Discussion Papers in Economics

Gradualism in Aid and Reforms.

Parimal Kanti Bag and Prabal Roy Chowdhury

March 2016

Discussion Paper 16-06

Indian Statistical Institute, Delhi
Economics and Planning Unit
7, S. J. S. Sansanwal Marg, New Delhi 110016, India

Gradualism in Aid & Reforms*

Parimal Kanti Bag† Prabal Roy Chowdhury‡

March 11, 2016

Abstract

This paper analyzes a dynamic framework involving strategic interactions between an international donor and a recipient government in a bid to review the efficacy of aid conditionality in ensuring governance reforms in LDCs. We find that irrespective of whether the donor can fully commit to the aid program or not, for maximal improvement in governance the aid should be disbursed in increments with each subsequent tranche being conditional on prior reforms, demonstrating *aid gradualism*. While the attraction of future aid incentivizes the decision makers to implement reforms, these reforms in turn also make aid diversion less feasible. Further, under full commitment, the optimal aid package may involve offering scope for *interim aid diversion* to the elites, so that long-term improvements in governance can entail tolerating some aid diversion in the short run. With only partial commitment (so that *time consistency* requires the donor to reconfigure aid in each round), it is shown that (a) interim aid diversion is no longer viable, and (b) both the aid and implemented reforms exhibit *strong gradualism*, or what is known as *starting small and grow later* principle in commitment models. Also, in this case the initial aid can help screen the recipient's type, so that conditionality can possibly open the gates to selectivity.

***JEL* Classification:** H8, O2.

Key Words: Reforms, budget support, political elites, aid diversion, screening, transparency, governance, time consistency, commitment, gradualism, start small grow later (SSGL) principle.

***Acknowledgements:** We thank two anonymous referees for many thoughtful suggestions on an earlier version of this manuscript, and the co-editor for helpful guidance. Neng Qian and Sayoudh Roy provided research assistance. Support received from the Singapore Ministry of Education Academic Research Fund Tier 1 (grant no. R-122-000-151-112) is gratefully acknowledged. Any mistakes are ours.

†Department of Economics, National University of Singapore, Faculty of Arts and Social Sciences, AS2 Level 6, 1 Arts Link, Singapore 117570; E-mail: ecsbpk@nus.edu.sg

‡Economics and Planning Unit, Indian Statistical Institute, New Delhi, India; E-mail: prabalrc@isid.ac.in

1 Introduction

The doctrine of 'aid conditionality' constitutes a well-known theme in the economics of international relations. Our objective is to review the usefulness of aid conditionality in ensuring governance reforms in less developed countries. To that end we develop a theory of how conditional aid should be structured in a dynamic model of strategic interaction between an international donor and a recipient government. Inter alia, we examine if long-term improvements in governance can entail tolerating some aid diversion in the short run. Further, we shall examine if conditionality can help identify countries that are more promising candidates for receipt of aid.

The notion of governance is of course multi-faceted. It refers to not just institutional aspects like the rule of law (and its implementation), transparency and freedom of press, but also to economic ones such as opening up of markets to competition, both external and domestic. Good governance not only protects the basic rights of citizens and institutions, allowing them to function normally, it also ensures that the higher echelons of government and the elite cannot misuse their political power for either unfair advantage or extortion. It is therefore not too implausible to attribute a broad range of governance reform objectives to international donors, be it the World Bank, the IMF, or developed countries.[1]

Based on policies adopted by aid agencies as well as individual governments, a large empirical literature has developed on the subject of what contributes to the success or failure of aid initiatives. For instance, using data for 56 developing countries over 1970–1993 Burnside and Dollar (2000) have argued that "aid has a positive impact on growth in developing countries with good fiscal, monetary, and trade policies but has little effect in the presence of poor policies."[2] Despite some arguments to the contrary,[3] the connection between the quality of the recipient's domestic policies/governance and aid effectiveness appears to be reasonably well accepted now.

Following on, policy recommendations with regard to aid take two different approaches. One school advocates ***conditionality***[4] – giving aid in tranches, with the release of later tranches

[1]See further discussion of this issue in the 'sub-section' on *aspects of modelling strategy* (p. 5).

[2]The results are robust to alternative specifications such as inclusion or exclusion of middle income countries, adjusting for outliers, whether policies as assumed to be exogenous or endogenous.

[3]Easterly (2003) and Easterly et al. (2004), who use the same specifications but incorporate additional data, find that this linkage might not be significant. However, Burnside and Dollar (2004) uses a new data set focusing only on the 1990s to argue that the results noted in Burnside and Dollar (2000) do go through. Svensson (1999) also argues that aid impacts positively on growth in countries with with better democracy.

[4]Mosley (1992) (p. 129) observes that there is extensive use of aid conditionality by international financial institutions (IFIs) since the establishment of the new international order at the Bretton Woods conference in 1944. Among the IFIs, the World Bank, as well as the IMF often demanded a long list of reforms in governance in return for aid. Santiso (2001a), Santiso (2001b), and Kapur and Webb (2000) document

made conditional on specific reforms.[5] In the context of aid to sub-Saharan African countries, it has been shown that conditional aid was largely successful in implementing governance improvements, opening up to foreign competition in this case (see Morrissey (2004a)). The other school of thought focuses on *selectivity*, i.e. giving aid to countries that have good governance/policies to begin with. Our theoretical analysis and results fall within the first strand, but we also contribute to the selectivity debate. In particular, we will show that conditionality can be a tool for identifying which countries have better reform potential due to more efficient governance, generally a pre-condition for selective provision of aid.

Why do aid programs typically involve conditionality? While the literature has identified several problems intrinsic to non-conditional aid that may make conditionality useful,[6] in this paper we trace the need for conditionality to the political economy of the recipient countries, in particular the fact that the interest of the elite who control the government and that of the general public need not coincide. This is in consonance with the "institutional failures" approach, first developed in the context of the African economy, e.g. Adam and O'Connell (1999), Bates (1981), Sandbrook (1986) and Collier (1990). To quote Bates (1983) (p. 165):

> "...recent experiences in Africa and elsewhere make it clear that the preferences of governments often bear little correspondence to any idealization of the public interest. Rather, governments engage in bureaucratic accumulation and they act so as to enhance the wealth and power of those who derive their incomes from the public sector... They engage in economic redistribution, often from the poor to the rich."

Further, as Adam and O'Connell (1999) argue, such institutional failures lead to economic failures, so that aid becomes a necessity.

We analyze the issue of aid conditionality in a framework where the political tension discussed above is formalized as an ex-post moral hazard problem so that a part of the aid can be diverted by the privileged class, and may not reach the intended beneficiaries.[7] The extent of such diversion depends on the level of governance – with better governance less can

World Bank conditionality, whereas Guitian (1995) documents the use of conditionality in the IMF. See also Koberle (2003), and an OED (2004) report. Conditionality also relates to an important debate in the aid literature: transformation versus gradualism; see Temple (2010) (p. 4422).

[5] Conway (2003) remarks: "The credit available in an IMF arrangement is disbursed in tranches. The tranches are disbursed according to a set timetable on the request of the borrowing country. Later tranches can only be disbursed if the country has met the conditions defined in the Letter of Intent associated with the arrangement."

[6] Banerjee and He (2008) suggest that given the leakages in aid due to lack of monitoring, transparency and governance, provision of aid requires careful program evaluations. Banerjee (2007) discusses alternative ways of enhancing the effectiveness of aid, including making governments compete for aid through transparency.

[7] Thus we follow Boone (1996), McGuire and Olson (1996) and Adam and O'Connell (1999) in assuming that the government is fully captive. Drazen (2000) calls such recipient governments 'elitist'.

be diverted. To allow a role for conditionality, we examine a dynamic framework where the total aid amount can be disbursed in three tranches, if so desired. Thus the donor, who cares about governance, can condition the release of future tranches on governance improvement. Given that the lack of commitment could be a key reason why aid programs might fail (Kanbur (2000)), we consider two scenarios. The first scenario, where the donor can commit to the aid program from the beginning, leads to the *full conditional commitment* (or FCC) program. A second case is where the donor cannot commit fully to the aid package over the entire three periods, but instead will have to solve for a *partial conditional commitment* (or PCC) program.[8]

This paper makes three key contributions. The first is a foundational one, in that we provide a simple formalization of the idea of *aid gradualism*. Should aid be staggered at all, with future tranches being dependent on past performance? Clearly any requirement that the recipient engages in costly reforms will go unheeded if the entire aid is given up front. So disbursing the aid over two rounds, with the second round being conditional on governance improvement, certainly seems to be a better alternative. But when it comes to extending this logic to more than two rounds, the issue is less clear-cut.

One problem with extending the above logic to more than two rounds is precisely the issue of commitment mentioned earlier. If the donor cannot commit to the aid program at the beginning, i.e. solves a PCC program rather than an FCC program, then such splitting up may not be time consistent, and hence not implementable. As to other potential issues with staggering aid over more than two periods, consider an aid scheme which is staggered over three periods, say. Note that transferring some of the aid amount from later rounds to an earlier round may allow the donor to provide greater incentives, as the recipient elite can appropriate a larger fraction of any aid amount early on, with some of the planned reforms yet to materialize. This may not only allow for greater reforms early on, such reforms would relax the incentive constraints down the line, which may make eventual improvements in governance levels possible.

Our results are as follows. We find that irrespective of the commitment abilities, the optimal aid delivery mechanism involves gradualism, conditioning the continuation of future aid upon intermediate reforms. Given gradualism, a related question is the issue of starting slow versus a big bang approach. We find that in case full commitment is not possible, aid and reforms exhibit an increasing pattern, that is, start with small improvements and

[8]Commitment can of course be a double-edged sword. If a promise is believed to be a promise, Barro and Gordon (1983) have shown that monetary authorities can disinflate less than the promised rate and surprise the market to achieve output gains. Staiger and Tabellini (1987) consider the problem of time consistency that arise with commitment in the context of trade policy.

push for bigger gains in the later stages, thus conforming with the insights of conditional commitment models, e.g., Watson (1999), Watson (2002), Klimenko et al. (2008). At this point we should emphasize that we allow for linear costs, which do not bias the optimal mechanism towards gradualism or starting small. Thus neither of these results can be traced to purely technological reasons.

A second issue of interest is that of *interim aid diversion*, i.e. whether, at an intermediate stage, the optimal scheme should allow for aid over and above that required to cover the costs of reforms with a view to incentivizing the elite to deliver on governance. Interestingly we find that if the donor can fully commit to its aid program then (a) interim aid diversion is more likely when the total aid is larger, and (b) as the volume of aid goes up, interim aid diversion increases, both absolutely and as a fraction of total aid. This result is of interest since it shows that achieving long-term goals, governance in our case, may conceivably entail loss of immediate economic benefits.

Following on the last point, any program that allows for interim aid diversion may well be considered a failure if preventing aid diversion happens to be one of the criteria for success (either directly or indirectly), whereas in terms of tangible improvement in governance, such a program may be considered to be a success. This intuition is related to Drazen (2000)'s argument that one reason why aid, even conditional aid may appear to under-perform, is because the donors may not be concerned with immediate economic benefits. Similarly, Morrissey (2004a) argues that while conditionality was largely successful in ensuring governance reforms in sub-Saharan Africa, the results regarding export performance is decidedly mixed. Thus, our result may provide a partial explanation of an observation in the literature, notably by Dollar and Svensson (1998), that loan conditionality has no impact on the success of loans.[9]

We then examine a scenario where there is asymmetric information regarding the recipient's ability to pursue reform. Our third key point is that in such cases conditionality can be used to screen out countries with low reform potential. This involves schemes where the magnitude of total aid can be conditioned on an initial performance on the governance front. Thus conditionality can provide a gateway to selectivity, as identifying the governance levels of countries is key to the provision of selective aid. This is interesting given that the data suggests that selectivity is not that widely practiced by donors.[10] One possible reason could be that donors are not really sure about the future reform potential of the recipient.[11]

[9] Dollar and Svensson (1998) examine the performance of World Bank loans over the period 1980–1995. While they find that political variables are important in explaining loan success, conditionality and other World Bank inputs seem to have no impact.

[10] See Alesina and Weder (2002), and Dollar and Levin (2004).

[11] Rodrik (1989) pointed out why, when reforms can be aborted or reversed, the government may have to

■ ***Aspects of modelling strategy.*** In this sub-section we discuss certain aspects of our modelling strategy. First, and coming back to an issue briefly discussed earlier, why focus on governance? The first justification of course is that "it is there". Temple (2010) (p. 4464) mentions the case of the Millennium Challenge Account, launched by the US Government in 2004, linking aid to political freedoms, control of corruption, and respect for civil liberties and the rule of law.[12] Why are donors interested in governance though? Governance improvements can be an objective in itself, as they promote the liberal values that the donors value. Further such governance improvements have an instrumental purpose, in that reforms help in facilitating development. For example, Kaufmann et al. (1999) document empirical evidence of a strong causal relationship between governance and development.

Second, one key feature of the analysis is to recognize that aid plays a critical role in covering the costs of reforms, either direct or indirect. Direct costs are "... often due to administrative and capacity weakness", see Morrissey (1999). Whereas, as pointed out by Bakoup (2013), indirect costs may involve "... i) transfers either between the various socioeconomic categories or between the government and such socioeconomic categories; ii) government staffing expenditures; iii) tax expenditures; iv) net depreciation of non-tax revenues; v) public sector net lending; and vi) debt service variations." Also, for reforms spanning several fiscal years, the political feasibility of the needed funds transfer could be an issue, so that budget support becomes relevant.[13] Not surprisingly, Morrissey (2004a) argues that "Donors can assist with technical and financial support", while in the context of gains and losses from trade reforms, Morrissey (2004b) notes that reform "... is an inherently slow process, and aid can compensate for the costs."

Third, we consider a framework where reforms as well as the related costs can be verifiably measured.[14] Temple (2010) (section 9.1 on aid conditionality, p. 4469) discusses how aid can be conditioned on reform efforts that can be verified directly, or indirectly, and how reforms might improve the welfare of the poor. Bakoup (2013) also suggests that the measurability issues are not insurmountable. He points out that in recent decades ***budgetary aid*** (or what is often called, ***budget support***), aimed at offsetting the negative financial effects of reforms in the short to medium term, is becoming a significant component of European Union and African Development Bank (ADB) aid operations geared towards economic and structural

engage in excessive reforms to signal its type. In contrast, our formulation assumes reforms to be permanent.

[12] Temple (2010) finds that the history of aid shows that governance was one of the focus in the 1990s and 2000s (see Table 1 on p. 4424). In a similar vein, Morrissey (2004b) argues, "Conditionality ... tend to be many and wide ranging, applying not only to most areas of economic policy but also to aspects of governance and political processes."

[13] The government's inter-temporal budget may start from a negative balance during initial years of reform to a positive balance in subsequent years. See the discussion in Bakoup (2013).

[14] Drazen (2000) flags observability of outcomes as a potential issue.

reforms.[15],[16]

■ *Literature review.* Before proceeding to the formal model we relate our work to some of the other works in the literature. The role of gradualism has been examined in other economic applications, from dynamic prisoner's dilemma and contribution games (Marx and Matthews (1998), Lockwood and Thomas (2002), etc.), to international trade (Staiger (1995), Bond and Park (2002), Chisik (2003), etc.), among others. In the context of finance, Neher (1999) examines a holdup problem facing a venture capitalist where future loans are given in stages based on past performance.

The paper closest to ours is Scholl (2009), who studies a dynamic neoclassical growth model involving self-enforcing contracts between a donor and a recipient country. The donor wants a reduction of distortionary taxes to maximize household welfare, but the government may use the tax proceeds and aid money for unproductive consumption. The donor is aware of this issue, and designs the conditional aid program to stimulate the economy. Scholl (2009) however finds that conditional aid is not very cost effective as more aid must be given to less benevolent political regimes. Our paper is very different with its focus on reforms (as opposed to household welfare), the dynamic structure of aid and issues of interim aid diversion. In addition, our recipient's type may be unknown to the donor.[17]

In a couple of influential papers, Svensson (2000a) and Svensson (2000b) examine the role of donor commitment in the presence of the Samaritan's dilemma. Svensson (2000a) suggests that the delegation of aid disbursal to agencies with less aversion to poverty can, by mitigating the Samaritan's dilemma, lead to better enforcement of aid conditionality and improve the welfare of the poor. Svensson (2000b) argues that by lowering the cost of deviating from the optimal path, aid may incentivize the recipient to deviate, anticipating

[15] The European Commission defines *budget support* as "policy dialogue, financial transfers to the national treasury account of the partner country, performance assessment and capacity building based on partnership and mutual accountability." See EC (2011). (Source: Bakoup (2013).)

[16] 25% of European Union's 13 billion EURO external aid over the 2003–2009 period is budgetary aid; 31.3% of the 11.3 billion Units of Account of the operations of the ADB window and 21.8% of the 10.9 billion UC operations of the ADF windows respectively at the ADB over the period 1999-2009 (source: Bakoup (2013)).

[17] Two other papers using a repeated agency framework are Cordella et al. (2003) and Kletzer (2005), where the first authors focus on the connection between imperfectly enforceable conditional aid and policy outcomes, while the latter stresses the credibility of aid sanctions.

that bad behavior will not be punished.[18,19,20]

We start with the basic framework in the next section. In section 3 we analyze the simpler two-period problem. Then in section 4 we present the three-period problem, with the analysis and formal results followed up in sections 5–7, and conclusions in section 8. The important proofs are included in Appendix A, while proofs of several lemmas and Propositions 4 and 10′, appear in supplementary material B. A separate supplementary file contains Mathematica simulations on interim aid diversion.

2 Basic framework

The framework comprises a donor who has an aid budget $\mathcal{A}$ earmarked for a specific LDC. The purpose of this aid is to incentivize some target level of ***reform***, aimed at improving ***governance*** in the LDC. Such an objective reflects a belief on the part of the donor that the lack of development in the LDC can essentially be attributed to a lack of governance – transparency, right institutions, efficient administration, among others. Further, it reflects a belief that any reforms in governance (democracy, increased transparency, opening up markets) are relatively hard to reverse, the reason being that once a reform takes place it creates its own constituency, people who are beneficiaries of such reform and would consequently oppose any rollback. This is the other side of the observation that reforms in governance are hard to implement (Morrissey (2004b) notes that reform "... is an inherently slow process"), precisely because the existing elite may have an interest in perpetuating the status quo.

We assume that the magnitude of $\mathcal{A}$ has been already determined by internal considerations in the donor organization, and is thus exogenous to our problem.[21] Once the magnitude of aid is decided, the donor simply wants to maximize its objective by disbursing the available budget. We assume that the donor can credibly commit to cancelling the aid program in case some of the aid conditionalities are not being adhered to.[22] Further, the aid is of

[18] Further, donor officials may have an incentive to carry through with the disbursement, once it is approved (see Mosley et al. (1995), Svensson (1998)). While the issue of time consistency is beyond the scope of this paper, it would be of interest to examine if there are any possible links between gradualism and the issue of commitment.

[19] Another focus in the literature is on aid disbursement due to informational problems, e.g., Federico (2001), Murshed and Sen (1995), Azam and Laffont (2003) and Tornell and Lane (1999). These papers do not deal with the issue of conditionality, the central theme of the present paper.

[20] Mosley (1992) examines conditionality in the context of loans.

[21] The donor might be a rich country or an organization with international influence such as the World Bank, the IMF, the United Nations, or even a charity. Aid budget may be determined for a region first, then broken down for specific countries.

[22] We do not model credible commitment – an explicit modelling would require multiple recipients. See, for instance, Svensson (2003).

general nature in that it is not tied to any specific project, and it is up to the recipient country as to how it spends the grant.

The LDC is governed by an elite class who determine whether some key reforms should take place or not. One can think of the elite class as comprising of, among others, key political decision makers. The objective of the elite is to maximize *aid diversion*, i.e. the amount of aid money that it can appropriate. Thus in any period t, in case a surplus fund of S is available, the elite can appropriate a fraction $\alpha(g_t)$ of this fund, where g_t denotes the level of *governance* in the LDC at period t. The fraction of S remaining after elite appropriation, $(1 - \alpha(g_t))$, is used for consumption by the poor (hospitals, schools, food and shelter to the flood affected victims, and the likes), which we call *developmental benefits*. Ascribing such uncharitable motives to the elite may not be too far from reality as evidenced by the experience of aid agencies dealing with the LDCs.

The level of governance may be thought of as simply the inverse of some corruption index. It is an aggregative summary of how a country is perceived by the outside world. Governance may be improved by passing better laws (including, in some instances, simplification in tax rules), tightening enforcement, bringing in more transparency, improving the freedom of the press, opening up the economy to competition, both domestic as well as foreign, etc. All of these ensure, directly or indirectly, that the public duties carried out by the government will be more efficient and less prone to corruption and undue influence. This also means that any development initiative undertaken with the help of international aid money will be less likely to be diverted. In the context of *MCC* (Millennium Challenge Corporation) aid, which is an USA initiative, one finds that countries are tracked on "control of corruption" and "democratic rights".[23]

Improving the level of governance from g to g' involves a cost of $C(g'-g)$. While this cost can be interpreted as a purely technological one, arising out of the need to upgrade facilities and services (such as more record keeping of accounts, complaints, procedural scrutiny, committee works, etc.), we interpret it more broadly. In particular, this cost can be attributed to the necessary compensatory measures, including direct transfers, that the incumbent government must make to those who lose out from reforms (see Temple (2010), paragraph 2, section 9.1 on 'Policy conditionality', p. 4429). Replacing labor intensive bureaucracy by machines would incur direct costs and sometimes indirect costs due to resistance from trade unions; filling up the posts of vacant judges for better law enforcement is costly; and so on. These reforms must come at the expense of the government's alternative expenditure plans, or financed out of aid money. We will assume that the country in question is cash strapped,

[23]In the Concluding section, we provide a sample tracking document for Indonesia. Similar documents can be found for several other recipients; see https://www.mcc.gov/where-we-work.

so that reforms become easier with external help.

We shall maintain the following assumptions regarding $\alpha(g)$ and $C(.)$:

Assumption 1. **(i)** $\alpha(g)$ *is strictly decreasing, and twice differentiable in* g.
(ii) $C(.)$ *is strictly increasing, weakly convex and twice differentiable, with* $C(0) = 0$.

We shall consider a framework where the total aid, i.e. $\mathcal{A}$, is time-bound and to be spent over a given calender time, normalized to 1. The donor's problem is to decide whether to distribute this given amount in two tranches, or three, all equally spaced.

3 Two-period horizon

In this section we analyze the optimal aid design when the donor decides to disburse the total amount $\mathcal{A}$ in two tranches. Let the discount factor applicable to the two-period problem be denoted by Δ, where $\Delta = e^{-r/2}$, where r is the real rate of interest.

The LDC has an initial level of governance g_1 determined by its extant institutions, in particular rules and enforcement machinery. The donor is interested in maximizing the level of governance at the end of the planning horizon:

[**Two-period aid–governance program**] Given a fixed aid budget $\mathcal{A} > 0$, the donor provides an aid $\mathcal{A}_1 \geq 0$ to the LDC at $t = 1$ asking it to engage in reforms so as to improve its governance level to g_2. At $t = 2$, the donor provides a further aid of $\mathcal{A}_2 \geq 0$ provided the governance level at $t = 2$ is at least g_2. In case the governance level falls short of g_2, no further aid is provided. Hence

$$\mathcal{A}_1 + \mathcal{A}_2 = \mathcal{A}. \tag{1}$$

The donor's problem. The donor's problem is to design an appropriate aid program so as to maximize g_2 subject to (a) the incentive constraint that the LDC delivers g_2, and (b) the financial constraint that the available funds are sufficient to cover the costs of reforms. Further, the donor can *commit* to this conditional aid program, in that it cannot renege on its aid commitment of $\mathcal{A}_2$ in period 2 provided the governance improvement has been implemented. This is important since clearly part of the last-period aid, $\mathcal{A}_2$, will be diverted by the elites without generating any governance improvements.[24]

[24] In later sections we shall examine the implications of different degrees of commitment for the three-period problem.

The *incentive compatibility constraint* (in short, IC) facing the LDC elite at $t = 1$ is given by:

$$\alpha(g_1)[\mathcal{A}_1 - C(g_2 - g_1)] + \Delta \cdot \alpha(g_2)\mathcal{A}_2 \geq \alpha(g_1)\mathcal{A}_1,$$

where the RHS denotes the immediate payoff to the elite if they decide to renege on the contract and appropriate the maximal possible payoff immediately, whereas the LHS is the sum of its payoff at $t = 1$ in case it decides to invest in improving governance to g_2. Simplifying, we obtain:

$$\frac{\Delta \cdot \alpha(g_2)\mathcal{A}_2}{\alpha(g_1)} \geq C(g_2 - g_1). \tag{2}$$

The LDC also faces a *financial constraint* (or FC) in that the aid available at $t = 1$ must be enough to fund the suggested reforms:

$$\mathcal{A}_1 \geq C(g_2 - g_1). \tag{3}$$

Definition 1. *Given an aid budget of $\mathcal{A} > 0$, an aid–governance program $(\tilde{\mathcal{A}}_1, \tilde{g}_2; \tilde{\mathcal{A}}_2)$ will be referred to as a 2-round optimal program if it maximizes g_2 subject to (1), (2) and (3).*

We start with an observation on the ideal aid–governance program.

Lemma 1. *The optimal aid–governance program must involve gradualism in the sense that* (**i**) *the donor spreads the aid over two periods, i.e. $\tilde{\mathcal{A}}_1 > 0$, $\tilde{\mathcal{A}}_2 > 0$, and* (**ii**) *the level of governance improves over time, i.e. $\tilde{g}_2 > g_1$.*

The result is intuitive, since either giving away the entire aid in the first round, or delaying all aid until the second, will achieve zero reform.

Next a more complete characterization of the optimal aid program follows.

Proposition 1 (**Two rounds better than one**)**.** *In the optimal solution both the incentive constraint (2) and the financial constraint (3) must hold with equality, determining the maximum implementable governance, $\tilde{g}_2 > 0$, uniquely. Further,*

(i) *In period 1, aid is just enough to cover the expenses towards governance improvement, i.e., $\tilde{\mathcal{A}}_1 = C(\tilde{g}_2 - g_1) > 0$.*

(ii) *The aid in period 2, $\tilde{\mathcal{A}}_2 > 0$, is not spent on governance improvement at all. Out of the aid $\tilde{\mathcal{A}}_2$, only $(1 - \alpha(\tilde{g}_2))\tilde{\mathcal{A}}_2$ trickles down to the poor and $\alpha(\tilde{g}_2)\tilde{\mathcal{A}}_2$ is diverted by the elite.*

(iii) *$\tilde{\mathcal{A}}_2 > \tilde{\mathcal{A}}_1$, i.e., the donor must allocate a greater share of the aid for period 2 for the elites to accept the demand for reforms.*

An interesting aspect of the optimal program is that the aid trajectory is rising. The reason is simple – improving governance means the elites can divert less of the aid and so they must be compensated in the form of higher future aid so as not to divert the maximum possible amount, $\alpha(g_1)\mathcal{A}_1$. Another point to be noted is that aid diversion happens only in the last period. Offering a slack in period 1 financing is not optimal. Since improvement in governance happens only in period 1, it is better to leave the carrot till the improvement has already taken place so that aid diversion would be minimal. In contrast, in section 5 we will see that in a three-period problem leaving a slack in period 2 is a plausible aid policy.

Let us now write the incentive and financial constraints, (2) and (3), as equalities:

$$\frac{\Delta \cdot \alpha(g_2)\mathcal{A}_2}{\alpha(g_1)} = C(g_2 - g_1), \tag{4}$$

$$\mathcal{A}_1 = C(g_2 - g_1). \tag{5}$$

Denote the resulting g_2 by $\tilde{g}_2$.

From (4), (5) and the budget balancing condition, we have $\Delta \cdot \alpha(g_2)[\mathcal{A} - C(g_2 - g_1)] = C(g_2 - g_1)$. By totally differentiating this equation and using Assumption 1, we obtain:

$$\Delta\Big\{\alpha'(g_2)\big[\mathcal{A} - C(g_2 - g_1)\big]\mathrm{d}g_2 + \alpha(g_2)\big[\mathrm{d}\mathcal{A} - C'(g_2 - g_1)\mathrm{d}g_2\big]\Big\} = \alpha(g_1)C'(g_2 - g_1)\mathrm{d}g_2$$

$$\text{i.e.,} \qquad \frac{\mathrm{d}g_2}{\mathrm{d}\mathcal{A}} = \frac{\alpha(g_2)}{\Delta\Big\{-\alpha'(g_2)\big[\mathcal{A} - C(g_2 - g_1)\big] + C'(g_2 - g_1)\big[\alpha(g_1) + \alpha(g_2)\big]\Big\}} > 0.$$

Thus, we have the following result.

Proposition 2. *The implementable governance level $\tilde{g}_2$ is strictly increasing in the overall aid budget, $\mathcal{A}$.*

4 Three-period horizon

In the rest of this paper we focus on the case where aid is provided in three equally spaced tranches. We thus extend the donor's aid horizon from two periods to three, so that $t = 1, 2, 3$. Let the discount factor applicable to the three-period problem be denoted by δ, where $\delta = e^{-r/3}$, where we recall that r is the real rate of interest. We examine if the donor would optimally like to spread the aid over more than two periods. Related issues of interest are the time profile of aid and the pattern of reforms – whether the approach should be one of starting with small improvements and growing, or getting the major boost early on and then slowing down. While for $C(g' - g)$ strictly convex (i.e. increasing marginal cost) efficiency considerations may favour a staggered approach, when $C(g' - g)$ is linear this effect is no longer at play. For a significant part of our analysis we are going to assume $C(g' - g)$ to be weakly convex (thus admitting linearity), as in Assumption 1. Also, we will analyze the implications of full and partial commitment with regard to future aid.

■ *Feasibility, incentive compatibility and some definitions.* The donor's objective is to maximize g_3 subject to the various incentive and financial constraints at both $t = 1$ and $t = 2$. The donor offers a *conditional commitment* contract whereby the pattern and quantum of per period aid is conditional on the LDC's performance on the governance front. Further, the donor commits to terminating the contract should the recipient fail to deliver on the promise of governance improvement.

[**Three-period aid–governance program**] The donor provides an aid of $\mathcal{A}_1$ to the LDC at $t = 1$, asking it to implement reforms immediately so that the governance level improves to g_2 in the next period. At $t = 2$, provided the achieved governance level is at least g_2 the donor provides a further aid of $\mathcal{A}_2$ asking the LDC to improve its governance level to g_3 in the next period. If the governance level at the end of period 2 is at least g_3, then the donor provides a further aid of $\mathcal{A}_3$ in period 3. In case the governance level at the end of period t, $t = 1, 2$, is less than g_{t+1}, then the aid program is immediately stopped.

Thus an aid–governance program can be denoted by the profile

$$(\mathcal{A}_1, g_2; \mathcal{A}_2, g_3; \mathcal{A}_3).$$

We next turn to formalizing the various incentive and financial constraints. Given that the whole of the aid amount $\mathcal{A}$ is to be distributed over the three periods, we have

$$\mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 = \mathcal{A}. \tag{6}$$

The IC constraint at $t = 2$ is given by:

$$\alpha(g_2)[\mathcal{A}_2 - C(g_3 - g_2)] + \delta\alpha(g_3)\mathcal{A}_3 \geq \alpha(g_2)\mathcal{A}_2,$$

where the RHS is the payoff to the elite if it decides to renege on the contract and appropriate the maximal possible payoff immediately, whereas the LHS is the sum of its payoff at $t = 2$

in case it decides to invest in improving governance to g_3. Simplifying, we obtain

$$\frac{\delta\alpha(g_3)\mathcal{A}_3}{\alpha(g_2)} \geq C(g_3 - g_2). \tag{7}$$

The LDC also faces a financial constraint in that the aid available at $t = 2$ must be enough to fund the suggested reforms:

$$\mathcal{A}_2 \geq C(g_3 - g_2). \tag{8}$$

Next, the IC constraint at $t = 1$ is given by

$$\alpha(g_1)[\mathcal{A}_1 - C(g_2 - g_1)] + \delta\alpha(g_2)[\mathcal{A}_2 - C(g_3 - g_2)] + \delta^2\alpha(g_3)\mathcal{A}_3 \geq \alpha(g_1)\mathcal{A}_1.$$

Re-organizing, we have

$$\frac{\delta\alpha(g_2)[\mathcal{A}_2 - C(g_3 - g_2)]}{\alpha(g_1)} + \frac{\delta^2\alpha(g_3)\mathcal{A}_3}{\alpha(g_1)} \geq C(g_2 - g_1). \tag{9}$$

And the financial constraint at $t = 1$ is given by

$$\mathcal{A}_1 \geq C(g_2 - g_1). \tag{10}$$

Definition 2. *Given g_1 and the aid budget $\mathcal{A}$, a program $(\mathcal{A}_1, g_2; \mathcal{A}_2, g_3; \mathcal{A}_3)$ is* feasible *if it satisfies the budget, incentive and financial constraints (6)–(10).*

We shall focus on two classes of programs, both equally natural depending on the context, with different degrees of commitment on the part of the donor. Under *full conditional commitment* (henceforth FCC), the donor can commit to the period 2 and period 3 programs at $t = 1$ itself. This implies that the aid–governance program under FCC need not be time consistent, in the sense that the optimal program prescribed from $t = 2$ onwards need not maximize donor utility standing at $t = 2$. We shall also examine a scenario with *partial conditional commitment* (henceforth PCC), where the donor cannot commit to the program from $t = 2$ onwards. What the donor can commit to is that, subject to the governance level being at least g_2 at $t = 2$, it will offer another conditional contract for the remaining two periods, where the total aid budget is going to be $\mathcal{A} - \mathcal{A}_1$. Clearly the PCC program involves time consistency.

Whether FCC or PCC is more natural depends on the donor's specific charter, i.e., whether it is capable of making a commitment. If the donor is a durable financial institution such as the World Bank or the IMF, and moreover is not directly answerable to the public,[25] then FCC would be natural.[26] If, however, the donor is a country's incumbent government, its power to fully plan out the aid sequence may be limited especially if the aid horizon outlasts the government's tenure, in which case requiring time consistency, leading to a PCC program, might be more plausible. In fact if aid is bilateral – from one country to another – the threat of conditional aid with intermittent re-evaluation while asking the recipient country to make the necessary reforms is more credible, as exemplified by recent discussions in the UK about their aid policies.[27] After all, the donor country's government is answerable to its electorate.

We next consider these two alternative programs in more detail.

Given g_1 and the aid budget $\mathcal{A}$, an aid–governance program $(\mathcal{A}_1, g_2; \mathcal{A}_2, g_3; \mathcal{A}_3)$ is optimal under full conditional commitment if it solves the following problem:

$$\begin{aligned} & \max \; g_3 \\ \text{subject to:} \quad & \text{feasibility conditions } (6) - (10). \end{aligned}$$

Denote the optimal FCC program by $(\hat{\mathcal{A}}_1, \hat{g}_2;\ \hat{\mathcal{A}}_2, \hat{g}_3;\ \hat{\mathcal{A}}_3)$.

Definition 3. *An aid–governance program $(\mathcal{A}_1, g_2; \mathcal{A}_2, g_3; \mathcal{A}_3)$ is said to be* time consistent *if $(\mathcal{A}_2, g_3; \mathcal{A}_3)$ maximizes donor utility starting in period 2, with a starting aid $\mathcal{A} - \mathcal{A}_1$.*

We operationalize the notion of time consistency by observing that the donor's problem in period 2 is identical to the two-period problem analyzed in section 3. Thus from Proposition 2, *time consistency* requires that the financial constraint for $t = 2$, i.e. (8), will be binding along with the IC at $t = 2$, i.e., condition (7). These equations will yield, uniquely, the solutions $g_3(g_2, \mathcal{A}_1)$ and $\mathcal{A}_2(g_2, \mathcal{A}_1)$. We are now in a position to define optimality under PCC.

[25] Consider Kanbur (2000)'s account of how the World Bank caved in to external pressures for not enforcing conditionality clauses regarding African debt. While this shows the difficulty of commitment itself, an issue also underlined by Svensson (2003), it suggests that organizations like the World Bank are perhaps more likely to fit in the FCC mold.

[26] Conway (2003) says: "Pakistan has a long history as a user of IMF credit facilities. This history can be broken into two parts: the initial generation (1958-1978) of non-cancelled and nearly completely disbursed arrangements, and a subsequent generation (1980-2000) of arrangements with limited disbursement and frequent cancellation."

[27] As recently as January 2012, politicians in the UK were calling for greater checks on foreign aid: "Aid given by the UK to countries with a history of fraud and corruption should be 'conditional' on them improving their governance, MPs have said." Source: http://www.bbc.co.uk/news/uk-politics-16410677.

Given g_1 and the aid budget $\mathcal{A}$, an aid–governance program $(\mathcal{A}_1, g_2; \mathcal{A}_2, g_3; \mathcal{A}_3)$ is optimal under partial conditional commitment, if it solves the following problem:

$$\begin{aligned} & \max g_3 \\ \text{subject to :} \quad & \text{feasibility constraints } (6) - (10), \text{and time consistency}. \end{aligned}$$

Denote the optimal PCC program by $(\mathcal{A}_1^*, g_2^*;\ \mathcal{A}_2^*, g_3^*;\ \mathcal{A}_3^*)$.

Before proceeding with the formal analysis, let us fix a few concepts that form the principal themes of this paper.

Definition 4. *In the three-period problem, an aid–governance scheme* $(\mathcal{A}_1, g_2; \mathcal{A}_2, g_3; \mathcal{A}_3)$ *satisfies* gradualism *if and only if it involves*

1. *gradualism in aid, i.e.* $\mathcal{A}_i > 0, i = 1, 2, 3$*, and*

2. *gradualism in governance, i.e.* $g_3 > g_2 > g_1$*.*

An aid–governance plan is to be referred as a *quick-fix* if for any target governance improvement g, the authorities adopt a two-period approach, i.e. $\mathcal{A}_3 = 0$ and $g_2 = g$.

It may be noted that, in principle, there is nothing to rule out $g_3 > g_2 > g_1$ with an accompanying aid program of $\mathcal{A}_1 > 0, \mathcal{A}_2 = 0$ and $\mathcal{A}_3 > 0$. Such a scheme must incentivize the recipient to carry on with the improvement of governance from g_2 to g_3 even if no interim aid, $\mathcal{A}_2 > 0$, is provided. But for this to happen, the aid recipient must carry forward some $\mathcal{S} > 0$ surplus fund from period 1 efforts to period 2 activities rather than stealing it. But then the aid authority should be able to understand the incentive, and thus hold back the $\mathcal{S}$ amount from $\mathcal{A}_1$ to alter second-period aid from $\mathcal{A}_2 = 0$ to $\mathcal{A}_2 = \mathcal{S}$. Our definitions of gradualism in governance improvement and aid gradualism, therefore, go hand-in-hand: one implies the other.

Definition 5. *An aid–governance mechanism* $(\mathcal{A}_1, g_2; \mathcal{A}_2, g_3;\ \mathcal{A}_3)$ *satisfies the "starting small and grow later" principle (henceforth* SSGL*) if and only if*

1. *it involves* SSGL *in aid, i.e.* $\mathcal{A}_1 < \mathcal{A}_2 < \mathcal{A}_3$*, and*

2. *it involves* SSGL *in governance, i.e.* $g_3 - g_2 > g_2 - g_1 > 0$*.*

Aid gradualism does not necessarily imply SSGL in aid, i.e. the starting small principle in aid disbursement. Nor does gradualism in governance imply SSGL in governance improvement. SSGL is thus a stronger form of gradualism.

5 Full commitment

We begin by showing that for the same overall aid and allotted time span, an aid program achieves better governance if it involves three rounds, rather than two. Further, the improvement is done with definite progress made in each stage rather than in one single push, suggesting that a gradual approach is a better strategy for aid givers.

Proposition 3 (**Three rounds better than two**). *Suppose Assumption 1 holds. Fixing the available aid budget at $\mathcal{A} > 0$, the optimal aid program under full commitment will exhibit the following characteristics:*

(i) *A three-period program of governance improvement dominates the quick-fix solution involving only two periods. That is, the final governance level achievable is strictly higher in a three-period program.*

(ii) *In the three-period setup, governance improvement occurs* gradually *with $\hat{g}_3 > \hat{g}_2 > g_1$ and $\hat{\mathcal{A}}_i > 0, i = 1, 2, 3$.*

The idea is as follows. Note that in the optimal two-period program the aid promised for the final round, $\tilde{\mathcal{A}}_2$, does not contribute to governance improvement. With an extended three-period horizon, part of this amount $\tilde{\mathcal{A}}_2$ can be moved to period 3, conditioning the release of this amount on further reforms in the second period. Intuitively, this generates a trade-off. First, since the aid amount is split up into smaller tranches, the gain from deviating in any one period is smaller. On the other hand, the gain from complying also gets smaller, since, because of governance improvements, aid diversion gets more difficult. However, in the optimal contract the first effect necessarily dominates since the donor can always design the associated reforms in an appropriate fashion. Consequently the reform agenda can be pushed further. Interestingly, this intuition follows purely from the logic of gradualism, and does not require the cost function $C(.)$ to be convex.

Formally, one can split the aid amount earmarked for period 2 under the optimal two-period program among period 2 and 3 while maintaining the original g_2 and setting $g_3 = g_2$ (i.e. no real improvement in the final round), because there is no discounting and thus the combined aid diversion opportunity (period 2 and period 3 together) remains the same. With positive aid retained in period 2, g_2 can now be lowered slightly (to some g_2') that is made up for by requiring an equivalent g-improvement in period 3 to maintain $g_3 = g_2$. And this is possible without prompting default through maximal aid diversion from the set governance objectives in period 1 or period 2, due to a combination of weak convexity of the cost function $C(.)$, strictly decreasing $\alpha(g)$, and interim aid diversion opportunity in period

2 (see Step 3 of the proof). Finally, as g_2 is lowered all of the constraints become slack so that the final g can be increased at the end of period 3 beyond the original g_2.[28]

As a counterpart to Proposition 3, a relevant question is the implementation of a target level of governance and an appropriate time frame for it.[29] It is plausible that for any given time frame, two periods or three periods, the higher the aid budget the higher the implementable g, giving rise to an upward-sloping function $g(\mathcal{A}, \tau), \tau = 2, 3$. Then by Proposition 3, $g(\mathcal{A}, 3) > g(\mathcal{A}, 2)$ over the range of g implementable under both time horizons. Put differently, for any target $\bar{g}$ implementation costs would be strictly lower with a gradual approach compared to that with a quick-fix solution, i.e., $\mathcal{A}(\bar{g}, 3) < \mathcal{A}(\bar{g}, 2)$.

Thus from a purely financial point of view, the lower implementation costs may justify aid agencies adopting a measured wait-and-see approach to governance reforms and keeping the recipient country on a tight leash. This result should remain valid under a moderate degree of discounting.

The next result is technical, suggesting how the various constraints impact on the optimal aid program. Part of the characterization will be used below (in Proposition 5) to determine the optimal time structure of aid.

Proposition 4 (**Optimal FCC program**). *Suppose Assumption 1 holds. With the feasibility requirements (6)–(10), the donor's optimal program can be partially characterized as follows:*

- **(i)** *Both period 1 financial constraint, (10), and period 2 incentive constraint, (7), must be binding.*
- **(ii)** *At least one of the two constraints – period 2 financial constraint, (8), and period 1 incentive compatibility constraint, (9) – must bind.*
- **(iii)** *If the equilibrium governance improvement program satisfies a weaker version of* SSGL, *i.e.* $\hat{g}_3 - \hat{g}_2 \geq \hat{g}_2 - g_1$, *then the constraint (9) would bind.*

Intuitively, the period 1 financial constraint binds because, with the aid program starting in period 1, any slack in period 1 cannot be used to incentivize either current or future reforms, and is just a waste of aid money. Similarly, if the incentive constraint in the second period is slack, this means that the aid allocated to the last period, i.e. $\hat{\mathcal{A}}_3$, can be reallocated as follows. Let us reduce $\hat{\mathcal{A}}_3$ by a small amount (this has no financial implications

[28] While the argument relies on "small changes" and thus might appear to be a local analysis, the dominance through gradualism is of a general nature. It only suggests that the dominance is possible without necessarily exploiting the best possible mechanism.

[29] A related question is how the governance level is affected as the recipient becomes more patient. This we address later in Proposition 6.

as there is no reform in period 3), and re-distribute the amount so freed up between $\widehat{\mathcal{A}}_1$ and $\widehat{\mathcal{A}}_2$ without affecting the incentive constraint in period 2. This relaxes all the constraints so that the reforms agenda can be pushed further.

Given part **(ii)**, any of three combinations can characterize the optimal aid contract: both (8) and (9) binding; only (8) binding and (9) non-binding; and (8) non-binding and (9) binding. The first possibility yields all constraints binding, and the second possibility keeps default incentives in check with the final round aid diversion, $\alpha(g_3)\mathcal{A}_3$, strictly dominating earlier round opportunities. The third possibility offers an interesting policy direction, analyzed below under the heading: interim aid diversion.

Proposition 4 is a partial characterization. More structure can be put on the optimal aid scheme under the additional assumption that $C(.)$ is additively separable. Specifically, a partial ranking of $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ is now possible.

Assumption 2. *For any $g' > g$, $C(g' - g) = C(g') - C(g)$.*

Assumption 2 allows for linear cost functions so long as we retain Assumption 1 and thus $C(0) = 0$.

Proposition 5 (**Time structure of aid**). *Suppose both Assumptions 1 and 2 hold.*

(i) *In the solution of the FCC problem, period 1 incentive compatibility constraint (9) must bind.*

(ii) $\alpha(g_1)\widehat{\mathcal{A}}_1 = \delta\alpha(\hat{g}_2)\widehat{\mathcal{A}}_2 \geq \delta^2\alpha(\hat{g}_3)\widehat{\mathcal{A}}_3$ *and* $\widehat{\mathcal{A}}_1 < \widehat{\mathcal{A}}_2$.

Remark 1. Given Proposition 5**(ii)**, it is natural to ask if it is possible to rank $\widehat{\mathcal{A}}_3$ vis-a-vis $\widehat{\mathcal{A}}_2$. Unfortunately, numerical simulations show that the results can go either way, so that the time profile of aid could be non-monotonic. Consider the case where $\alpha(g) = 1 - g$, and $g_1 = 0$. Table A.2 in the Appendix shows that when $C(x)$ is linear, $\widehat{\mathcal{A}}_3 > \widehat{\mathcal{A}}_2$ for all $\mathcal{A} \in \{1.1, 1.2, \dots, 2\}$, whereas Table A.3 shows that when $C(x) = x^2$, $\widehat{\mathcal{A}}_3 > \widehat{\mathcal{A}}_2$ for $\mathcal{A} \in \{1, 1.1, \dots, 1.5\}$, and $\widehat{\mathcal{A}}_3 < \widehat{\mathcal{A}}_2$ for $\mathcal{A} \in \{1.6, 1.7, \dots, 2\}$ (see the segments coded in red).

Another related issue is whether the optimal aid program follows SSGL, i.e. $\hat{g}_3 - \hat{g}_2 > \hat{g}_2 - g_1$, or not. Again numerical simulations suggest that the answer could go either way. Consider the case where $\alpha(g) = 1 - g$, and $g_1 = 0$. Table A.2 shows that when $C(x)$ is linear, SSGL holds for $\mathcal{A} \in \{1.1, 1.2, \dots, 1.6\}$, but it does not hold for $\mathcal{A} \in \{1.8, 1.9, 2\}$, whereas Table A.3 shows that when $C(x) = x^2$, SSGL is never satisfied for any $\mathcal{A} \in \{1.1, 1.2, \dots, 2\}$. This ambiguity arises solely because of non-binding (8).

In contrast under time consistency to be studied in the next section, we will see that both aid and governance improvement will have strictly upward trajectories.

■ *Interim aid diversion under FCC.* Below we consider an aid scheme where the donor may knowingly tolerate aid diversion by recipient country political elites. This connects with some stylized facts on conditional aid. As discussed in the introduction, while aid conditionality does seem to have a positive impact on policy outcomes, the impact on economic outcomes is not that straightforward (recall the discussion on trade policy reform in sub-Saharan Africa (Morrissey (2001), Morrissey (2004b))). While the literature has argued that this gap between improvements in governance and economic outcomes can be attributed to shocks (Morrissey (2004a)), as well as administrative weaknesses (Morrissey (1999)), our analysis suggests a possible alternative reason, namely the fact that maximizing long-term governance improvements may involve neglecting immediate economic gains.

The last point above brings our attention to a warning in January 2012 in the UK by their International Development Secretary, Andrew Mitchell, who said: "We make absolutely clear to countries that transparency and good governance are vital. We are prepared to withhold funding to governments when our standards are not met, as we have done in Malawi." (Source: http://www.bbc.co.uk/news/uk-politics-16410677). To put this warning in perspective, despite the tough talk our analysis suggests that it may be wise to adopt a more pragmatic approach when dealing with corrupt governments: *provide a slack to powerful political elites along the way to improve governance in the long run*.

Formally, we shall argue that the optimal contract may involve leaving a rent for the recipient in the second period whenever the aid $\mathcal{A}$ is large enough. Moreover, such rent is increasing in $\mathcal{A}$. Next, turning to the analysis, recall that under the FCC program, the IC at $t = 2$, i.e. (7), as well as the FC at $t = 1$, i.e. (10), bind (see Proposition 4). From (6), (7), (8) and (10), we can then eliminate $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$, to obtain (11). Similarly, we can use (6), (7), (9) and (10) to obtain (12):

$$F(g_2, g_3; \mathcal{A}) \equiv \mathcal{A} - C(g_2 - g_1) - \frac{\alpha(g_2)C(g_3 - g_2)}{\delta\alpha(g_3)} - C(g_3 - g_2) \geq 0, \tag{11}$$

$$I(g_2, g_3; \mathcal{A}) \equiv \mathcal{A} - C(g_2 - g_1) - \frac{\alpha(g_2)C(g_3 - g_2)}{\delta\alpha(g_3)} - \frac{\alpha(g_1)C(g_2 - g_1)}{\delta\alpha(g_2)} \geq 0. \tag{12}$$

Thus, (11) (respectively (12)) captures the financial constraint at $t = 2$ (respectively the incentive constraint at $t = 1$), subject to both the financial constraint at $t = 1$, and the incentive constraint at $t = 2$ binding. The inequalities (11) and (12), determining the set of feasible outcomes (g_2, g_3) under the FCC program, together constitute the key to the subsequent analysis. Denote the partial and cross partial derivatives by:

$$\big(F_i(g_2, g_3; \mathcal{A}), F_{ij}(g_2, g_3; \mathcal{A});\ I_i(g_2, g_3; \mathcal{A}), I_{ij}(g_2, g_3; \mathcal{A}) \big),$$

where the subscripts i and j refer to the i-th and j-th arguments.

It is easy to see that $F(g_2, g_3; \mathcal{A})$ and $I(g_2, g_3; \mathcal{A})$ are both strictly decreasing in g_3 (see Lemma 2 below). Hence for any given g_2, there can be at most one g_3 that solves $F(g_2, g_3; \mathcal{A}) = 0$. Likewise at most one g_3 can solve $I(g_2, g_3; \mathcal{A}) = 0$.

Definition 6. *Let* $f(g_2, \mathcal{A})$ *solve* $F(g_2, f(g_2, \mathcal{A}); \mathcal{A}) = 0$, *and* $i(g_2, \mathcal{A})$ *solve* $I(g_2, i(g_2, \mathcal{A}); \mathcal{A}) = 0$.[30]

The following lemma states several properties of $F(g_2, g_3; \mathcal{A})$, $I(g_2, g_3; \mathcal{A})$, $f(g_2, \mathcal{A})$ and $i(g_2, \mathcal{A})$ that will be needed in the analysis to follow.

Lemma 2. *Consider* $F(g_2, g_3; \mathcal{A})$, $I(g_2, g_3; \mathcal{A})$, $f(g_2, \mathcal{A})$ *and* $i(g_2, \mathcal{A})$.

(i) $F(g_2, g_3; \mathcal{A})$ *and* $I(g_2, g_3; \mathcal{A})$ *are both strictly decreasing in* g_3. *Moreover,* $F_2(g_2, g_3; \mathcal{A}) < I_2(g_2, g_3; \mathcal{A}) < 0$.

(ii) $I_1(g_2, g_3; \mathcal{A}) < F_1(g_2, g_3; \mathcal{A})$.

(iii) *Whenever either* $g_3 - g_2 \geq g_2 - g_1$, *or* $C(.)$ *is linear, it follows that* $F_1(g_2, g_3; \mathcal{A}) > 0$, *and* $f(g_2, \mathcal{A})$ *is strictly decreasing in* g_2.

(iv) $i(g_1, \mathcal{A}) > f(g_1, \mathcal{A})$.

(v) $I_{12}(g_2, g_3; \mathcal{A}) > 0$.

(vi) $I_{11}(g_2, g_3; \mathcal{A}) < 0$, *whenever* $\alpha(.)$ *and* $C(.)$ *are both linear.*

Assumption 3 below specifies more general versions of properties stated in Lemma 2**(iii)** and 2**(vi)**. Our numerical simulations (see the Appendix) suggest that Assumption 3**(ii)** will be satisfied whenever $\alpha(.)$ and $C(.)$ are both linear.

Assumption 3. **(i)** $F_1(g_2, g_3; \mathcal{A}) > 0$ *and* $I_{11}(g_2, g_3; \mathcal{A}) < 0$.

(ii) *Let there exist* $\tilde{g}(\mathcal{A}) > g_1$ *such that* $i(g_2, \mathcal{A}) > g_2$ *if and only if* $g_2 < \tilde{g}(\mathcal{A})$.

$g_2^*(\mathcal{A})$ defined below will play a key role in the analysis:

$$g_2^*(\mathcal{A}) = \underset{g_1 \leq g_2 \leq \tilde{g}(\mathcal{A})}{\operatorname{argmax}} \; i(g_2, \mathcal{A}). \tag{13}$$

For ease of exposition let us assume that $g_2^*(\mathcal{A}) < \tilde{g}(\mathcal{A})$. Note that Assumption 3**(ii)** then ensures that $i(g_2^*, \mathcal{A}) > g_2^*(\mathcal{A})$.

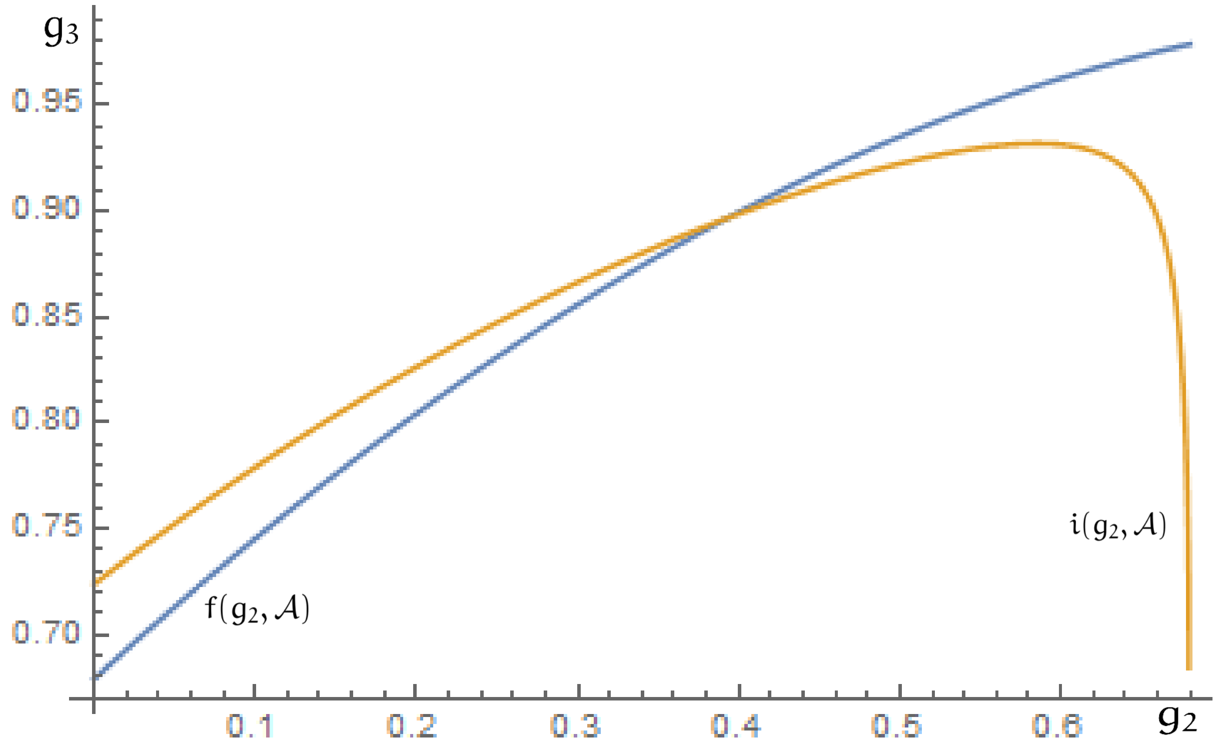

Figure 1: Sample Mathematica plot (interim aid diversion), $\alpha = 1 - g, C(x) = x^2, g_1 = 0, \mathcal{A} = 1.9$; g_3 maximized at $g_2 = 0.585275$

The following lemma is a key step in our analysis of interim aid diversion. Lemma 3**(i)** ensures that $i(g_2, \mathcal{A})$ has a unique peak at g_2^*. We then use this result to provide a characterization of interim aid diversion in part **(ii)** (see Figure 1). Lemma 3**(ii)** is, in turn, used in Proposition 6 to argue that the level of aid has a close connection with both the existence and magnitude of aid diversion.

Lemma 3. *Suppose Assumption 3 holds. Then the following are true:*

(i) $i(g_2, \mathcal{A})$ *is strictly increasing for all* $g_2 < g_2^*$, *and strictly decreasing for all* $g_2 > g_2^*$.

(ii) *There will be interim aid diversion if and only if* $i(g_2^*, \mathcal{A}) < f(g_2^*, \mathcal{A})$.

What is the reasoning behind the slack in period 2 financial constraint, i.e. (8)? Intuitively, the goal is to incentivize reforms improvement in period 1. This would lead to a higher g_2 that, due to the synergy between g_2 and g_3 as far as incentives are concerned (formally $I_{12} > 0$, recall Lemma 2**(v)**), will translate into a higher g_3. With a slack in period 2, the elite has a greater incentive to deliver on expected improvement in period 1 and avail

[30]While, in the optimal program, both (11) and (12) must hold, the solution to (11) need not satisfy the IC at $t = 1$, and the solution to (12) need not satisfy the FC at $t = 2$.

of the slack. But, why should the donor create a slack in period 2, rather than shifting the amount to period 3? This is because a transfer of Δ dollars from the final round reward $\mathcal{A}_3$ to $\mathcal{A}_2$, increases the recipient's second-period margin by $[\alpha(g_2)-\alpha(g_3)]\Delta > 0$. This, in turn, will ease the first-period financial constraint (9) because $\alpha(g_2) > \alpha(g_3)$. Given this slack created in the first-period financial constraint, the donor can now use the amount so freed up to push up g_3. This allows the donor to sustain greater reforms overall.

We next argue that there will be interim aid diversion for sufficiently large aid. Also, we establish two comparative statics results – how governance improvement is affected as the level of aid and the discount factor changes.

Proposition 6 (**Interim aid diversion**). *Suppose Assumption 3 holds.*

1. *The optimal governance in period 3, i.e. g_3^*, is strictly increasing in the amount of total aid $\mathcal{A}$, as well as the discount factor δ.*
2. *Further, suppose that with an increase in g_2, the incentive constraint in period 1 is relaxed to a significant extent, more formally $I_{12}(g_2, g_3; \mathcal{A}) > \frac{I_{11}(g_2,g_3;\mathcal{A})F_2(g_2,g_3;\mathcal{A})}{F_1(g_2,g_3;\mathcal{A})}$. Then the following are true:*

   (i) *There exists $\mathcal{A}^* \geq 0$ such that there will be interim aid diversion if and only if $\mathcal{A} > \mathcal{A}^*$.*

   (ii) *The amount of interim aid diversion increases with $\mathcal{A}$.*

We already know that interim aid diversion is a way of increasing g_2, and eventually g_3. However, why should a large volume of aid lead to interim aid diversion, and further, why should an increase in $\mathcal{A}$ lead to an increase in its magnitude? This has to do with the assumption that an increase in g_2 relaxes the incentive constraint at $t = 1$ to a *significant* extent, so that aid diversion becomes a potentially important channel for increasing g_3 as well. The cost of allowing interim aid diversion is of course that the amount left over in period 3 will be smaller, which will make it harder to incentivize an increase in g_3 in the second period. With $\mathcal{A}$ large, the marginal cost of adopting such an approach is comparatively small, explaining both the diversion and its magnitude.

Aid diversion percentage. Following Proposition 6-*2* **(ii)**, it may be natural to ask if the *percentage* of aid diversion is also increasing in $\mathcal{A}$. While an analytical answer is hard to establish, based on numerical simulation presented in Tables A.2 and A.3 in the Appendix we can state the following:

Let $g_1 = 0$ and $\alpha(g) = 1 - g$:

(i) *Consider a linear cost function, i.e.* $\mathsf{C}(\mathsf{x}) = \mathsf{x}$. *Then,* $\forall \mathcal{A} \in [1.4, 1.5, \ldots, 2]$, *there is interim aid diversion. Further, the proportional slack, i.e.* $\frac{\hat{\mathcal{A}}_2 - (\hat{g}_3 - \hat{g}_2)}{\mathcal{A}}$, *is strictly increasing in* $\mathcal{A}$ *for these values of* $\mathcal{A}$.

(ii) *Consider a convex cost function, in particular* $\mathsf{C}(\mathsf{x}) = \mathsf{x}^2$. *Then,* $\forall \mathcal{A} \in [1.0, 1.1, \ldots, 2]$, *there is interim aid diversion. Further, the proportional slack, i.e.* $\frac{\hat{\mathcal{A}}_2 - (\hat{g}_3 - \hat{g}_2)}{\mathcal{A}}$, *is strictly increasing in* $\mathcal{A}$ *for these values of* $\mathcal{A}$.

6 Partial commitment

■ *Aid structure under PCC.* Adding renegotiation in the aid contract following each round, subject to satisfactory progress in governance, alters the donor's feasible program since it weakens the donor's hand relative to the full commitment program. As a result, the implementable level of governance can only (weakly) worsen. We first analyze the optimal aid mechanism under time consistency, then compare it with the mechanism under full commitment.

The following result is a summary presentation of Lemma A.1 reported in the Appendix.

Proposition 7 (**Characterization**)**.** *Suppose Assumption 1 holds.*

(i) *Then in the optimal PCC program both pairs of incentive and financial constraints, (7)–(10), will bind.*

(ii) *There will be no interim aid diversion.*

Interestingly, and in contrast to the result under FCC, we find that the aid package under PCC does not involve interim aid diversion. This is so because time consistency requires the aid package to be freshly calculated after round 1. Since the idea behind interim aid diversion under FCC is to incentivize reforms in period 1, interim aid diversion becomes redundant in period 2, and will be dispensed with under PCC. This suggests, for example, that interim aid diversion is more likely to be used whenever the donor country has a stable government which is likely to outlast the aid horizon, or the aid department is independent and can commit to a long-term program.

Building on the above characterization and Lemma A.1 we obtain our main result under PCC:

Proposition 8 (**Strong gradualism**)**.** *Suppose Assumption 1 holds. In the three-period PCC problem, the optimal aid mechanism involves gradualism as well as the* SSGL *principle:*

(i) $g_3^* > g_2^* > g_1$ *and* $\mathcal{A}_3^* > \mathcal{A}_2^* > \mathcal{A}_1^* > 0$.

(ii) $g_3^* - g_2^* > g_2^* - g_1 > 0$.

(iii) *The final governance level achievable under a three-period setup is higher than under a two-period setup.*

Note that despite the fact the donor faces greater constraints under PCC because of the time consistency requirement, the three-period program continues to prevail over a two-period scheme. This confirms the power of aid gradualism to deliver on improvements.

Further, both aid levels and governance improvements exhibit the *SSGL* property, or what might be termed as ***strong gradualism***. The principle of "starting small and grow later" for conditional commitment has been previously noted in other applications, for instance, Watson (2002) and Klimenko et al. (2008). Its application in the context of aid conditionality should thus be of interest.

Both these results are especially noteworthy as Assumption 1 imposes no restrictions on the curvature of the governance improvement costs, $C(.)$. In particular, for linear costs and zero discounting, purely efficiency considerations suggest that one should expect equal improvements in governance in each round.

Recall that *SSGL* may not obtain under FCC, suggesting that it is time consistency that drives this strong property. *SSGL* in aid can be understood as follows. That $\mathcal{A}_3^* > \mathcal{A}_2^*$ follows due to the reason stated following Proposition 1. Absence of financial slack in period 2 under sequential rationality means the recipient must be incentivized entirely through aid diversion in the last period, so how $\alpha(g_1)\mathcal{A}_1$, $\alpha(g_2)\mathcal{A}_2$ and $\alpha(g_3)\mathcal{A}_3$ compare determines the default decision. To avoid default, aid diversion in the initial rounds must be dominated by one in the final round, and sequential rationality implies $\alpha(g_2)\mathcal{A}_2 = \alpha(g_3)\mathcal{A}_3$ (see (4) and (5)). So $\alpha(g_1)\mathcal{A}_1 \leq \alpha(g_2)\mathcal{A}_2$, thus $\mathcal{A}_2^* > \mathcal{A}_1^*$ follows. Finally, *SSGL* in governance is a direct implication of there being no interim aid diversion under sequential rationality and $\mathcal{A}_2^* > \mathcal{A}_1^*$. Under full commitment, the possibility of interim aid diversion again causes the *SSGL* property to fail.

An Example. In this example we verify that the three-period problem dominates the two-period problem under PCC. Consider $C(x) = x$ and $\alpha(g) = 1 - g$. Moreover, let $g_1 = 0$. We first solve for the two-period case. From (4) and (5), and putting $g_2 = \tilde{g}$, we have

$$\mathcal{L}(\tilde{g}) := (1 - \tilde{g})(\mathcal{A} - \tilde{g}) = \tilde{g}. \tag{14}$$

Note that the LHS of (14) is convex, initially decreasing and then increasing in $\tilde{g}$, with $\mathcal{L}(0) > 0$. Thus the solution involves (the negative root is chosen as otherwise $\tilde{g} > 1$):

$$\tilde{g} = 2 + \mathcal{A} - \sqrt{(2 + \mathcal{A})^2 - \mathcal{A}}.$$

For the three-period case, putting $g_3 = g^*$, we obtain

$$(1 - g^*)(\mathcal{A} - g^*) = [1 - (1 - g^*)(\mathcal{A} - g^*)][g^* - (1 - g^*)(\mathcal{A} - g^*)], \tag{15}$$

using the incentive and financial constraints. Comparing (14) and (15), we find that $g^* > \tilde{g}$. This follows as the LHS of both equations have the same functional form, and the RHS of (15) is less than the RHS of (14). ||

■ *Comparing FCC and PCC optimal programs.* The difference between full and partial commitment mechanisms could be due to one of two reasons – whether in the FCC program the second-period financial constraint, (8), binds or the first-period incentive compatibility constraint, (9), binds. When either constraint fails to bind, clearly the FCC program will strictly dominate the PCC program; otherwise the two programs yield an identical aid package. As Proposition 6-2(i) demonstrates (and is confirmed in Tables A.2 and A.3 in the Appendix), whenever the total aid $\mathcal{A}$ is large, constraint (8) may be slack, which has a simple economic explanation: interim aid diversion helps in improving governance.

7 Endogenous aid

So far the sanctioned budget for aid, i.e. $\mathcal{A}$, has been assumed to be fixed and exogenous. Here we consider an extension of our basic framework where the level of total aid will be determined endogenously depending on the aid recipient's private *type*, i.e. its efficiency in implementing reforms. Introduction of recipient types converts the aid design problem into one with incomplete information. While this makes the model more complex, it allows one to address an important issue in dynamic aid allocation – the interaction between the level of aid, selectivity of the better type, and its impact on the implemented level of governance. Further, this relates to the policy debate on whether to opt for conditionality, or selectivity discussed earlier. Given that the data suggests that selectivity, despite Burnside and Dollar (2000), is not really followed by all donors (Alesina and Weder (2002) and Dollar and Levin (2004)), it is conceivable that one reason is such uncertainty about recipient types. Our analysis shows that by identifying the better recipients, conditional aid schemes with endogenous aid can pave the way for informed use of selectivity. That is, selectivity and conditionality can be complementary tools for effective aid provision, rather than as alternatives, as they are often perceived to be.

For ease of exposition, the formal analysis is carried out with $\delta = 1$, so that there is no discounting of the future by the recipient.[31] Given that any additional aid is determined

[31] The results go through qualitatively for $0 < \delta < 1$.

only after the initial screening, it seems natural to assume that the continuation aid program will be optimal from that point on. We therefore analyze the PCC case.

Let the recipient be one of two potential types, $\tau \in \{\lambda, 1\}$ where $0 < \lambda < 1$, depending on its cost of improving governance: for $\tau = \lambda$, governance improvement cost is given by $\lambda C(g' - g)$, whereas for $\tau = 1$ governance improvement will cost $C(g' - g)$. That is, λ-type is (more) efficient, and we call $\tau = 1$ the inefficient type. The recipient knows its type but the donor does not. Let $p = \Pr(\tau = \lambda)$.

In the above scenario how should the donor disburse aid? Assuming that the continuation fund for aid is easier to be approved if early performance indicates superior cost-effectiveness by the recipient, we analyze a familiar screening mechanism: choose initial aid to determine the recipient's type and then vary future aid based on the information gained.

If the recipient is revealed to be of type $\tau = 1$, the total aid is $\mathcal{A}$ whereas if $\tau = \lambda$, then the total aid is $\mathcal{A} + \Upsilon$, where exogenous $\Upsilon > 0$ is awarded to efficient government.

Write the aid constraint over three periods as

$$\mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 = \mathcal{A}, \text{ for } \tau = 1, \quad \text{whereas } \mathcal{A}_1 + \mathcal{A}_2' + \mathcal{A}_3' = \mathcal{A} + \Upsilon, \text{ for } \tau = \lambda. \tag{16}$$

Here the donor starts with an initial aid and then proceeds with different aid trajectories after separation of types.

Given that in the PCC program aid has to be configured afresh in each period, incentive compatibility and financial constraints for any type can be taken to bind if the continuation aid after period 1 is designed optimally – our two-period analysis and Proposition 1 will apply. And if the aid has to be designed optimally, for a given type, for the entire three-period duration, again incentive and financial constraints will be binding (Proposition 7).

We now propose more formally a separating incentive scheme, $\mathcal{M}$:

The donor offers an aid in period 1 equal to $\breve{\mathcal{A}}_1 < \mathcal{A}$. While giving this aid the donor also sets a unique pair of targets, $\breve{g}_2$ and $\tilde{g}_2$ to bind period 1 financial constraints for the respective types:

$$\breve{\mathcal{A}}_1 = C(\breve{g}_2 - g_1), \qquad \breve{\mathcal{A}}_1 = \lambda C(\tilde{g}_2 - g_1). \tag{17}$$

Given $\lambda < 1$, it should be clear that $\tilde{g}_2 > \breve{g}_2$.

If the level of governance at the end of period 1 is g_2 such that $\breve{g}_2 \leq g_2 < \tilde{g}_2$, the donor offers an overall aid $\mathcal{A}$ with the balance $\mathcal{A} - \breve{\mathcal{A}}_1$ disbursed as in the reconfigured optimal PCC program for the inefficient type.[32] *If the achieved improvement is $\tilde{g}_2$ or higher, the donor offers an overall aid $\mathcal{A} + \Upsilon$ with $\mathcal{A} + \Upsilon - \breve{\mathcal{A}}_1$ then disbursed optimally as in the two-period*

[32]Reconfiguration is necessary as $\breve{g}_2$ is not necessarily the same as g_2^*. What is important is that the aid and target governance scheme in the continuation game satisfy the properties stated in Proposition 7.

program for the efficient type. If the level of governance falls below $\breve{g}_2$, the donor cuts off the aid completely. $\|$

We will not solve for what should be the optimal level of $\breve{\mathcal{A}}_1$. Given that $0 < p < 1$, it makes sense to screen types with $\breve{\mathcal{A}}_1$ somewhere between the optimal initial aid for the inefficient and efficient types.

It should be clear that under the above scheme the inefficient type will not be able to mimic the efficient type; it would ensure improvement to just $\breve{g}_2$ and from then the offered continuation program $(\breve{\mathcal{A}}_2, \breve{g}_3; \breve{\mathcal{A}}_3 = \mathcal{A} - \breve{\mathcal{A}}_1 - \breve{\mathcal{A}}_2)$ would satisfy IC and FC, binding both constraints. So we need to ensure only the efficient type's incentive compatibility and financial constraints, to which we turn next.

The IC constraint for (the already revealed) $\tau = \lambda$ at $t = 2$ is given by:

$$\alpha(\tilde{g}_2)[\tilde{\mathcal{A}}_2 - \lambda C(\tilde{g}_3 - \tilde{g}_2)] + \alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3 \geq \alpha(\tilde{g}_2)\tilde{\mathcal{A}}_2, \tag{18}$$

where $\tilde{\mathcal{A}}_3 = \mathcal{A} + \Upsilon - \breve{\mathcal{A}}_1 - \tilde{\mathcal{A}}_2$.

The financial constraint for $\tau = \lambda$ at $t = 2$ is given by:

$$\tilde{\mathcal{A}}_2 \geq \lambda C(\tilde{g}_3 - \tilde{g}_2). \tag{19}$$

Both (18) and (19) will bind, as observed in the two-period problem.

Stepping back and writing the IC constraint for the efficient type at $t = 1$ involves three different incentive considerations. Recipient must be screened for its true type as well as it must not run away with the initial or interim aid. We write these incentives in a single condition as follows:

$$\begin{aligned}
&\alpha(g_1)[\breve{\mathcal{A}}_1 - \lambda C(\tilde{g}_2 - g_1)] + \alpha(\tilde{g}_2)[\tilde{\mathcal{A}}_2 - \lambda C(\tilde{g}_3 - \tilde{g}_2)] + \alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3 \\
&\geq \max\big\{\alpha(g_1)[\breve{\mathcal{A}}_1 - C(\breve{g}_2 - g_1)] + [C(\breve{g}_2 - g_1) - \lambda C(\breve{g}_2 - g_1)] \\
&\qquad +\alpha(\breve{g}_2)[\breve{\mathcal{A}}_2 - C(\breve{g}_3 - \breve{g}_2)] + [C(\breve{g}_3 - \breve{g}_2) - \lambda C(\breve{g}_3 - \breve{g}_2)] \ + \ \alpha(\breve{g}_3)\breve{\mathcal{A}}_3, \\
&\qquad \alpha(g_1)[\breve{\mathcal{A}}_1 - C(\breve{g}_2 - g_1)] + [C(\breve{g}_2 - g_1) - \lambda C(\breve{g}_2 - g_1)] + \alpha(\breve{g}_2)\breve{\mathcal{A}}_2, \\
&\qquad \alpha(g_1)\breve{\mathcal{A}}_1\big\}.
\end{aligned} \tag{20}$$

Lemma 4. *Condition (20) can be reduced to*

$$\alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3 - \alpha(\breve{g}_3)\breve{\mathcal{A}}_3 \geq (1 - \lambda)[C(\breve{g}_2 - g_1) + C(\breve{g}_3 - \breve{g}_2)]. \tag{21}$$

Let us now write (7) and (8) as equalities,

$$\alpha(\tilde{g}_2)[\tilde{\mathcal{A}}_2 - \lambda C(\tilde{g}_3 - \tilde{g}_2)] + \alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3 = \alpha(\tilde{g}_2)\tilde{\mathcal{A}}_2,$$
$$\tilde{\mathcal{A}}_2 = \lambda C(\tilde{g}_3 - \tilde{g}_2),$$

using which yields

$$\alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3 = \alpha(\tilde{g}_2)\tilde{\mathcal{A}}_2, \tag{22}$$

which can further be rewritten as

$$\alpha(\tilde{g}_3)[\mathcal{A} + \Upsilon - \breve{\mathcal{A}}_1 - \lambda C(\tilde{g}_3 - \tilde{g}_2)] = \alpha(\tilde{g}_2)\lambda C(\tilde{g}_3 - \tilde{g}_2),$$
$$\text{or,} \quad \alpha(\tilde{g}_3)[\mathcal{A}_{23}(\Upsilon) - \lambda C(\tilde{g}_3 - \tilde{g}_2)] = \alpha(\tilde{g}_2)\lambda C(\tilde{g}_3 - \tilde{g}_2), \tag{23}$$

where $\mathcal{A}_{23}(\Upsilon) \equiv \mathcal{A} + \Upsilon - \breve{\mathcal{A}}_1$.

Our first result is about the effectiveness of endogenous conditional aid, assuming enough disposable funds.

Proposition 9 (**Aid disbursement through screening**). *Fix the efficient type's productivity, $0 < \lambda < 1$. Suppose further that the donor wants to retain the inefficient type recipient while screening, applying the incentive scheme $\mathcal{M}$. Then,*

(i) *For Υ sufficiently large, the reduced version of period 1 IC, as in condition (21), will be satisfied;*

(ii) *An exogenous aid increment (Υ) awarded to the efficient type in the separating mechanism will lead to better governance: $\tilde{g}_3 > \breve{g}_3$;*

(iii) *For any Υ inducing separation of types, any higher aid increment $\Upsilon' > \Upsilon$ would also enable screening of types and lead to a higher overall governance for the efficient type compared to the initial $\tilde{g}_3$.*

Thus, whenever the rewards for efficiency Υ is large enough, the donor can identify the two types of recipients by disbursing differential overall aid based on initial performance. It also follows that higher aid leads to a higher level of governance through more efficient reform implementation. Note that while the result in part **(iii)** about higher governance is an intuitive comparative static observation with respect to the level of aid *for* the efficient type, that in part **(ii)** is a comparison across the recipient's *types*.

We next present a negative result to highlight the limitations of conditional aid in case the reward for efficiency, i.e. Υ, is restricted.

Proposition 10 (Impossibility of screening). *Suppose that there exists $\bar{g} \geq 0$ such that $\lim_{g \to \bar{g}} \alpha(g) \to 0$. Then $\exists \bar{\Upsilon} > 0$ such that whenever the conditional additional aid Υ is small, i.e. $\Upsilon < \bar{\Upsilon}$, the mechanism $\mathcal{M}$ will fail to separate the types in the initial period.*

Thus, the type-dependent aid plan might fail for lack of enough funds due to the familiar *ratchet effect* in models of adverse selection: should the efficient type's productivity be revealed, the donor would run the the efficient type down by by setting a very high target level of governance. This would take away any potential gain, i.e. rent, that the efficient type could hope to extract. As a result, the efficient type would rather project as an inefficient type and take advantage of cost savings.

Remark 2. Note that Proposition 10 states that separation of types may be expensive in terms of the additional resource requirement, i.e. Υ. An implication is that separation is not feasible whenever there is a significant difference in efficiency between the two types:

Proposition $10'$. *Suppose that $\lim_{g \to \infty} \alpha(g) \to 0$. Then $\exists \bar{\lambda}$, $0 < \bar{\lambda} < 1$, such that whenever the efficient type's productivity is very high, i.e. $\lambda < \bar{\lambda}$, then for any given $\Upsilon > 0$, the incentive mechanism $\mathcal{M}$ will fail to separate the types in the initial period.*

Further, for any $\Upsilon < \bar{\Upsilon}$, we conjecture that there is a pooling aid contract that involves offering both the types the same contract, to be precise the PCC program for the inefficient type. Interestingly, in any such aid program, the additional amount Υ is never utilised along the equilibrium path. What the negative results in Propositions 10 and $10'$ demonstrate is a possible limitation of conditional aid, one that arises if the aid programme is not sufficiently responsive to the recipient's productivity.

8 Conclusion and discussion

In future works, our present framework can be expanded to allow for competing recipients. In such a general multi-lateral aid setting additional issues of interest arise. First, a donor may have goals other than, or in addition to, governance, e.g. developmental/welfare goals. In addition, the donor's objective may differ from one country to another – governance improvement and reform in one, and development objectives, such as the eradication of child malnutrition, disease control and the likes, in another. But even without such complexities, the single-recipient case poses interesting issues and theoretical subtleties that, moreover, has interesting implications for the empirical and policy debate on conditionality. Our work should be viewed as a necessary first step in understanding and analyzing the additional issues that arise in a more general setting.[33]

Although the predominant focus of this work has been theoretical in nature, we would like to conclude by noting several aspects about a major US aid initiative, the *Millennium Challenge Corporation* (or, *MCC*), established in 2004. As reported in *The Economist*,[34] under MCC programs aid is conditional on implementation of various closely monitored reforms. MCC states on its website (https://www.mcc.gov/who-we-fund): "For a country to be selected as eligible for an MCC assistance program, it must demonstrate a commitment to just and democratic governance, investments in its people and economic freedom as measured by different policy indicators." Further it states, "When considering a country for a subsequent compact, the Board also takes into consideration evidence of improved scorecard policy performance and a commitment to sectoral reform."

MCC monitors all aid recipient countries, publishing year-by-year indices of their performance (typically over five years) under the three broad categories of "Economic Freedom," "Investing in People," and "Ruling Justly". Based on these measures, a country can be deemed to have either passed or failed with respect to "Control of Corruption",[35] and "Democratic Rights". A sample tracking of performance is summarized in Table 1 for five countries (see also Fig. 2). Interestingly, many of the 'governance' indicators appear to exhibit an increasing pattern. But there are also instances, Indonesia and Philippines for example, where performance indicators show repeated failures in the category of "Control of Corruption" and yet these countries continued to receive MCC aid (see Table 1, Table 2, and Fig. 2 below). This suggests that adherence to strict conditions along important dimensions for future eligibility of aid was not always followed, despite claims to the contrary (see the last paragraph above).

Next, in Table 2, we briefly allude to some empirical facts on the temporal distribution of MCC aid money for ten selected countries.[36] It is interesting that over the first few years of

[33] A related work is Carter et al. (2015), who analyze a problem involving multiple aid recipients where a single donor chooses a dynamic aid plan to maximize a weighted sum of the recipient countries' welfare. The focus is on the effect of differential aid absorption ability of the recipients, rather than aid diversion or aid conditionality.

[34] See the report, *Conditional aid,* that appeared on March 2nd, 2013 (http://www.economist.com/-news/international/21572754-controversial-approach-helping-poor-countries-seems-work-after-all-carrots-all).

[35] "This indicator measures the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as 'capture' of the state by elites and private interests. It also measures the strength and effectiveness of a country's policy and institutional framework to prevent and combat corruption." See https://www.mcc.gov/who-we-fund/indicator/control-of-corruption-indicator.

[36] See https://www.mcc.gov/where-we-work. Country-by-country information can be accessed at https://assets.mcc.gov/. Note that most compacts allow for some flexibility among projects and some diversion of funds from one project to another, so that our focus on aggregate outlay in a country rather than on specific projects is not too unreasonable.

Table 1: **Scorecard by MCC**

Country	Ruling Justly	FY12	FY13	FY14	FY15	FY16	Pass/Fail
Benin	Political Rights	93%	95%	93%	92%	92%	**FAILED**
	Civil Liberties	98%	100%	100%	100%	100%	in
	Freedom of Information	92%	91%	93%	92%	92%	"Control of
	Government Effectiveness	80%	85%	82%	83%	81%	Corruption"
	Rule of Law	66%	64%	75%	74%	75%	category
	Control of Corruption	54%	62%	49%	51%	58%	in FY14
Georgia	Political Rights	45%	38%	42%	46%	52%	**Never FAILED**
	Civil Liberties	52%	47%	54%	68%	70%	in
	Freedom of Information	50%	47%	54%	61%	67%	"Control of
	Government Effectiveness	97%	97%	100%	100%	100%	Corruption"
	Rule of Law	62%	72%	77%	79%	89%	or "Democratic
	Control of Corruption	76%	84%	88%	89%	89%	Rights" category
Indonesia	Political Rights	66%	69%	69%	71%	74%	Consistently
	Civil Liberties	48%	44%	46%	43%	48%	**FAILED**
	Freedom of Information	50%	50%	50%	46%	52%	in
	Government Effectiveness	66%	66%	58%	61%	78%	"Control of
	Rule of Law	34%	34%	35%	46%	52%	Corruption"
	Control of Corruption	31%	28%	31%	39%	48%	category
Madagaskar	Political Rights	17%	16%	16%	51%	67%	**FAILED**
	Civil Liberties	59%	65%	62%	60%	67%	"Half
	Freedom of Information	39%	42%	38%	49%	56%	Overall"
	Government Effectiveness	54%	51%	40%	38%	29%	in
	Rule of Law	58%	58%	51%	49%	56%	FY14 & FY15
	Control of Corruption	95%	95%	71%	66%	54%	
Philippines	Political Rights	52%	50%	50%	50%	52%	**FAILED**
	Civil Liberties	52%	47%	65%	64%	37%	in
	Freedom of Information	62%	63%	65%	64%	45%	"Control of
	Government Effectiveness	83%	91%	85%	82%	89%	Corruption"
	Rule of Law	38%	41%	42%	54%	56%	category
	Control of Corruption	24%	19%	46%	61%	56%	in FY12–14

Table 2: **Multi-year Aid program under MCC (in US$)**

Country	Year signed	Y1	Y2	Y3	Y4	Y5
Benin	2006	32,383,808	63,130,808	98,300,808	90,386,808	23,095,808
Georgia	2013	15,635,071	33,900,156	39,150,844	36,972,152	10,991,777
Ghana	2006	74,811,000	100,033,000	149,242,000	155,611,000	67,312,000
Indonesia	2011	113,079,374	116,890,423	114,711,816	126,699,269	116,619,088
Madagascar	2005	26,749,000	48,247,000	26,589,000	8,188,000	
Mozambique	2007	56,469,313	65,738,629	125,448,545	151,285,110	107,982,456
Nicaragua	2006	20,400,000	41,300,000	52,500,000	47,500,000	13,300,000
Philippines	2010	43,560,000	89,740,000	93,320,000	105,710,000	76,520,000
Tanzania	2008	57,745,000[a]	155,481,000	227,252,000	167,395,000	90,263,000
Vanuatu	2006	11,420,000	23,510,000	26,430,000	2,850,000	1,520,000

[a] Year 1 amount for Tanzania is inclusive of CIF (Compact Implementation Funding).

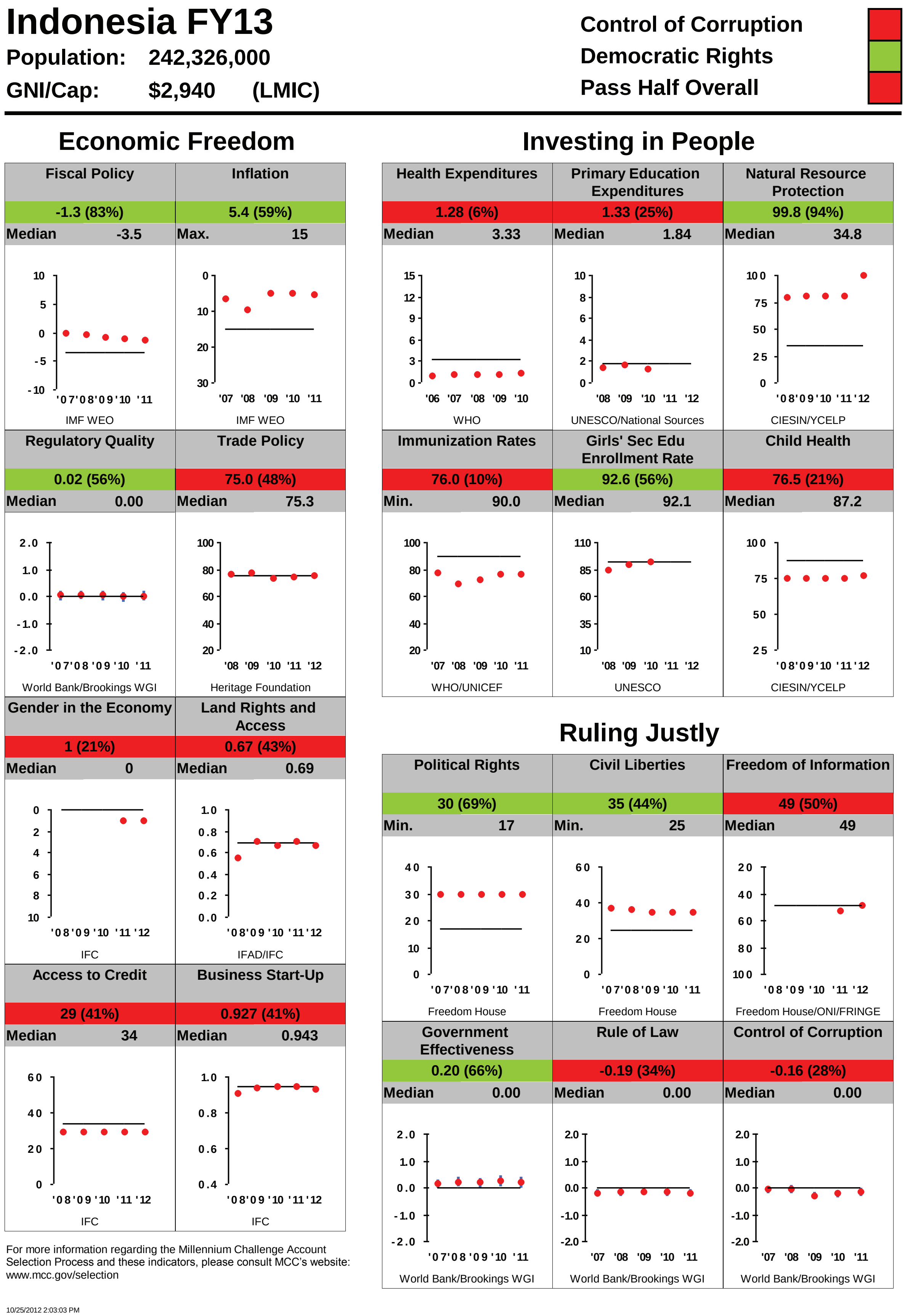

Figure 2: Indonesia: MCC Compact summary of performance in the Year 2013

the program, the aid pattern seems to be broadly increasing. Under the assumption that the bulk of the targeted work was done in the first few years, with the later years being mainly devoted to program evaluation etc.,[37] such a pattern is consistent with the increasing aid pattern that we find when commitment is partial (see Proposition 8).[38] This is of course only suggestive, and we are by no means claiming that the reported data vindicates our model. This is even more true given that the underlying aid objectives are often multi-dimensional (growth, poverty reduction, governance improvement/reforms, etc.), rather than just focused on governance as under our theoretical framework.

The preceding discussion on MCC suggests some further avenues for future work. Given our discussion of the seeming divergence between aid conditions and the actual implementation of policies, it would be of interest to analyze what are the factors that go into a decision regarding the magnitude of aid, and whether aid is continued or not, and, relatedly, whether repeat aid is provided or not. Consider MCC aid to Gerogia, Indonesia and Philippines (all with agreements starting between 2010 and 2013). Note that Indonesia and Philippines received much more aid than Georgia despite (a) Georgia having performed much better in the the categories of corruption and human rights, and (b) the gross national income of Georgia and Indonesia are not too disparate.[39] The aid decision involving any country is therefore more complex than just basing it on one or two performance measures. To be able to formulate any plausible hypothesis would require a separate analysis that allows for multiple recipients. Further, an empirical analysis of these issues should also be rewarding.

A Appendix

Proof of Proposition 1. First we claim that in the optimal program either (2) or (3) must hold with equality. If not, it is possible to increase g_2 slightly and still satisfy the incentive and financial constraints. But this would improve governance at $t = 2$, a contradiction.

Next we show that both constraints, (2) and (3), must bind. We do this in two steps.

[37] Nicaragua compact, point 5 of Exhibit A (Notes): "Although most Project Activities will take place from Year 1 through Year 4 (except for roads), the five-year Compact Term will allow additional time to ensure that Project Activities are completed. Monitoring and Evaluation will continue after the completion of the Project Activities."

[38] Also, recall from our discussion earlier that the PCC formulation is perhaps more natural when aid is bilateral.

[39] In fact, Georgia's per-capita gross national incomes (GNI) during FY12-FY16 are \$2690, \$2860, \$3280, \$3570 and \$3720 respectively, whereas similar figures for Indonesia and Philippines are, respectively \$2580, \$2940, \$3420, \$3580, \$3650, and \$2050, \$2210, \$2470, \$3270, \$3440. Thus while relatively much lower per-capita GNI of Philippines could be one reason for greater aid, income figures for Indonesia and Georgia are very similar.

Step 1. First suppose that (2) binds, while (3) does not. Then construct another program $(\mathcal{A}_1', \mathcal{A}_2')$ such that $\mathcal{A}_1' = \tilde{\mathcal{A}}_1 - \epsilon$ and $\mathcal{A}_2' = \tilde{\mathcal{A}}_2 + \epsilon$, $\epsilon > 0$. Given $\tilde{A}_1 > 0$ (from Lemma 1), such a program exists for ϵ small. Moreover, for ϵ small, both the constraints hold with strict inequality. Thus g_2 can be increased slightly.

Step 2. Next suppose that (3) binds, while (2) does not. Then construct another program $(\mathcal{A}_1', \mathcal{A}_2')$ such that $\mathcal{A}_1' = \tilde{\mathcal{A}}_1 + \epsilon$ and $\mathcal{A}_2' = \tilde{\mathcal{A}}_2 - \epsilon$, $\epsilon > 0$. Given $\tilde{\mathcal{A}}_2 > 0$ (from Lemma 1), such a program exists for ϵ small. Moreover, for ϵ small, both the constraints hold with strict inequality. Thus g_2 can be increased slightly.

Parts **(i)** and **(ii)** follow from the fact that both the financial and incentive constraints bind in the optimal solution. For a proof of **(iii)**, observe that since both the constraints bind we have:

$$\frac{\Delta \cdot \alpha(\tilde{g}_2)}{\alpha(g_1)} = \frac{\tilde{\mathcal{A}}_1}{\tilde{\mathcal{A}}_2}.$$

The result now follows as $\Delta < 1$, $\tilde{g}_2 > g_1$ (from Lemma 1), and $\alpha(g)$ is strictly decreasing. Q.E.D.

Proof of Proposition 3. (i) The proof relies on the following steps.

Step 1. Starting from any initial g_1, take the optimal two-period incentives $(\tilde{\mathcal{A}}_1, \tilde{g}_2; \tilde{\mathcal{A}}_2)$ where $(\tilde{\mathcal{A}}_1, \tilde{\mathcal{A}}_2) \gg 0$ and $\tilde{g}_2 > g_1$ and plug these values into the IC and financial constraints for the three-period problem. That is, we set $\mathcal{A}_1 = \tilde{\mathcal{A}}_1$, $\mathcal{A}_2 = \mathcal{A} - \tilde{\mathcal{A}}_1 = \tilde{\mathcal{A}}_2$, $\mathcal{A}_3 = 0$, $g_2 = g_3 = \tilde{g}_2$. Now check that all four constraints in the three-period problem are satisfied:

$$\begin{array}{llcll}
(7) & \delta\dfrac{\alpha(g_3)\mathcal{A}_3}{\alpha(g_2)} & \geq & C(g_3 - g_2) & \\
 & \text{i.e.,} \quad \delta\mathcal{A}_3 & = & 0 = C(g_3 - g_2) & (\text{by } g_3 = g_2, \mathcal{A}_3 = 0); \\
(8) & \mathcal{A}_2 & \geq & C(g_3 - g_2) & \\
 & \text{i.e.,} \quad \mathcal{A}_2 & > & 0 = C(g_3 - g_2) & (\text{by } g_3 = g_2, \tilde{\mathcal{A}}_2 > 0, \text{ and } \mathcal{A}_2 = \tilde{\mathcal{A}}_2); \\
(9) & \delta\alpha(g_2)[\mathcal{A}_2 - C(g_3 - g_2)] + \delta\alpha(g_3)\mathcal{A}_3 & > & \alpha(g_1)C(g_2 - g_1) & \\
 & \text{i.e.,} \quad \delta\alpha(g_2)\mathcal{A}_2 & > & \alpha(g_1)C(g_2 - g_1) & (\text{by } \mathcal{A}_2 > 0, \Delta < \delta, g_3 = g_2, \mathcal{A}_3 = 0, \text{ and } (4)); \\
(10) & \mathcal{A}_1 & = & C(g_2 - g_1) & (\text{true by } (5)).
\end{array}$$

Starting from the above inequalities, next we will construct another program with a higher g_3 where all the inequalities are satisfied.

Step 2. Reduce $\mathcal{A}_2$ by ϵ to $\mathcal{A}_2'$, and increase $\mathcal{A}_3$ by ϵ to $\mathcal{A}_3'$, where ϵ is sufficiently small.

Thus from the above set of inequalities we can write, for ϵ small,

$$\begin{aligned} \delta\alpha(g_3)\mathcal{A}_3' &> 0 = \alpha(g_2)C(g_3 - g_2) && (7.1) \\ \mathcal{A}_2' &> 0 = C(g_3 - g_2) && (8.1) \\ \delta\alpha(g_2)[\mathcal{A}_2' - C(g_3 - g_2)] + \delta^2\alpha(g_3)\mathcal{A}_3' &> \alpha(g_1)C(g_2 - g_1) && (9.1) \\ \mathcal{A}_1 &= C(g_2 - g_1). && (10.1) \end{aligned}$$

Step 3. Lower g_2 slightly to g_2' while maintaining g_3 at the original g_2-value, to obtain:

$$\begin{aligned} \delta\alpha(g_3)\mathcal{A}_3' &> \alpha(g_2)C(g_3 - g_2') && (7.2) \\ \mathcal{A}_2' &> C(g_3 - g_2') && (8.2) \\ \delta\alpha(g_2')[\mathcal{A}_2' - C(g_3 - g_2')] + \delta^2\alpha(g_3)\mathcal{A}_3' &> \alpha(g_1)C(g_2' - g_1) && (9.2) \\ \mathcal{A}_1 &> C(g_2' - g_1). && (10.2) \end{aligned}$$

The inequalities follow from continuity of expressions (on both sides of inequalities) in g_2.

Step 4. Let us now increase g_3 slightly to $g_2 + \epsilon'$. Since ϵ' is very small and the expressions on both sides of (7.2)–(10.2) are continuous in g_3, all the inequalities will continue to be maintained, implying $g_2 + \epsilon'$ is achievable under the three-period program, thus proving the claim.

(ii) To establish gradualism first we show that $\hat{g}_3 > \hat{g}_2$, then we eliminate the possibility that $\hat{g}_2 = g_1$.

Suppose, to the contrary, $\hat{g}_3 = \hat{g}_2$. If $\hat{\mathcal{A}}_3 = 0$, it is a two-period program which is not optimal as shown in part **(i)**. So let $\hat{\mathcal{A}}_3 > 0$ and construct an alternative program $(\mathcal{A}_1', g_2'; \mathcal{A}_2', g_3'; \mathcal{A}_3')$ such that

$$\mathcal{A}_1' = \hat{\mathcal{A}}_1,\ \mathcal{A}_2' = \hat{\mathcal{A}}_2 + \hat{\mathcal{A}}_3,\ \mathcal{A}_3' = 0,\ g_2' = \hat{g}_2,\ g_3' = \hat{g}_3.$$

This is clearly feasible (i.e., satisfy (6)–(10)) but effectively a two-period program which, by part **(i)**, can be strictly dominated by a three-period program. Hence, $\hat{g}_3 > \hat{g}_2$.

Now consider the possibility $\hat{g}_2 = g_1$. If $\hat{\mathcal{A}}_1 = 0$, effectively it is a two-period program which, by part **(i)**, cannot be optimal. So let $\hat{\mathcal{A}}_1 > 0$. Now, starting from this one, construct another program $(\mathcal{A}_1', g_2'; \mathcal{A}_2', g_3'; \mathcal{A}_3')$ such that

$$\mathcal{A}_1' = 0,\ \mathcal{A}_2' = \hat{\mathcal{A}}_1 + \hat{\mathcal{A}}_2,\ \mathcal{A}_3' = \hat{\mathcal{A}}_3,\ g_2' = \hat{g}_2,\ g_3' = \hat{g}_3.$$

Again, this is effectively a feasible two-period program, which can be strictly dominated by a three-period program. Hence it must be that $\hat{g}_2 > g_1$.

We can therefore conclude that the optimal three-period program must exhibit gradualism: $\hat{g}_3 > \hat{g}_2 > g_1$. Q.E.D.

Proof of Proposition 5. **(i)** In the proof of Proposition 4**(iii)**, the key step was to show that $Z'(0) < 0$. We follow the same method here.

Under Assumption 2, we have

$$Z'(0) \equiv Z'(\epsilon)|_{\epsilon=0} = C'(\hat{g}_2 - g_1)[1 - \frac{\alpha(\hat{g}_2)}{\alpha(\hat{g}_3)}] + \frac{\alpha'(\hat{g}_2)}{\alpha(\hat{g}_3)} C(\hat{g}_3 - \hat{g}_2) < 0,$$

given that $\hat{g}_3 > \hat{g}_2$ (by Proposition 3), $C(.)$ is strictly increasing and $\alpha(g)$ is strictly decreasing. The rest of the argument is the same as in Proposition 4**(iii)**.

(ii) We have

$$\alpha(g_1)C(\hat{g}_2 - g_1) = \alpha(g_1)\hat{\mathcal{A}}_1 = \delta^2\alpha(\hat{g}_3)\hat{\mathcal{A}}_3 + \delta\alpha(\hat{g}_2)\hat{\mathcal{A}}_2 - \delta\alpha(\hat{g}_2)C(\hat{g}_3 - \hat{g}_2) = \delta\alpha(\hat{g}_2)\hat{\mathcal{A}}_2,$$

where the first equality follows from (10) binding, the second from (9) binding, and the last one from (7) binding. Given that (i) $\hat{g}_2 > \hat{g}_1$ (Proposition 3) and (ii) $\alpha(g)$ is strictly decreasing, this in turn yields that $\hat{\mathcal{A}}_1 < \hat{\mathcal{A}}_2$.

Finally, from the second equality above we have $\alpha(g_1)\hat{\mathcal{A}}_1 \geq \alpha(\hat{g}_3)\hat{\mathcal{A}}_3$, since $\hat{\mathcal{A}}_2 \geq C(\hat{g}_3 - \hat{g}_2)$. Q.E.D.

Proof of Proposition 6. **1.** We first claim that g_3^* is increasing in $\mathcal{A}$. We shall prove this by showing that $f(g_2, \mathcal{A})$ and $i(g_2, \mathcal{A})$ are both increasing in $\mathcal{A}$. Invoking the implicit function theorem,

$$\frac{\partial f(g_2, \mathcal{A})}{\partial \mathcal{A}} = -\frac{\frac{\partial F(g_2, g_3; \mathcal{A})}{\partial \mathcal{A}}}{\frac{\partial F(g_2, g_3; \mathcal{A})}{\partial g_3}}.$$

Hence $\frac{\partial f(g_2,\mathcal{A})}{\partial \mathcal{A}} > 0$, since $\frac{\partial F(g_2,g_3;\mathcal{A})}{\partial \mathcal{A}} = 1$ and $\frac{\partial F(g_2,g_3;\mathcal{A})}{\partial g_3} < 0$ (from part **(i)** of Lemma 2).

Similarly,

$$\frac{\partial i(g_2, \mathcal{A})}{\partial \mathcal{A}} = -\frac{\frac{\partial I(g_2, g_3; \mathcal{A})}{\partial \mathcal{A}}}{\frac{\partial I(g_2, g_3; \mathcal{A})}{\partial g_3}} > 0,$$

since $\frac{\partial I(g_2, g_3; \mathcal{A})}{\partial \mathcal{A}} = 1$ and $\frac{\partial I(g_2, g_3; \mathcal{A})}{\partial g_3} = -\frac{\alpha(g_2)C'(g_3 - g_2)}{\delta\alpha(g_3)} + \frac{\alpha(g_2)C(g_3 - g_2)\alpha'(g_3)}{\delta(\alpha(g_3))^2} < 0.$

Clearly the result holds irrespective of whether there is interim aid diversion, or not.

We claim that g_3^* is increasing in δ. We shall prove this by showing that $f(g_2, \mathcal{A})$ and

$i(g_2, \mathcal{A})$ are both increasing in δ. By the implicit function theorem,

$$\frac{\partial f(g_2, \mathcal{A})}{\partial \delta} = -\frac{\frac{\partial F(g_2,g_3;\mathcal{A})}{\partial \delta}}{\frac{\partial F(g_2,g_3;\mathcal{A})}{\partial g_3}}.$$

Hence $\frac{\partial f(g_2,\mathcal{A})}{\partial \delta} > 0$, since $\frac{\partial F(g_2,g_3;\mathcal{A})}{\partial \delta} > 0$ and $\frac{\partial F(g_2,g_3;\mathcal{A})}{\partial g_3} < 0$ (from part **(i)** of Lemma 2).

Similarly, $\frac{\partial i(g_2,\mathcal{A})}{\partial \delta} = -\frac{\frac{\partial I(g_2,g_3;\mathcal{A})}{\partial \delta}}{\frac{\partial I(g_2,g_3;\mathcal{A})}{\partial g_3}} > 0$, since $\frac{\partial I(g_2,g_3;\mathcal{A})}{\partial \delta} > 0$ and $\frac{\partial I(g_2,g_3;\mathcal{A})}{\partial g_3} < 0$ (see verification of Lemma 3**(ii)**).

Clearly, the result holds irrespective of whether there is interim aid diversion, or not.

2. Recall that interim aid diversion happens whenever $i(g_2^*, \mathcal{A}) < f(g_2^*, \mathcal{A})$ (see Lemma 3**(ii)**). Thus the claim in part **(i)** is equivalent to showing that if $i(g_2^*, \mathcal{A}) < f(g_2^*, \mathcal{A})$ for some $\mathcal{A}$, then the inequality holds for all higher values of aid. Formally, we need to show that

$$\frac{dF(g_2^*(\mathcal{A}), i(g_2^*(\mathcal{A}), \mathcal{A}); \mathcal{A})}{d\mathcal{A}} \geq 0.$$

Further, this also establishes the claim in part **(ii)**.

Next note that

$$\begin{aligned}\frac{dF(g_2^*(\mathcal{A}), i(g_2^*(\mathcal{A}), \mathcal{A}); \mathcal{A})}{d\mathcal{A}} &= \frac{\partial F(g_2^*, g_3^*; \mathcal{A})}{\partial \mathcal{A}} + \frac{\partial F(g_2^*, g_3^*; \mathcal{A})}{\partial g_2^*}\frac{dg_2^*}{d\mathcal{A}} \\ &\quad + \frac{\partial F(g_2^*, g_3^*; \mathcal{A})}{\partial g_3^*}\frac{\partial i(g_2^*, \mathcal{A})}{\partial \mathcal{A}} + \frac{\partial F(g_2^*, g_3^*; \mathcal{A})}{\partial g_3^*}\frac{\partial i(g_2^*, \mathcal{A})}{\partial g_2^*}\frac{dg_2^*}{d\mathcal{A}} \\ &= 1 + \frac{\partial F}{\partial g_2^*}\frac{dg_2^*}{d\mathcal{A}} + \frac{\partial F(g_2^*, g_3^*; \mathcal{A})}{\partial g_3^*}\frac{\partial i(g_2^*, \mathcal{A})}{\partial \mathcal{A}},\end{aligned}$$

since $\frac{\partial i(g_2^*,\mathcal{A})}{\partial g_2^*} = 0$ (recall (13)) and $\frac{\partial F(g_2^*,g_3^*;\mathcal{A})}{\partial \mathcal{A}} = 1$.

Next consider the term $\frac{dg_2^*}{d\mathcal{A}}$. Differentiating both sides of

$$I(g_2, i(g_2, \mathcal{A}), \mathcal{A}) \equiv 0$$

with respect to g_2 and evaluating $\frac{\partial i(g_2,\mathcal{A})}{\partial g_2}$ at $g_2 = g_2^*$ yields:

$$\frac{\partial i(g_2^*, \mathcal{A})}{\partial g_2^*} = -\frac{I_1(g_2^*, g_3^*; \mathcal{A})}{I_2(g_2^*, g_3^*; \mathcal{A})} = 0;$$

similarly obtain: $\frac{\partial i}{\partial \mathcal{A}} = -\frac{I_3}{I_2} = -\frac{1}{I_2} > 0$, given that $I_3 = 1$ and $I_2 < 0$.

Next, totally differentiating $I_1(g_2^*, i(g_2^*, \mathcal{A}); \mathcal{A}) = 0$, we have

$$I_{11} dg_2^* + I_{12} \Big[\underbrace{\frac{\partial i}{\partial g_2}}_{=0} dg_2^* + \frac{\partial i}{\partial \mathcal{A}} d\mathcal{A} \Big] + \underbrace{I_{13}}_{=0} d\mathcal{A} = 0,$$

or, $$\frac{dg_2^*}{d\mathcal{A}} = -\frac{I_{12} \frac{\partial i}{\partial \mathcal{A}}}{I_{11}} = \frac{I_{12}}{I_{11}} \frac{1}{I_2} > 0. \quad \text{(by parts } \textbf{(ii)} \text{ and } \textbf{(v)} \text{ of Lemma 2 and Assumption 3}\textbf{(i)}\text{)}$$

Hence,

$$\frac{dF(g_2^*(\mathcal{A}), i(g_2^*(\mathcal{A}), \mathcal{A}); \mathcal{A})}{d\mathcal{A}} = 1 - \frac{\partial F}{\partial g_2^*} \frac{I_{12} \frac{\partial i}{\partial \mathcal{A}}}{I_{11}} + \frac{\partial F}{\partial g_3^*} \frac{\partial i}{\partial \mathcal{A}} = 1 - \frac{F_1}{I_2} \Big[- \frac{I_{12}}{I_{11}} + \frac{F_2}{F_1} \Big] > 0,$$

given that (i) $I_2 < 0$, (ii) $I_{12} > 0$, $I_{11} < 0$ and $\left| \frac{I_{12}(g_2, g_3; \mathcal{A})}{I_{11}(g_2, g_3; \mathcal{A})} \right| > \left| \frac{F_2(g_2, g_3; \mathcal{A})}{F_1(g_2, g_3; \mathcal{A})} \right|$ (by hypothesis), and (iii) $F_1 > 0$ given Assumption 3**(i)**. Q.E.D.

Numerical simulation on aid diversion. Using Mathematica (see the supplementary file), governance g_3 is maximized under the assumption that $g_1 = 0$, $\alpha(g) = 1 - g$ and two alternative cost (of governance improvement) assumptions – linear ($C(x) = x$) and quadratic ($C(x) = x^2$). The results are reported in Tables A.2 and A.3.

As one can see, under the optimal program the constraint (8) in most cases is in slack, i.e. $\hat{\mathcal{A}}_2 - C(\hat{g}_3 - \hat{g}_2) > 0$. Further, the column (8)/$\mathcal{A}$ indicates that in percentage terms interim aid diversion can be significant, as high as 38% (refer Table A.3); also, *the aid diversion percentage tends to increase with the aid budget for both Tables.*

$\mathcal{A}$	$\hat{\mathcal{A}}_1$	$\hat{g}_2$	$\hat{\mathcal{A}}_2$	$\hat{g}_3$	$\hat{\mathcal{A}}_3$	(6′)	(7)	(8)	(8)/$\mathcal{A}$
1	0.227777	0.227777	0.294963	0.52274	0.47726	1	3.02014E-008	3.49962E-009	–
1.1	0.240472	0.240472	0.316607	0.557078	0.542922	1.09999998825149	3.56359317610355E-008	9.49248957216042E-010	–
1.2	0.251799	0.251799	0.336539	0.588337	0.611663	1.2	0	4.996E-016	–
1.3	0.263458	0.263458	0.357695	0.616717	0.678847	1.3	-5.55112E-017	**0.00443544**	–
1.4	0.289326	0.289326	0.407116	0.642875	0.703558	1.4	-5.55112E-017	**0.0535668**	**3.82%**
1.5	0.313953	0.313953	0.457626	0.667253	0.728421	1.5	9.49517E-010	**0.104326**	**6.95%**
1.6	0.337126	0.337126	0.508583	0.689943	0.754291	1.6	2.45603E-008	**0.155766**	**9.73%**
1.7	0.359195	0.359195	0.560537	0.711042	0.780268	1.7	-1.11022E-016	**0.20869**	**12.27%**
1.8	0.380132	0.380132	0.613246	0.730642	0.806623	1.8	-1.66533E-016	**0.262735**	**14.59%**
1.9	0.4	0.4	0.666667	0.748837	0.833333	1.9	0	**0.317829**	**16.72%**
2	0.418861	0.418861	0.720759	0.765718	0.86038	2	1.11022E-016	**0.373903**	**18.69%**

$\mathcal{A}$	(9)	(10)
1	7.23845E-008	-2.17419E-010
1.1	7.91897344643022E-008	5.84201764741721E-009
1.2	-5.55112E-017	0
1.3	-5.55112E-017	0
1.4	-5.55112E-017	0
1.5	5.18476E-008	-5.36185E-009
1.6	1.51515E-008	3.55657E-008
1.7	-5.55112E-017	0
1.8	0	0
1.9	5.55112E-017	0
2	5.55112E-017	0

Table A.2: Simulation results: $C(x) = x$

$\mathcal{A}$	$\hat{\mathcal{A}}_1$	$\hat{g}_2$	$\hat{\mathcal{A}}_2$	$\hat{g}_3$	$\hat{\mathcal{A}}_3$	(6′)	(7)	(8)	(8)/$\mathcal{A}$	(9)	(10)
1	0.198828	0.445901	0.358831	0.822418	0.442341	1	0.000000008	**0.217066**	**21.7%**	2.57873E-008	-3.59719E-010
1.1	0.216793	0.46561	0.405683	0.841707	0.477524	1.1	1.26257E-008	**0.264234**	**24%**	2.69114E-009	-2.2499E-009
1.2	0.234299	0.484045	0.454108	0.85855	0.511593	1.2	-2.55361E-009	**0.313854**	**26.15%**	1.84779E-008	3.69984E-009
1.3	0.251265	0.501263	0.503802	0.873313	0.544933	1.3	-3.64195E-009	**0.365381**	**28.1%**	0.000000025	-5.33398E-009
1.4	0.267717	0.517414	0.554756	0.886296	0.577527	1.4	-2.91145E-009	**0.418682**	**29.9%**	2.77482E-008	-2.9198E-009
1.5	0.283647	0.532585	0.606843	0.897745	0.60951	1.5	0	**0.473502**	**31.56%**	0	0
1.6	0.299107	0.546907	**0.660144**	0.907866	**0.640749**	1.6	1.43339E-008	**0.529852**	**33.11%**	4.2321E-008	-7.31378E-010
1.7	0.314246	0.560577	**0.715134**	0.916836	**0.670619**	1.7	1.10953E-008	**0.588214**	**34.6%**	3.31976E-008	-3.71714E-010
1.8	0.328725	0.573346	**0.770472**	0.9248	**0.700803**	1.8	-1.49228E-009	**0.646952**	**35.94%**	2.85905E-008	-8.61455E-009
1.9	0.342547	0.585275	**0.825962**	0.931886	**0.731491**	1.9	0.000000023	**0.705823**	**37.14%**	7.63415E-008	2.23736E-007
2	0.356105	0.596745	**0.883076**	0.938202	**0.760819**	2	-2.77556E-017	**0.766483**	**38.32%**	5.55112E-017	0

Table A.3: Simulation results: $C(x) = x^2$

Remark 3. The entry, E, refers to exponential function. (6′) is computed for the value of $\hat{\mathcal{A}}_1 + \hat{\mathcal{A}}_2 + \hat{\mathcal{A}}_3$. (7)–(10) are computed for the surplus values, LHS-RHS of the constraints. Although all surplus values are not exactly zeros, for most cases the values are very close to zeros so that we could take the constraint as "approximately" binding. In particular, there are no inconsistencies between Table A.2 entries showing at several places $\hat{\mathcal{A}}_1 = \hat{g}_2$ and yet the surplus values in constraint (10) not showing exactly zeros, because both $\hat{\mathcal{A}}_1$ and $\hat{g}_2$ (just like other choice variables) are approximations that are equal up to a certain number of decimal points. Finally, for the blue colored cells the surplus values are significantly positive, so the constraint (8) can be taken to be slack. Finally, in the Mathematica program we have written the budget constraint as weak inequality, as follows:

$$\mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 \leq \mathcal{A}. \tag{6′}$$

However, as one can see from Tables A.2 and A.3, in almost all cases the optimal aid profile binds the budget constraint.

Note that Table A.3 numbers for $\hat{g}_2$ and $\hat{g}_3$ for any given $\mathcal{A}$ can be generated using simple Mathematica plot for quadratic cost specification as in Fig. 1 (which set $\mathcal{A} = 1.9$). ||

Proposition 7 is a shortened version of the following lemma.

Lemma A.1. *Suppose Assumption 1 holds. Then the optimal PCC program can be characterized as follows:*

(i) *Constraints (7), (8), and (10) all bind.*

(ii) $\alpha(g_1)\mathcal{A}_1^* \leq \delta^2\alpha(g_3^*)\mathcal{A}_3^* = \delta\alpha(g_2^*)\mathcal{A}_2^*$, *which implies that* $\mathcal{A}_3^* \geq \mathcal{A}_2^* \geq \mathcal{A}_1^*$.

(iii) $g_3^* - g_2^* \geq g_2^* - g_1$.

(iv) *Period 1 incentive compatibility constraint (9) will bind.*

(v) $\mathcal{A}_1^* > 0$.

(vi) $g_2^* - g_1 > 0$.

The proof of this lemma is included in supplementary material B.

Proof of Proposition 8. **(i)** From Lemma A.1, $0 < \alpha(g_1)\mathcal{A}_1^* \leq \alpha(g_2^*)\mathcal{A}_2^* = \delta\alpha(g_3^*)\mathcal{A}_3^*$. It must therefore be that $\mathcal{A}_1^*, \mathcal{A}_2^*, \mathcal{A}_3^* \gg 0$. And since $\mathcal{A}_1^* = C(g_2^* - g_1)$ and $\mathcal{A}_2^* = C(g_3^* - g_2^*)$, it follows that $g_3^* > g_2^* > g_1$ (apply the fact that $C(0) = 0$ and $C(.)$ is strictly increasing). Given that $\alpha(g_1)\mathcal{A}_1^* \leq \alpha(g_2^*)\mathcal{A}_2^* = \delta\alpha(g_3^*)\mathcal{A}_3^*$, this in turn implies that $\mathcal{A}_3^* > \mathcal{A}_2^* > \mathcal{A}_1^*$.

(ii) The result follows from $\mathcal{A}_2^* = C(g_3^* - g_2^*) > \mathcal{A}_1^* = C(g_2^* - g_1)$ and the result above that $\mathcal{A}_2^* > \mathcal{A}_1^* > 0$, given that $C(.)$ is strictly increasing.

(iii) Suppose not so that $\tilde{g}_2 \geq g_3^*$, where $\tilde{g}_2$ is the maximal implementable g in the two-period setup. Now consider the following three-period program: $(\mathcal{A}_1 = 0, g_2 = g_1; \mathcal{A}_2 = \tilde{\mathcal{A}}_1, g_3 = \tilde{g}_2; \mathcal{A}_3 = \tilde{\mathcal{A}}_2)$, where $(\tilde{\mathcal{A}}_1, \tilde{g}_2; \tilde{\mathcal{A}}_2)$ is the optimal two-period program. Given the initial governance level g_1 and first-period aid and target governance $(\mathcal{A}_1 = 0, g_2 = g_1)$, clearly $(\mathcal{A}_2 = \tilde{\mathcal{A}}_1, g_3 = \tilde{g}_2; \mathcal{A}_3 = \tilde{\mathcal{A}}_2)$ is time consistent in the subgame starting in period 2. Stepping back to period 1 it is easy to see that the three-period program $(\mathcal{A}_1 = 0, g_2 = g_1; \mathcal{A}_2 = \tilde{\mathcal{A}}_1, g_3 = \tilde{g}_2; \mathcal{A}_3 = \tilde{\mathcal{A}}_2)$ is feasible as per Definition 2. But then it cannot be optimal in the three-period program as it fails to satisfy the optimality requirement $g_3^* > g_2^* > g_1$ as shown in part **(i)**. This contradicts our hypothesis that $g_3^* \leq \tilde{g}_2$. Q.E.D.

Proof of Proposition 9. **(i)** The proof follows in several steps.

Step 1. Recall that $\mathcal{A}_{23}(\Upsilon) \equiv \mathcal{A} + \Upsilon - \breve{\mathcal{A}}_1$. With $\breve{\mathcal{A}}_1$ bounded above by $\mathcal{A}$, we then have $\mathcal{A}_{23}(\Upsilon) \to \infty$ as $\Upsilon \to \infty$.

Step 2. Rewrite (23) as

$$\mathcal{A}_{23}(\Upsilon) = \lambda C(\tilde{g}_3 - \tilde{g}_2)\left[1 + \frac{\alpha(\tilde{g}_2)}{\alpha(\tilde{g}_3)}\right].$$

Recall that from step 1, the LHS goes to infinity, as Υ goes to infinity. Hence, to maintain equality in the preceding equation, the RHS must go to infinity, which in turn implies that $\tilde{g}_3 \to \infty$.

Step 3. Since from step 2 we know that $\tilde{g}_3$ goes to infinity as Υ goes to infinity, then from (23) the RHS of (23) $\to \infty$, because $\tilde{g}_2$ is bounded above (follows from the second equation in (17) and the fact that $\breve{\mathcal{A}}_1 < \mathcal{A}$).[40] This implies the LHS of (23), i.e. $\alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3$ goes to infinity.

Step 4. Note that steps 1-3 establish that $\alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3 \to \infty$ as $\Upsilon \to \infty$, which can then be applied to (21), to establish our claim.

(ii) In the continuation game after first-period separation, for each type $\tau = 1, \lambda$, both the IC and FC constraints must bind as in (4) and (5):

[40] $\breve{\mathcal{A}}_1$ is bounded above by the necessity of having to keep the inefficient type in play.

For $\tau = 1$,

$$\frac{\alpha(\breve{g}_3)\breve{\mathcal{A}}_3}{\alpha(\breve{g}_2)} = C(\breve{g}_3 - \breve{g}_2), \quad \breve{\mathcal{A}}_2 = C(\breve{g}_3 - \breve{g}_2);$$

for $\tau = \lambda$,

$$\frac{\alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3}{\alpha(\tilde{g}_2)} = \lambda C(\tilde{g}_3 - \tilde{g}_2), \quad \tilde{\mathcal{A}}_2 = \lambda C(\tilde{g}_3 - \tilde{g}_2).$$

We can now write:

$$\begin{aligned} \lambda C(\tilde{g}_3 - \tilde{g}_2)\left[1 + \frac{\alpha(\tilde{g}_2)}{\alpha(\tilde{g}_3)}\right] = \tilde{\mathcal{A}}_2 + \tilde{\mathcal{A}}_3 &= \mathcal{A} + \Upsilon \\ &> \mathcal{A} = \breve{\mathcal{A}}_2 + \breve{\mathcal{A}}_3 = C(\breve{g}_3 - \breve{g}_2)\left[1 + \frac{\alpha(\breve{g}_2)}{\alpha(\breve{g}_3)}\right]. \end{aligned} \tag{A.1}$$

By design, $\tilde{g}_2 > \breve{g}_2$, hence $\alpha(\tilde{g}_2) < \alpha(\breve{g}_2)$. Suppose now, contrary to our claim, $\tilde{g}_3 \leq \breve{g}_3$. Then

$$\alpha(\tilde{g}_3) \geq \alpha(\breve{g}_3),$$

so that

$$\frac{\alpha(\tilde{g}_2)}{\alpha(\tilde{g}_3)} < \frac{\alpha(\breve{g}_2)}{\alpha(\breve{g}_3)}.$$

This would lead to a contradiction of (A.1), since $\lambda < 1$ and our contrarian hypothesis implies $\tilde{g}_3 - \tilde{g}_2 < \breve{g}_3 - \breve{g}_2$.

(iii) Start from a separating incentive scheme as proposed, with the incremental aid Υ. Consider now a higher incremental aid Υ'. Maintaining initial aid at $\breve{\mathcal{A}}_1$, initial target governance at $\breve{g}_2$ and $\tilde{g}_2$, and second-period aid at $\breve{\mathcal{A}}_2$ and $\tilde{\mathcal{A}}_2$, let period 3 aid for the efficient type be increased to $\tilde{\mathcal{A}}_3 + [\Upsilon' - \Upsilon]$.

Since the incentives for the inefficient type has remained unchanged, so long as the modified incentives separate the two types in period 1, the inefficient type's incentives will be time consistent in the remaining two periods binding IC and FC constraints. We thus only need to ensure that the λ-type's period 1 IC will be satisfied and fulfill the time consistency requirement in the remaining periods.

Going back to (21) which must be satisfied as our starting incentive scheme is separating, with the additional aid $(\Upsilon' - \Upsilon)$ clearly the following will hold:

$$\alpha(\tilde{g}_3)[\tilde{\mathcal{A}}_3 + (\Upsilon' - \Upsilon)] - \alpha(\breve{g}_3)\breve{\mathcal{A}}_3 > (1-\lambda)[C(\breve{g}_2 - g_1) + C(\breve{g}_3 - \breve{g}_2)]. \tag{A.2}$$

Further, period 2 IC and FC that were previously binding, as in (22) and (23), can now be modified as:

$$\alpha(\tilde{g}_2)[\tilde{\mathcal{A}}_2 - \lambda C(\tilde{g}_3 - \tilde{g}_2)] + \alpha(\tilde{g}_3)[\tilde{\mathcal{A}}_3 + (\Upsilon' - \Upsilon)] > \alpha(\tilde{g}_2)\tilde{\mathcal{A}}_2, \tag{A.3}$$

$$\tilde{\mathcal{A}}_2 = \lambda C(\tilde{g}_3 - \tilde{g}_2). \tag{A.4}$$

Let us define

$$L(x) = \alpha(\tilde{g}_3(x))[\tilde{\mathcal{A}}_3 + (\Upsilon' - \Upsilon) - x] - \alpha(\tilde{g}_2)[\tilde{\mathcal{A}}_2 + x],$$

where $x \in [0, \Upsilon' - \Upsilon]$ and $\tilde{g}_3(x)$ solves

$$\tilde{\mathcal{A}}_2 + x = \lambda C(\tilde{g}_3(x) - \tilde{g}_2). \tag{A.5}$$

Note that both $\tilde{g}_3(x)$ and $L(x)$ are continuous in x. Further, since $L(0) > 0$ by (A.3) and $L(\Upsilon' - \Upsilon) < 0$ (follows from (22), because $\tilde{g}_3(\Upsilon' - \Upsilon) > \tilde{g}_3 \equiv \tilde{g}_3(0)$ and $\alpha(.)$ is strictly decreasing), by the intermediate value theorem there exists some $x^* \in (0, \Upsilon' - \Upsilon)$ such that $L(x^*) = 0$, i.e.,

$$\alpha(\tilde{g}_3(x^*))[\tilde{\mathcal{A}}_3 + (\Upsilon' - \Upsilon) - x^*] = \alpha(\tilde{g}_2)[\tilde{\mathcal{A}}_2 + x^*]. \tag{A.6}$$

It now remains to verify that modifying the efficient type's incentives in the last two periods to

$$(\tilde{\mathcal{A}}_2 + x^*,\ \tilde{g}_3(x^*);\ \ \tilde{\mathcal{A}}_3 + (\Upsilon' - \Upsilon) - x^*),$$

while leaving the rest of the original incentives unchanged, will satisfy the efficient type's IC constraint in period 1. (Note that, by construction, the modified incentives already <u>bind</u> the efficient type's IC and FC constraints in period 2 as per (A.6) and (A.5), and thus satisfy the time consistency requirement.)

Using (A.6), write:

$$\alpha(\tilde{g}_3(x^*))[\tilde{\mathcal{A}}_3 + (\Upsilon' - \Upsilon) - x^*] > \alpha(\tilde{g}_2)\tilde{\mathcal{A}}_2,$$

which combined with (22) yields

$$\alpha(\tilde{g}_3(x^*))[\tilde{\mathcal{A}}_3 + (\Upsilon' - \Upsilon) - x^*] > \alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3.$$

Now use (21) to conclude that

$$\alpha(\tilde{g}_3(x^*))[\tilde{\mathcal{A}}_3 + (\Upsilon' - \Upsilon) - x^*] - \alpha(\breve{g}_3)\breve{\mathcal{A}}_3 > (1-\lambda)[C(\breve{g}_2 - g_1) + C(\breve{g}_3 - \breve{g}_2)],$$

thus satisfying period 1 IC constraint.

Finally, by construction,

$$\tilde{g}_3(x^*) > \tilde{g}_3. \qquad \text{Q.E.D.}$$

Proof of Proposition 10. Consider the continuation aid program from the second period onwards. For any λ', where $\lambda' \in \{\lambda, 1\}$, let $(\mathcal{A}_2(\lambda'), \mathcal{A}_3(\lambda'), g_3(\lambda'))$ solve the IC at $t = 2$, as well as the FC at $t = 2$ with both binding, for the case when $\Upsilon = 0$. Using (4) and (5) with the obvious modifications, we therefore have:

$$\alpha(g_3(\lambda'))\mathcal{A}_3(\lambda') = \alpha(g_2)\mathcal{A}_2(\lambda'), \quad \lambda' \in \{\lambda, 1\}. \tag{A.7}$$

It is straightforward to argue that $g_3(\lambda) > g_3(1)$, i.e. the governance level is higher if the recipient is more efficient. Hence $\alpha(g_3(\lambda)) < \alpha(g_3(1))$, since $\alpha(.)$ is a decreasing function. Using this fact, (A.7), and the fact that the starting aid amount at $t = 2$ is the same for both types of recipient, it follows that $\mathcal{A}_2(\lambda) < \mathcal{A}_2(1)$. Using this, and (A.7), we then have

$$\alpha(g_3(\lambda))\mathcal{A}_3(\lambda) = \alpha(g_2)\mathcal{A}_2(\lambda) < \alpha(g_2)\mathcal{A}_2(1) = \alpha(g_3(1))\mathcal{A}_3(1). \tag{A.8}$$

Note that $\lim_{\Upsilon \to 0}(\tilde{\mathcal{A}}_2(\Upsilon), \tilde{\mathcal{A}}_3(\Upsilon)) = (\mathcal{A}_2(\lambda), \mathcal{A}_3(\lambda))$. From (A.8), this implies that as Υ becomes very small, the LHS of (21) becomes negative while the RHS is bounded away from zero (since it is independent of Υ). Hence (21) cannot be satisfied for Υ small. Q.E.D.

References

Adam, C.S. and S.A. O'Connell, "Aid, taxation and development in Sub-Saharan Africa," *Economics & Politics*, 1999, *11*, 225–253.

Alesina, A. and B. Weder, "Do corrupt governments receive less foreign aid?," *American Economic Review*, 2002, *92*, 1126–1137.

Azam, J-P. and J-J. Laffont, "Contracting for aid," *Journal of Development Economics*, 2003, *70*, 25–58.

Bakoup, F., "Promoting economic reform in developing countries: Rethinking budgetary aid?," Working paper no. 167, African Development Bank Group (http://www.afdb.org/en/news-and-events/article/country-economic-thematic-review-meeting-promoting-economic-reforms-in-developing-countries-rethinking-budgetary-aid-11554/), 2013.

Banerjee, A.V., "Making aid work: How to fight global poverty effectively," *in Making Aid Work,* Boston Review/MIT Press (also available at http://iis-db.stanford.edu/evnts/4741/MakingAidWork.pdf), 2007.

— **and R. He**, "Making aid work," *in W. Easterly (Ed.), Reinventing foreign aid,* MIT Press, Cambridge, Massachusetts, 2008, pp. 44–92.

Barro, R. and D.B. Gordon, "Rules, discretion and reputation in a model of monetary policy," *Journal of Monetary Economics*, 1983, *12*, 101–121.

Bates, R.H., *Markets and states in tropical Africa: The political basis of agricultural policies*, University of California Press, Berkeley, 1981.

— , *Governments and agricultural markets in Africa*, in D.G. Johnson and G.E. Schuh, eds. The Role of Markets in the World Food Economy, West View Press, Boulder, CO., 1983.

Bond, E.W. and J.H. Park, "Gradualism in trade agreements with asymmetric countries," *Review of Economic Studies*, 2002, *59*, 367–397.

Boone, P., "Politics and the effectiveness of foreign aid," *European Economic Review*, 1996, *40*, 289–329.

Burnside, C. and D. Dollar, "Aid, policies and growth," *American Economic Review*, 2000, *90*, 847–868.

— **and** — , "Aid, policies, and growth: Reply," *American Economic Review*, 2004, *94*, 781–784.

Carter, P., F. Postel-Vinay, and J. Temple, "Dynamic aid allocation," *Journal of International Economics*, 2015, *95*, 291–304.

Chisik, R., "Gradualism in free trade agreements: a theoretical justification," *Journal of International Economics*, 2003, *59*, 367–397.

Collier, P., "Africa's external economic relations 1960-1990," *African Affairs*, 1990, pp. 339–356.

Conway, P., "Endogenous IMF conditionality: Theoretical and empirical implications," Department of Economics, University of North Carolina, Working paper, (http://www.unc.edu/~pconway/dload/ec_conway.pdf), 2003.

Cordella, T., G. Dell'Ariccia, and K.M. Kletzer, "Conditional aid, sovereign debt, and debt relief," Mimeo, UCL, Santa Cruz 2003.

Dollar, D. and J. Svensson, "What explains the success or failure of structural adjustment programs?," Policy Research Working Paper, no. 1938, 1998.

— **and V. Levin**, "The increasing selectivity of foreign aid, 1984-2002," The World Bank Policy Research Working Paper No. 3299, (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=610344), 2004.

Drazen, A., *Political economy in macroeconomics*, Princeton University Press, 2000.

Easterly, W., "Can foreign aid buy growth," *Journal of Economic Perspectives*, 2003, *17*, 23–48.

— **, R. Levine, and D. Roodman**, "New data, new doubts: A comment on Burnside and Dollar's 'Aid, policies, and growth' (2000)," *American Economic Review*, 2004, *94*, 774–780.

EC, "Communication from the commission to the European parliament, the council, the European economic and social committee of the regions: The future approach to the EU budget support to third countries. Brussels, 13.10.2011 COM(2011) 638 final," European Commission 2011.

Federico, G., "Samaritans, rotten kids and policy conditionality," Mimeo, Nuffeld College, Cambridge 2001.

Guitian, M., "Conditionality: Past, present, future," *Staff Papers: International Monetary Fund*, 1995, *42*, 792–835.

Kanbur, R., "Aid conditionality and debt in Africa," *in* Finn Tarp (Ed.), *Foreign aid and development: Lessons learnt and directions for the future,* Routledge Press, 2000.

Kapur, D. and R. Webb, "Governance-related conditionalities of the international financial institutions," New York and Geneva: United Nations Conference on Trade and Development, G-24 Discussion Paper Series 6, (https://ideas.repec.org/p/unc/g24pap/6.html), 2000.

Kaufmann, D., A. Kraay, and P. Zoido-Lobaton, "Governance matters," Policy Research Working Paper 2196, The World Bank, (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=188568), 1999.

Kletzer, K.M., "Aid and sanctions," Working Paper Series, Department of Economics, UCL, SC 2005.

Klimenko, M., G. Ramey, and J. Watson, "Recurrent trade agreements and the value of external enforcement," *Journal of International Economics*, 2008, *74*, 475–489.

Koberle, S., "Should policy based lending still involve conditionality?," *The World Bank Research Observer*, 2003, *18*, 249–273.

Lockwood, B. and J.P. Thomas, "Gradualism and irreversibility," *Review of Economic Studies*, 2002, *69*, 339–356.

Marx, L. and S. Matthews, "Dynamic voluntary contributions to a public project," *Review of Economic Studies*, 1998, *67*, 327–358.

McGuire, M.C. and M. Jr. Olson, "The economics of autocracy and majority rule: The invisible hand and the use of force," *Journal of Economic Literature*, 1996, *34*, 72–96.

Morrissey, O., "Political economy dimensions of economic policy reforms," *in* M. McGillivray and O. Morissey (eds.), *Evaluating Economic Liberalisation,* London: McMillan, 1999.

_ , "Implementation and outcomes of African trade policy reforms in the 1990s," *in* D. Belshaw and I. Livingstone eds., *Renewing Development in Sub-Saharan Africa: Policy, Performance and Prospects,* London: Routledge, 2001, pp. 339–353.

_ , "Alternatives to conditionality in policy based lending," Mimeo, Development Policy Forum, Paris 2004a.

_ , "Conditionality and aid effectiveness re-evaluated," *World Economy*, 2004b, *27*, 153–171.

Mosley, P., "A theory of conditionality," *in* P. Mosley, ed., *Development Finance and Policy Reform: Essays in the Theory and Practice of Conditionality in Less Developed Countries,* New York: St. Martin's Press, 1992.

_ , J. Harrigan, and J. Toye, *Aid and Power*, Vol. 1, London: Routledge, 1995.

Murshed, M. and S. Sen, "Aid conditionality and military expenditure reduction in developing countries: Models of asymmetric information," *The Economic Journal*, 1995, *105*, 498–505.

Neher, D.V., "Staged financing: An agency perspective," *Review of Economic Studies*, 1999, *66*, 255–274.

OED, "2003 Annual review of development effectiveness: The effectiveness of Bank support for policy reform," *in* Washington DC: World Bank Operations Effectiveness Department, 2004.

Rodrik, D., "Promises, promises: Credible policy reform via signaling," *The Economic Journal*, 1989, *99*, 756–772.

Sandbrook, R., "The state and economic stagnation in tropical Africa," *World Development*, 1986, *14*, 319–332.

Santiso, C., "Good governance and aid effectiveness: The World Bank and conditionality," *The Georgetown Public Policy Review*, 2001a, *7*, 1–22.

_ , "International co-operation for democracy and good governance: Moving toward a second generation?," *European Journal of Development Research*, 2001b, *13*, 154–180.

Scholl, A., "Aid effectiveness and limited enforceable conditionality," *Review of Economic Dynamics*, 2009, *12*, 377–391.

Staiger, R.W., "A theory of gradual trade liberalization," *in* J. Levinsohn, A.V. Deardorff and R.M. Stern (Eds.), *New Directions in Trade Theory,* University of Michigan Press Ann Arbor, MI, 1995, pp. 249–284.

_ **and G. Tabellini**, "Discretionary trade policy and excessive protection," *American Economic Review*, 1987, *77*, 823–837.

Svensson, J., "Reforming donor institutions: Aid tournaments," World Bank Working Paper, 1998.

_ , "Aid, growth, and democracy," *Economics and Politics*, 1999, *11*, 275–297.

_ , "Foreign aid and rent seeking," *Journal of International Economics*, 2000a, *51*, 437–461.

_ , "When is foreign aid policy credible? Aid dependence and conditionality," *Journal of Development Economics*, 2000b, *61*, 61–84.

_ , "Why conditional aid does not work and what can be done about it?," *Journal of Development Economics*, 2003, *70*, 381–402.

Temple, J.R.W., "Aid and conditionality," *in* D. Rodrick and M. Rosenzweig (Eds.), *Handbook of Development Economics*, 2010, *5*, 4415–4518.

Tornell, A. and P.R. Lane, “The voracity effect,” *American Economic Review*, 1999, *89*, 22–46.

Watson, J., “Starting small and renegotiation,” *Journal of Economic Theory*, 1999, *85*, 52–90.

_ , “Starting small and commitment,” *Games and Economic Behavior*, 2002, *38*, 176–199.

B Supplementary material (not for publication)

Proof of Lemma 1. Suppose that contrary to the claim in part (**ii**) of this lemma, $\tilde{g}_2 = g_1$. We shall argue that in that case it is possible to construct a feasible aid–governance scheme with a higher level of governance. Fixing $\mathcal{A} > \epsilon_1 > 0$, $\epsilon_2 > 0$, by the continuity of $C(g_2 - g_1)$ and $\alpha(g_2) - \alpha(g_1)$ we can choose $g_2' - g_1 = \eta > 0$ appropriately small such that $0 < C(\eta) < \epsilon_1$ and $\alpha(g_1) - \alpha(g_2') < \epsilon_2$.

Now letting $\mathcal{A}_1' = \epsilon_1$, $\mathcal{A}_2' = \mathcal{A} - \epsilon_1$, we can write

$$\Delta \cdot \frac{\alpha(g_2')}{\alpha(g_1)} \frac{\mathcal{A} - \epsilon_1}{\epsilon_1} \quad > \quad \Delta \left[\frac{\alpha(g_1) - \epsilon_2}{\alpha(g_1)}\right] \left[\frac{\mathcal{A} - \epsilon_1}{\epsilon_1}\right] = \Delta \left[1 - \frac{\epsilon_2}{\alpha(g_1)}\right] \left[\frac{\mathcal{A} - \epsilon_1}{\epsilon_1}\right] > 1,$$

where the last inequality holds whenever ϵ_1 and ϵ_2 are sufficiently small.

Hence, $\Delta \cdot \alpha(g_2')(\mathcal{A} - \epsilon_1) > \alpha(g_1)\epsilon_1 > \alpha(g_1)C(g_2' - g_1)$, satisfying the incentive compatibility condition (2). Finally, by construction, $\mathcal{A}_1' > C(\eta) = C(g_2' - g_1)$, thus satisfying the financial constraint (3). Thus, we have constructed a feasible aid package $(\mathcal{A}_1', \mathcal{A}_2')$ leading to $g_2' = g_1 + \eta$ that dominates the initial level of governance g_1. A contradiction.

Next given that $\tilde{g}_2 > g_1$, from (2) we have $\tilde{\mathcal{A}}_2 > 0$, and from (3) we have $\tilde{\mathcal{A}}_1 > 0$ (using the fact that $C(0) = 0$ and $C(.)$ is strictly increasing). Q.E.D.

Proof of Proposition 4. **(i)** Suppose period 1 financial constraint does not bind, so (10) holds with strict inequality. Then $\mathcal{A}_1$ can be reduced slightly, and $\mathcal{A}_2$ and $\mathcal{A}_3$ can be increased slightly so that $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ still sum up to A. This relaxes (7), (8) and (9), so that g_3 can be increased slightly, given that $\alpha(g)$ and $C(.)$ are continuous functions. This yields a contradiction.

Suppose period 2 incentive constraint is not binding, so (7) holds with strict inequality. Now take away some $\zeta > 0$ from $\hat{\mathcal{A}}_3$ such that (7) continues to hold, and add this ζ to $\hat{\mathcal{A}}_2$ such that both (8) and (9) hold with strict inequalities (strict inequality in (9) follows, since $\hat{g}_3 > \hat{g}_2$, which in turn follows from Proposition 3, implies $\zeta\alpha(\hat{g}_2) > \zeta\alpha(\hat{g}_3)$). Next take away some $\epsilon > 0$ from $\hat{\mathcal{A}}_2 + \zeta$ (where $\epsilon < \zeta$) and add this to $\hat{\mathcal{A}}_1$ such that conditions (8) and (9) continue to be satisfied (given that both were satisfied with strict inequalities due to ζ transfer), and condition (10) now holds with strict inequality. As a result, the solutions $(\hat{g}_3, \hat{g}_2)$ are unaffected and yet period 1 financial constraint is *not* binding, contradicting the fact that at the optimal outcome (10) binds (as shown above).

(ii) Suppose not, so that both (8) and (9) hold with strict inequality. Then take away a small $\eta > 0$ from $\hat{\mathcal{A}}_2$ and add it to $\hat{\mathcal{A}}_1$ such that (8) and (9) continue to be satisfied and condition (10) holds with strict inequality. Again, the first part of (**i**) is contradicted.

(iii) Suppose to the contrary that (9) does not bind. Then, given **(i)** and **(ii)** in this Proposition, the donor's problem can be written as:

$$\begin{aligned}
\max\quad g_3, &\quad \text{subject to} \\
\delta\mathcal{A}_3 &= \frac{\alpha(g_2)}{\alpha(g_3)}C(g_3-g_2), \qquad (7.3)\\
\mathcal{A}_2 &= C(g_3-g_2), \qquad (8.3)\\
0 &< -\alpha(g_1)C(g_2-g_1)+\delta\alpha(g_2)[\mathcal{A}_2-C(g_3-g_2)]+\delta^2\alpha(g_3)\mathcal{A}_3, \quad (9.3)\\
\mathcal{A}_1 &= C(g_2-g_1). \qquad (10.3)
\end{aligned}$$

Having solved the above problem, below we will return to our (hat) notations for the optimal incentives under the FCC program.

Let us increase g_2 to $\hat{g}_2+\epsilon$, $\epsilon>0$ small. Then define

$$Z(\epsilon)=\frac{\alpha(\hat{g}_2+\epsilon)}{\alpha(\hat{g}_3)}C(\hat{g}_3-\hat{g}_2-\epsilon)+C(\hat{g}_2+\epsilon-g_1).$$

Thus $Z(.)$ is the sum of the RHS of (7.3) and (10.3), taking the increase in g_2 into account. Next observe that

$$Z'(0)\equiv Z'(\epsilon)|_{\epsilon=0}=-\frac{\alpha(\hat{g}_2)}{\alpha(\hat{g}_3)}C'(\hat{g}_3-\hat{g}_2)+C'(\hat{g}_2-g_1)+\frac{\alpha'(\hat{g}_2)}{\alpha(\hat{g}_3)}C(\hat{g}_3-\hat{g}_2).$$

Given that $C(.)$ is strictly increasing and weakly convex, $\alpha(g)$ is strictly decreasing (Assumption 1), $\hat{g}_3\geq\hat{g}_2$, and $\hat{g}_3-\hat{g}_2\geq\hat{g}_2-g_1$ (by hypothesis), it follows that

$$Z'(0)<0. \tag{B.1}$$

Next note that for ϵ sufficiently small, conditions (7), (8), and (9) will all hold with strict inequality; however, (10) will be violated.[41] We next argue that $\mathcal{A}_1,\mathcal{A}_2,\mathcal{A}_3$ can be adjusted in such a way that all the required weak inequalities will hold (including equalities).

Let X',X,Y respectively denote the absolute values of the changes in the RHS of (7.3), (8.3), and (10.3) respectively as a result of increasing g_2 to $\hat{g}_2+\epsilon$. Given (B.1), for ϵ small,

$$X'>Y.$$

[41]Here while we refer to the original feasibility constraints in section 4, the reader may as well think in terms of (7.3)–(10.3), for ease of visualization.

Next consider $\mathcal{A}_1', \mathcal{A}_2', \mathcal{A}_3'$ such that

$$\mathcal{A}_1' = \hat{\mathcal{A}}_1 + Y, \quad \mathcal{A}_2' = \hat{\mathcal{A}}_2, \text{ and } \quad \mathcal{A}_3' = \hat{\mathcal{A}}_3 - Y.$$

Observe that with the above reallocation of the aid budget (from period 3 to period 1), all the inequalities in (7), (8), (9) will continue to hold strictly (as after the g_2 increase), and now constraint (10) binds. We only need to explain how (9) will hold strictly as the rest are straightforward. The overall change in the RHS of (9.3) is

$$\begin{aligned} = \quad & -\alpha(g_1)[C(\hat{g}_2 + \epsilon - g_1) - C(\hat{g}_2 - g_1)] + \delta\alpha(\hat{g}_2 + \epsilon)[\hat{\mathcal{A}}_2 - C(\hat{g}_3 - \hat{g}_2 - \epsilon)] \\ & -\delta\alpha(\hat{g}_2)[\hat{\mathcal{A}}_2 - C(\hat{g}_3 - \hat{g}_2)] + \delta^2\alpha(\hat{g}_3)[\mathcal{A}_3' - \hat{\mathcal{A}}_3], \end{aligned}$$

which can be kept suitably small due to continuity of the expression in ϵ, to preserve the strict inequality of (9.3). Thus, g_3 can be increased by a small enough amount to $\hat{g}_3 + \gamma(\epsilon)$ (i.e., $\gamma(\epsilon) > 0$) so that all the strict inequalities in (7), (8), (9) continue to hold and (10) binds, which would be a contradiction. Q.E.D.

Proof of Lemma A.1. (i) That both (7) and (8) must bind follows from sequential rationality and Proposition 1.

Suppose (10) does not bind, so (10) holds with strict inequality. Then $\mathcal{A}_1$ can be reduced slightly, and $\mathcal{A}_2$ and $\mathcal{A}_3$ can be increased slightly. We then argue that one can increase g_3 in a manner that (7) and (8) binds (so that sequential rationality is satisfied), and the other constraints hold, which would be a contradiction.

Now let us increase g_3 by $\epsilon > 0$ but small. Then the corresponding change in $\mathcal{A}_3$ so as to keep (7) binding is

$$\delta\alpha(g_3^*)\, \mathrm{d}\mathcal{A}_3 = \alpha(g_2^*)C'(g_3^* - g_2^*)\, \mathrm{d}g_3 - \delta\alpha'(g_3^*)\mathcal{A}_3^*\, \mathrm{d}g_3.$$

Similarly, the change in $\mathcal{A}_2$ so as to keep (8) binding is

$$\mathrm{d}\mathcal{A}_2 = C'(g_3^* - g_2^*)\, \mathrm{d}g_3.$$

Substituting the above two in (9) the LHS becomes

$$\begin{aligned} &= \frac{\delta\alpha(g_2^*)}{\alpha(g_1)}\left[C'(g_3^* - g_2^*)\, \mathrm{d}g_3 - C'(g_3^* - g_2^*)\, \mathrm{d}g_3\right] + \frac{\delta^2}{\alpha(g_1)}\left[\alpha(g_3^*)\, \mathrm{d}\mathcal{A}_3 + \alpha'(g_3^*)\mathcal{A}_3^*\, \mathrm{d}g_3\right] \\ &= \frac{\delta\alpha(g_2^*)}{\alpha(g_1)}C'(g_3^* - g_2^*)\, \mathrm{d}g_3 > 0, \end{aligned}$$

thus satisfying the constraint. The contradiction is finally established.

(ii) Part **(i)** together with the IC condition (9) imply that $\alpha(g_1)\mathcal{A}_1^* \leq \delta^2\alpha(g_3^*)\mathcal{A}_3^* = \delta\alpha(g_2^*)\mathcal{A}_2^*$. The ranking of $\mathcal{A}_i^*$'s follows from the fact that $\alpha(g_1) \geq \alpha(g_2^*) \geq \alpha(g_3^*)$ and $1 \geq \delta \geq 0$.

(iii) Follows from the previous steps, applying that $C(.)$ is strictly increasing: $C(g_3^* - g_2^*) = \mathcal{A}_2^* \geq \mathcal{A}_1^* = C(g_2^* - g_1)$.

(iv) Suppose to the contrary that (9) does not bind. Then, given part **(i)**, the donor's problem can be written as:

$$
\begin{aligned}
&\max \quad g_3, \quad \text{subject to} \\
\delta\mathcal{A}_3 \quad &= \quad \frac{\alpha(g_2)}{\alpha(g_3)}C(g_3 - g_2), \qquad (7.3) \\
\mathcal{A}_2 \quad &= \quad C(g_3 - g_2), \qquad (8.3) \\
\alpha(g_1)\mathcal{A}_1 \quad &< \quad \alpha(g_1)[\mathcal{A}_1 - C(g_2 - g_1)] + \delta\alpha(g_2)[\mathcal{A}_2 - C(g_3 - g_2)] + \delta^2\alpha(g_3)\mathcal{A}_3, \quad (9.3) \\
\mathcal{A}_1 \quad &= \quad C(g_2 - g_1). \qquad (10.3)
\end{aligned}
$$

Having solved the above problem, below we will return to our (star) notations for the optimal incentives under the PCC program.

Let us increase g_2 to $g_2^* + \epsilon$, $\epsilon > 0$ small. Then define

$$Z(\epsilon) = \frac{\alpha(g_2^* + \epsilon)}{\alpha(g_3^*)}C(g_3^* - g_2^* - \epsilon) + C(g_2^* + \epsilon - g_1).$$

Thus $Z(.)$ is the sum of the RHS of (7.3) and (10.3), taking the increase in g_2 into account. Next observe that

$$Z'(0) \equiv Z'(\epsilon)|_{\epsilon=0} = -\frac{\alpha(g_2^*)}{\alpha(g_3^*)}C'(g_3^* - g_2^*) + \frac{\alpha'(g_2^*)}{\alpha(g_3^*)}C(g_3^* - g_2^*) + C'(g_2^* - g_1).$$

Given that $C(.)$ is strictly increasing and weakly convex, $\alpha(g)$ is strictly decreasing (Assumption 1), $g_3^* \geq g_2^*$, and $g_3^* - g_2^* \geq g_2^* - g_1$ (by part **(iii)** above), it follows that

$$Z'(0) < 0. \tag{B.2}$$

Next note that for ϵ sufficiently small, conditions (7), (8), and (9) will all hold with strict inequality; however, (10) will be violated.[42] We next argue that $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ can be adjusted

[42]Here while we refer to the original feasibility constraints in section 4, the reader may as well think in terms of (7.3)–(10.3), for ease of visualization.

in such a way that all the required weak inequalities will hold (including equalities).

Let X', X, Y respectively denote the absolute values of the changes in the RHS of (7.3), (8.3), and (10.3) respectively. Given (B.2), for ϵ small,

$$X' > Y.$$

Next consider $\mathcal{A}_1', \mathcal{A}_2', \mathcal{A}_3'$ such that

$$\mathcal{A}_1' = \mathcal{A}_1^* + Y, \quad \mathcal{A}_2' = \mathcal{A}_2^*, \quad \text{and} \quad \mathcal{A}_3' = \mathcal{A}_3^* - Y.$$

Observe that with the above reallocation of the aid budget (from period 3 to period 1), all the inequalities in (7), (8), (9) will continue to hold strictly (as after the g_2 increase), and now constraint (10) binds.[43] We only need to explain how (9) will hold strictly as the rest are straightforward. The overall change in the LHS of (9.3) is

$$\begin{aligned} = \quad & -\alpha(g_1)[C(g_2^* + \epsilon - g_1) - C(g_2^* - g_1)] + \delta\alpha(g_2^* + \epsilon)[\mathcal{A}_2^* - C(g_3^* - g_2^* - \epsilon)] \\ & -\delta\alpha(g_2^*)[\mathcal{A}_2^* - C(g_3^* - g_2^*)] - \delta^2\alpha(g_3^*)[C(g_2^* + \epsilon - g_1) - C(g_2^* - g_1)], \end{aligned}$$

which can be kept suitably small due to continuity of the expression in ϵ, to preserve the strict inequality of (9.3). Thus, g_3 can be increased by a small enough amount to $g_3^* + \gamma(\epsilon)$ (i.e., $\gamma(\epsilon) > 0$) so that all the strict inequalities in (7), (8), (9) continue to hold and (10) binds, which would be a contradiction.

Given that we require sequential rationality, we need to be slightly more careful as adjusted $\mathcal{A}_2'$ and $\mathcal{A}_3'$ above do not bind (7) and (8).

First note that $\lim_{\epsilon \to 0} \gamma(\epsilon) = 0$. Starting from the PCC deviation program $(\mathcal{A}_1', \mathcal{A}_2', \mathcal{A}_3', g_2^* + \epsilon, g_3^* + \gamma(\epsilon))$, we will now construct another program with (7) and (8) holding with equality and a final governance level higher than g_3^*.

For any $\epsilon > 0$, consider $\rho(\epsilon) > 0$ and $\beta(\epsilon) > 0$ such that (7) and (8) are satisfied with equality:

$$\begin{aligned} \delta\frac{\alpha(g_3^* + \gamma(\epsilon))}{\alpha(g_2^* + \epsilon)}(\mathcal{A}_3' - \rho(\epsilon)) &= C(g_3^* + \gamma(\epsilon) - g_2^* - \epsilon), \\ \mathcal{A}_2' - \beta(\epsilon) &= C(g_3^* + \gamma(\epsilon) - g_2^* - \epsilon). \end{aligned}$$

Since

$$\lim_{\epsilon \to 0} \beta(\epsilon) = 0 = \lim_{\epsilon \to 0} \rho(\epsilon)$$

[43] Verification that (10) binds would be as follows: $\mathcal{A}_1' = \mathcal{A}_1^* + Y = \mathcal{A}_1^* + C(g_2^* + \epsilon - g_1) - C(g_2^* - g_1) = C(g_2^* + \epsilon - g_1)$.

(because in the limit we should get back (7.3) and (8.3), along with $\lim_{\epsilon \to 0} \gamma(\epsilon) = 0$), for some $\bar{\epsilon} > 0$ small enough the following must be true

$$\begin{aligned}&\delta\alpha(g_2^* + \bar{\epsilon})(\mathcal{A}_2' - \beta(\bar{\epsilon}) - C(g_3^* + \gamma(\bar{\epsilon}) - g_2^* - \bar{\epsilon}) + \delta^2\alpha(g_3^* + \gamma(\bar{\epsilon}))(\mathcal{A}_3' - \rho(\bar{\epsilon}))\\ &\geq \alpha(g_1)C(g_2^* + \bar{\epsilon} - g_1),\end{aligned}$$

because otherwise we will have contradicted (9) using continuity of both sides of the above inequality (in ϵ).

This last inequality (satisfying (9)), for the revised program

$$(\mathcal{A}_1' + \rho(\bar{\epsilon}) + \beta(\bar{\epsilon}),\ \mathcal{A}_2' - \beta(\bar{\epsilon}),\ \mathcal{A}_3' - \rho(\bar{\epsilon}),\ g_2^* + \bar{\epsilon}, g_3^* + \gamma(\bar{\epsilon})),$$

establishes that the overall governance level is improved to $g_3^* + \gamma(\bar{\epsilon})$ while satisfying time consistency – a contradiction. Note that while arriving at the contraction constraint (10) remains slack: $\mathcal{A}_1' + \rho(\bar{\epsilon}) + \beta(\bar{\epsilon}) > C(g_2^* + \bar{\epsilon} - g_1)$; time consistency (or sequential rationality) requires only (7) and (8) to bind.[44] This completes the argument.

(v) Suppose not so that $\mathcal{A}_1^* = 0$. This implies $g_2^* = g_1$ since (10) is binding. So, (9) can be re-written after using (10) as:

$$\alpha(g_3^*)\mathcal{A}_3^* \geq 0.$$

Clearly, $\mathcal{A}_3^* > 0$, otherwise the program reduces to an one-shot one. Thus, under the contraposition hypothesis (9) can be written as

$$\alpha(g_3^*)\mathcal{A}_3^* > 0,$$

assuming $\alpha(g) > 0, \forall g$. That is, (9) is non-binding. But this contradicts the result in part **(iv)**.

(vi) Follows from part **(v)** and the fact that (10) is binding as shown in part **(i)**. Q.E.D.

Proof of Lemma 2. (i)-(ii) Follows by straightforward differentiation of the expressions of $F(g_2, g_3; \mathcal{A})$ and $I(g_2, g_3; \mathcal{A})$, i.e. (11) and (12).

(iii) Recalling that $F_1(g_2, g_3; A) = \frac{\partial F(g_2, g_3; \mathcal{A})}{\partial g_2}$, we have

$$\frac{\partial F(g_2, g_3; \mathcal{A})}{\partial g_2} = C'(g_3 - g_2) - C'(g_2 - g_1) + \frac{\alpha(g_2)C'(g_3 - g_2)}{\delta\alpha(g_3)} - \frac{\alpha'(g_2)C(g_3 - g_2)}{\delta\alpha(g_3)},$$

which is positive whenever either $g_3 - g_2 \geq g_2 - g_1$ (recall that $C''(.) \geq 0$), or $C(.)$ is linear.

[44] That constraint (10) should bind is a requirement of the optimal PCC program as shown in part **(i)** above, but not for sequential rationality.

By the implicit function theorem, we have

$$\frac{\partial f(g_2, \mathcal{A})}{\partial g_2} = -\frac{\frac{\partial F(g_2, g_3; \mathcal{A})}{\partial g_2}}{\frac{\partial F(g_2, g_3; \mathcal{A})}{\partial g_3}}.$$

Next observe that

$$\frac{\partial F(g_2, g_3; \mathcal{A})}{\partial g_3} = -C'(g_3 - g_2) + \frac{\alpha(g_2)\alpha'(g_3)C(g_3 - g_2)}{\delta(\alpha(g_3))^2} - \frac{\alpha(g_2)C'(g_3 - g_2)}{\delta\alpha(g_3)},$$

which is negative. The proof now follows from the signs of $\frac{\partial F(g_2, g_3; \mathcal{A})}{\partial g_2}$, $\frac{\partial F(g_2, g_3; \mathcal{A})}{\partial g_3}$ and the expression for $\frac{\partial f(g_2, \mathcal{A})}{\partial g_2}$.

(iv) First, using $i(g_2, \mathcal{A})$ in (12) and evaluating it at $g_2 = g_1$ together with the fact that $\mathcal{A} > 0$ implies that $i(g_1, \mathcal{A}) > g_1$. Now given that $i(g_1, \mathcal{A}) > g_1$, we have, using (11) and (12),

$$F(g_1, i(g_1, \mathcal{A}); \mathcal{A}) = I(g_1, i(g_1, \mathcal{A}); \mathcal{A}) - C(i(g_1, \mathcal{A}) - g_1) < I(g_1, i(g_1, \mathcal{A}); \mathcal{A}) = 0,$$

where the inequality follows since $C(i(g_1, \mathcal{A}) - g_1) > 0$. Finally, given that $F(g_1, i(g_1, \mathcal{A}); \mathcal{A}) < 0$, $F(g_1, f(g_1, \mathcal{A}); \mathcal{A}) = 0$, and $F(g_2, g_3; \mathcal{A})$ is strictly decreasing in g_3, we will have $i(g_1, \mathcal{A}) > f(g_1, \mathcal{A})$.

(v) Straightforward differentiation yields that

$$\begin{aligned} I_{12}(g_2, g_3; \mathcal{A}) = & -\frac{\alpha'(g_2)C'(g_3 - g_2)}{\delta\alpha(g_3)} + \frac{\alpha'(g_2)\alpha'(g_3)C(g_3 - g_2)}{\delta\alpha^2(g_3)} \\ & + \frac{\alpha(g_3)C''(g_3 - g_2)}{\delta\alpha(g_3)} - \frac{\alpha(g_2)\alpha'(g_3)C'(g_3 - g_2)}{\delta\alpha^2(g_3)} > 0, \end{aligned}$$

since $\alpha'(.) < 0$ and $C''(.) \geq 0$.

(vi) For the case when $C(.)$ and $\alpha(.)$ are both linear, by differentiating we have:

$$\begin{aligned} I_{11}(g_2, g_3; \mathcal{A}) = & \frac{2\alpha'(g_2)C'(g_3 - g_2)}{\delta\alpha(g_3)} + \frac{\alpha(g_1)C'(g_2 - g_1)\alpha'(g_2)}{\delta\alpha(g_2)} \\ & - \frac{2\alpha(g_1)C'(g_2 - g_1)(\alpha'(g_2))^2}{\delta\alpha^3(g_2)} < 0, \end{aligned}$$

since $\alpha'(.) < 0$. Q.E.D.

Proof of Lemma 3. The argument presented below should be studied with Figs. 3 and 4.

(i) Suppose not. There are two cases to consider.

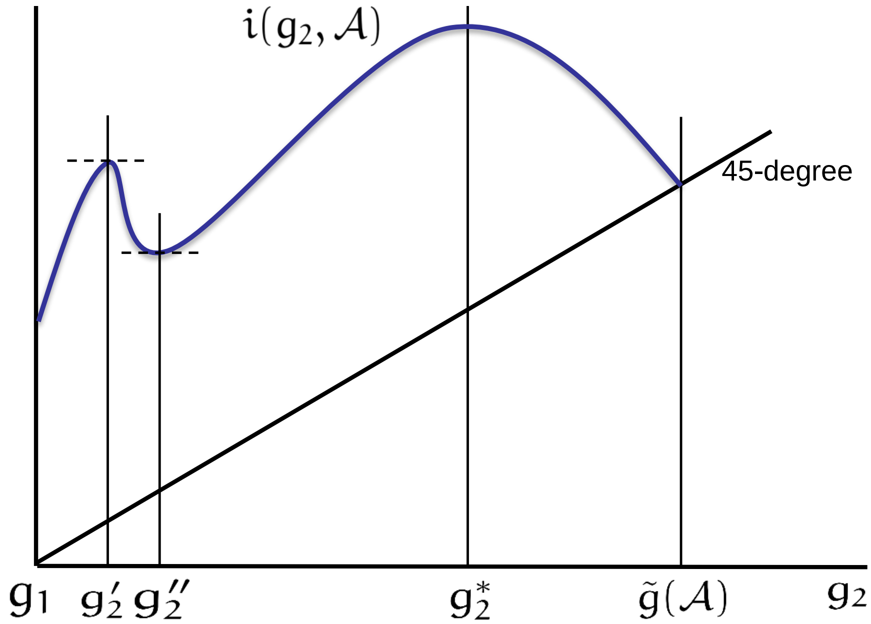

Figure 3: Proof of Lemma 3(i), Case (a)

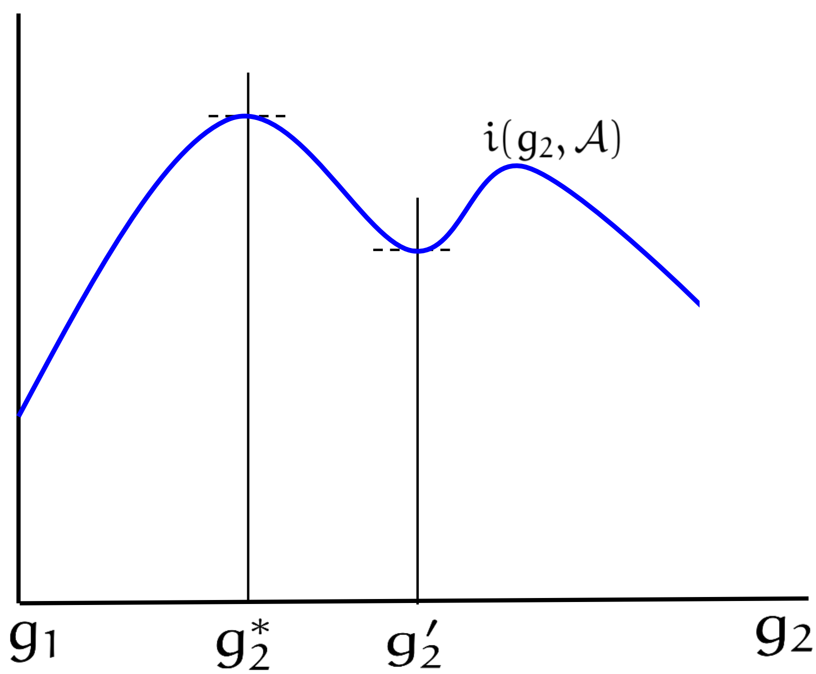

Figure 4: Proof of Lemma 3(i), Case (b)

Case (a). First, suppose to the contrary that $i(g_2, \mathcal{A})$ is not strictly increasing for $g_2 < g_2^*$. Then there exists $g_2' < g_2^*$ such that $i'(g_2', \mathcal{A}) = 0$, so that $I_1(g_2', i(g_2', \mathcal{A}); \mathcal{A}) = 0$. We first observe that $i'(g_2, \mathcal{A}) = -\frac{I_1(.)}{I_2(.)}$. Given that $I_2 < 0$ (from Lemma 2(**i**)), and $I_1(g_2, g_3; \mathcal{A})|_{g_2=g_1} = -\frac{\alpha'(g_2)C(g_3-g_1)}{\delta\alpha(g_3)} + \frac{\alpha(g_1)C'(g_3-g_1)}{\delta\alpha(g_3)} > 0$, it follows that $i'(g_1, \mathcal{A}) > 0$. Given that $i'(g_1, \mathcal{A}) > 0$, it is sufficient to consider the case where g_2' is a local maximizer of $i(g_2, \mathcal{A})$, and hence there exists g_2'', where $g_2' < g_2'' < g_2^*$, and g_2'' is a local minimizer of $i(g_2, \mathcal{A})$ such that $i(g_2'', \mathcal{A}) < i(g_2', \mathcal{A})$ and $I_1(g_2'', i(g_2'', \mathcal{A}); \mathcal{A}) = 0$. Hence

$$\begin{aligned}
0 &= I_1(g_2', i(g_2', \mathcal{A}); \mathcal{A}) \\
&> I_1(g_2'', i(g_2', \mathcal{A}); \mathcal{A}) \quad \text{(since } g_2' < g_2'', \text{ and, from Assumption 3(i), } I_{11} < 0) \\
&> I_1(g_2'', i(g_2'', \mathcal{A}); \mathcal{A}) \quad \text{(since } i(g_2'', \mathcal{A}) < i(g_2', \mathcal{A}), \text{ and, from Lemma 2(v), } I_{12} > 0),
\end{aligned}$$

which contradicts the fact that g_2'' is a local minimizer.

Case (b). Next, suppose to the contrary that $i(g_2, \mathcal{A})$ is not strictly decreasing for $g_2 > g_2^*$. Then there exists $g_2' > g_2^*$ such that $I_1(g_2', i(g_2', \mathcal{A}); \mathcal{A}) = 0$. Hence

$$\begin{aligned}
0 = I_1(g_2^*, i(g_2^*, \mathcal{A}); \mathcal{A}) &> I_1(g_2', i(g_2^*, \mathcal{A}); \mathcal{A}) \quad \text{(since } g_2' > g_2^*, \text{ and } I_{11} < 0) \\
&\geq I_1(g_2', i(g_2', \mathcal{A};)\mathcal{A}) \quad \text{(since } i(g_2', \mathcal{A}) \leq i(g_2^*, \mathcal{A}), \text{ and } I_{12} > 0),
\end{aligned}$$

which is a contradiction.

(ii) Suppose $i(g_2^*, \mathcal{A}) < f(g_2^*, \mathcal{A})$ (see Fig. 3). Given that $F(g_2, g_3; \mathcal{A})$ is strictly decreasing in g_3, this implies $F(g_2^*, i(g_2^*, \mathcal{A}); \mathcal{A}) > 0$, so that $(g_2^*, i(g_2^*, \mathcal{A}))$ satisfies both the constraints (11) and (12). Finally, recalling that $i(g_2, \mathcal{A})$ has a unique maximizer g_2^* (from part **(i)** of this lemma), and the set of feasible g_3, from (12), is $\{g_3 | g_3 \leq i(g_2, \mathcal{A})\}$, the optimal solution must involve $g_3^* = i(g_2^*, \mathcal{A})$.

We next prove the causality in the reverse direction. Suppose the outcome involves interim aid diversion so that the optimal g_2, call it $\hat{g}_2$, satisfies $i(\hat{g}_2, \mathcal{A}) < f(\hat{g}_2, \mathcal{A})$. We next argue that it must be that $\hat{g}_2 < g_2^*$. Suppose to the contrary that $\hat{g}_2 > g_2^*$. Then, for ϵ positive but small, from part **(i)** of this lemma and Assumption 3(**i**) earlier it follows that $(\hat{g}_2 - \epsilon, i(\hat{g}_2 - \epsilon, \mathcal{A}))$ is feasible and yields a higher level of g_3, which contradicts the optimality of $\hat{g}_2$. Hence $\hat{g}_2 < g_2^*$. But, then for ϵ positive but small, from part **(i)** of this lemma it follows that $(\hat{g}_2 + \epsilon, i(\hat{g}_2 + \epsilon, \mathcal{A}))$ is feasible and yields a higher level of g_3 (see Figure 5). This is a contradiction. Q.E.D.

Proof of Lemma 4. The left-hand side of (20) is the payoff to the recipient from acting true to its type λ, whereas the right-hand side is the payoff from projecting as type $\tau = 1$

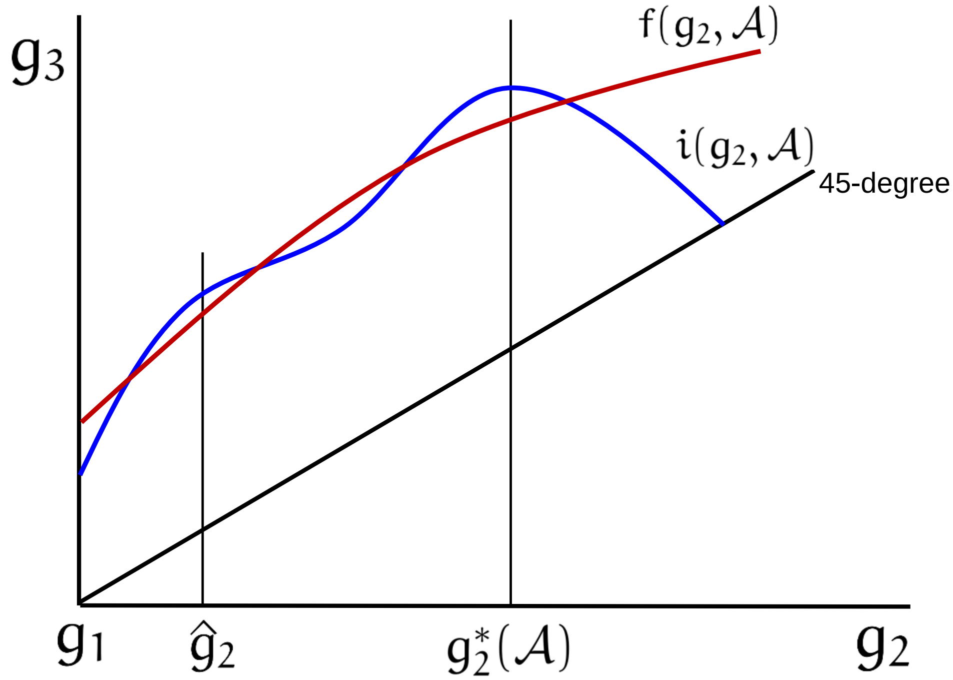

Figure 5: Proof of Lemma 3(ii); increase in g_2 leads to improved g_3

and continuing to deliver on expected governance improvement or simply divert the aid on offer (without any attempt at improving governance) either in period 2 or as early as in period 1.[45]

IC for type $\tau = 1$ at $t = 1$ requires that

$$\alpha(g_1)[\breve{\mathcal{A}}_1 - C(\breve{g}_2 - g_1)] + \alpha(\breve{g}_2)[\breve{\mathcal{A}}_2 - C(\breve{g}_3 - \breve{g}_2)] + \alpha(\breve{g}_3)\breve{\mathcal{A}}_3$$
$$\geq \max\left\{\alpha(g_1)[\breve{\mathcal{A}}_1 - C(\breve{g}_2 - g_1)] + \alpha(\breve{g}_2)\breve{\mathcal{A}}_2,\ \alpha(g_1)\breve{\mathcal{A}}_1\right\},$$

which can be reduced to

$$\alpha(\breve{g}_3)\breve{\mathcal{A}}_3 = \max\left\{\alpha(\breve{g}_2)\breve{\mathcal{A}}_2,\ \alpha(g_1)\breve{\mathcal{A}}_1\right\},$$

because IC and FC for type $\tau = 1$ must be binding in period 2 in the optimal program and

[45] One aspect worthwhile to note is that we allow the recipient to keep the entire cost savings due to its better type and not subject it to imperfect diversion that depends on the accumulated level of governance. The justification for this would be that the recipient can always do the necessary cost-padding up to the inefficient type's costs that the country's governance cannot control. While this might be an extreme assumption, we like to see to what extent this rampant opportunism can limit the effectiveness of aid meant for the efficient type.

$\breve{\mathcal{A}}_1 = C(\breve{g}_2 - g_1)$ by design. This further implies:

$$\alpha(\breve{g}_2)\underbrace{[\breve{\mathcal{A}}_2 - C(\breve{g}_3 - \breve{g}_2)]}_{=0} + \underbrace{[C(\breve{g}_3 - \breve{g}_2) - \lambda C(\breve{g}_3 - \breve{g}_2)]}_{>0} + \alpha(\breve{g}_3)\breve{\mathcal{A}}_3 > \max\left\{\alpha(\breve{g}_2)\breve{\mathcal{A}}_2,\ \alpha(g_1)\breve{\mathcal{A}}_1\right\}.$$

So (20) reduces to

$$\begin{aligned}
&\alpha(g_1)[\breve{\mathcal{A}}_1 - \lambda C(\tilde{g}_2 - g_1)] + \alpha(\tilde{g}_2)[\tilde{\mathcal{A}}_2 - \lambda C(\tilde{g}_3 - \tilde{g}_2)] + \alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3 \\
&\geq \alpha(g_1)\underbrace{[\breve{\mathcal{A}}_1 - C(\breve{g}_2 - g_1)]}_{=0} + [C(\breve{g}_2 - g_1) - \lambda C(\breve{g}_2 - g_1)] \\
&\qquad + \alpha(\breve{g}_2)\underbrace{[\breve{\mathcal{A}}_2 - C(\breve{g}_3 - \breve{g}_2)]}_{=0} + [C(\breve{g}_3 - \breve{g}_2) - \lambda C(\breve{g}_3 - \breve{g}_2)] \\
&\qquad + \alpha(\breve{g}_3)\breve{\mathcal{A}}_3,
\end{aligned}$$

which, in turn, can be simplified to

$$\alpha(\tilde{g}_3)\tilde{\mathcal{A}}_3 - \alpha(\breve{g}_3)\breve{\mathcal{A}}_3 \geq (1-\lambda)[C(\breve{g}_2 - g_1) + C(\breve{g}_3 - \breve{g}_2)].$$

(Recall, $\tilde{\mathcal{A}}_2 = \lambda C(\tilde{g}_3 - \tilde{g}_2)$ due to binding financial constraint in period 2, and $\tilde{\mathcal{A}}_1 = \lambda C(\tilde{g}_2 - g_1)$ by design.) That is, the efficient type should be left enough to divert from the final period aid to compensate for what it could obtain through cost savings and aid diversion by projecting as the inefficient type. Q.E.D.

Proof of Proposition 10′. Recalling (23), we have:

$$\frac{\alpha(\tilde{g}_3)}{\alpha(\tilde{g}_2)}[\mathcal{A} + \Upsilon - \breve{\mathcal{A}}_1 - \lambda C(\tilde{g}_3 - \tilde{g}_2)] = \lambda C(\tilde{g}_3 - \tilde{g}_2).$$

Note that the RHS of the preceding equation is increasing, and the LHS is decreasing in $\tilde{g}_3$, so that $\tilde{g}_3$ is decreasing in λ. Further, given that $\lim_{g\to\infty}\alpha(g) \to 0$, $\tilde{g}_3$ increases without bounds as λ goes to zero.

Next let $\bar{\lambda}$ be such that $\forall\, \lambda < \bar{\lambda}$, it is the case that

$$\alpha(\tilde{g}_3)[\mathcal{A} + \Upsilon - \breve{\mathcal{A}}_1 - \lambda C(\tilde{g}_3 - \tilde{g}_2)] < \alpha(\breve{g}_3)\breve{\mathcal{A}}_3 + (1-\lambda)[C(\breve{g}_2 - g_1) + C(\breve{g}_3 - \breve{g}_2)].$$

Such a $\bar{\lambda}$ exists since (a) $\breve{g}_3$, $\breve{g}_2$, $\breve{\mathcal{A}}_3$ and $\breve{\mathcal{A}}_1$ do not depend on either λ, or Υ, (b) the RHS of the preceding inequality is decreasing in λ, and (c) the LHS goes to zero as $\tilde{g}_3$ becomes sufficiently large. But then for all such λ, the incentive constraint at $t = 1$ for the efficient type, i.e. (21), will be violated. Q.E.D.