling' the same or a similar license. A simple version could be

$$U(x, X) = xX(1 - \varepsilon X) - \delta X^2 + \alpha(1 - x), \tag{11}$$

where the first term describes the financial benefit from bribes, the second the public good externality due to living in a corrupt society ($\delta > 0$) and the third term refers to the utility of being clean (no worries about getting caught, good conscience etc., $\alpha > 0$). The parameter $0 \leq \varepsilon \leq 1$ determines the payoff in a 100% corrupt society from an individual corrupt act, because this payoff equals $1 - \varepsilon$. Hence, normalizing ε to 1 implies that individual payoffs from corruption are entirely eroded if everyone is willing to accept any bribe (perfect competition, thus no profits anymore). Fig. 4 shows an example ignoring again the external social costs of corruption since they are irrelevant for the individual optimization due to their public good characteristic.

The introduction of more complicated payoffs is only justified if it can provide additional insights. To start with the negative, to find a further complexity of indeterminacy - more precisely, both eigenvalues of the Jacobian to be negative (or have negative real parts) - is impossible even if moving beyond the assumed pure adjustment cost. The reason is that (at least local) indeterminacy, i.e. existence of a stable steady state, requires proper interaction between the individual control and the economy wide externality (see Wirl, 2007). Indeed, it is reasonable to assume that the costs of expanding individual corruption depends on the level of corruption within the society. Actually, the two conceivable arguments, network effects and moral, point in the same direction: (i) as corruption spreads, the corresponding networks expand and it becomes easier to join them and (ii) the moral costs are also reduced

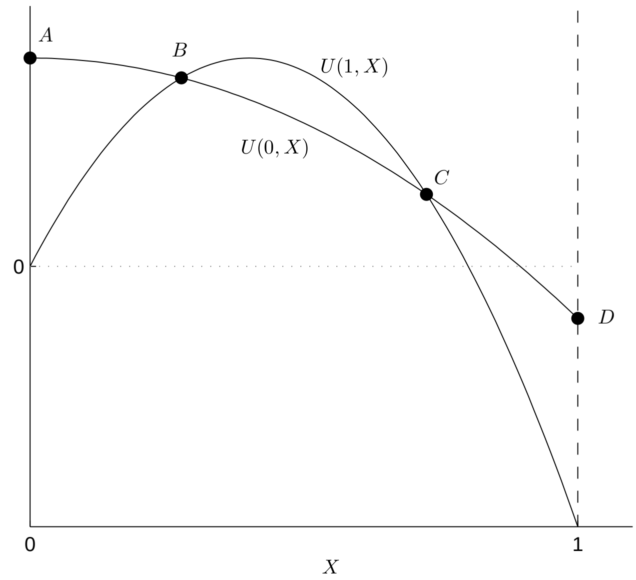

Figure 4: Utility function $U(0, X)$ and $U(1, X)$ in the Schelling diagram

if everybody is already doing it. Therefore, it is reasonable to impose that

$$\frac{\partial k}{\partial u \partial X} < 0.$$

This negative mixed derivative of the costs turns positive with respect to the total payoff so that indeterminacy can be ruled out (Proposition 5.7 Wirl, 2007) in the case of corruption of the Schelling type. Given this result, we continue using the simple adjustment cost framework. On the positive side, quadratic payoffs allow for interior and saddlepoint stable steady states. In addition, it allows us to make two formal points: (i) to question the usual classification in the literature that history is sufficient to determine the future in the case of an unstable node; (ii) non-monotonicity of competitive outcomes in spite of a single state framework, even if invoking additional criteria like Pareto-dominance in the case of multiple equilibria.

The corresponding Hamiltonian

$$\mathcal{H} = xX(1 - \varepsilon X) - \delta X^2 + \alpha(1 - x) - \gamma \frac{u^2}{2} + \lambda u,$$

implies the necessary conditions,

$$u^* = \frac{\lambda}{\gamma},$$

$$\dot{\lambda} = r\lambda - X(1 - \varepsilon X) + \alpha,$$

so that we have

$$\dot{u} = ru - \frac{1}{\gamma}\left(x(1-\varepsilon x) + \alpha\right),$$

and the following interior steady states (they correspond to the intersection of the static curves in Fig. 4 and so they are indexed accordingly),

$$\hat{x}_C = \frac{1+\sqrt{1-4\varepsilon\alpha}}{2\varepsilon}, \qquad \hat{u}_C = 0, \qquad \hat{\lambda}_C = 0, \tag{12}$$

$$\hat{x}_B = \frac{1-\sqrt{1-4\varepsilon\alpha}}{2\varepsilon}, \qquad \hat{u}_B = 0, \qquad \hat{\lambda}_B = 0. \tag{13}$$

The determinant of the Jacobian matrix evaluated at the steady states is $\det J(\hat{x}) = \frac{1}{\gamma}(1 - 2\varepsilon\hat{x})$, and the eigenvalues are $\xi_{1,2} = \frac{1}{2}(r \pm \sqrt{r^2 + \frac{8\hat{x}\varepsilon - 4}{\gamma}})$.

The first steady state is admissible if $0 \le \hat{x}_C \le 1$ and $1 - 4\varepsilon\alpha > 0$, or

$$1 - \varepsilon \le \alpha \le \frac{1}{4\varepsilon}.$$

Looking at the determinant of the Jacobian at the first steady state, which is $\det J(\hat{x}_C) = -\sqrt{\frac{1}{\gamma}(1-4\varepsilon\alpha)} < 0$, we see that this point is always a saddle point.

It can be easily shown that the second steady state is admissible if

$$0 \le \alpha \le \frac{1}{4\varepsilon}.$$

The determinant of the Jacobian at this steady state is $\det J(\hat{x}_B) = \sqrt{\frac{1}{\gamma}(1-4\varepsilon\alpha)} > 0$, the trace of the Jacobian matrix is $\mathrm{tr}J(\hat{x}_B) = r > 0$. Thus, the steady state is always unstable, and it is a node if $r^2 + \frac{1}{\gamma}(8\hat{x}_B\varepsilon - 4) \ge 0$. Inserting $\hat{x}_B$, we find that the steady state is a node if and only if $\frac{1}{4\varepsilon} \ge \alpha \ge \frac{1}{4\varepsilon} - \frac{r^4\gamma^2}{64}$.

Note that while for certain parameters we can find an admissible interior saddle point $\hat{x}_C$ above the unstable steady state $\hat{x}_B$, it is not possible to find one below. The positivity of β implies that the marginal contribution to the objective of corruption is negative for x positive and sufficiently small, i.e.

$$\mathcal{H}_x = X(1-\varepsilon X) - \beta < 0 \text{ for } X = x \text{ small enough}$$

This implies that left of the smallest interior steady state it is optimal to converge to the origin, hence instability of this steady state.

By looking at the Lagrangian function

$$\mathcal{L} = \mathcal{H} + \nu_1 x + \nu_2(1-x),$$

we find the following steady states at the boundary of the admissible region:

$$\hat{x}_A = 0, \qquad \hat{u}_A = 0, \qquad \hat{\lambda}_A = 0, \qquad \hat{\nu}_{A1} = \alpha, \qquad \hat{\nu}_{A2} = 0,$$
$$\hat{x}_D = 1, \qquad \hat{u}_D = 0, \qquad \hat{\lambda}_D = 0, \qquad \hat{\nu}_{D1} = 0, \qquad \hat{\nu}_{D2} = 1 - \varepsilon - \alpha,$$

where the one with $\hat{x}_D = 1$ is not feasible if $\varepsilon = 1$. Note, that the boundary steady state $\hat{x}_D$ is only admissible if $\alpha < 1 - \varepsilon$, thus the interior steady state with the higher level of corruption $\hat{x}_C$ and this boundary steady state cannot be admissible at the same time.

Thus, when assuming $\frac{1}{4\varepsilon} - \frac{r^4\gamma^2}{64} \geq 1 - \varepsilon$, we obtain the following scenarios:

- Region I: $\alpha > \frac{1}{4\varepsilon}$: one admissible steady state with no corruption.
- Region II: $1 - \varepsilon \leq \alpha \leq \frac{1}{4\varepsilon}$: two admissible interior steady states and one at the boundary with no corruption ($\hat{x}_A$). The interior steady state with the higher level of corruption is always a saddle point, while
  - $\frac{1}{4\varepsilon} - \frac{r^4\gamma^2}{64} \leq \alpha \leq \frac{1}{4\varepsilon}$: interior steady state with lower level of corruption is an unstable node,
  - $1 - \varepsilon \leq \alpha \leq \frac{1}{4\varepsilon} - \frac{r^4\gamma^2}{64}$: interior steady state with lower level of corruption is an unstable focus.
- Region III: $0 \leq \alpha < 1 - \varepsilon$: three admissible steady states, where one is in the interior of the admissible region (unstable focus), and two at the boundary with no and maximal corruption, respectively.

Quadratic benefits allow for (saddlepoint) stable interior steady states and more complicated dynamics, such as the one shown in Fig. 5. In this example, the no corruption outcome is globally attainable by a competitive equilibrium while the high corruption equilibrium (below 100% now) requires sufficient initial corruption; again the overlap, i.e. the interval $[\bar{x}, 1]$, is large. However, the most important reason to study this extension to nonlinear (here just quadratic) payoffs is to correct common perceptions. Firstly, it demonstrates that an unstable node allows also for an overlap and Fig. 6 shows a corresponding example: Although the unstable steady state is a node, history is insufficient to determine the longrun outcome. The reason is that it is possible to pick either of the two saddlepoint branches for any $x_0 = X_0 \in (0.4368,\ 0.4771)$. Therefore, *the common rule found and applied in the related literature - if a node then history is determinant, if a focus then history is insufficient and expectations or coordinations are needed to determine the longrun outcome - is wrong.* The unfortunate consequence of this finding is that tedious calculations are necessary to discriminate between these two different and highly policy relevant cases instead of the relatively simple check node or focus.

The second point concerns the time paths of intertemporal competitive equilibria if the unstable steady state is a spiral. Consider the example shown in Fig. 7. This allows to choose within the overlap, i.e., $x_0 = X_0 \in [0, \bar{x}]$, from a wide range of intertemporal competitive equilibria going either to the left or the right (in this example, it is the high corruption that is globally reachable, but only by a single path). Given this multiplicity within the overlap, e.g., $x_0 = X_0 = 0.13$, the agents may find the 'envelopes' as focal points (to use another famous reference to Schelling), because they converge monotonically to one of the two (stable) steady states. However, they may choose any point from inside along the spiral

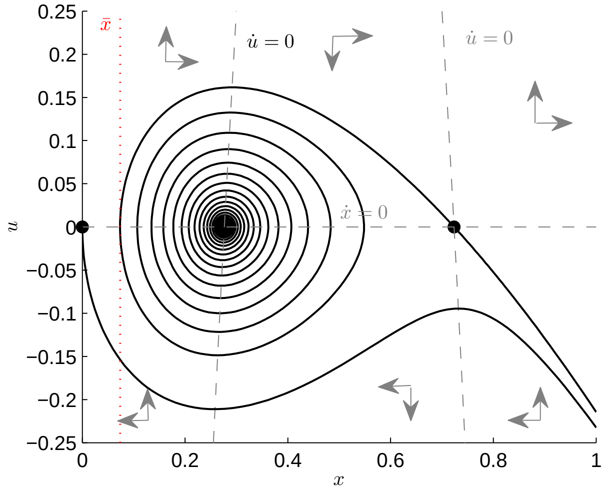

Figure 5: Phase diagram for $r = 0.04$, $\gamma = 1$, $\delta = 0.25$, $\alpha = 0.2$, $\varepsilon = 1$

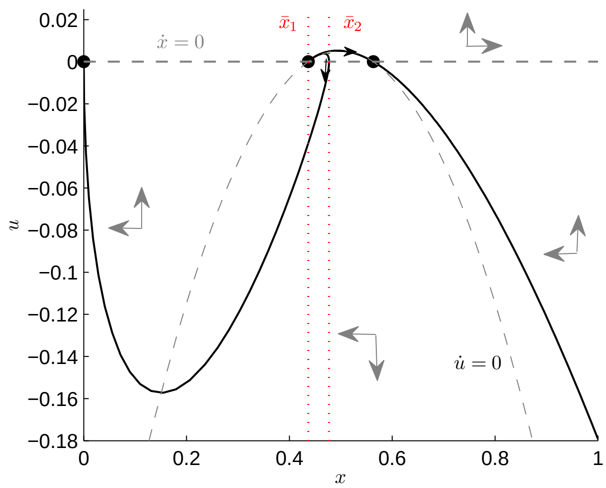

Figure 6: Phase diagram for $r = 0.75$, $\gamma = 0.25$, $\alpha = 0.246$, $\varepsilon = 1$

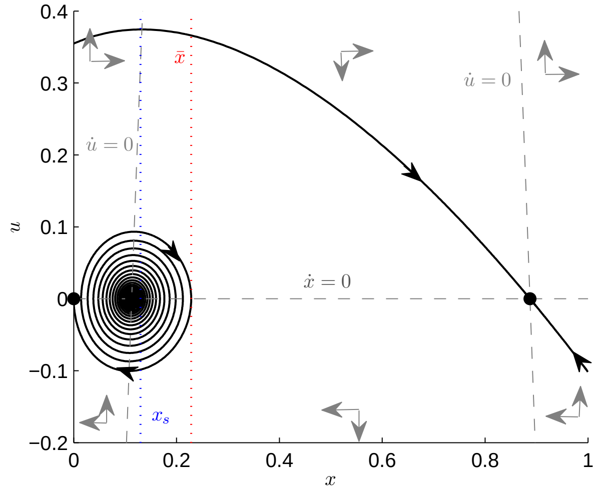

Figure 7: Phase diagram for $r = 0.04$, $\gamma = 1$, $\delta = 0.25$, $\alpha = 0.1$, $\varepsilon = 1$

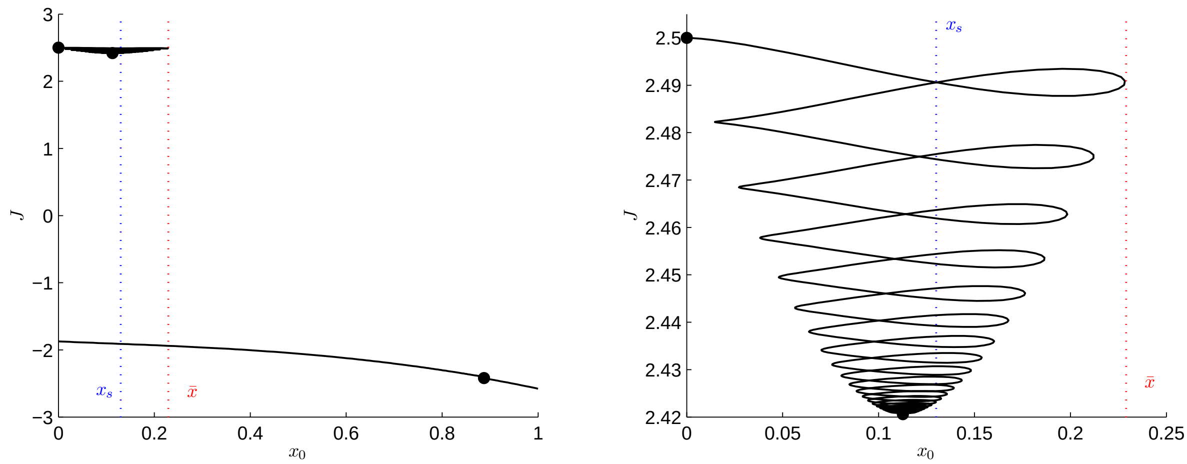

Figure 8: Objective value evaluated for initial points lying on the stable path leading to the different steady states. The right panel is a zooming of the left panel depicting the objective value along the spiral seen in Fig. 7.

which leads to indeterminacy (not to be confused with indeterminacy due two eigenvalues having negative real parts), to non-monotonic solution paths, and convergence to the left hand - no corruption - equilibrium. However, the choice of the monotonic paths is much less natural than it seems, because starting at an interior part of the spiral can Pareto-dominate the monotonic (and thus not focal) alternative. Indeed, considering the example from Fig. 8 and $x_0 = X_0 = x_s = 0.13$, then choosing an initial condition inside the spiral that takes one turn around leads to higher a payoff than the direct route to the origin. This suggests that (transient) cycling can be the most efficient competitive outcome in a one-state variable model. This makes that a dynamic competitive equilibrium model can have a qualitatively totally different outcome compared with a one-state optimal control model, because the latter always has a monotonic trajectory as optimal solution (Hartl, 1987).

5 Final remarks

This paper puts Schelling's model of corruption into a dynamic competitive equilibrium framework. As in Krugman (1991), multiple equilibria are characteristic and their determination can depend on history or history plus a coordination of agents' expectations. Yet, while Krugman distinguishes these two cases along the local stability characteristics of the unstable steady state - node versus focus - it is shown that even a node can render the need for coordination. Furthermore, in the case of a focus, we show that transient cycling (thus non-monotonic in the state) can Pareto-dominate the direct monotonic approach of the longrun equilibrium. This is striking, because for the case of an optimal control model with one state we know that optimal trajectories are always monotonic (see Hartl, 1987). This stresses at the same time how crucial social interactions can be.

Natural extensions of the model are in the direction of more states along the lines sketched above, or including a dynamic externality about how corruption evolves or harms the economy. Another alternative is to apply the individual payoff as sketched in Schelling (and similar to Kuran, 1989) in a political context to explain stability or instability of political institutions like dictatorships with a topical application to recent and surprising revolutions. A good starting point is the paper of Kuran that considers a similar payoff structure for the individual (to 'falsify' his true political preference) in a static context but adds a second and slowly moving threshold to obtain a sudden, radical switch to support the opposition.

Acknowledgments

This research was supported by the Austrian Science Fund (FWF) under Grant P21410-G16.

References

Aidt, T. S. (2003). Economic analysis of corruption: a survey. *The Economic Journal*, 113(491):F632–F652.

Andvig, J. C. (1991). The economics of corruption: A survey. *Studi economici*, 43(1):57–94.

Andvig, J. C. and Moene, K. O. (1990). How corruption may corrupt. *Journal of Economic Behavior and Organization*, 13(1):63–76.

Dong, B. and Torgler, B. (2011a). Corruption and social interaction: Evidence from china. Fondazione Eni Enrico Mattei Working Papers 560.

Dong, B. and Torgler, B. (2011b). Democracy, property rights, income equality, and corruption. Fondazione Eni Enrico Mattei Working Papers 559.

Epstein, J. M. (2002). Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences*, 99:7243–7250.

Feichtinger, G., Rinaldi, S., and Wirl, F. (1998). Corruption dynamics in democratic systems. *Complexity*, 5(3):53–64.

Feichtinger, G. and Wirl, F. (1994). On the stability and potential cyclicity of corruption within governments subject to popularity constraints. *Mathematical Social Sciences*, 28(2):113–131.

Friedrich, C. J. (1972). *The Pathology of Politics, Violence, Betrayal, Corruption, Secrecy, and Propaganda*. Harper & Row, New York.

Fukao, K. and Benabou, R. (1993). History versus expectations: A comment. *The Quarterly Journal of Economics*, 108(2):535–542.

Glaeser, E. L., Sacerdote, B. I., and Scheinkman, J. A. (1996). Crime and social interactions. *The Quarterly Journal of Economics*, 111(2):507–548.

Glaeser, E. L., Sacerdote, B. I., and Scheinkman, J. A. (2003). The social multiplier. *Journal of the European Economic Association*, 1(2–3):345–353.

Grass, D., Caulkins, J. P., Feichtinger, G., Tragler, G., and Behrens, D. A. (2008). *Optimal Control of Nonlinear Processes: With Applications in Drugs, Corruption and Terror*. Springer, Heidelberg.

Hartl, R. F. (1987). A simple proof of the monotonicity of the state trajectories in autonomous control problems. *Journal of Economic Theory*, 40(1):211–215.

Karp, L. and Thierry, P. (2007). Indeterminacy with environmental and labor dynamics. *Structural Change and Economic Dynamics*, 18(1):100–119.

Krugman, P. (1991). History versus expectations. *The Quarterly Journal of Economics*, 106(2):651–667.

Kuran, T. (1989). Sparks and prairie fires: A theory of unanticipated political revolution. *Public Choice*, 61(1):41–74.

Lui, F. T. (1986). A dynamic model of corruption deterrence. *Journal of Public Economics*, 31(2):215–236.

Popper, K. R. (1966). *The Open Society and its Enemies*, volume 1, The Spell of Plato. Routledge, London.

Rose-Ackerman, S. (1975). The economics of corruption. *Journal of Public Economics*, 4:187–203.

Rose-Ackerman, S. (1978). *Corruption: A Study in Political Economy.* Academic Press, New York.

Rose-Ackerman, S. (2010). The law and economics of bribery and extortion. *Annual Review of Law and Social Sciences*, 6:217–238.

Schelling, T. (1973). Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities. *Journal of Conflict Resolution*, 17(3):381–428.

Schelling, T. C. (1978). *Micromotives and Macrobehavior.* Norton, New York.

Shleifer, A. and Vishny, R. W. (1993). Corruption. *The Quarterly Journal of Economics*, 108(3):599–617.

Wirl, F. (1998). Socio-economic typologies of bureaucratic corruption and implications. *Journal of Evolutionary Economics*, 8(2):199–220.

Wirl, F. (2007). Social interactions within a dynamic competitive economy. *Journal of Optimization Theory and Applications*, 133(3):385–400.

Wirl, F. (2011). Conditions for indeterminacy and thresholds in neoclassical growth models. *Journal of Economics*, in press.

Wirl, F. and Feichtinger, G. (2006). History versus expectations: Increasing returns or social influence? *The Journal of Socio-Economics*, 35(5):877–888.

6 Appendix

6.1 Reformulation as a degenerate game

To clarify the theoretical setting of the problem we formally consider two decision makers. One is the (representative) individual agent considering the individual level of corruption $x(\cdot) \in [0,1]$ and the other is the aggregate population with its corruption level $X(\cdot) \in [0,1]$. Let $v(\cdot)$ denote the aggregate actions of the population, then the individual faces the following problem

$$\max_{u(\cdot)} \left\{ \int_0^\infty \overset{-rt}{\mathrm{e}} \left[U(x(t), X(t)) - k(u(t, x(t), X(t)) \right] \mathrm{d}t \right\} \tag{14a}$$

$$\text{s.t.} \quad \dot{x}(t) = u(t) \tag{14b}$$

$$x(0) = X(0). \tag{14c}$$

Considering the dynamics of the aggregate population

$$\dot{X}(t) = v(t)$$

we immediately find from assumption (5) and the initial condition (14c) that

$$v(t) = u(t), \ t \geq 0 \tag{14d}$$

has to hold. Therefore the problem of the aggregate population can also be formulated as an optimization problem yielding

$$\min_{v(\cdot)} \left\{ \int_0^\infty \overset{-rt}{\mathrm{e}} |u(t) - v(t)| \, \mathrm{d}t \right\} \tag{14e}$$

$$\text{s.t.} \quad \dot{X}(t) = v(t) \tag{14f}$$

$$X(0) = X_0. \tag{14g}$$

This reformulation shall help to clarify the mathematical foundation of the economically motivated problem. It also helps to understand the reason for the "deviating" behavior of the solutions, simply because the problem is not an optimal control problem at all; it is rather a degenerate dynamic game. Moreover it justifies the derivation of the necessary optimality conditions as it is done in the paper.