Decentralized Aid and Democracy

Joaquín Morales Belpaire*

September 17, 2013

Abstract

The last three decades have seen an important surge of the non-governmental sector in the provision of foreign aid. Using NGOs to deliver aid can be a solution to bypass corrupt authorities, avoiding that aid resources are captured by local elites. However NGOs may also act as surrogates for governmental provision of public goods. This implies that citizens make their own governments less accountable. In democratic countries, this can reduce electoral support for provision of public services by the state and harm the poor that don’t directly benefit from the NGOs’ projects. We develop a theoretical model of vote over public finances to analytically characterise the effect of decentralised aid on welfare. We find that non-governmental aid can harm the poor, weaken governance and aggravate inequalities by crowding-out governmental expenditures. These inefficiencies occur even in a flawless institutional context. We also find that the crowding-out effect can be mitigated if NGOs target countries with low income inequalities or by focusing on humanitarian-oriented missions.

JEL Classification: D72, H44, L31, L33, O19.
Keywords: democracy, governance, NGOs, nonprofits, foreign aid, taxation, public goods.

*University of Namur (CRED), Rempart de la Vierge 8, 5000, Namur, Belgium. E-Mail: jmorales@fundp.ac.be. Tel. +32-81-72 48 14
1 Introduction

The involvement of the non-governmental sector in foreign aid delivery has substantially expanded in the last thirty years. OCDE’s Official Development Aid (ODA) channelled through NGOs raised from a negligible amount in the early 80’s to over $3 bn. by the early 2000’s (around 6% of ODA) and up to $18 bn. in 2010 (18% of ODA, OCDE 2010). Staggeringly, there is no yet a clear assessment on the overall impact of decentralized aid on development. While impact studies at the project level become increasingly sophisticated and reliable, still not much is known about their effects on the polity.

Few economists have looked into this matter. Conversely, scholars from other social sciences have produced an abundant literature on the political repercussions of aid provision by NGOs. Far from reaching consensus, most of this literature is extremely critical towards decentralized aid (see for example Arellano-López and Petras, 1994; Petras, 1997; Lorgen, 1998; Mercer, 2002). Amongst the most vocal, it is worth quoting sociologist James Petras (1997):

“The important political point is that the NGOs depoliticized sectors of the population, undermined their commitment to public employees, and co-opted potential leaders in small projects. . . . In practice, ‘non-governmental’ translates into anti-public-spending activities, freeing the bulk of funds for neoliberals to subsidize export capitalists while small sums trickle from the government to NGOs.”

Decentralized aid, according to this author, crowds-out governmental expenditure, weakens political representation of the poor and is instrumental to the rich and powerful. The gravity of this accusation makes worth exploring if the argument withstands formal modelling. Is it true that ‘non-governmental’ translates into ‘anti-public-spending’?

Our analysis finds that this is indeed the case in a quite general theoretical specification. We show that in a democracy where citizens vote over public finances, decentralized aid can reduce public spending and aggravate inequalities. Our results do not depend on assumptions over corrupt authorities, devious NGO managers or imperialist agendas. Instead, we find that aid provided by NGOs can be inefficient even under assumptions of sound institutionality. In our model, the crowding-out effect appears under assumptions of full commitment by politicians, no corruption, healthy public finances and successful NGO projects. Although these assumptions are

---

1Exceptions include Yontcheva and Masud (2005); Dreher et al. (2007); Koch (2007).
naive, this leads us to the important conclusion that good institutions are not enough to guarantee aid efficiency.

The mechanism by which NGOs crowd-out government expenditures is the following. When services provided by an NGO are substitutable to governmental services, any voter benefiting from the project would favor a low taxation-low expenditure policy platform. This happens because while NGO projects are financed by foreign donors, governmental services require taxation. Any beneficiary of an NGO would find optimal to update her optimal taxation level downwards for a given provision of similar services by the NGO. By skewing the distribution of political preferences toward smaller government, such behaviour pulls the median voter away from poor non-beneficiaries of the project. As a result, the preferences of the poor non-beneficiaries become more distant from the policy implemented, hindering their representation and reducing redistribution.

We find that this effect is larger in unequal or polarized societies, and that it is mitigated in more homogeneous societies. We also find that humanitarian NGOs, by targeting fewer but poorer agents, limit the extent of governmental contraction, because a limited number of agents update their electoral preferences. The paper makes the case that humanitarian projects are thus desirable not only because they help the poorest, but also because they limit the extent of the externality inflicted on other poor people.

The paper unfolds as follows. Section 2 reviews relevant literature. Section 3 sets up the model, the equilibrium of which is characterized in Section 4. Section 5 studies the welfare implications of the new equilibrium at the individual and aggregate levels. Section 6 concludes.

2 Literature review

Surprisingly few studies on aid efficiency categorize aid effects by type of provider, NGO or official channels\(^2\). This is problematic because there are serious reasons to believe that official and decentralised aid are systematically different. Failure to differentiate them would induce misguided policy advice.

There is an evident reason why literature on NGOs’ efficiency is so exiguous: cross-country data on NGO activities is scarce, dispersed and of debatable quality. The few existing studies build databases on meagre informational resources. A notable example is provided by Koch et al. (2009) who have built the most comprehensive (to our knowledge) cross-section database on NGOs. They find that when NGOs’ decide upon the geographic location of their projects they often follow official aid. Therefore, countries receiving

\(^2\)With the notable exception of Yontcheva and Masud (2005)
more development assistance also host more NGOs’ projects. The correlation resulting from this complementarity implies that untangling the different effects by type of aid becomes necessary for sound aid-effectiveness analysis.

Getting around the issue of data scarcity justifies the use of an applied theoretical model. For this purpose, it is important to review a large scope of literature so that the model is coherent with the state of the art. There are three main streams that are relevant for our purposes. First, we want to review the theoretical and empirical micro-foundations that describe the NGOs’ behaviour, so as to justify our assumptions at the agency level. Secondly, studies from sociology or political sciences suggest that NGOs’ projects have wider impacts on society than the project itself, making them into potential sources of externalities, namely crowding-out governmental expenditure. Finally, we examine a few cross-country studies that have looked at the impact of aid on inequality and governance, to see if empirical macro findings are consistent with the aforementioned externality.

There is a widespread consensus that aid has the potential of being successful at the project level (see e.g. Mosley, 1986; Arndt et al., 2010; Howes et al., 2011). Given the general improvement of project evaluation techniques, in particular through the surge of randomized controlled trials, these successes are rigorously tested and identified (Banerjee and Duflo, 2011). In this paper we do not question project efficiency in itself to have a clear picture of the aggregate effects at the polity level. This being said, not all NGO behave in the same manner. Recent studies based on local-level surveys enable us to better understand the behaviour of NGOs and other aid agencies (Easterly, 2002; Fruttero and Gauri, 2005; Brass, 2012). These seem generally confronted to a pragmatic versus humanitarian tradeoff. NGOs have to compromise between helping those that need it most, in the most adverse and risky situations, versus catering to many in safer, easier cases. Empirical evidence shows that NGOs can favour either approach in different world locations. Our model we allow for variation in the type of project, whether humanitarian or pragmatic.

Having this in mind, it may be the case that successful projects affect agents other than the direct beneficiaries. This, as we stated in the introduction, has been the focus of most of the sociological and political sciences literature. A sizeable number of authors have claimed that NGOs have indirectly harmed the poor by crowding-out state-provided public services (see e.g. Arellano-López and Petras, 1994; Petras, 1997; Lorgen, 1998; Mercer, 2002). They claim that, during the structural reforms of the eighties, NGOs have eroded governmental accountability by becoming surrogates for public services. This loss of accountability has allegedly resulted in less social inclusion, weakened governance and reduced fairness. While economists are
generally sceptical toward this rhetoric, it’s worth noting that Collier and Dollar (2004) have stated the following

“By detaching the wellbeing of the population from the actions of the government, ... NGOs can undermine democratic accountability. [If] the government is democratic, then offsetting the direct benefits of the project are unquantifiable externalities as accountability is undermined”.

Interestingly, these adverse effects become apparent at the macroeconomic level. Cross-country literature finds evidence that aid negatively impacts equity and governance. We cannot, however, disentangle which type of aid causes these effects, because official aid is correlated to NGOs’ presence. It cannot be ruled out that NGOs indeed have and adverse impact on inequality and governance. In what regards to inequality, although Chong et al. (2009) find weak evidence that aid reduces inequality when the quality of institutions is considered, Layton and Fuller (2008); Bjørnskov (2010) and Herzer and Nunnenkamp (2012) find robust evidence that it actually increases it. In particular, Bjørnskov finds that aid worsens inequalities only in democratic countries, assumed to have better institutions than non-democracies. In relation to governance, Collier and Dollar (2004) claim that aid has a negligible impact on governance quality. However, Rajan and Subramanian (2007) find that the negative effect of aid on institutions is important by objectively measuring the share of industries depending on good governance (industries depending on numerous transactions). This share is noticeably lower in aid-recipient countries, implying poor governance levels. In the same vein, Djankov et al. (2008) find that foreign aid has a negative impact on the Polity IV measures of democracy (that include representativeness of citizens) and on the number of official checks required before a policy is implemented. For these authors, ‘aid is a bigger curse than oil’: they find that the negative effects of aid on these indicators are far greater than those of natural resource windfalls.

Summing up, while projects implemented by NGOs seem to be successful at the local level, they might cause externalities that erode governmental accountability at the aggregate level. Citizens would rely less on the state for provision of public services and favour a reduction of government size. This contraction might translate into less redistribution and narrowed political representation of poor citizens. Next section synthesizes the aforementioned concepts into a single mechanism.

---

3If, in turn, reduced governance and equality have adverse effects on growth, this mechanism could be at the heart of what has been called the Micro-Macro Paradox by Mosley (1986). The paradox states that despite project successes at the local level, aid
3  Set-up of the Model

Agents in this model are citizens and an NGO. Citizens decide upon a simple majority vote on the taxation level aimed at financing a public good. The NGO can also provide a local public good but it is not accountable to all citizens.

3.1  Citizens’ choices over Public Finances

We depart from the Downsian framework (Downs, 1957). Suppose an electorate composed of \( N \) agents indexed \( i \in N \). Each agent’s exogenous income is \( y_i \in \mathbb{R}_+ \). \( N \) is assumed to be sufficiently large such that we can draw a continuous distribution over an interval normalised to one. Denote its cumulative distribution function (CDF) \( F_Y(y_i) \), which is supposed unimodal, and let \( \bar{y} \) be the average income. We assume that the government can provide an amount \( G \) of public good (PG) according to a simple budget constraint \( G = \tau \bar{y} \) where \( \tau \) is a flat tax rate over income. We assume that since this PG is financed by all of the electorate, every tax-payer is entitled to access to it.

Under this setting citizens vote upon a level \( G \) of PG to be provided by the government. Any citizen indexed \( i \) has the following quasi-linear preferences:

\[
U_i(c_i, G) = c_i + H(G)
\]

where \( c_i \) is her own private consumption in the numéraire, which is a tradable good, and where function \( H(G) \) is increasing and concave in \( G \). The quasi-linear function implies that attention is focused on public goods for which the is a tradable (imperfect) substitute.

The saturated budget constraint of any agent is \( c_i = (1 - \tau)y_i \). This, combined to the government’s budget constraint and the assumption of proportional taxation, allows us to write the agent’s maximization problem as

has no effect on growth at the aggregate level. The existence of this paradox has been
sometimes substantiated (Mosley, 1986; Rajan and Subramanian, 2008; Doucouliagos and Paldam, 2009) and some other times invalidated (Burnside and Dollar, 2000; McGillivray et al., 2006; Arndt et al., 2010; Mekasha and Tarp, 2011). It should be noted that because we focus in a political economics setting, we do not look at effects of reduced governance or inequity on growth.

4The use of the quasilinear utility function implies that the quantity of the PG has no effect on the price of traded good (the numéraire), which considerably simplifies the setting. An important share of the literature on public goods assumes this specification, see among others Diamond (1973), Sandmo (1980), Mas-Colell et al. (1995, p.360), Persson and Tabellini (2002, pp. 48, 70, 98, 119). A generalization is nonetheless provided in appendix A.
follows

\[
\max_G \left( 1 - \frac{G}{\bar{y}} \right) y_i + H(G). \tag{1}
\]

By a slight abuse of notation, the bliss point \(G^*_i\) for any agent \(i\) is

\[
G^*_i = H^{-1}_G \left( \frac{y_i}{\bar{y}} \right). \tag{2}
\]

Because \(H^{-1}_G(\cdot)\) is monotonically decreasing on \(y_i\), the preferred level of public good decreases as income increases. Following Persson and Tabellini (2002, p.21-23), given that (1) is single peaked and has the single-crossing property the median voter theorem (Black, 1948) can be applied.

**Proposition 1**  Given the properties of equation (1) and by application of the median voter theorem the outcome of the election corresponds to the preferred policy of the agent with median income \(y_m\) defined by \(F_Y(y_m) \equiv \frac{1}{2}\). The resulting implementation in terms of PG is

\[
G_m = H^{-1}_G \left( \frac{y_m}{\bar{y}} \right). \tag{3}
\]

**Proof** Denote \(G_m\) is the bliss point of the median voter. For any candidate campaigning for a platform different from \(G_m\), it is a dominant strategy for her opponent to campaign for \(G_m\) to get the majority. It is never profitable to deviate from proposing \(G_m\) for any candidate. \(\blacksquare\)

It should be noticed that this equilibrium does not necessarily result in a social optimum, unless the median voter is also the average voter. To see this, consider the Benthamite social welfare function, that is the aggregation all utilities

\[
W(G) = \int_{y_i \in Y} \left( 1 - \frac{G}{\bar{y}} \right) y_i + H(G) \, dF(y) = \bar{y} + H(G) - G. \tag{4}
\]

This expression is maximized when the electoral outcome is \(G^* = H^{-1}_G(1)\). By equation (2) and the median voter theorem, we will have \(G_m = H^{-1}_G(1)\) if and only if \(y_m = \bar{y}\). This social optimum has interesting properties. We

\[5\]We implicitly assume here that politicians’ motivations are driven by an ego rent obtained when elected and that they don’t have political considerations of their own. This naive assumption is taken to clearly identify the implications of our mechanism in the standard voting model.
know that the mean of a distribution is the point that minimizes the sum of the distances between each observation and the mean itself. Mathematically that is
\[ \bar{y} = \arg \min_{y_0} \int_{y_i \in Y} |y_i - y_0| dy_i \]

The social optimum implies that each individual position is as close as possible to the implemented policy for given moments of the income distribution. When the policy implemented deviates from the optimal bliss point, we interpret this as degradation of representativity and therefore of governance.

To simplify the subsequent analysis, we map income levels into a single-dimensional distribution of political preferences. Let the \( \theta \) be the type of the agent, such that any agent with the same preferred policy belongs \( \theta \). It is defined by
\[ \theta \equiv \frac{G_i}{G_1} \]
where \( G_1 \) is the bliss point for \( y_1 = \min\{y_i \in Y\} \). Its domain is the set
\[ \Theta \equiv \{\theta : 0 \leq \theta \leq 1, \theta \in \mathbb{R}_+\} \]

(5)

In this unit interval, types are ranked from low to high requirements in terms of needs: the higher the type, the needier an agent is. This rescaling will be useful in the following subsection. Given the continuity, differentiability and monotonicity of equation (2) relative to income, and by the change of variables theorem (see e.g. Kaplan, 1984, pp. 238-245), it is possible to map the distribution of incomes \( y \) into a distribution density function \( f(\theta) \) of types \( \theta \in \Theta \).

The support of this distribution is \( \Theta \), where \( \lim_{\theta \to 0} F_{\Theta}(\theta) = 0 \) and \( F_{\Theta}(1) = 1 \). We define also the median type by \( \theta_m \) where \( F_{\Theta}(\theta_m) = \frac{1}{2} \). The level of public good that will be provided is therefore \( G_m = \theta_m G_1 \).

Nonetheless, for any \( \theta < \theta_m \) there is over-provision of the PG and would like that the fiscal burden was reduced. Conversely, for higher types, the state is considered as under-providing public services. The next section studies how NGOs intend to tackle these unmet needs.

### 3.2 Modelling NGOs

We look for an empirical foundation justifying the manner in which NGOs are modelled. Two recent studies, one by Fruttero and Gauri (2005) in Bangladesh and another by Brass (2012) in Kenya, highlight which are the factors determining NGOs’ location choices. Much of the observed behaviour is similar to that of official aid bureaucracies described by Easterly (2007). In
a principal-agent set-up, agencies are accountable to donors plagued by imperfect information, making agencies adopt several strategies to get funded. To secure funding two types of strategic location choices are observed.

First, agencies choose to cluster together around specific geographical regions. Because effort is not perfectly monitored by donors, NGOs might regroup in order to free-ride on other NGOs’ efforts to cater to a ‘common pool of beneficiaries’. Donors might only observe the aggregated results and be unable to disentangle which result is attributable to which agency. The second location decision depends upon human needs. Whilst it is expected that aid targets areas with higher needs, this kind of mission could also be riskier. NGOs might choose to locate in less poverty-stricken or peaceful areas, increasing the probability that a project is successful. In these areas, it is also easier to reach more beneficiaries, and headcounts are a cheap way of signalling effort to principals. Nonetheless, some other NGOs might instead target agents confronted with hardship, implying a higher risk of mission failure. This might be done for purely humanitarian reasons or for strategic reasons to signal zeal to donors.

Evidence for both pragmatic motivations and humanitarian motivations are found in this burgeoning literature. In our model we allow for agencies to vary on the degree of motivation and pragmatism. We introduce an NGO able to provide a local public good (LPG) to a fraction of the population. Let this decentralized aid agency be an organisation whose funding is external to the recipient country’s budget. Although its activities are legal, this agency does not use the structure of the central administration to channel aid.

Suppose that the NGO’s program reaches a fraction $\Delta \leq 1$ of the population. This fraction is a continuum of individuals along the $\theta$-type dimension where $\theta_h$ is the highest type amongst beneficiaries of the project and $\theta_l$ the lowest. This implies $\Delta = F(\theta_h) - F(\theta_l)$. We can therefore define the set of beneficiaries as

$$\Theta_b = \{\theta : \theta_l < \theta \leq \theta_h\}$$

This assumes continuity of the targeted section of the population.\(^6\) In order to cater to this continuous array of types $\Theta_b$, the agency has to choose an ideal type $\theta_t$ it targets such that $\theta_t \in \Theta_b$, that is it chooses the typical level of need that will be attended. All agents belonging to $\Theta_b$ benefit thus from a LPG considered as ideal for $\theta_t$ in addition to the public good $G$ provided by the state. The LPG provided by the NGO is then $g = \theta_t G_1$, the bliss

\(^6\)Although it could be argued that the agency provides instead parcelled programs to different ranges of types, or equivalently that several agencies are providing a multitude of programs, we focus on the case of a unique continuum of beneficiaries. The aforementioned fact that NGOs often cluster implies that this simplification is not too far from reality.
point of the targeted type. It follows that any agent’s utility level becomes $U_i(c, G + gb)$ where $b$ is an indicator variable equal to one if the agent is a recipient of the program.

We set the utility function of the agency to be

$$U^A(\Delta, \theta_t) = \Delta \theta_t^{\beta/1-\beta} \quad \text{where} \quad 0 < \beta < 1. \quad (7)$$

The utility level of the agency increases in the number of beneficiaries and the extent to which it is able to target higher types, that is, the neediest. Since the agency’s targeting is made in the bounded set $\Theta$, we allow for increasing marginal utility of the humanitarian tasks. When $\beta$ tends toward one the power of $\theta_t$ in the utility function tends toward infinity. This implies that the whole priority is put on helping the neediest, at the expense of coverage. Conversely, if parameter $\beta$ tends toward zero, the agency prioritizes only the number of beneficiaries. We sketch a simple, unidimensional distinction amongst different types of decentralized aid agencies:

**Definition** Depending on parameter $\beta$ the decentralized aid agency can be either:

1. **Coverage Prioritizing (CP)** when $\beta$ tends toward zero. If $\beta$ equals zero the agency is a pure maximizer of coverage.

2. **Benthamite Utilitarian (BU)** when $\beta = \frac{1}{2}$. All agents are weighted equally and the agency maximizes aggregated utility.

3. **Needs Prioritizing (NP)** when $\beta$ tends toward one. If $\beta$ equals one, the agency is Rawlsian, in the sense that it only targets the worse-off individual.

The program is assumed to be financed by the agency’s own limited funds, which are allocated exogenously by donors from the North. The budget constraint is assumed to be $B \geq c_g g + c_\Delta \Delta$ where $B$ is its available budget and parameters $c_g$ and $c_\Delta$ are costs of producing $g$ and of providing coverage to $\Delta$ beneficiaries, respectively\(^7\). These costs depend on technologies, characteristics of the mission and local conditions to which the agency is confronted to. Other logistical costs (office supplies, wages, travel expenditures...) are assumed away from $B$ which thus consist of the net resources devoted to the program.

The fact that the cost of provision increases in the number of targeted beneficiaries is straightforward. The idea that the cost of provision increases

\(^7\)For simplicity reasons, linearity of the budget constraint is assumed. Assuming more complex cost structures does not change the main results in an informative way.
with the targeted type is somewhat debatable. We assume that operations targeting hard situations are costlier because they require more efforts, they are riskier and they involve higher accessibility costs.

The utility maximisation problem of the agency is then

$$\begin{align*}
\text{Max}_{\theta_t, \Delta} & \quad \Delta \theta_t^{\beta_1-\beta} \\
\text{s.t} & \quad B \geq c_g \theta_t G_1 + c_\Delta \Delta.
\end{align*}$$

(8)

The solution to this utility maximization problem is given by

$$
g^* = \theta_t^* G_1 = \beta \frac{B}{c_g} \quad \text{and} \quad \Delta^* = (1 - \beta) \frac{B}{c_\Delta}.
$$

Denote by $\Delta_C = B/c_\Delta$ the maximum number of beneficiaries that a particular agency can cover. Similarly let $G_1 \theta_C = B/c_g$ be the highest feasible amount of LPG. We should therefore take this ‘maximum capacity’ vector $(\theta_C, \Delta_C)$ as parameters reflecting technical and financial endowments of the agency, dependent on exogenous characteristics $\{B, c_g, c_\Delta\}$.

The solution of the agency is summarized by a vector $p$ such that

$$
p = \begin{bmatrix}
\theta_t^* \\
\Delta^*
\end{bmatrix} = \begin{bmatrix}
\beta & 0 \\
0 & 1 - \beta
\end{bmatrix} \begin{bmatrix}
\theta_C \\
\Delta_C
\end{bmatrix}
$$

(9)

Let us call $p$ the project vector. This simple system characterizes two dimensions of the agency’s optimal solution. First, newly defined parameters $\theta_C$ and $\Delta_C$ in the ‘maximum capacity’ vector define the material and technical capacities characterizing which degree of need can the agency satisfy and for how many people. Secondly, parameter $\beta$ defines the prioritization criteria in the matrix pre-multiplying the maximum capacity vector. This prioritization criteria originates in the internal decision process of the organization, on its ethical values, strategic choices and features inherent to its mission.

Before moving into the next section, we need to make an additional technical assumption such that $\Theta_b \subseteq \Theta$.

**Assumption 1** For any given program vector $p$ we assume there is an allocation rule such that:

$$
\begin{align*}
F(\theta_l) &= F(\theta_l)(1 - \Delta) \\
F(\theta_h) &= F(\theta_l)(1 - \Delta) + \Delta
\end{align*}
$$

(10)
This assumption allows computational simplicity and ensures that the agency does not target types that don’t exist. Nonetheless, any other allocation rule such that $\Theta_b \subseteq \Theta$ would provide results similar to the ones presented here.

Given the agency’s optimal solution, we now set out to explore the project’s impacts on electoral outcomes.

4 Characterization of the Equilibrium

The timing of this model is simple: an NGO implements a program, beneficiaries update their electoral preferences and the outcome of subsequent elections defines the equilibrium. We provide analytical and graphical tools to characterize its properties.

4.1 Crowding-out of Governmental Expenditures

Suppose that prior to elections program $p = (\theta_t, \Delta)$ is implemented. Recipients will benefit from the LPG and will update their preferred governmental policy platform $G_i$. From equation (1) their maximization problem becomes

$$\max_G \left(1 - \frac{G}{y}\right) y + H(G + gb)$$

where $b = 1$ if $\theta \in \Theta_b$.

The solution to this maximization problem, which is analogous to the benchmark case, is $G_i = H^{-1}(y_i/y) - gb$ with $g = \theta_t G_i$. Expressed in ex-ante terms of types $\theta$, the updated type of the agents is defined by

$$\tilde{\theta} = \min\{0, \theta - \theta_t b\}.$$

The mass of voters with political preferences distributed over the segment $\Theta_b$ shifts to the left (directionally) by a distance $\theta_t$. Figure 1 depicts how the distribution of types shifts following the introduction of an NGO project.

Appendix A shows that this mechanism is not defined by the choice of the utility function of the citizens. We maintain our current specification for purposes of exposition.
Let us define three sub-domains $\Theta_k$ where $k = \{a, b, c\}$ and $k$ designs a binary variable such that
\[
k = 1 \text{ if } \theta \in \Theta_k.
\]
It is associated to the following subdomains
\[
\begin{align*}
\Theta_a & \equiv \{ \theta : \theta \leq \theta_h - \theta_t \} \\
\Theta_b & \equiv \{ \theta : \theta_t < \theta \leq \theta_h \} \\
\Theta_c & \equiv \{ \theta : \theta > \theta_h \}
\end{align*}
\] (12)
These are respectively: the set that includes all beneficiaries ex-post, the set of all beneficiaries ex-ante, and all those who are of a type higher than any beneficiary.

We will denote $J(\theta)$ the ex-post cumulative distribution function (CDF), which can be expressed in terms of $F(\theta)$. Figure 2 depicts function $J(\theta)$ for a given project $p$ in a $1 \times 1$ box and compares it to function $F(\theta)$ from which it is derived.

Insert figure 2 here

**Theorem 1** Given a distribution of types $F(\theta)$, a project with characteristics $p = (\theta_t, \Delta)$, the allocation rule given by assumption 3.3, the definition of indicator variables $k$ and the median voter theorem, the ex-post median voter $\tilde{\theta}_m$ is defined by
\[
J(\tilde{\theta}_m) = \frac{1}{2}
\]
where $J(\theta)$ is the ex-post CDF. It is a continuous piecewise function such that
\[
J(\theta) = 
\begin{array}{c}
\begin{bmatrix}
a \\
1 - a - b \\
b \\
c
\end{bmatrix}'
\end{array}
\begin{bmatrix}
F(\theta) + F(\theta_t + \theta) - F(\theta_t)(1 - \Delta) \\
F(\theta) + \Delta \\
F(\theta_t)(1 - \Delta) + \Delta \\
-\Delta
\end{bmatrix}.
\] (13)
This function is continuously non-decreasing and greater or equal than $F(\theta)$. The proof of this theorem is given in Appendix.

Three important corollaries stem from Theorem 1.

**Corollary 1** Reduction of government spending.
All portions of the ex-post CDF (detailed in Appendix B) are such that $J(\theta) \geq F(\theta)$ for all $\theta$. This inequality is strict if $\Delta > 0$ for all $\theta < \theta_h$. Under this conditions we have
\[
J(\theta_m) \geq F(\theta_m) \quad \text{and} \quad F(\theta_m) = J(\tilde{\theta}_m) = \frac{1}{2}
\]
\[
\Rightarrow \quad J(\theta_m) \geq J(\tilde{\theta}_m)
\]
\[
\Leftrightarrow \quad \theta_m \geq \tilde{\theta}_m.
\]
It follows that, if $\theta_m < \theta_h$, the outcome of elections always shifts in such a way as to reduce government expenditures. Denote this shift
\[ s \equiv \theta_m - \hat{\theta}_m \text{ where } 0 \leq s \leq \theta^m. \tag{14} \]

Simply put, beneficiaries whose ex-ante type laid to the right of the median type shift to its left. A new median-voter emerges that can gather more support for a reduced expenditure platform.

**Corollary 2 Full Disruption/Status Quo regimes** Notice that if $\theta = 0$, then $J(0) = \Delta F(\theta_t)$. Therefore, if $\Delta F(\theta_t) \geq \frac{1}{2}$ we have $s = \theta_m$: the outcome of the election is a complete disruption of the PG. Conversely, for any $\theta \geq \theta_h$, $J(\theta) = F(\theta)$. If $F(\theta_h) \leq \frac{1}{2}$ there is no shift in electoral outcomes ($s = 0$). This is the Status Quo situation.

We have that for both $\theta_t$ and $\Delta$ high enough, the provision of the PG by the state will be discontinued. On the contrary, for low levels of these variables, the program has a small or no impact on electoral outcomes.

**Corollary 3 Negative and positive externalities.** Let us focus on the subset of non-beneficiaries when $s > 0$. Any ex-ante type $\theta > \theta_h$ is made unambiguously worse-off by the project. Indeed, for these agents the ex-ante PG was already considered under-provided by the state. Further reductions in governmental expenditures only makes them worse-off. Conversely, any ex-ante type $\theta \leq \hat{\theta}_m$ is made unambiguously better-off by the crowding-out effect.

Let us now analyse in detail how the size of the shift in political outcomes is affected by the characteristics of $p$ and the distribution of types. Let us focus on solutions that lie outside of the Status Quo and Full Disruption cases. From equations (9) and (13) we can rewrite $J(\theta_m) = J(\theta_m - s) = \frac{1}{2}$ as
\[
\begin{bmatrix}
    a \\
    1 - a - b \\
    b \\
    c
\end{bmatrix}' \begin{bmatrix}
    F(\theta_m - s) + F(\theta_t + \theta_m - s) - F(\theta_t)(1 - \Delta) \\
    F(\theta_m - s) + \Delta \\
    F(\theta_t)(1 - \Delta) + \Delta \\
    -\Delta
\end{bmatrix} = \frac{1}{2}. \tag{15}
\]

Denote $b$ the vector of indicator variables pre-multiplying vector $F$, which contains the different segments of the function such that $b'F = \frac{1}{2}$. From
this it is possible to analyse the impact of project $p$ on shift $s(p)$ trough comparative statics. By the implicit function theorem this yields the following gradient:

$$\nabla s(\theta_t, \Delta) = -\left[ \frac{\partial b'F}{\partial s} \right]^{-1} \left[ \frac{\partial b'F/\partial \theta_t}{\partial b'F/\partial \Delta} \right]$$  \hspace{1cm} (16)$$

The explicit form of this gradient is provided in Appendix. Intuitive explanations of comparative statics are summarized in what follows. To facilitate the analysis we define three regimes (additional to the Status Quo and Full disruption ones).

**Needs Prioritization regime**

If $a = b = 0$ the gradient is

$$\nabla s(\theta_t, \Delta) = \begin{bmatrix} 0 \\ \theta_h - \theta_t \end{bmatrix}$$

Here $\theta_h - \theta_t < \tilde{\theta}_m < \theta_l$: the ex-post median voter is not a beneficiary. All that matters to her is the number of recipients $\Delta$ who have changed types, this is, who have been transferred from $\Theta_b$ to $\Theta_a$ and are thus willing to vote for her preferred policy, irrespective of the the level of LPG provided by the project.

**Coverage Prioritizing regime**

If $a = b = 1$ we have

$$\nabla s(\theta_t, \Delta) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

This is the domain where $\theta_l < \tilde{\theta}_m < \theta_h - \theta_t$ and therefore $\Theta_b \cap \Theta_a \neq \emptyset$. The ex-post median voter is a beneficiary of the program, therefore the level of LPG provided has an impact on her electoral preferences. When increasing $\Delta$, a higher-type agent moves from $\sup \Theta_b$ to $\sup \Theta_a$ which remains to the right of the ex-post median voter, implying no effect of additional beneficiaries (the converse is true for $\inf \Theta_b$).
Utilitarian regime

If $a \neq b = 0$ we have

$$\nabla s(\theta_t, \Delta) = \begin{bmatrix} > 0 \\ > 0 \end{bmatrix}$$

Finally, when $\tilde{\theta}_m < \min\{\theta_h - \theta_t, \theta_l\}$ the ex-post median voter is among ex-post beneficiaries. Her preferences change with the LPG received. Additionally, an increase in the number of beneficiaries transfers agents from $\inf \Theta_b$ to $\inf \Theta_a$, from the right to the left of the ex-post median voter. Hence both the number of beneficiaries and the level of the LPG have an impact on the size of the shift.

Having characterized how the size of the shift changes with the different values of vector $p$, we turn now to describe the effects of dispersion and skewness in the following proposition. Because we want to keep the most general version of distribution $F(\theta)$ we cannot proceed by standard comparative statics. We use pairwise comparisons for any single-crossing pair of distribution functions.

**Proposition 2 Effects of dispersion and skewness**

Taking $p$ given, the shift $s$ in electoral outcomes has the following characteristics:

1. The size of the shift $s$ increases with the level of inequality in society.
2. The size of the shift $s$ decreases with negative skewness, that is when poor become the majority. The converse is true.

Proof is provided in the appendix.

Interpretations from proposition 4.5 are relatively straightforward but important in what relates to policy. The simple intuition is that depending on the relative size of the upper-class relative to the middle- or lower-class, the electoral impact of any given project will be different. If the rich are relatively numerous, they will largely welcome and benefit from the new political stand of project beneficiaries. The pulling factor is lesser if there
is a large, non-beneficiary lower or middle class that can easily withstand changes from a few electors.

This proposition has important repercussions on the aid selectivity debate. If aid projects are willing to mitigate the crowding-out effect, aid should target countries with poor, homogeneous majorities. By contrast, targeting unequal or polarized societies can bring the undesirable repercussion of aggravating already tense situations. We will see, in the next section, that negative effects of the project can be mitigated depending on the choice of the NGO’s prioritization criterion \( \beta \), preventing the exclusion of polarized counties from aid.

We investigate in what follows the interaction between the agency’s project and political outcomes by proposing a graphical representation of political outcomes depending on the characteristics of the agency’s project.

### 4.2 Impact of NGO characteristics on Electoral Outcomes

Let us study how any combination \((\theta_t, \Delta)\) changes the size of \(s\). In order to do this, take a 1 × 1 box where \(\theta_t\) is represented in the X-axis and \(\Delta\) is represented in the Y-axis.

Consider the case of complete disruption of centrally provided PG, which occurs if \(s = \theta_m\). This happens when \(\Delta F(\theta_t) \geq \frac{1}{2}\). Then any project of coordinates \((\theta_t, \Delta)\) belongs to the Full Disruption regime if it lies above hyperbola

\[
\Delta : [\theta_m; 1] \rightarrow [\frac{1}{2}; 1], \quad \Delta(\theta_t | s = \theta_m) = \frac{1}{2 F(\theta_t)}. \tag{17}
\]

Similarly, political outcomes are not modified when \(s = 0\), which implies \(F(\theta_h) \leq \frac{1}{2}\). Inverting axes, we can trace a function

\[
\theta_t : [0; \frac{1}{2}] \rightarrow [0; \theta_m], \quad \theta^t(\Delta | s = 0) = F^{-1}\left(\frac{1}{2} - \Delta \frac{1}{1 - \Delta}\right). \tag{18}
\]

Any vector \((\theta_t, \Delta)\) below this hyperbola has no impact on electoral outcomes. Finally, we have to distinguish the regimes where the public good provision is reduced but does not disappear. The change in behaviour of function \(J(\theta_m) = \frac{1}{2}\) depends on the binary variables \(\{a, b\}\). By Theorem 1 we have:

\(^8\text{Variable } c \text{ is excluded and valued zero since we know that it corresponds to the full disruption case.}\)
1. If \( a = b = 0 \) (Needs Prioritizing regime), the function changes behaviour at \( \theta_h - \theta_t \), defining the following the cut-off line:

\[
F(\theta_h - \theta_t) + \Delta = \frac{1}{2},
\]

which is continuously increasing in \( \theta_t \) and yields \( \Delta = \frac{1}{2} \) when \( \theta_t = 1 \).

2. If \( a = b = 1 \) (Coverage Prioritizing regime), the function changes behaviour at \( \theta_t \), defining the following the cut-off line:

\[
F(\theta_t + \theta_l) = \frac{1}{2},
\]

which is also increasing in \( \theta_t \) and yields \( \theta = \theta_m \) when \( \Delta = 1 \).

Proof of the two previous statements is provided in the appendix. The interception between these two cut-off lines occurs when \( \theta_h - \theta_t = \theta_l \). Using any cut-off line, it is easy to see that this yields \( F(\theta_h) = \frac{1}{2} \), which is the cut-off line of the Status Quo case. Therefore all three lines intercept at this point.

Finally, regime \( a = 1 \neq b \) is the set in the box that remains undefined: it is the Utilitarian Regime. These regimes are depicted in figure 3.

Insert figure 3 here

It is easy to see that the preference profile of the ex-ante median voter, \( \theta_m \), and hence the skewness of \( F(\theta) \), crucially determine how much of the above box is covered by each regime. The likelihood that a program will lead to a shift in political outcomes increases with positive skewness (\( \theta_m \) tends to zero), and is reduced in societies where the poor form the majority, in line with proposition 2.

Having described the possible regimes, we can now represent the impact of the project on electoral outcomes. First let us define as ‘iso-electoral outcome curves’ level sets such that

\[
L_{\bar{\theta}_m}(h) = \{ (\theta_t, \Delta) \mid h(\theta_t, \Delta) = \bar{\theta}_m \}, \quad \text{where} \quad J(\bar{\theta}_m) = \frac{1}{2}.
\]

Given gradient (16), an iso-electoral outcome map can be easily traced in the above box. Figure 4 depicts this mapping. The grey thin lines represent the iso-electoral outcome levels. Notice how the direction of the gradient changes depending on the regime. For a given budget constraint, and high or low values of the prioritization criterion, \( \beta \), the effect on the size of the shift is small, while it is big for intermediate values of \( \beta \). A Needs Prioritizing agency (following the definitions given in section 3.3) will target fewer, worse-off agents, so that the number of electors shifting electoral preferences is
small. Conversely, a Coverage Prioritizing agency will target a larger set of beneficiaries, providing a low level of the PG. Even if many voters benefit from it, their preferences shift will be small. It is the Benthamite Utilitarian agency that creates the biggest shift: balanced compromise between the level of LPG provided and number of beneficiaries makes several agents ask for an important tax reduction.

Insert figure 4 here

Insert figure 5 here

To be able represent the size of the shift, we need to add a third dimension to this typology representing the ex-post median voter. In Figure 5 the thick iso-electoral outcome curve allows us to explicitly represent \( s \) for a given project \((\theta_t, \Delta)\). These graphical tools will be useful in the next section, where equity and welfare implications are discussed.

## 5 Welfare Implications

In this section we explore how the change in political outcomes affects welfare both at the individual and at the aggregate levels. This distinction is necessary to understand the effect in terms of fairness and in terms of efficiency.

### 5.1 Welfare implications at the Individual Level

To avoid complications linked to the concavity of \( H(\cdot) \) we propose the following approximation. We use the first-order Taylor Series of \( H(\cdot) \) around \( G_m \), and plug it back into the utility function of citizen \( i \), that is

\[
U(c_i, G) = c_i + H(G_m) + H_G(G_m)(G - G_m) + R
\]

where \( R \) is the remainder of the expansion. This remainder barely affects results, we will therefore neglect it. Using the optimal solution of the citizens, the characteristics of the NGO’s project and the outcomes of the elections the utility function at equilibrium values can be approximated to

\[
V_i(y_i, y_m, \bar{y}, s(\Delta, g) \mid b, q) = \left(1 - \frac{G_m - qsG_1}{\bar{y}}\right)y_i + H(G_m) + \frac{y_m}{\bar{y}}q(gb - sG_1)
\]

where \( q \) is an indicator variable valued 1 if the project is implemented.

Form an individual point of view, the project is beneficial if
\[ V_i(y_i, y_m, \bar{y}, s(\Delta, g) \mid b, q = 1) \geq V_i(y_i, y_m, \bar{y}, s(\Delta, g) \mid b, q = 0). \]

That is when

\[ sy_i + y_m(\theta_t b - s) \geq 0 \text{ where } b = 1 \text{ if } \theta_t \in \Theta_b. \]  

(19)

The first term is related to the advantages of a reduced tax burden. The second term evaluates the effect of change in PG and LPG provision for beneficiaries and non-beneficiaries. To go forward with the analysis, we need to state the following lemma:

**Lemma 1** For any shift \( s \) and program \( \theta_t \) we have \( \theta_t \geq s \): the reduction of the centrally provided PG is never larger than the LPG received. In particular, under the coverage prioritizing regime, \( \theta_t = s \).

This is proven in Appendix.

**Proposition 3** Effect of the policy at the individual level: Given Lemma 6.1, Figure 4 and equation (19),

1. Since \( \theta_t \geq s \) any beneficiary is better-off with the project than without it.

2. Non-beneficiaries who are poorer (richer) than the ex-ante median voter are unambiguously worse-off (better-off) as a result of the project. The poorer (richer) an agent, the stronger the impact of the shift.

Notice that it is in the interest of rich agents to allow and encourage decentralised aid projects, even when they do not benefit directly from them. At the same time, it is the poorer non-beneficiaries that would be most sanctioned by the externality caused by these projects. The question remains as to whether in terms of aggregate welfare it is worthwhile that some agents endure this externality. This is the question that we address below.

### 5.2 Welfare Implications at the Aggregate Level

Social Welfare Analysis requires that we compare the impact of the project compared to the situation in which the same budget is allocated to the government via official development aid. Let us first explore the impact of aid in form of direct budget support. Suppose that, instead of implementing a decentralized project, the agency donates its budget \( B \) to the government.
The government’s fiscal balance becomes then $G = \tau \bar{y} + B$. For any voter $i$, the new optimization programme is

$$\max_G \left(1 - \frac{G - B}{\bar{y}}\right) y_i + H(G)$$

The bliss point for agent $i$ is $G^*_i = H^{-1}_G(y_i/\bar{y}) = G_m$, which is identical to the ex-ante choice (equation (3)) and yields therefore the same electoral outcome. The only perceptible difference is a reduction in taxes paid per individual.

In terms of social welfare, aggregate utility becomes

$$W(G_m) = \bar{y} + H(G_m) - G_m + B$$

In this context, budget support is equivalent as a lump-sum transfer that has no political repercussions. This extreme level of aid fungibility is unrealistic, but it specifies the context where decentralized aid should be the most desirable because it earmarks more fund to specific beneficiaries (Collier and Dollar, 2004). In other words, if budget support policies dominate decentralized aid in this context, they should perform even better with lower levels of fungibility.

Using the same Taylor Series approximation than in the previous section, we can state that the project is beneficial for society if

$$\int_Y V_i(y_i, y_m, \bar{y}, s(\Delta, g) \mid b, q = 1)dF(y) \geq \int_Y V_i(y_i, y_m, \bar{y} \mid q = 0)dF(y) + B$$

The left hand side of this inequality is

$$\bar{y} - G_m + sG_1 + H(G_m) + \left(\int_{\Delta} (g - sG_1)dF(y) + \int_{1-\Delta} (-sG_1)dF(y)\right) \frac{y_m}{\bar{y}}$$

$$= \bar{y} - G_m + sG_1 + H(G_m) + (\Delta g - sG_1)\frac{y_m}{\bar{y}}$$

(21)

By substracting (20) from (21) we can establish which form of aid is preferable. Decentralized aid should be preferred if

$$(\Delta g - sG_1)\frac{y_m}{\bar{y}} + sG_1 - B > 0$$

(22)

For decentralized aid to be unambiguously preferred to direct budget support it is required that $\Delta g \geq sG_1 \geq B$. In words, the size of the shift has to be small enough so that the benefits (beneficiaries times LPG provided)
of the project are larger than the reduction of PG provision. Simultaneously, the extent of the tax reduction has to be larger than the budget used by the agency. Notice that the project might be beneficial in terms of welfare if only one of the inequalities holds: either the project brings benefits in itself ($\Delta g > sG_1$) or it brings benefits in the form of tax reduction ($sG_1 > B$).

Supposing that inequality (22) holds, it is possible to find a set of optimal prioritization criterion $\beta$ for the agency. Knowing that the shift $s$ varies with $\beta$, denote its first derivative $s'(\beta)$. Using the optimal solution of the agency in (9), the first order condition for optimality is

$$
\left( (1 - 2\beta^*)\Delta_c\theta_c - s'(\beta) \right) \frac{y_m}{y} + s'(\beta) = 0,
$$

which leads to

$$
\beta^* = \frac{1 + s'(\beta^*) \left( 1 - \frac{\bar{y}}{y_m} \right)}{2}
$$

(23)

where, using the graphic intuition provided in the previous section,

$$
s'(\beta^*) < 0 \quad \text{if } \beta^* \gg \frac{1}{2},
$$

$$
s'(\beta^*) > 0 \quad \text{if } \beta^* \ll \frac{1}{2}.
$$

(24)

We can thus state the following proposition

**Proposition 4** By condition (23) there are different types of local optima according to the regimes described in Figure 3 and in ex-ante conditions. These are:

1. In the Benthamite Utilitarian regime, the optimal prioritization is criterion is $\beta^* = \frac{1}{2}$. This maximizes welfare but also the extent of the externality. It is the only maximum if income distribution is symmetrical (if $y_m = \bar{y}$).

2. If rich are the majority ($\bar{y} < y_m$) there exists a local optimum $\beta^* \ll \frac{1}{2}$ in the Coverage-prioritizing regime.

3. If poor are the majority ($\bar{y} > y_m$) there exists a local optimum $\beta^* \gg \frac{1}{2}$ in the Needs-prioritizing regime.

When $\beta^* \approx \frac{1}{2}$ we have by Figure 4 that $s'(\beta) = 0$, so that the shift is maximized. We therefore have a local optimum that maximizes aggregate welfare but also the externality. This creates an opposition between efficiency and fairness.

Interestingly, this opposition emerges under assumptions of sound institutionality. The model thus reconciles the idea that in contexts with good
institutions, aid can promote economic efficiency, hence growth (Burnside and Dollar, 2000; Collier and Dollar, 2004; McGillivray et al., 2006) and simultaneously aggravate inequalities (Layton and Fuller, 2008; Herzer and Nunnenkamp, 2012; Bjørnskov, 2010).

At the same time, the distance from the average bliss point to the median bliss point increases. Consistent with the definition of representativeness that we used in section 3.1, we interpret this as a degradation of governance in line with Rajan and Subramanian (2007) and Djankov et al. (2008).

Combine now propositions 2 and 4. The first proposition describes how the shape of the income distribution affects the size of the shift. We have seen that if a population is homogeneous or has a large poor majority, the size of the shift is small. There is no much harm done in implementing a utilitarian policy; projects aiming at this type of countries can more efficiently direct their efforts at maximizing aggregate utility. Conversely, in polarized or very unequal societies targeting the slender middle class can potentially lead to political turmoil. Notice that in equation (23) when poor are the majority in a polarized society, the Needs Prioritization can hold a local social optimum with $\beta^* \ll 1/2$. This choice reduces the extent of the externality, helps those in need and is socially optimal (locally): humanitarian missions can be a particularly useful instrument in unequal societies.

The model provides theoretical foundations for pro-poor projects in developing countries that go beyond ethical foundations or marginally decreasing returns: they can simultaneously represent an efficient outcome in economic terms and contain political externalities hazardous to vulnerable agents.

6 Conclusion

This paper has developed a Downsian model of public finances interacted with the provision of public services by an NGO. We assume that NGO projects are substitutable to governmental provision of public goods. Under a very general setting we find that decentralized aid would lead to a reduction of state-provided public goods. The reason is that aid makes beneficiaries willing to support lower-taxation platforms during elections. This crowding-out translates into an externality for non-recipients of aid, implying that poor non-recipients are made worse-off. We find that this crowding-out effect is smaller in homogeneous societies with large poor majorities compared its effect in unequal or polarized societies. We model NGOs as confronted to a trade-off between prioritizing need or coverage. Depending on its optimal choices, we characterize a variety of outcomes affecting individual and aggregate welfare. We find that agencies balancing coverage and needs (thus
maximizing aggregate utility) have the highest impact both in improving aggregate welfare and provoking state disengagement. Coverage-prioritizing agencies might improve welfare, but only through the alleviation of the tax burden, while the project itself does not hold much interest for beneficiaries. Finally, Needs-prioritizing agencies, which cater to the poorest despite the additional risks and costs involved, can improve welfare and limit the extent of the inequalities created, providing a good compromise between efficiency and fairness. Combining the effects of income distribution and type of project, the paper concludes that pragmatism should be addressed to homogeneous societies with large poor majorities and that a stronger humanitarian accent should be put on unequal or polarized societies.

The model rests on deliberately simplifying assumptions to show that, even without institutional failures at the government level, decentralized aid can harm the poor. It sheds light on the problem of lack of coordination between NGOs and government. The crowding-out problem can be easily overcome if NGOs provide services that do not exceed what the government should provide, but this requires substantial coordination between agencies and government. The main limits of the paper are two-fold. First, it is too optimistic on the institutional side. These naive assumptions are made to isolate the effect described in this paper from other institutional failures. Secondly, it uses a partial equilibrium setting. In general equilibrium, NGOs might develop human capital or increase productivity, improving overall market conditions. Conversely, agents might react to the externality through revolt and protest, harming economic performance and political stability. These considerations should be taken into account for further developments of the model.
References


Howes, S., Otor, S., and Rogers, C. (2011). Does the world bank have a micro-macro paradox or do the data deceive? Development Policy Centre Discussion Papers 1105, Development Policy Centre, Crawford School of Public Policy, The Australian National University.


Appendix A: Robustness Checks

We show that our main result, that is decentralized aid crowds-out government spending, is robust to most generalizations.

Generalized Utility Function

Suppose a citizen maximizes $U_i(c_i, G, g)$ subject to $G = \tau \bar{y}$ and $c_i = (1 - \tau)y_i$ and benefits from a project sized $g$. The associated Lagrangian is:

$$L(c_i, G) = U_i(c_i, G, g) + \lambda \left( c_i - \left( 1 + \frac{G}{\bar{y}} \right) y_i \right)$$

The marginal rate of substitution of the public good to get an additional unit of private consumption evaluated is

$$MRS_{CG}(g) = \frac{y_i}{\bar{y}}$$

rich agents are more willing to substitute public goods for private consumption. Using the the marginal utilities of $c_i$ and $G$ we can explore how this rate changes with $g$:

$$\frac{dMRS_{CG}(g)}{dg} = \frac{1}{(U_i/G)^2} \left( \frac{\partial^2 U}{\partial c_i \partial G} \frac{\partial U}{\partial G} - \frac{\partial^2 U}{\partial G \partial g} \frac{\partial U}{\partial G} \right) > 0$$

If we assume utility supermodular in $c, g$ and submodular in $G, g$, that is if the project complements private consumption and substitutes state-provided public goods. We could imagine a converse case, where an advocacy NGO encourages state provision of public good and discourages consumption of some private goods.

NGOs internalize their impact

Suppose that NGOs realize they are exerting an externality and wish to limit the political impact of their actions. They decide to satisfy needs of people only above what the state provides to begin with. Suppose that the target of the NGO is now $\theta_t = \tilde{\theta} + \theta_d$ where $\tilde{\theta}$ is the ex-post median voter and the NGO only complements part $\theta_d$ of the needs. The NGO problem becomes:

$$\max_{\theta_d, \Delta} \Delta(\tilde{\theta} + \theta_d)^{\beta/1-\beta}$$

s.t $B \geq c_g \theta_d G + c_\Delta \Delta$. 

28
The NGO targets then $\theta_t = (1 - \beta)(\theta_C + \tilde{\theta})$ which is higher than what it would have targeted would it have ignored the provision by the state. Beneficiaries new problem becomes:

$$\max_G \left( 1 - \frac{G}{\bar{y}} \right) y + H(G(1 - \beta) + g).$$

This solves to

$$G^*_i = \frac{1}{1 - \beta} \left[ H^{-1}_G \left( \frac{y_i}{\bar{y}(1 - \beta)} \right) - g \right]$$

which is decreasing in $g$. Notice, by using specification $H(\cdot) = \ln(\cdot)$ and by denoting $G^{TA}_i$ the bliss point of agent $i$ when the agency internalised political impact and $G^{NA}_i$ when it does not, optimality conditions yield

$$\frac{1 - \beta}{g + G^{TA}_i(1 - \beta)} = \frac{1}{g + G^{NA}_i} = \frac{y_i}{\bar{y}}$$

$$\Leftrightarrow G^{NA}_i = g \frac{\beta}{1 - \beta} + G^{TA}_i.$$

This result states that there can be a greater crowding out effect when the NGO takes into account its political impact. Because voters know that the NGO is willing to fill the gaps where the state does not intervene, they deliberately create these gaps.

**Progressive taxation or informal sector**

We relax the assumption that all agents pay a flat tax. Assume that the tax rate any voter $i$ is proportional to its income. Similarly, this can represent populations with an important informal sector, assuming that income and formality are positively correlated.

For simplicity assume any agent budget constraint is

$$c_i = (1 - \tau y_i) y_i$$

where $\tau y_i \leq 1$. Fiscal balance becomes

$$G = \int_y \tau y_i^2 f(y_i) = \tau (\bar{y}^2 + \sigma^2)$$

where $\sigma^2$ is the variance of income distribution. Each voter maximizes then
Max $G \left( 1 - \frac{G y_i}{\bar{y}^2 + \sigma^2} \right) y_i + H(G + g)$

Yielding

$$G_i^* = H^{-1}_G \left( \frac{y_i^2}{\bar{y}^2 + \sigma^2} \right) - g.$$ 

which has the same properties than the original model, however lower-income and upper-income political positions become exacerbated.

**Appendix B: Proofs of Theorems and Propositions**

**Theorem 1**

Function (13) can be constructed piecewise as follows:

1. Piece 1 is defined over $\Theta_a$ therefore $a = 1$ and $b = c = 0$. It includes non-recipients and ex-post recipients

$$J(\theta' | \theta' \in \Theta_a) = \int_0^\theta f(\vartheta)d\vartheta + \int_{\theta_t - \theta}^\theta f(\theta_t + \vartheta)d\vartheta = F(\theta) + F(\theta_t + \theta) - F(\theta_t)$$

2. Piece 2 depends on whether $\theta_h - \theta_t < \theta_l$ or not. If $\theta_h - \theta_t < \theta_l$ then $a = b = c = 0$, therefore:

$$J(\theta | \theta' \in [\theta_h - \theta_t; \theta_l]) = F(\theta_h - \theta_t) + F(\theta_h) - F(\theta_l) + \int_{\theta_h - \theta_t}^\theta f(\vartheta)d\vartheta = F(\theta) + \Delta$$

If $\theta_h - \theta_t > \theta_l$ then $a = b = 1$ and $c = 0$, therefore:

$$J(\theta | \theta' \in [\theta_l; \theta_h - \theta_t]) = F(\theta_t + \theta_l) + \int_{\theta_t}^\theta f(\theta_t + \vartheta)d\vartheta = F(\theta_t + \theta)$$

3. Piece 3 is defined over $\Theta_b \setminus \Theta_a$ (thus $a = c = 0$ and $b = 1$) which is empty ex-post. We get

$$J(\theta | \theta' \in \{\Theta_b \setminus \Theta_a\}) = F(\theta_h)$$
4. Piece 4 is defined over $\Theta_c$ (therefore $c = 1$ and $a = b = 0$) such that

$$J(\theta | \theta \in \Theta_c) = F(\theta_h) + \int_{\theta_h}^{\theta} f(\vartheta) d\vartheta = F(\theta)$$

Equation (13) syntheses these four pieces under a single expression through the subset-defined system of binary variables.

**Comparative Statics for project vector $p$**

Equation (16) can be rewritten as

$$\nabla s(\theta_t, \Delta) = \begin{bmatrix}
af(\theta_t + \hat{\theta}_m) + (b - a)f(\theta_t)(1 - \Delta) \\
af(\theta_t + \hat{\theta}_m) + (1 - b)f(\hat{\theta}_m) \\
1 - a + (a - b)F(\theta_t) \\
af(\theta_t + \hat{\theta}_m) + (1 - b)f(\hat{\theta}_m)
\end{bmatrix}.$$  

Remark that if $a = 1$ and $b = 0$ the sign of the numerator of the first element of the gradient seems ambiguous. However we can prove that

$$f(\theta_t + \hat{\theta}_m) - f(\theta_t)(1 - \Delta) > 0 \quad \text{(A1)}$$

Assume $f(\theta)$ is unimodal, and denote its mode $\theta^M$, such that $f'(\theta) > 0$ if $\theta < \theta^M$ and $f'(\theta) < 0$ if $\theta > \theta^M$. If $\theta_t + \hat{\theta}_m \leq \theta^M$, (A1) holds without ambiguity.

Suppose now that $\theta_t + \hat{\theta}_m > \theta^M$. Since we are in the regime $a = 1$ and $b = 0$ we have $\hat{\theta}_m \leq \min\{\theta_t, \theta_h - \theta_t\} \Rightarrow \theta_t + \hat{\theta}_m \leq \theta_h$. By assumption (10) if $\Delta = 0$ then $\theta_t = \theta_h$: the inequality holds for $\Delta = 0$.

Taking its first derivative with respect to $\Delta$ we get

$$f'(\theta_t + \hat{\theta}_m) \frac{\partial \hat{\theta}_m}{\partial \Delta} + f(\theta_t) > 0$$

Since $\partial \hat{\theta}_m / \partial \Delta = \partial(\theta_m - s) / \partial \Delta = -\partial s / \partial \Delta < 0$ by (16). Therefore (A1) holds for any $\Delta$. 

31
Proof of Proposition 2

Take two functions denoted \( F(\theta) \) and \( \hat{F}(\theta) \) such that

\[
\begin{align*}
F(\theta) &< \hat{F}(\theta) \quad \text{if} \quad 0 < \theta < \vartheta \\
F(\theta) &> \hat{F}(\theta) \quad \text{if} \quad \vartheta < \theta < 1 \\
F(0) &= \hat{F}(0), \quad F(1) = \hat{F}(1), \quad F(\vartheta) = \hat{F}(\vartheta)
\end{align*}
\]

Where \( \vartheta \) is some value where these two functions intercept. Let \( \hat{J}(\theta) \) be the ex-post CDF derived from \( \hat{F}(\theta) \). By equation (15) we have

\[ J(\theta) \leq \hat{J}(\theta) \quad \text{for all} \quad \theta \leq \vartheta \]

Let \( \hat{s} \) be the shift associated to \( \hat{F}(\theta) \) such that \( \hat{J}(\theta_m - \hat{s}) = \frac{1}{2} \). We have:

\[
\begin{align*}
J(\theta_m - s) &= \hat{J}(\theta_m - \hat{s}) \quad \text{by equation (15)} \\
\Rightarrow J(\theta_m - s) &\geq J(\theta_m - \hat{s}) \quad \text{if} \quad \theta_m - \hat{s} \leq \vartheta \\
\Leftrightarrow s &\leq \hat{s} \quad \text{since} \quad J(\cdot) \text{ non-decreasing}
\end{align*}
\]

The crossing point \( \vartheta \) will allow us to understand the impact of the second and third moments of the distribution.

1. Suppose \( \vartheta = \theta_m \) and both \( F(\theta) \) and \( \hat{F}(\theta) \) are S-shaped. Then \( \hat{F}(\theta) \) is more dispersed than \( F(\theta) \). Since \( \hat{F}(\theta) \) is associated with a larger shift, more variance leads to a larger crowding-out effect.

2. Suppose \( \vartheta = 1 \). Then \( \hat{F}(\theta) \) is more positively skewed than \( F(\theta) \). Therefore positive skewness leads to a bigger shift.

Behaviour of the Cut-off lines

By assumption 2.3 we have \( \frac{\partial \theta_h}{\partial \theta_t} > 0, \frac{\partial \theta_l}{\partial \theta_t} > 0 \) and \( \frac{\partial \theta_l}{\partial \Delta} < 0 \). Remark that the cut-off line between the NP and UB regimes is such that \( F(\theta_h - \theta_t) + \Delta = F(\theta_m - s) + \Delta \leftrightarrow \theta_h = \theta_m - s + \theta_t \). By appendix D this implies that

\[
\frac{\partial \theta_h}{\partial \theta_t} = \frac{f(\theta_t)(1 - \Delta)}{f(\theta_h)} < 1
\]

Properties of \( F(\theta_h - \theta_t) + \Delta = \frac{1}{2} \)

By the implicit function theorem and allocation rule (10) the important characteristics are:

\[ \frac{\partial \Delta}{\partial \theta_t} = f(\theta_h - \theta_t) \left( 1 - \frac{\partial \theta_h}{\partial \theta_t} \right) > 0 \]

32
\[ \theta_t = 0 \iff F(\theta_h) = \Delta \iff \Delta = \frac{1}{4} \]

\[ \theta_t = 1 \iff \theta_h = \theta_t \iff \Delta = \frac{1}{2} \]

Properties of:
\[ F(\theta_t + \theta_l) = \frac{1}{2} \]

Are the following:
\[ \frac{\partial \Delta}{\partial \theta_t} = -\frac{f(\theta_t + \theta_l)(1 + \partial \theta_t / \partial \theta_l)}{f(\theta_t + \theta_l)(\partial \theta_l / \partial \Delta)} > 0 \]

\[ \Delta = 0 \iff \theta_t = \theta_l \iff \theta_t = \frac{\theta_m}{2} \]

\[ \Delta = 1 \iff \theta_l = 0 \iff \theta_t = \theta_m \]

**Proof of Lemma 1**

Consider the three following regimes:

1. Coverage prioritizing regime: the shift is defined by \( F(\theta_t + \theta_m - s) = F(\theta_m) \iff \theta_t = s \).

2. Benthamite utilitarian regime: in figure 3 it lays to the right of the previous regime. Notice that in appendix D it is shown that \( 0 < \partial s / \partial \theta_t < 1 \). Therefore as \( \theta_t \) increases departing from the coverage prioritization scenario, \( s \) increases less than proportionally. Therefore \( \theta_t > s \).

3. Needs prioritizing regime: by the previous case we know that at the cut-off line between the Needs and Utilitarian regimes \( \theta_t > s \). Any displacement downwards implies a reduction in \( s \) since at the needs regime \( \partial s / \partial \Delta > 0 \). Therefore at any point below the cut-off line \( \theta_t > s \).
Figure 1: Ex-ante and ex-post density distributions.

Figure 2: Ex-ante and ex-post cumulative density distributions.
Figure 3: Delimitations by regime.

Figure 4: Iso-electoral outcome curve and optimal choice of the NGO.
Figure 5: Size of the shift relative to characteristics of the project.