Can good projects succeed in bad communities?

Asim Ijaz Khwaja[*]
Harvard University
December 2007

Abstract

The lack of "social capital" is increasingly forwarded as an explanation for why communities perform poorly. Yet, to what extent can these community-specific constraints be compensated? I address this question by examining determinants of collective success in a costly problem in developing economies - the upkeep of local public goods. One difficulty is obtaining reliable outcome measures for comparable collective tasks across well-defined communities. In order to resolve this I conduct detailed surveys of community-maintained infrastructure projects in Northern Pakistan. The findings show that while community-specific constraints do matter, they can be compensated by better project design. Inequality, social fragmentation, and lack of leadership in the community do have adverse consequences but these can be overcome by changes in project complexity, community participation and return distribution. Moreover, the evidence suggests that better design matters even more for communities with poorer attributes. Using community fixed effects and instrumental variables offers a significant improvement in empirical identification over previous studies. These results offer evidence that appropriate design can enable projects to succeed even in "bad" communities.

JEL Classification: O12; Z13; D7; H41; D02; D23
Keywords: Public goods; Development; Collective action; Social capital; Institutional design

---

[*] Harvard University, Cambridge MA 02138, USA. Email: akhwaja@ksg.harvard.edu. I would like to thank Alberto Alesina, Abhijit Banerjee, Alexandra Cirone, Jishnu Das, Esther Duflo, Oliver Hart, Brian Jacob, Sehr Jalal, Michael Kremer, Atif Mian, Carolina Sanchez, T.N. Srinivasan, Chris Udry, Jeffrey Williamson and seminar participants at Harvard, KSG, NEUDC (Cornell), UC-San Diego, U-Michigan, UT-Austin, Williams College, and Yale for comments. I gratefully acknowledge financial support from the Social Science Research Council and the MacArthur Foundation and am indebted to the Aga Khan Rural Support Program Pakistan staff for providing support, feedback and company during the fieldwork. All remaining errors are my own.

I. Introduction

In the past decade there has been a mushrooming of literature on social capital and its role in economic development. Whether it is civic engagement in community groups affecting government quality across regions in Italy (Putnam, 1993), ethnic fragmentation in US cities determining the provision of local public goods and services (Alesina, Baqir and Easterly, 1999), or the level of trust in a country influencing its growth, judicial quality and corruption (Knack and Keefer, 1997; La Porta et al., 1997), there is increasing emphasis that such group-specific factors are important (see Baland and Platteau 1996; Durlauf and Fafchamps, 2004; Alesina and La Ferrara, 2005 for reviews).

Yet how essential are such group-attributes in determining collective success? In particular, can they be compensated by better design? The literature on social capital has largely ignored this question. There are several empirical challenges posed in trying to answer it. First, one has to have comparable collective tasks across groups. Second, reliable outcome measures that reflect collective performance have to be constructed. Third, to the extent that groups can influence task design, identifying the design effect separately from the group's attributes becomes a challenge. Sampling multiple tasks from a given group to exploit within group variation can help, but then one has to ensure group boundaries are well-defined. Finally, one is confronted with a whole spectrum of possible social capital measures. This paper addresses these challenges and attempts to fill this gap in the literature by examining the relative importance of design versus group-specific factors in determining collective success. The results are especially promising in that they show that even communities with low social capital can do well in collective tasks if these tasks are well designed.

Specifically, I collect primary data to show that design factors can indeed compensate for "poor" group attributes. The particular context is the upkeep of community-maintained public infrastructure projects in rural communities in a developing economy, Pakistan. Restricting analysis to similar infrastructure projects provides comparable collective tasks across communities. A further restriction is to projects whose upkeep is solely the communities' responsibility. This, together with detailed technical and functional estimates of project condition, provides reliable collective outcome measures. Finally, wherever possible, multiple such projects are surveyed in a given community, providing the crucial within-community outcome variation needed for identifying project design effects. Community boundaries were carefully identified to make this feasible and ensure that only the selected community was responsible for the projects. A downside is that our final sample – 132 such projects in 99 communities, with only a third having multiple projects – is admittedly small. While such sample sizes are not atypical in this literature given that each observation is a village, and small sample asymptotics do not pose a threat, one may exante nevertheless raise concerns about the statistical precision and generalizability of the findings. As such, care was taken that neither was compromised: The smaller sample stems from a desire for cleaner comparisons and measurement, and is borne out in the resulting precision of the estimates. The substantial variation in attributes across our sampled communities and the similarity of our community attribute estimates to those in the literature reassure that external validity is not a significant concern.

The first, albeit casual, empirical observation is that there is as much variation in collective outcomes *within* a given community as there is *between* communities. This is significant since within-community variation is not readily attributable to social capital based explanations, as social capital is not considered to vary within a community – While there is considerable debate on what constitutes a group's social capital or how it is measured, it is nevertheless regarded as an attribute *specific* to the *group*, whether manifested in the norms, level of trust, or networks present in the group or its degree of inequality and fragmentation.

I take advantage of this substantial within-community outcome variation to identify the effect of project design factors by using community fixed effects. The identification issue is that design factors are not randomly assigned but are chosen or influenced by communities. To the extent that we cannot control for all such community factors that also directly influence collective performance, estimates for the design factors will be biased. For example, if "better" communities choose more challenging design factors, then not accounting for this results in a downward bias in the design factor estimates. The community fixed effect estimation avoids this bias by only comparing design factors *within* the same community group. Therefore the estimates are no longer confounded by unobserved level differences across communities.

The fixed effects specification shows that design factors have large effects: A standard deviation increase in the degree of project *Complexity* (indexed by higher capital/skilled labor requirements and lower experience) of a project leads to a 13 percentage point drop in maintenance (measured on a 0-100 scale).*Community participation* in various aspects (construction, usage and maintenance) of a given project is not always beneficial: While community participation in project *decisions* that are *non-technical* in nature (i.e. selecting project type, usage rules etc.) is beneficial, community participation, in the *technical decisions* that have to be made in the same project (i.e. deciding project design, scale etc.) is detrimental – a 10 percentage points increase in community participation in non-technical decisions raises maintenance by 6 percentage points but the same increase for participation in technical decisions lowers maintenance by 4 percentage points. *Inequality* in the distribution of project returns has U-shaped relationship with maintenance, with initial increases in inequality hurting maintenance but eventually improving it. Projects made as extensions of existing community projects (*continuation projects*) have 42 percentage points higher maintenance levels than new projects. Similarly, communities better maintain (24 percentage points higher) projects initiated by *non-governmental organizations* (NGOs) than those by the local government.

While the above estimates could still be biased due to unobserved project design factors that are correlated with these design variables, all the results do include separate fixed effects for each of the seven types of infrastructure projects considered (irrigation channels, roads, etc.) and a variety of project controls, such as project age, cost, etc. This ensures that we are comparing similar projects and that the estimates indeed reflect design rather than community attributes. Moreover, additional robustness checks are run for a variety of design factors to address potential endogeneity issues.

I then compare these project design estimates to community factors in two ways. First, under plausible assumptions, the estimated community fixed effect itself provides a measure of the aggregate impact of *all* community-specific factors, observed *and* unobserved. This suggests that even a few design factors can compensate for poor group-level attributes. A hypothetical comparison shows that the best-designed project (in terms of the above design factors) in the "worst" community does just as well as the worst-designed project in the "best" community. Best and worst communities are simply those with the best and worst average performance, i.e., the highest and lowest community fixed effect.

Second, I also estimate the impact of particular community-specific factors emphasized in the social capital literature: A standard deviation increase in *social heterogeneity* (determined by heterogeneity in clan, religious, and political groups) adversely effects maintenance by lowering it 5 percentage points. *Land inequality* in the community has a U-shaped effect, and the presence of a *project leader* is associated with 7.5 percentage points higher maintenance. It is worth noting that the leadership effect, while anecdotally present in the literature, has not been previously carefully identified. By instrumenting for leadership with attributes of hereditary-leader households such as whether the household has a healthy male member between 25 to 50 years old, this paper is able to do so. While these results are robust and both statistically and economically significant, a comparison of their magnitudes also bears out the earlier claim that their impact can be compensated for by the project-design factors.

Finally, while not the primary objective of the paper, particularly since sample limitations make it difficult to do so, I also examine interactions between community-specific factors and project design. Results suggest that the communities with poorer attributes are even more sensitive to project design: Communities with higher inequality respond even more positively to both lowering project complexity and if the project is made as an extension of an existing project rather than as a new one. The impact of community participation is dampened in unequal communities – not surprising once we recognize that greater inequality may translate into less representative and hence less effective member participation. Similarly, communities that lack leadership are more sensitive to project-design factors such as project complexity.

Together these results show that even communities with poor attributes such as low social capital can perform well if tasks are appropriately designed. In fact, the importance of project design holds even more so for communities with poorer attributes and lower social capital.

Finally, consider the potential concerns regarding generalizability of these results. Since our preferred estimates for the design factors are from the smaller sample of communities with multiple projects that meet our selection criteria, could these communities be qualitatively different from those where we could not sample multiple comparable projects? Our results show that this is not the case. Not only do these two sets of communities look very similar along various characteristics, but OLS estimates are similar in both — for example, it is the community fixed effect that changes estimates not the sample restriction. Moreover, since the sampled communities are randomly selected from essentially the universe of communities in the area of study, our sample is representative of this region. Finally, should one be concerned that communities in this region differ in a way that would affect generalizing the results? While this cannot really be tested without analyzing data from other regions/countries, such concerns are unlikely: The sampled region has comparable community characteristics to rural communities in other developing countries and our estimates on community factors such as inequality and heterogeneity are similar to those from studies in other regions.

The broader import of this work is significant: At a time when the literature is concerned with the importance of group-level factors (such as a group's social capital) in determining group performance, this paper observes that not only is there significant variation in performance within groups, but that this variation can be explained by features related to the design of the task. In doing so it also relates to on an older literature that emphasized appropriate design/mechanisms as a means of solving collective problems (Hirschman, 1967; Clarke, 1971; Groves, 1973; Stewart, 1978; Betz et al., 1984). More generally, it shows that while there may be reason to be concerned about declining social capital in the US (Putnam, 1995, 2000) or ethnic fragmentation in Africa (Easterly and Levine, 1997), there is also reason to believe that careful design can compensate for these poor group attributes. Moreover, the specific problem examined in the paper is important, as public investments in developing economies rarely last their expected lifetimes. Estimates by multilateral development agencies show that in the last decade alone $12 billion in regular maintenance expenditure could have prevented the $45 billion spent on road reconstruction in Africa (World Bank, 1994).

This paper also makes several specific contributions by presenting cleaner estimates and novel results for determinants of collective performance.[1] The use of group-level fixed effects allows significantly better identification of the impact of design factors on collective performance. The paper also presents one of the few clean empirical estimates of the effect of leadership on group performance (see Jones and Olken 2005 for an alternate instrumentation strategy). In addition, existing work on determinants of collective success in local public projects prevalent in developing economies has consisted mostly of anecdotal evidence and case studies and only recently, econometric analyses of collective action determinants (Wade, 1987; Lam, 1998; Dayton-Johnson, 1999; Agarwal, 2001; Miguel, 2001). In contrast to these studies (Olken 2005 is a notable exception), this paper not only develops detailed outcome measures but also exploits outcome and project design variation within communities.

---

[1] While the result on community participation not being beneficial in all project decisions is novel, this is the focus of a previous article and so is not stressed here. This shorter article (Khwaja 2004) develops a theoretical model to explain how participation in all project decisions may not be beneficial, and then shows how the community participation results support the model. The shorter article does not examine any of the other results on project design or community factors detailed in this paper, nor compare the relative import of the two, the focus of this paper.

II. What determines collective success?

What are all the factors that affect how a group is able to work together and achieve success in collective tasks? While this question continues to elicit considerable interest, the goal of this paper is more modest: to examine the relative importance of particular types of factors. Specifically, how design factors compare to those that are considered to be inherent attributes of a group, such as its "social capital". This section introduces a simple framework to help structure this question and explain the empirical methodology used. A more formal treatment of how particular community and design features influence collective success is developed in Khwaja (2006).

In order to motivate the distinction between design and group-invariant factors, note that while the extensive literature on social capital may not have reached a consensus on how to define let alone measure social capital, generally all these definitions view a group's social capital as inherent/specific to that group at a given time.[2] Putnam attributes the improved performance of Northern over Southern Italy to the attributes specific to the former, such as higher participation in associational activities that are reflective of greater social capital. Similarly, when Alesina and Le Ferrara (2000) consider the determinants of social capital in US communities, they examine the extent to which this is determined by community-level income inequality and ethnic heterogeneity.

Whatever factors determine a group's social capital,(to the extent that they are specific to the group), the distinction this paper makes between group-specific factors and those such as task design that vary *within* a given group helps inform us of the relative importance of social capital. For example, the level of income inequality in a group at a given point in time is the same for the group regardless of what collective task the group performs. Thus inequality is a group-level attribute and potentially part of what constitutes the group's social capital. In contrast, the degree of complexity of a collective task is a potential determinant of the groups' collective success that is not specific to the group but can vary across different tasks taken up by the group. The eventual goal is to evaluate how large an effect design factors have relative to potentially adverse group-level factors such as inequality, socio-economic fragmentation, weak leadership, and low wealth/skills, etc. on a group's collective performance. Are groups with poor social capital at a severe disadvantage, or can these factors be compensated through task design that may be more readily altered for a given group?

Moreover, most of the literature that emphasizes group-level attributes as determinants of collective success posits such attributes have a direct/additive effect on outcomes (see Durlauf and Fafchamps 2004 for

---

[2] Putnam [1995] views social capital as the network of social relations that link a group of agents and facilitate coordination and cooperation. The World Bank defines social capital as including, but not limited to social networks and associated norms. Coleman [1988] considers it as social structures that facilitate actions and not within actors nor in physical implements of production. Frank [1993] treats it as the amalgam of formal and informal mechanisms for

a detailed review): Whether the group-level determinant of interest is group inequality, social heterogeneity, size, wealth, skills, or any other group demographic, it is assumed to directly raise/lower the outcome, regardless of the collective task. Papers that find different levels of trust in countries impact their growth (Knack and Keefer 1997), or ethnic fragmentation in US cities affects the provision of local public goods and services (Alesina, Baqir and Easterly 1999) all allude to or estimate such a specification. This will dictate the primary specification used in this paper as well, and the additional contribution is to recognize and separately identify the impact of design factors.

In addition, to the extent that group-factors interact with design in determining collective success, this paper will also estimate specifications that allow for such interactions. While sample size considerations limit such results, they also have implications for evaluating the relative importance of design and group factors. Specifically, if we find that the impact of group-specific factors varies by task design, then the importance of these group-specific factors, such as social capital, potentially diminishes. This is especially so if, as our results will show, design factors matter even more so for groups with lower social capital.

Methodology

One of the primary empirical challenges this paper deals with is to identify the impact of (project/task) design factors and to compare them to the impact of group/community-specific factors such as its social capital. The following specification, based on the previous discussion, makes this distinction and identification issues clear:

$$Y_{ij} = \alpha + \mathbf{P_{ij}}\beta + \mathbf{X_i}\gamma + \eta_i + \varepsilon_{ij} \tag{1}$$

Y_{ij} is the outcome of a particular collective task/project j undertaken by group/community i. $\mathbf{P_{ij}}$ is a set of specific attributes of the collective task (henceforth "design" factors), $\mathbf{X_i}$ is a vector of group-specific factors, η_i represents unobserved group characteristics and ε_{ip} is a general error term.[3]

Since it is likely that groups choose or influence design factors, a simple OLS estimate of (1) would yield biases estimates of the design factors $\mathbf{P_{ij}}$ i.e. Corr($\mathbf{P_{ij}}$, η_i) $\neq 0$. The direction of the bias depends on how groups influence design – if "better" groups choose/are assigned more difficult (easy) design factors then OLS estimates of the design factors would be downward (upward) biased.

Our strategy to address this bias is to estimate specification (2) that introduces fixed effects at the community/group-level:

$$Y_{ij} = \alpha_i + \alpha_{j-type} + \mathbf{P_{ij}}\beta + \varepsilon_{ij} \tag{2}$$

---

solving collective action problems. See Durlauf and Fafchamps [2004] for a detailed discussion on the range of definitions.

[3] A more general specification would also allow for variables to vary over time. However, given our analysis makes use of cross-sectional data, we simplify by ignoring the time dimension. The specification does allow for non-linear effects and interactions within the two types (group or task-specific) of factors i.e. both inequality and its squared terms can be included in $\mathbf{X_i}$, as well as interactions between group inequality and group wealth etc.

where α_i is the group/community fixed effect[4] and $\alpha_{j\text{-}type}$ is a separate intercept for each type of collective task. To the extent that the relevant omitted variables are at the group level, this specification will provide unbiased estimates of the design factors by ensuring that such factors are only compared *within* a given community.[5] By not comparing across communities we are therefore able to deal with the concern that different communities may pick different design factors.

Note that the specification also includes separate fixed effects for the type of the collective task where type is one of seven different types of infrastructure projects ranging from irrigation channels to roads (see Table 1). This partly allays further identification concerns that arise from unobserved factors specific to the collective task/project that directly affect performance and are correlated with the design factors considered. For example, it could be that road projects are generally harder to maintain than irrigation channels but also have more capital requirements. Not taking this into account would result in an upward bias in the estimate of the capital requirement design factor. In addition, while we only examine infrastructure projects, one may still be worried that outcome measures may differ systematically across project types. By forcing our comparisons only *within* a given project-type we eliminate such concerns. We also include a variety of other project level attributes such a project age, construction costs, etc. as controls. Finally, we run further checks in cases where we still suspect particular identification concerns. For example, in the community participation design factor, halo effects are a potential concern: for example, community members may (falsely) report higher participation in more successful projects. We directly address this by using individual level data to show that halo effects are unlikely.

The inclusion of project-type fixed effects also improves (policy) interpretation of our design factor results. In contrast to the type of a project which is determined more by need, our design factor estimates are for task-specific features that can be varied for any given project type, such as the extent of group participation in task decisions etc. These factors can be manipulated while addressing a given need and therefore have more policy relevance. For example, it makes little sense to argue that if irrigation projects generally perform better than road projects, one should only "design" the former. However, if one finds that allowing for simpler project design or soliciting greater group participation improves performance, then such design improvements can be introduced in any project type.

Estimating (2) poses two measurement challenges. First, one has to ensure that more than one collective outcome be measured for the *exact* same group – otherwise using group fixed effects will not address the

---

[4] With two projects per-community, the community fixed effect here is analogous to running a first-difference specification (e.g. the LHS/RHS variables corresponding to project 2 are differenced from those for project 1).

[5] Glaeser, Laibson and Sacerdote (2002) employ a related strategy by examining how little variation group-level (state, PSU, religion) fixed effects explain for the number of types of groups an individual in the US belongs to. However, their motivation is to show that such measures of social capital (i.e. individual group membership) are not only determined by group-level attributes but also reflect individual characteristics. Since our outcome and data is not at the individual level, our interpretation is that such within-group variation reflects differences in the design of the collective task and its interactions with community-level factors.

identification concerns arising from omitted group factors discussed above. Surveys rarely sample multiple collective tasks for a given group. What is even harder is to avoid incorrectly attributing multiple projects to a group, since a closer examination often reveals these projects in fact do not fully overlap in terms of the responsible/beneficiary group. Second, one wants an outcome measure that is of direct and measurable value to the group members and that can be sensibly compared across projects. An advantage of collecting primary data was that great care was taken to address both these concerns. Groups were carefully delineated and multiple projects only sampled if they shared exactly the same group. In addition, restricting to externally initiated infrastructure projects allowed for comparable outcome measures – the *current state* (of maintenance and functionality) of the project – and precise project details such as project costs and identity of the group responsible for project upkeep.

We will discuss the above restrictions in more detail later. A cost that is important to mention here is this restriction leads to a drop in sample size to a third of the original community sample and half of the original sampled projects.[6] While the potential decrease in precision is addressed by reducing measurement error through careful survey work and using multiple sources to construct and triangulate main outcome variables, an empirical concern that remains is whether the sample drop can produce biased estimates of the design factors in the overall population.

We address this concern in several ways. First, the sample drop, while systematic (we are more likely to obtain multiple projects of the required types in larger communities), is unlikely to bias our results since the selected communities look remarkably similar across other community and project design measures. Table 2 compares the means of various community and project level measures and shows that communities which have multiple projects meeting our selection criteria really only differ in that they are somewhat larger and have more roughly 1.5 more projects – factors that can be controlled for and which (it turns out) have little impact on project performance. However, along all characteristics such as community inequality and our outcome measures, these communities are not different. Second, we can compare the specification (2) results to specifications without the group fixed effects both in the full and reduced sample. When we do so (Table 3) we find that the design estimates are not affected by change in sample size, but rather (as we would expect) by the inclusion of the group fixed effect.

While the above addresses issues of internal consistency, a remaining concern may be external validity. Since the 99 communities in our sample are randomly selected from essentially the universe of communities in the area of study, a state in Northern Pakistan, our sample is likely to be representative of this region. A

---

[6] Given our restriction to specific types of infrastructure projects, our original sample of 132 projects is already roughly a fifth of all public projects (schools, community halls, public graveyards, etc.) in the sampled communities. However, there is little reason for concern about this restriction since it does not eliminate any community, applies equally across all communities, and is driven more by measurement and outcome comparability concerns. Moreover, a decomposition of outcome variation between and within communities is similar whether one uses the 132 or full 651 public projects.

harder concern to address in this or any similar study is whether communities in this region differ in a way that would affect generalizing the qualitative results – specifically, how important design and group factors may be in determining group success. While it is not possible to directly address this without having similar data from other regions/countries, I argue that the sampled region is likely to be representative of rural communities in developing countries. This is both because observed community characteristics in the sample are comparable to rural communities in other developing countries, and because our estimates of community factors such as community inequality and heterogeneity are similar to those from studies in other regions.

In addition to estimating the impact of design factors, one would also like to compare them to that of group-specific factors. We adopt two strategies for doing so. First, one can compare the estimates of design factors to the magnitude of the group fixed effect estimated in specification (2). Specifically, the group fixed effect can provide a comparison of the "worst" to the "best" group. The difference between these two groups' performance can then be compared to a combination of the design factors. The advantage of this method is that the group fixed effect includes all group-invariant factors, observable and unobservable. Since a lot of factors identified as contributing to a group's social capital may be hard to define or measure, the group fixed effect has the advantage of capturing their combined effect without actually having to define or measure them. A disadvantage is that one cannot separately look at the effect of different group factors.

A related issue is interpretational. To the extent that group factors also influence design, these factors have both a direct and indirect (though design) effect on the outcome. By controlling for design factors, the group fixed effect in specification (2) captures only the direct effect (and any indirect effects due to unobserved design factors). While this is desirable, one may well argue that certain design factors will always be influenced by a group and therefore the indirect effects should also be included in judging the group factors' overall impact. This can be done by estimating the group fixed effect in specification (2) without any design factors included. However, as our results will show, the indirect effects generally work in *reducing* the overall group effects (since better groups tend to pick/be assigned more challenging designs). Since this paper argues that design factors are as important, I prefer to stick to the larger (direct) estimates of the group-factors as this makes the comparison with design factors even more compelling.

An alternate comparison strategy is to estimate the effect of particular group factors that have been emphasized in the (social capital) literature, such as group inequality and socio-ethnic fragmentation. One can do so by estimating specification (1) directly which allows one to include a set of group-specific factors. While this approach has the advantage of estimating the impact of group factors separately, to the extent the group unobservable or design factors in specification (1) are correlated with these group factors, estimates will be biased. Moreover, the direction of the bias is hard to judge a priori. We minimize such concerns but focus on group factors that we believe are more likely to be exogenous. For example, given almost non-existent land markets in our area of study, land inequality is likely to be based on initial settlements patterns in a community and not be affected by more recent factors such as project design or outcomes. In addition, in the case of estimating the impact of leadership, we use instrumental variable techniques to improve identification. However, since this approach only estimates the impact of a few community factors, we prefer the first (community fixed-effect) comparison.

Equation (1) can also be extended by allowing for interaction effects between group and design factors. This is particularly relevant if one believes that group factors may not have the same level effect on each collective task but that the effect may be magnified or attenuated through task design. Since omitted group factors are an issue as in equation (1), I will estimate interaction effects in the reduced sample that allows the inclusion of group fixed effects:

$$Y_{ij} = \alpha_i + \alpha_{j-type} + \mathbf{P}_{ij}\beta + \mathrm{I}_{ij}\delta + \varepsilon_{ij} \tag{3}$$

where $\mathbf{I}_{ij}$ includes interactions between the group-specific factors, $\mathbf{X}_i$ and collective task-specific factors, $\mathbf{P}_{ij}$.. Given the small sample size, a note of caution is that examining such interaction effects in detail is really beyond the scope and not the focus of this paper. The purpose of this exercise is only to highlight a few and fairly tentative interaction effects – particularly those that shed further light on the relative importance of design factors by asking whether they matter even more (or less) in lower social capital groups.

III. Data

This paper uses primary data on infrastructure projects and rural communities collected from Baltistan, a state in the Himalayan region of Northern Pakistan (Figure 1). Given the empirical methodology, it was important to restrict the study-sample to comparable collective tasks for which detailed *outcome and task-specific measures* could be obtained. As a result, the sample is restricted to externally initiated infrastructure projects provided to well-defined groups (usually a cohabitating community like a village/sub-village) and the outcome of interest is the *current state* of maintenance and functionality of the project.

Since measuring collective success for any task is non-trivial, restricting the task to physical projects simplifies matters considerably as reasonably precise and comparable engineering-based measures of current project state/condition are possible. The restriction to externally funded, while not essential, enables one to obtain accurate project cost, age and characteristics data through these external agencies (government, non-governmental organizations). Moreover, the outcome and current state of the project provide an accurate measure of a *group's* collective abilities as compared to other outcomes like project completion/construction success, since the external agencies considered only provided support in initial project construction – its maintenance thereafter was solely the group's responsibility.[7]

A second consideration, highlighted in the methodology above, is to ensure that in cases where multiple infrastructure projects did exist, they were only sampled provided that the *same* group was solely responsible

---

[7] Projects where a group was not solely responsible either because its upkeep was (partially) the state's/other agency's responsibility or because another group also shared in the responsibility were excluded.

for the project's maintenance. Previous studies in the literature rarely sample multiple-tasks for a group or take care in ensuring the underlying group remains the same for such multiple-projects. Although the cost of doing so has been a reduced sample size, this cost is justified by an improvement in measures and estimation strategies.[8]

There are also several features of the environment chosen that make it particularly suitable for the study. First, not only is there a high importance attached to infrastructure projects such as roads and irrigation channels given the remoteness of the region,[9] but harsh weather conditions[10] lead to rapid degradation of projects unless regular maintenance is undertaken. It is therefore common to see projects lying damaged and poorly maintained even within the first year of their construction. As a consequence, the state of a project primarily depends on how much maintenance effort the beneficiary group has exerted and provides a good measure of the group's collective abilities.

Second, the relatively low population density in the region implies that settlements are fairly distinct making it easier to separately identify groups responsible for the upkeep of each project and obtain accurate measures for the attributes of each group. For the most part these groups are the community that forms a village or, in the case of larger villages, a smaller habitation in the village. The communities sampled range from small and remote pastoral settlements of 10 households (with 6-8 members per household) to larger ones with 200 or more households, and altitudes range from 7,000 to 12,000 feet above sea level. Thus in addition to being able to identify distinct groups, there is significant variation across groups/communities, lessening external validity concerns. Of particular relevance is having both remote/small rural communities to which the literature often attributes high social cohesion and larger semi-urban settlements that may have less cohesion. This variation is important given that we want to estimate the degree to which group-level differences affect collective performance and argue that our results are not just specific to this environment.

Finally, the majority of infrastructural projects (irrigation and road projects) involve two main external organizations – the government's Local Bodies and Rural Development (LB&RD) department and a non-governmental organization (NGO), the Aga Khan Rural Support Program (AKRSP).[11] Both initiate similar projects and only provide technical and financial support during construction. After construction, these

---

[8] Nevertheless, in comparison to the empirical literature that looks at such public goods in communities, even a sample of a hundred communities is a relatively large sample.

[9] Two major constraints to development in the region are scarcity of irrigation water and poor road access: With low annual precipitation at 150-200mm (Whiteman, 1985), the primary sources of water for the predominantly agriculture-based economy are glacial melt and rivers. The regional capital, Skardu, offers the only road connection to the rest of the country, but this link is often disrupted for weeks due to frequent landslides. Most villages remain unconnected to Skardu and the rocky and steep terrain makes for very slow movement even on the few metalled roads; a 40-mile journey can take up to 4 hours in a jeep.

[10] Temperatures vary from -20^0 Celsius in the winter to 40^0 Celsius in the summer. Floods, landslides, and avalanches occur frequently, and with very damaging consequences.

[11] The AKRSP has been working in the region since 1984 and is involved with other development interventions such as agricultural support services, micro-credit and enterprise development. However, the majority of the NGO's resources are spent on helping construct infrastructure projects.

organizations' role ends and only the community is responsible for maintenance. So not only are we able to compare similar projects, we can also exploit variation in design between the two agencies' approaches.

In addition to the above factors which make the region suitable for this study, it is worth highlighting other features of the environment that will be useful in the empirical analysis by improving causal inferences for the group-level determinants.

The primary measure of group inequality used is based on land holdings across households in the community. Baltistan underwent land reforms in the 1970s that transferred ownership rights from local rulers to tenants (Dani, 1989). As a result, 79% of farmers own their land and only 2.2% are tenants (1981 Census). Therefore land-holdings in this primarily agrarian setting provide a good measure of wealth distribution. Moreover, as is common in such low-income countries, land markets are virtually non-existent and land distribution, frozen since the reforms, is based on settlement history, household structure, and inheritance (MacDonald, 1994). This makes it plausible that land distribution is exogenous to project outcomes.

The other measure of fragmentation the literature has emphasized is socio-ethnic. The main lines of social differentiation in Baltistan are along clans and religio-political groups. Clans are generally unique to the community and trace a common ancestor. While the population is predominantly Muslim, there are various (Shia, Nur Bakhshi, Sunni, and Ismaili) sects. Even political affiliations are based on religious and familial associations rather than party platforms, resulting in little movement across party lines. Thus social fragmentation in a community is also likely to be exogenous to the outcome of projects in the community.

Finally, community management is based on the *panchayat* system prevalent in South Asia. A group of elders, headed by a hereditary leader (*trampa*), is responsible for community affairs: "The position of *trampa* – no longer formally recognized by the government but practically recognized by villagers, usually passes, upon death, with its attendant obligations, duties and privileges from father to eldest son" (MacDonald, 1994). This hereditary position will be useful in our analysis by providing a potential instrument for leadership in the group, another factor that has been stressed in the literature (Olson, 1965).

A. Data Description

In order to obtain precise estimates given the small samples used in our analysis, great care was taken in the survey work that measurement error be minimized. I describe the survey tools and variable construction in some detail below to explain how this was done.

Detailed community, household and project level surveys were carried out in 99 communities in Baltistan. These communities, either villages or *mohallahs* (sub-villages), were randomly selected from the population of communities where the AKRSP works. Since this NGO has over 90% coverage in the region, this population is representative of the region. Figure 2 shows a map of the region, indicating the selected communities.

While communities had an average of slightly over 6 public projects each, and we obtained very basic information regarding these (651) projects, detailed project level surveys using engineering teams were only conducted in the smaller sample of projects (132) that satisfied our selection criteria described earlier – that the project be a comparable infrastructure project, and that *only* the sampled community be responsible for its upkeep. An additional project satisfying these criteria was found in 33 communities.

In terms of information gathered, primary surveys and secondary sources provide project state measures (current physical and operational condition, and level of maintenance work), community-level variables (community land, income, education, level of development, inequality, social divisions, wages, migration, conflict, hereditary leadership, natural disasters), project variables (project type, scale, expenditure and construction details, age, complexity, skill requirements, design flaws, external organization details), project net benefits/need (level and distribution), and participation in project selection, planning, construction and maintenance decisions. The primary surveys consisted of four separate questionnaires.[12] Trained enumerators administered the first three and a team of engineers undertook the fourth.

The first, a *group questionnaire,* gathers information on community demographics, and details of the project(s) selected from the community, such as the level and distribution of project costs and benefits, participation in project decisions and project maintenance. This questionnaire was administered to a balanced and representative group of community members. In addition, five households were randomly selected for the *individual questionnaire,* which explored sensitive issues such as community conflicts, fund mismanagement and individual participation in project decisions. The *hereditary leader questionnaire* was administered to an adult member of the hereditary leader household and gathered demographic information on the household. The *technical survey* consisted of site visits by engineers to assess the project's physical condition, maintenance system and operational state. Questions were tailored to each project type. As an example, for irrigation channels, questions were asked regarding bed seepage, side-wall breaks, and discharge. For electricity projects, questions ranged from checking the turbine blades to noting the condition of the head-pipe. Any financial constraints and design flaws in project construction were also noted.

Table 1 gives the break-up of the sample by type of project. Table 2 presents the descriptive statistics for variables of primary interest. The construction of the primary measures is described below.

The main outcome variable, the current state of maintenance of a project, is captured by three complementary measures: its *physical, functional and maintenance-work scores. Physical score* estimates the percentage of the project that is in its initial physical state. A score of 70 means the project is in 70% of its initial condition, or alternately, that it requires 30% of the initial (real) investment to restore it to the initial condition. Functional and maintenance-work scores have similar interpretations: *Functional score* captures the percentage of initial project purpose satisfied (e.g. what percentage of the area to be irrigated is currently

---

[12] Questionnaires are available upon request.

receiving water), and *maintenance-work score* estimates the percentage of required maintenance needs carried out.[13] The latter two, while more subjective, provide useful complements to physical score. Since each of these measures are highly correlated (correlations ranging from 0.73 to 0.94), and our results are robust to using any one, we will exploit their informational content by presenting results combining them. As factor analysis reveals roughly equal loadings on the three measures when combined into the single principal component, a simple average of the three, *Total score*, is used as the primary outcome measure.[14]

It is important to emphasize the care with which these scores are constructed represents an enormous improvement over previous studies which usually rely on self-reported and simple rankings of projects. This is partly why, despite not having large sample sizes, we are able to make statistically reliable inferences. Moreover, each of the three measures is constructed using multiple questions and information sources to ensure validity and reliability. For example, an irrigation channel's physical score is constructed as follows. The initial score is based on 10 questions in the group questionnaire. The score is verified using the enumerator's site visit notes, and then averaged with the third independent source, the technical survey, administered at a different time from the enumerator survey. Nevertheless, it is reassuring to note that the correlation between these sources is more than 0.6 for all three scores. Note that the scores incorporate both community-reported and technical assessments of maintenance. Only using a technical measure ignores the community's perception. A technical assessment of an irrigation channel may wrongly assign a lower score for not carrying out a side wall repair with cement, even though members correctly decide the repair could just as effectively be carried out by mud and stones as the water pressure in the channel is low. On the other hand, only using a community reported measure suffers from possible community misreporting.

At the community level, while there are a multitude of measures available in the data I detail those that have been emphasized in the literature. *Land inequality* is constructed in a manner similar to a Gini coefficient except that it is based on grouped rather than individual data. For example, to calculate land inequality a standard Gini coefficient would require an estimate of how much land each household owns. However, individual land measures are costly to obtain in Baltistan as land is fragmented, spread over a large area, and in hazardous terrain. These considerations rule out the feasibility of this approach. Instead, the land inequality index is constructed using grouped data: Households with the maximum and minimum land holding in the community are identified. Using the two land holding sizes, three equal land holding bins are

---

[13] Maintenance needs vary for each project and this was taken into account using engineer-based technical judgements.

[14] Annual project returns are also estimated for each project. For example, for an irrigation project that irrigates new land, net benefit is estimated by considering the amount of new land cultivated under the project and then valuing the crops grown on that land using price and cost estimates obtained from local agricultural support departments. Since the benefit measure is at best a rough approximation to actual benefits and extremely noisy, it is not used in the analysis. However, it does correlate with the outcome measure used, total score, and provides a rough monetary return to improved project state: A 10 percentage point increase in total score is associated with an equivalent $26 annual household gain (per-capita GDP in Baltistan is $216 (Parvez, 1998) with 6-8 members per household).

created, and the number of households belonging to each bin noted. Since all households are distributed in one of the three bins, a grouped-Gini inequality index can be constructed.[15]

Social heterogeneity is an average of the fragmentation indices based on clan, religious and political divisions. The indices are constructed as is standard in the literature (Alesina and La Ferrara, 2000): Each index is the probability that two people randomly chosen from the community belong to a different (clan, religious or political) group. Mathematically, the index is $1-\sum_{k} s_k^2$ where s_k is the proportion of the kth group in the community. Higher values of the index represent greater heterogeneity.

Project leader is a binary variable that indicates whether the project has a leader or not, i.e., an individual selected by the community to manage the project. Care was taken in the interviews that the presence of a leader was not identified on the basis of project performance. While leadership can vary across different tasks within a community, a natural choice for a leader is the hereditary leader or traditional headman of the community (*trampa*) and therefore such leadership is a community-level attribute. Moreover, since the identification strategy used to estimate the effect of leadership in this paper only exploits community-level attributes (instrument using demographic characteristics of the hereditary leader household), for the purposes of this paper I will consider leadership as a community-specific factor. *Leader quality* is an average of the five community individuals' evaluation (good or bad) of the project leader's quality.

In terms of project-level variables, those that reflect design or attributes of a project (rather than simply its type) are of particular interest. *Project share inequality* is calculated in a manner analogous to the land inequality gini measure above using "grouped" data, and measures the inequality in *observed* division of project returns. For example, in an irrigation project, this measure captures the inequality in land holdings in the command area of the project. *Project Complexity* ranges from 0-3, where the index is increased by one each if: (i) the project has greater cash (for outside labor and materials) versus non-cash (local labor and materials) maintenance requirements, (ii) the community has had little experience with such a project, and (iii) the project requires greater skilled labor or spare parts relative to unskilled labor for project maintenance. *Project New* is a dummy variable indicating whether the project is a new instead of a continuation project, i.e. a project made as an extension to an existing community project. While extension work for a continuation project can be minimal, in general the existing project is a small community-made project and the external organization then spends substantial funds extending it. An example is modifying an existing mud-walled irrigation channel by cementing the bed, lining the walls with stones, and extending the channel's length.

---

[15] An alternative is to calculate a standard Gini coefficient using land-holdings of the five households in the individual questionnaires. While this was also done, and this individual-level Gini coefficient constructed correlates well with the group-level gini (0.89), the latter is preferred since it offers a more reliable, albeit coarser, measure of community land inequality; the group index includes *all* community members and is therefore less sensitive to outlyers. Moreover, the problem of mis-reporting or mis-measuring land size is less troublesome in the group-reported inequality measure since any bias is likely to be the same for all members, which would not be the case if they were asked individually.

Government/AKRSP project are dummy variables indicating the type of external organization involved in project construction. The primary external organization comparison is between the local government and AKRSP projects.[16] AKRSP emphasizes and, as supported in the data, elicits greater community participation. It does so by carrying out all project decisions, such as its identification, usage rules and disbursement of fund, through a village organization that has community representation. The LB&RD (Government agency) has no such emphasis as funds are usually disbursed on political considerations and through a representative appointed by local politicians. While both agencies work through elected bodies, the NGO process is at a more localized level and relies on a group of community members rather than a single community representative. Donor funding also implies that the NGO is required to maintain transparent accounts. The local government, on the other hand, has no clearly defined system of accountability.

Finally, the data attempts to measure *community participation* separately in various decisions that are made during a project's construction, usage and upkeep. While we solicited responses on participation in 20 such decisions, Table 2 presents summary statistics by grouping these decisions into two types for each project: non-technical and technical decisions.[17] The measures are constructed by averaging over the community participation in each decision included in the two categories. Community participation in a particular decision is measured as the fraction of the five randomly selected community respondents in the individual questionnaire who answered affirmatively to whether their household (directly or indirectly i.e. through a proxy) participated in the decision.

---

[16] There are a few other semi-governmental external agencies in the sample, but they have too few observations to allow any meaningful comparisons.

[17] See Khwaja (2004) for a model that justifies this grouping since it predicts that community participation in the two types of decisions may have opposite effects. The Non-technical decisions include: (i) Selecting project; (ii) Deciding level and distribution of community labor contribution in project construction; (iii) Deciding level and distribution of community non-labor (cash) contribution in project construction; (iv) Deciding wage to be paid for community labor used in project construction; (v) Deciding on any compensation paid for non-labor community resources used in project construction (e.g. land given up); (vi) Labor work for project construction; (vii) Monetary contribution for project construction; (viii) Deciding project usage/access rules (e.g. who gets to use the project when); (ix) Deciding sanction measures for project misuse (e.g. amount and nature of fines levied); (x) Raising Internal (to community) funds for project construction and maintenance; (xi) Deciding on distribution of project benefits (e.g. allocation of water, electricity across households); (xii) Deciding on maintenance system, policies and rules; (xiii) Deciding on level and distribution of community monetary contribution in project maintenance; (xiv) Deciding on level and distribution of community labor work towards project maintenance; and (xv) Deciding on nature, level and extent of any sanctions imposed for not participating in project maintenance. The technical decisions include: (i) Deciding project site; (ii) Deciding project scale (length, capacity); (iii) Deciding design of project; (iv) Deciding time-frame for project construction; and (v) Raising external (to community) funds for project construction and maintenance. Khwaja (2004) provides separate summary statistics for each measure.

IV. Results

A. Variation in Outcomes Between and Within Communities

Before providing the regression results, it is instructive to examine variation in the outcome of interest, project state, *between* and *within* communities. In particular, there is as much, if not more, variation in collective performance *within* communities as there is between them.

A simple way to illustrate this is to compare the standard deviation of the main outcome variable - total score – between and within communities in the set of 33 communities for which there are two sampled projects each. The standard deviation of average (over projects in a given community) total score for these communities is 16.3, reflecting outcome variation across communities. In contrast, the standard deviation of de-meaned (of the community average score) total score is 18.7, suggesting that if anything there is greater outcome variation within than across communities. While this result is only meant to be suggestive as it can arise for a variety of reasons, it does indicate that a simple explanation of collective performance based only on a community's inherent attributes such as its social capital is unlikely to explain a significant part of the variation in outcomes (within communities). Alternately, a regression of total score on community-fixed effects by themselves provides an R-squared of 43% as compared to an increase of an *additional* 50% once design factors are also included.

A concern with the above decomposition is that it is at the expense of a reduction in sample size and this reduction may introduce potential biases. While this concern will be addressed explicitly in the empirical analysis below, a useful, though admittedly coarse, comparison is to consider all 651 public projects in the surveyed communities. Since only a fifth of these projects met the project selection criteria,[18] they were not surveyed in any detail and only very rough community-reported outcome measures are available. Specifically, communities reported a physical state rank for each project as bad (1), slightly bad (2) and safe (3) and functional state rank as non-operational (1), partially operational (2) and operational (3). Given the poor outcome measures, failure to meet the selection criteria, and lack of detailed project attributes, we do not make use of this sample in the empirical analysis.[19] However, they still serve to illustrate the point that there is greater outcome variation within than across communities. A comparison of variation between and within communities shows that the standard deviation of project ranks is 0.40 and 0.36 *between* communities

---

[18] As explained in the methodology and data section before, projects were only surveyed if they met two important criteria – the group/community responsible for their management was well-defined and reasonable and comparable outcomes measures could be constructed.

[19] An added concern in addition to the crude outcome measures used is that these projects are not necessarily only the responsibility of the community and are often also managed by the government, external agency, or other villages. As such, the within community variation may not strictly be a within community comparison since it may involve looking at tasks which differ in the degree to which the community is involved. An important part of the selection process that justifies paying the costs of the smaller sample size used in the main analysis is that in does not suffer from this error.

for physical and functional ranks, but a larger 0.58 and 0.56 *within* communities for the two respective ranks. If one conceives of communities as "good" or "bad" in terms of their average performance, since more than one and a half times as much variation in a communities' performance comes from *within* the community across different tasks, this suggests that even bad communities often do well and vice versa.

Figures 1-2 in Appendix I present a more detailed picture of the above decomposition of outcome variation between and within communities and similarly illustrate that there is substantial, if not greater, variation *within* communities. Differences across the plotted points in the "between" variation graphs represent average performance differences across communities, while in the "within variation" graphs, differences in points for a given community reflect within community outcome variation.

This variation decomposition is nevertheless only suggestive, as greater variation of outcomes within communities could also arise on account of noisy data.[20] Alternately, a community may have limited "capacity" and so optimally choose to focus its efforts on one task at the expense of other tasks (a substitution effect) or different tasks may inherently have differential success rates. While the latter two explanations do acknowledge that community-specific factors are not of paramount important (i.e. the number of tasks, or type of tasks matters), what we are interested in is how much project "design" factors (rather than just type or number) matter relative to community-specific factors – i.e., ask whether changes in project design can counter any detrimental effects arising from poor community-level attributes. The next sections attempt to carry out this analysis.

B. Project Design Factors

Table 3 estimates specification (2) to identify the impact of project design factors. As detailed in the methodology, a concern is that design factors are correlated with omitted community-level factors. The use of community fixed effects allows for cleaner identification.

While the results presented are for linear models, a potential concern is that the outcome measure is a percentage rank and imposing cardinality may be an issue. However, all the results in Table 3 remain robust to using probability models such as ordered probits. I present the linear model results since the estimates are simpler to interpret.

Column (1) in Table 3 presents the primary and preferred specification. Columns (2)-(7) present various robustness checks. Since each design effect is interesting in its own regard, I will generally first discuss the results in Column (1), and then present the robustness checks. Note that in addition to community fixed effects, specifications include project-type fixed effects, i.e. outcomes are compared *within* the same type of project. This minimizes any concerns arising from differences in project type selection across communities.

Within community variation in the smaller sample captures the variation across projects for the *same* group of people that are fully responsible for the project, i.e. only their attributes matter in determining group-level effects.

[20] Between comparisons partly "average" out noise, or there may just be more noise in outcomes within communities.

The degree of *Project complexity* has a detrimental effect on project maintenance. Column (1) in Table 3 shows that a 1st to 3rd quartile increase in the complexity index is associated with a 25.5 percentage point drop in maintenance. Recall that project complexity captures whether the project requires cash inputs and skilled labor/material parts for its upkeep and the community's experience in maintaining such a project. Evidence from the field suggests this effect reflects the community's perceived risk of appropriation of its inputs rather than capital or resource shortages, or in other words. these projects are harder to maintain since community members are reluctant to contribute inputs such as capital to the project. Unlike one's own labor which naturally has a monitoring aspect it is difficult to ensure one's capital contributions are spent on project maintenance – a fraction of such contributions can and are appropriated. For a formal treatment of this effect see Khwaja (2006).

A stark illustration of this effect was apparent in a failed lift-irrigation project in one of the sample communities: While maintenance required cash inputs to maintain the lift-pump, cash constraints were an unlikely cause for failure since the community in question was well off and individuals successfully operated private pumps. Further inquiry revealed that community members were *unwilling*, not unable, to provide cash. Members explained they preferred not to give cash as they were unsure of how it was spent. The project had disagreements regarding where the contributions had been spent and whether the reported expenditure had been necessary or even taken place. The community maintained that projects that didn't require regular cash contributions would work better, even if they needed a cash-equivalent amount of labor.

Project share inequality, the measure of how the returns of the project are distributed amongst community members, has a U-shaped effect on maintenance: Increasing inequality by 0.1 units from perfect equality *lowers* maintenance by 24 percentage points (Column 1). The same increase at higher inequality (a Gini of 0.4), *raises* maintenance by 80 percentage points. Recall that these estimates include community fixed effects i.e. we compare how differences in the distribution of benefits between two different projects in the *same* community (and therefore same community level inequality) impact their maintenance.

The intuition for this non-linear relationship, modeled in Khwaja (2006) and supported from field evidence (also see Olson, 1965, Baland and Platteau, 1996, 1998; Alesina and La Ferrara, 2000, Baland et. al. 2006, and Bardhan, Ghatak and Karainov 2005 for related models), is as follows: The initial detrimental effect from increasing inequality arises from free-riding considerations – as a member's share increases in a project, they do not raise their own contribution as much as the member that loses share decreases their contribution. This leads to an overall drop in maintenance contributions and project state.[21] However, as a member obtains an even greater share of project returns, the project essentially becomes a private project and at this stage, instead of a public contribution game with members voluntarily contributing, the member with the large share is willing to undertake the greater cost of managing the project by using (outside) paid

---

[21] This assumes costs are increasing and convex in own contributions – a reasonable assumption in this context.

labor/inputs i.e. the project is effectively privatized and maintenance work (mostly) contracted out. While this may have poor distributional implications for the project, its maintenance improves. An examination of several projects from the field where a community member had a disproportionately large share indeed revealed exactly such a mechanism: Instead of the usual maintenance system where community members all voluntarily contribute to project upkeep, the member with the much larger share paid others to do such work.

One of the more surprising results is for *Community participation* in project decisions. Since the specification uses community fixed effects, the participation effect is not picking up features of the community which induce greater participation but rather aspects of project design such as the extent to which the project sought to solicit community participation. The literature primarily views community participation as an unqualified good (Narayan, 1995; Isham et al., 1996) but the results in Column (1) show otherwise: While community participation in *decisions* made in the project that are *non-technical* in nature (selecting project type, usage rules etc.) positively affects maintenance – a 10% increase is associated with a 5.5 percentage points improvement in maintenance – a similar increase of community participation in *technical decisions* made in the project (deciding project design, scale etc.) has a *negative* impact on maintenance, lowering it by 3.8 percentage points.[22] Khwaja (2004) develops a model based on the property rights literature to provide a theoretical foundation for understanding the effects of participation and why it may also have a negative impact. Specifically, once one recognizes that community participation is both a means of providing information and exerting control over a decision, and that project-improving effort may not be contractable, the optimal solution is to give control-rights to the party that has the more important investment to make. For technical decisions such investment effort is likely to come more from the external agency ('engineers') rather than the community.

While endogeneity arising from community unobservables inducing greater participation is not an issue given the estimation strategy, "halo effects" remain a concern: Since the participation measures are based on recall, even if a decision occurred prior to project maintenance, individuals may falsely report participation (no participation) if the project is currently doing well (poorly). Note that such halo effects would lead to an *overestimate* of the participation effect and while they may be a concern for the result on participation in non-technical decisions, for participation in technical decisions such a bias would make it harder to find the negative result. In any case, a potential solution to obtain unbiased estimates would be to instrument for participation. Unfortunately, plausible instruments are hard to come by. An alternate strategy is to check whether halo effects are present. I am able to do so since the five community members surveyed in the individual questionnaire were also asked to rank their perception of the project's current physical and operational state (but their response was not used to construct the project outcome measure used). If halo effects were significant, individuals who perceived the project to be in a better state relative to the others surveyed would also report relatively higher participation. Checking for such a positive correlation implied by halo effects, I find no such significant correlation for either the physical or operational measures of maintenance (correlations are -0.04 and 0.02 with significance levels of 32% and 55% respectively). Thus halo effects do not seem to be an issue.[23]

External organization type also has a significant effect as the NGO (AKRSP) initiated projects are associated with 23.6 % points higher maintenance compared to projects initiated by the local government (Column 1). The project type fixed effects and design complexity controls allay a concern that the effect is caused by the NGO constructing simpler projects. In fact, the data shows that if anything, the NGO constructs more complex projects. Another concern is that the effect could be an initial construction quality effect: The NGO constructs better projects and so, regardless of the community's maintenance effort, the projects remain in a better state.[24] We offer several reasons why this is unlikely and that our result is indeed capturing better community collective ability for the NGO-initiated projects. Column (5) in Table 3 shows that the NGO effect is robust to the inclusion of project construction quality measures such as the amount invested by the external organization,[25] and whether the community reported an initial design/construction problem in the project.

While a better test would use precise quality measures taken at the *time* of project construction, no such measure is available. Nevertheless, if the construction quality explanation holds merit, it must be that current physical score is determined primarily by construction quality, and therefore, project *physical score* provides an accurate measure of initial construction quality. Column (7) takes this rather severe view and regresses project *functional score* on project-specific factors *controlling for* physical score and shows that NGO projects still have (24 percentage points) higher functional scores: Not only do NGO projects outperform government ones in terms of overall state (*total score)*, but, controlling for current physical condition, NGO projects are also managed more effectively (i.e. meet a greater fraction of planned needs). These results offer strong evidence that the NGO effect is not just a construction quality or higher funds invested effect.[26] What remains unclear is why the NGO-initiated projects, even after the direct involvement of the NGO has been

---

[22] Similar results obtain if the decisions that are grouped in the non-technical and technical categories are considered individually (regressions in Khwaja (2004)).

[23] A concern that cannot be addressed arises since community participation is, after all, a choice. To the extent that the design factors that affect this choice are not controlled for, causality cannot be guaranteed. Thus, while community participation has a negative or positive effect on maintenance based on whether the decision is technical or not, causality can only be suggested by using community fixed effects and showing that halo effects are not important.

[24] This result would still be interesting since overall expenditure by the two agencies is controlled for. However, since such an effect is not a determinant of the community's collective success, it is not emphasized in this paper.

[25] This amount is underestimated for the NGO if it incurs greater overhead costs that are unaccounted for. While project-wise data on overheads was not possible to obtain, information collected at the NGO and government offices suggests this is not the case: While the two agencies may differ in the distribution of such costs, they do not in the total.

[26] Care must be taken in generalizing, since only one NGO is compared to a given local government. Moreover, the NGO, AKRSP, is an effective and high quality NGO, while the local government in Northern Pakistan is unlikely to be of above average quality.

removed, remain better sustained. While the data does not allow sharper tests to tease out potential explanations, some tentative hypotheses will be offered later in the paper.[27]

I also examine the impact of whether the project was built from scratch (*Project New*) or was a (substantial) *extension/*continuation of an existing community project. Ex ante it is not clear if there would be any difference between the two. For example, existing community projects that still need further external support could be problematic projects and therefore likely to fail. Alternatively, such projects may have already established management systems that lead to a greater likelihood of even a substantial extension of the project being well-maintained. The data supports the latter: Continuation (extension) projects are associated with 41.9 percentage points higher maintenance than new projects (Column 1). Since project-specific factors such as project type and complexity are controlled for, this result is not driven by new projects being of a different type or more complex. Nevertheless, an alternate explanation could simply be that continuation projects reflect greater community need than new projects. However, the effect is robust to the inclusion of perceived project need: Controlling for community members' rank (1 to 4) of project need does not lead to any change in the magnitude or significance of the effect (regression not reported). While it is not possible to entirely rule out such selection given the data, interaction effects presented later in the paper hint that continuation projects are better maintained for the reason suggested above: they have already set up the maintenance management systems and rules needed, an issue that new projects have yet to tackle.

Finally, while I do not include this in the main specification in Column (1), Column (2) also shows that having a project "leader" is weakly associated with 13 % points higher maintenance. A project leader was defined as an individual responsible (either directly or as the head of a committee) for managing the project. However, this factor is expected to be biased despite using community fixed effects, since it is likely that project success may affect whether a project reports having a leader or not. We will present a cleaner estimate later on using an instrument for leadership. Since the instrument is community-specific it cannot be used with community fixed effects.

Robustness Concerns:

While the above discussion was mostly based on the results in Column (1) of Table 3, several robustness checks are carried out in Table 3.

The foremost concern discussed in the methodology section arises from the sample drop that is entailed if community fixed effects are to be used. Does the restriction to communities where multiple projects

---

[27] While the local government employs more staff, it pays them lower salaries. This hints at lower incentives in the local government offices. Other explanations may be that the NGO is more aware of the local environment and needs, is less prone to corruption, has greater accountability and transparency, attracts a more dedicated staff, or elicits greater community participation. The data is unable to distinguish between these hypotheses. It supports some – NGO initiated projects are 35% more likely to have a project leader and have 20% higher community participation in project decisions but, as Table 3 shows, the NGO effect remains if leadership and participation are controlled for. Later on (Table 6) we will present results that suggest that NGOs may be better at setting up the management systems required for the newer and complex projects.

meeting the selection criteria were found likely to bias the design factor estimates? First of all, recall that in the data description we had already shown that these communities only differ in that they are generally larger in size. Columns (3)-(4) together show furthermore that the design factor estimates depend really on being able to correct for omitted community level variables rather than any sample change associated with employing such a correction.

Column (3) first estimates the same specification and sample as in Column (1) but excludes community fixed effects and shows that the coefficients estimates do indeed change significantly. The magnitude of the design factors is substantially smaller if community fixed effects are not included, suggesting that (unobservably) better communities indeed choose/are assigned more difficult design. Specifically, the results suggest more "able" communities are also more likely to choose more complex projects, and that NGOs are often more willing to work in less "able" communities. Moreover, a generalized Hausman test soundly rejects equality of coefficients between the two estimations with and without community fixed effects (Column 1 and 3). Thus Column (3) shows the importance of being able to use community fixed effects in correctly identifying design factor impacts.

Column (4) then asks the next question – does the sample restriction matter? It does so by running the same specification as in Column (3) (without community fixed effects) but in the full 132 community sample. Were the sample restriction important one would expect the estimates on the design factors to be significantly altered. However as Column (4) shows this is not the case. The larger sample essentially just reduces standard errors but a generalized Hausman test fails to find significant differences between coefficients estimated in column (4) from those in column (3) i.e. sample size restrictions do not matter.

Column (5) addresses a concern in using the simpler average of the three physical, functional and maintenance-work scores as the main outcome measure. In Column (5), an alternate outcome measure is constructed using factor analysis to combine the three component scores. As previously mentioned, doing so gives roughly equal weighting on each of these measures, and gives in similar results.

Column (6) re-estimates Column (1) but also adds measures to better control for initial project conditions (amount of funds invested by the external agency and an indicator of whether there was a problem in the initial design/construction of the project) to ensure that the results capture the community's effort in maintaining the project and are not an artifact of some initial external agency-specific condition. The results show such initial conditions are not a concern.

Finally, Column (7) carries out a similar but more stringent test of controlling for initial conditions in Column (6). In Column (7), the dependent variable is functional score, one of the components used to construct total score, the main outcome measure used in the other Columns. In addition, the second component, physical score is used as a control. If the results in Column (1) are driven by a projects initial physical state, then if one controls for the current physical state, our factors of interest will have no effect on how well a project is performing in terms of its *operation*. However, the results in Column (7) show that this is not a concern: The project-specific design factors are not just capturing the projects' current physical state, but also have similar impact on the current functional/operational state of the project conditional on its physical/construction quality.[28]

I already noted that the above results are robust to using models such as ordered probits which do not impose cardinality, in case there is concern that the outcome measure is a percentage rank. A related concern may be that the outcome measure is left censored (cannot have less than 0% state). However, in the data all projects have above 0% measure and so this is not an issue. Similarly, while measures above 100% are possible if the project improved upon its initial states (and there are a few instances of this) one may be concerned about right censoring. In any case, our results remain the same even when we explicitly allow for right (or both left and right) censoring and estimate tobit regressions.

C. Comparing Design to Community Factors

How large is the impact of design factors compared to community ones? As described in the methodology section, we provide two different comparisons – the first, with the magnitude of the community fixed effect and the second, by comparing to the effect of specific community factors.

The Community Fixed effect

The use of community fixed effects in the previous regressions allows one to compare the impact of the above design factors with the *combined* impact of all, observed and unobserved, community-specific factors i.e. the community fixed effect itself.

Figure 3 illustrates this by conducting the following thought experiment: If we could change project-factors, what is the best and worst maintenance outcome that could be obtained in each community *given* its attributes, for the same type and vintage of a project? The figure plots the *highest* and *lowest predicted total* score for a hypothetical irrigation project (the most common project) with an age of 8 years (the sample median age) in all sample communities by "setting" project complexity, community participation, external organization type, share inequality and whether the project is new or not, at their *best* and *worst within-sample* values. The predictions are based on estimates in Column (1) of Table 3 and I also include communities with only one project in our sample.

While these predictions ignore the relative costs of manipulating factors, the figure nevertheless illustrates the relative importance of the project-design factors compared to all community-specific factors: The difference between the highest and lowest project scores under a given design scenario (for example,

---

[28] It is interesting to note that the coefficients on the design factors are similar to those in Column 1. This suggest that a lot of the variation explained in our main outcome, total score, is in its functional score component. However, note that this does not mean that anyone of the three measures used to construct our main outcome measure are not by themselves affected by these factors. Separate regressions (not shown – available on request from author) using each of the three component scores, physical, functional and maintenance-work, as outcomes shows that these project-design factors all matter and have comparable impacts.

points 2w and 1w in the figure) gives the *largest* community effect. This difference in performance is the combined impact of *all* community-specific factors, observable *and unobservable* when comparing the "best" and "worst" community in terms of these attributes. This is compared with the combined effect of the project-specific design factors discussed above, i.e., the difference between maintenance of the best and worst project designs for a project in a *given* community (points 1b and 1w).

Thus the figure shows *the best-designed project in the worst community (1b) does as well as the worst-designed project in the best community (2w).* One can also compare less extreme values: The difference between the 3rd and 1st quartile community (in terms of their fixed effect) is 51 % points. In comparison a 1st to 3rd quartile change in project complexity, non-technical and technical decisions participation has a 25.5%, 14.7% and –18.5% points impact respectively on project state. Similarly, NGO-initiated projects and Continuation projects are associated with 23.6 and 41.9 % points higher project state.

Appendix Figure 3 performs a similar hypothetical comparison of best and worst projects. Instead of comparing a hypothetical 8 year old irrigation project in each community, I use the actual project type and age in the community and change the project-design factors as before. Thus communities with multiple projects will have a best and worst predicted score for each of the projects. The results are the same.

These figures illustrate the main observation this paper makes: That design factors have a comparable, if not larger, impact relative to all community-specific factors including observable ones like community inequality, fragmentation, and less tangible and unobservable ones like a community's degree of unity, societal norms of cooperation etc. This suggests that project-specific factors can indeed compensate for adverse community-specific factors such as lower social capital that are invariant to the community.[29]

Community Factors

While we have compared the combined impact of all community-specific factors to a few project-specific factors above, it is nevertheless instructive to estimate the impact of particular community-specific factors that have been emphasized in the literature, such as group inequality and heterogeneity (Alesina and LaFerrara 2000). Doing so allows a comparison of particular community-specific factors' impact to that of project-specific factors, and is also useful in examining interactions between the two in subsequent sections.

Columns (1)-(3) in Table 5 present the results from estimating equation (1) using 2SLS where the impact of project leadership is instrumented for using characteristics of hereditary leader households. Column (1) presents a sparse specification and Column (2) adds more project and community-specific factors as controls: in addition to the project specific factors used in Table 3, it includes measures of the community's attributes,

---

[29] A potential caveat to this interpretation is that the combined community-effect may appear smaller due to complicated interaction effects and aggregation issues. For example, Glaser et. al. (2002) argue that group social capital may not be a simple aggregation of individual social capital as returns to the capital of one individual may in fact induce negative returns on another. In our case, it could be that it just happens that desirable group attributes in our sample coincide with undesirable ones and so their combined impact is lower than the import of each attribute separately. While this seems unlikely, I will also estimate the impact of individual group-specific variables and obtain similar comparisons.

human and physical capital. Column (3) has the same set of controls as the specification in Column (2) but adds instruments for leader quality. The first stages for leadership presence and quality are given in Columns (4) and (5) respectively and are discussed below.

Before we examine the impact of salient community-specific factors, a note of caution to be raised is that these estimates are potentially biased since our preferred specification in Table 3 made use of community fixed effects to eliminate the likely bias arising from community unobservables. To the extent these unobservables are directly correlated with community-specific factors or indirectly through their impact on project-specific correlates of the community-factors, the estimates on the community-specific factors will also be biased.[30] We discuss only the community-factors of interest.

The literature has emphasized the impact inequality has on collective action, although the sign of the impact is at times ambiguous with models suggesting that inequality has no effect, (Bergstrom, Blume and Varian 1986; Warr 1983), is detrimental (Baland and Platteau 1997; Alesina and La Ferrara 2000), or has a positive impact at least at some parts of the distribution (Olson 1965; Baland and Platteau 1997). Khwaja (2006) suggests a quadratic effect of inequality for reasons similar to those posited for the impact of inequality in a project's share. The empirical results show that our measure of community inequality, *Land Inequality,* has a U-shaped effect on maintenance similar to, though smaller than, the effect of inequality in project share. Column (3) shows that a 0.1 unit increase in the land inequality index starting from perfect land equality is associated with a 21.7 percentage points *fall* in maintenance. The same increase at a higher inequality level of 0.4 (90th percentile) is associated with a 14.5 percentage point *rise* in maintenance. The result is robust to project and community-specific controls and, given the fixed land settlement patterns and non-existent land markets in Baltistan, land inequality is likely to be exogenous to project outcomes and therefore the bias in our estimates is low.

It has been argued that *Social Heterogeneity* makes it more difficult to sustain group cohesion due to resulting inequalities in access/benefit derived from the collective project (Khwaja 2006), preferences that lead one to favor one's own social groups (Alesina and La Ferrara, 2000), or by weakening social norms and sanction mechanisms. The results in Column (3) show that heterogeneity is indeed detrimental to project maintenance: An increase in the heterogeneity index from the 1st to 3rd quartile (0.25 to 0.43) is associated with a 7.1 percentage point drop in maintenance. The adverse impact of heterogeneity is robust to project and community-specific controls, particularly land inequality. As with land inequality, a household's social

---

[30] A potential alternate technique to estimate these effects is Hausman-Taylor (1981). However, the data requirements for this estimation to work are rather stringent and while the estimates on project-specific factors using this technique were similar to those in Table 3, the standard errors on the community-specific factors were too large to be meaningful. An alternate formulation that is used at times is to "decompose" the fixed effect through a two-stage estimation: In the first stage a specification as in Table 3 with community fixed effects is run and the second stage regresses the estimated community effects on community level variables. This yields similar results to our estimates.

grouping in Baltistan is likely to be invariant to the success (or failure) of community projects and therefore biases arising from reverse causal impacts are unlikely to be present.

While there is great interest and theoretical arguments as to why one would expect leadership in a group to have an impact on its collective performance (Olson 1965; Durlauf and Fafchamps 2004), the empirical literature on this is limited (Jones and Olken 2005 is a notable exception). Part of the problem arises from making a causal inference: Not only are community unobservables likely to impact both collective performance and whether the community chooses to have a leader or not (a bias that can be corrected by using community fixed effects as we did in Table 3), but it is likely that project performance will have an impact on whether the project has a "leader" or person primarily responsible for it. However, a feature in the Baltistan environment allows us to minimize this problem by instrumenting for leadership. As described previously, most communities in Baltistan have *trampas* (hereditary leaders). Hereditary leaders are not selected by the community but are by tradition a natural choice to lead. Therefore exogenous attributes of the hereditary leader household, such as whether it has a young and healthy male member (a "potential" leader), provide instruments for leadership as they are unlikely to be correlated to project outcomes but correlated with having a project leader or not through other channels. Note that while leadership can vary for each task undertaken by a given group, since our instruments will be community-invariant, we interpret any leadership effect as a community-specific effect.

Column (4) in Table 4 presents the first stage for the IV regressions in Columns (1)-(2). The instruments used are: (i) an indicator variable for whether the household has a healthy male member between the ages of 25 and 50, (ii) the average age of household members, and (iii) the average index of household members' presence in the community (1 = a lot, to 3 = very little). The first two variables are based on demographic "shocks" to the household and are therefore exogenous, while the third, conditional on community demographics, is also expected to be independent of project maintenance. The instruments are jointly significant at less than 1%. Columns (1)-(2) present the second stage and show that an increase from the 1st to the 3rd quartile in the predicted value[31] of having a project leader increases project score by 7.5 to 8.5 percentage points respectively.

Anecdotal evidence suggests that hereditary leaders may not be the best leaders and have lower than average quality. Column (3) in Table 4 instruments for both leadership presence and quality. The instruments used for leader quality are (i) whether the hereditary household has a healthy male member between 25 and 50 years old who has at least primary education and is always present in the community, (ii) whether the hereditary household is involved in an off-farming profession (a proxy for disinterest in community affairs), and (iii) the number of individuals in the community perceived as being "ideal" *potential* leaders. The

---

[31] Since the presence of a project leader is a binary variable, instrumenting for it results in a continuous predicted value between 0 and 1. The comparable effect to the change of the binary variable from 0 to 1 is estimated by considering an increase in the predicted value from its 1st to 3rd quartile.

instruments are jointly significant at less than 1% (Column 5). Column (3) shows that while the leader presence effect is reduced to 7 percentage points, a 1st to 3rd quartile increase in leader quality raises maintenance by an additional 7.6 percentage points. These results show that not only does leadership presence positively affect a group's collective success but that this effect increases in leader quality.[32]

The estimation in Table 4 also considers other community-specific factors but I will only mention some. First, note that *Number of public projects* does not matter suggesting that there is little concern that community resources are stretched if they more projects. *Community size* also has no significant effect once land inequality and social heterogeneity are controlled for suggesting that, contrary to some of the claims made in the previous literature, it is not size that matters per se, but the greater inequality and heterogeneity that is more likely to be present in larger groups that hinders collective action. *Community remoteness* measures (*walk and travel time*) and *Total cultivatable land* in the community have no significant effect, while *Single cropping zone* communities (those with one yearly harvest) are associated with 14.1-14.6 percentage points lower maintenance. Human capital measures in the community have expected though weak impacts that are not robust to leadership quality. The estimates in Columns (2)-(3) also included controls for a community's physical wealth (real estate, mechanized assets, off-farm income, etc.) and infrastructure (access to potable water, electricity, health facilities) and these factors have few significant effects.[33]

Thus, while community-specific factors such as inequality, heterogeneity and leadership do matter, their impact is moderate in comparison to that of the design/project-specific factors discussed previously. This is illustrated not only by a comparison of point estimates (Tables 4 and 5) but an F-test also reveals that the combined impact of the project-factors is significantly larger (at the 5% level) than that of the above

---

[32] An issue in the IV estimates is that not all communities have hereditary leaders (28 of the 99 communities do not). The instruments are suspect if these communities differ from those with hereditary leaders. Mean comparison tests show that the two types of communities do not significantly differ along community observables. In addition, observations suggest that hereditary leader presence is determined by where the hereditary leader was residing during the 1970 land reforms after which he no longer commanded formal authority over a set of villages but remained restricted to his village of residence. This difference is unlikely to determine project maintenance. Moreover, the second stage estimates presented for the full sample (the first stage interacts each variable with an indicator for hereditary leader presence) are similar to estimates in the restricted sample, which only includes communities with hereditary leaders (not shown).

[33] The survey also collected direct measures of "social capital" such as community trust/conflict. Since these measures are best interpreted as outcomes, OLS estimates of their effect on maintenance are biased upwards. Nevertheless, these estimates (not reported) support relatively smaller magnitudes of community-effects: *Trust* (do members trust each other) has no significant effect. However, communities that report high unity have 8 percentage points higher, and those with land disputes, 13 percentage points lower maintenance. Communities that do not report problems in raising cash for collective work have 10 percentage points higher maintenance, but there is no significant effect for problems in raising community labor. While the number of community organizations (normalized by community size) has no significant effect, a 1st to 3rd quartile increase (1 to 2.6) in the total (normalized) membership of community organizations is associated with 5 percentage points higher maintenance. The fraction of community households with temporary (seasonal) migrant members has no effect, but the analogous fraction for permanent migrants has a negative effect: A 1st to 3rd quartile increase (0 to 5%) worsens maintenance by 3 percentage points. A 1st to 3rd quartile increase (0 to 2%) in the fraction of community households that migrated recently into the community is associated with a 1 percentage points fall in maintenance. These estimates control for human/physical capital, and project-specific factors.

community-factors.[34] This is not surprising given the results depicted in Figure 3 which compared the impact of the selected project-specific factors to all community-specific factors.

D. Heterogeneity in the Impact of Project Design

While the above results support the main point this paper makes – that project-design factors have a comparable, if not larger impact, relative to community-specific factors – one may expect that there are interactions between the two. Rather than claiming that certain project-design factors improve performance in all types of communities, they may have different effects depending on the characteristics of the community. The presence of such interaction effects, while requiring a more nuanced interpretation, adds to the importance of design particularly if design factors are even more important in worse off communities. Given the sample size, a thorough exploration of such interactions effects would be asking too much of the data. Nevertheless, results in Tables 6-7 show that such interaction effects may indeed be important and worth further study.

Table 5 first considers the interactions of main interest – those between the project-specific and community specific variables highlighted previously. Instead of presenting all possible interactions, I only consider a few interactions of interest. Given the small sample, I will consider each potential interaction in a separate regression. All regressions in Table 5 use the same set of controls as in the main specification in Table 3 including community fixed effects, allowing for better identification.

Columns (1)-(4) consider interactions of various project-design features of importance in Table 3 with the land inequality measure. For simplicity the land inequality measure is recoded so communities above the sample median inequality are coded as being "unequal." Results are similar if a continuous measure is used instead. The findings in Columns (1)-(4) show significant and plausible interactions between community inequality and project design variables. For some design factors, community inequality worsens the problem: Column (1) shows that while communities with low levels if inequality also fare poorly if the project is a new rather than continuation one, the problem of new projects is really more severe and significantly so in unequal communities where such projects have 42 percentage points lower maintenance score than a continuation project. Similarly, Column (2) shows that while project complexity is a problem in relatively equal communities, unequal communities are more than twice as sensitive to increases in project complexity. Column (3) shows NGO-initiated projects outperform government initiated projects in both equal and unequal communities.

However, interestingly enough, in the case of community participation, Column (4) shows community inequality *dampens* both beneficial and detrimental effects of community participation. While a 10% points

---

[34] The test constructs a 95% confidence interval for the joint effect of 1st to 3rd quartile increases (discrete if variable is binary) in the main community-specific factors, and a similar confidence interval for the joint effect of project-specific factors. The lower bound of the project-specific interval lies above the upper bound of the community-specific interval.

increase in community participation in non-technical decisions leads to a 7 % points increase in project maintenance in an equal community, the same increase has barely any effect (a 1 % point increase) in an unequal community (the interaction is significant at a 16% confidence level). Similarly, while a 10% increase in community participation in technical decisions leads to a 5 % points drop in project maintenance in equal communities, there is again little effect (1.4% points) in unequal communities though the interaction term is only significant at 18%. While somewhat surprising at first, these interactions are very plausible if one considers that impact of inequality in a community is that the effectiveness of participation by a community is lessened i.e. for the same level of community participation, more unequal communities are less representative as some members are weighed more than others. Thus any effect of a community's participation, whether positive or negative, is dampened in more unequal communities.[35]

Columns (5)-(7) next explore interactions between leadership and project-design factors. As in Table 4 before, leadership is instrumented for by attributes of the hereditary leader household (Column (4), Table 4) so these results are not simply stemming from endogeneity of leadership to project performance. The results generally indicate that communities which lack leaders face greater sensitivity to project design factors: Column (5) shows that while complex projects remain a serious problem in communities without a leader, those that do have leaders face a significantly lesser problem in maintaining complex projects. Similarly, while government-initiated projects fare much worse (62% points lower) than NGO initiated ones in communities without leaders, there is no difference between projects initiated by the two agencies in communities with project leaders (although the interaction in Column (6) is only weakly significant at 18%). Finally, while the interaction term in Column (7) is not significant, it does have the expected sign i.e. leadership is likely to lessen the problem of maintaining new projects.

The results in Table 5 present a promising area for future research and highlight the importance of project-design even more so for communities with lower inherent abilities and social capital.

Table 6 concludes by examining various interactions of interest within project-specific factors. While the small sample size urges caution in reading too much into these interactions, it nevertheless adds to the main point this paper makes – the relative importance of project-design factors – by furthering our understanding of how these design factors matter.

First, we further explore the result that NGO-initiated projects fare better than government initiated ones: Columns (1) and (2) show that this poorer performance of *Government-initiated projects* really comes when considering complex projects - Column (1) - and projects that are made from scratch rather than built upon existing projects (Column 2 – interaction is significant at 13.5%). These results suggest that the NGO does better than the local governments in complex and new projects. Since new and complex projects require

---

[35] Unlike community inequality, community social-fragmentation did not show any robust interaction effects with these project-design factors (regressions not reported).

setting up management systems and transmitting new skills, this hints at the channels through which the NGO outperforms the local government. In addition, Column (3) shows that there are no significant interactions between *project complexity* and whether the project is *new* or not. This result, combined with the previous two, may suggest that continuation projects outperform new projects because new projects need to develop management systems and not because they require new skills. In other words, existing projects may already have the *project-specific* "management-capital" (i.e. the management rules and systems etc.) setup and an extension of these projects is able to make use of this existing capital.

V. Conclusion

Previous empirical work has examined the effect of group factors such as inequality and social heterogeneity on collective action. The results of this paper confirm these effects by showing they are robust to a larger set of community and project controls than used previously and also provide new evidence for the effect of leadership on collective success. However, the main empirical contribution is to carefully identify project-design effects and compare their impact to the community-specific effects. To our knowledge, this is the first such attempt to do so. This is particularly significant at a time when the debate on the collective performance has focused on a group's inherent attributes rather than the nature of the collective task and local public good provision failures in communities, whether in the United States or Sub-Saharan Africa, are often attributed to inherent failures in those communities.

The magnitude of estimated effects and large variation in maintenance within communities suggests rethinking the significance of community-specific factors, such as social capital, on collective performance. The estimates for all project-specific factors, except project leadership, are obtained by comparing the factor's impact on different projects in the *same* community. This provides the following two categories: Determinants that are inherent to the community (land inequality, social heterogeneity, and leadership) and determinants that are part of project and institutional design (project complexity, share inequality, community participation in project decisions, whether the project is new, and type of external organization). Comparisons show that the latter has larger effects than the former. This suggests that while social capital is indeed a stimulus to collective action, its scarcity can be compensated for by better project design.

Moreover, since social (capital) factors tend to persist over time (Putnam, 1993), they are best viewed as constraints rather than policy tools: while land redistribution can be used as a means of influencing community inequality, such reforms are notoriously hard to implement. Dividing the collective venture so that it involves more homogenous sub-groups can help, but doing so may increase unit costs substantially and suffer from the loss in diversity (Bowles and Gintis, 2002).

Therefore, rather than directly addressing the social capital constraint policy initiatives that emphasize project design may be more feasible and have better success in implementation. This paper offers several such project-design improvements: Designing projects that face fewer appropriation risks through better leadership and lower complexity, eliciting greater local information through the involvement of community members in project decisions, investing in simpler and existing projects, ensuring a more equitable distribution of project returns, and emulating NGOs can substantially improve project performance even in communities with low social capital.

This paper is admittedly limited both by sample size and coverage, lending a cautionary note to drawing policy inferences. Nevertheless, the contribution it hopes to make is a more modest one: To temper the current emphasis on group-level attributes such as social capital with a recognition that even groups that face possibly inherent and persistent constraints, such as a lack of social capital, can achieve success when offered well-conceived and carefully designed tasks.

References

Afridi, Banat Gul. (1988). Baltistan in History, Peshawar, Pakistan: Emjay Books International.

Agarwal, Arun and Sanjeev Goyal. (2001). "Group Size and Collective Action: Third Party Monitoring in Common-Pool Resources," *Comparative Political Studies*, Vol. 34, No. 1, pg 63-93.

Alesina, Alberto; Reza Baqir, and William Easterly. (1999). "Public goods and ethnic divisions," *Quarterly Journal of Economics,* 114**,** November.

Alesina, Alberto, and Eliana La Ferrara. (2000). "Participation in Heterogeneous Communities," *Quarterly Journal of Economics,* 115**,** August.

Alesina, Alberto, and Eliana La Ferrara. (2005). "Ethnic diversity and economic performance," *Journal of Economic Literature* 43(3), pg. 762-800.

Betz, Mathew; Patrick McGowan, and Rolf Wigand. (1984) Appropriate technology : choice and development, Durham: Duke University Press.

Baland, Jean-Marie and Jean-Philippe Platteau. (1996). Halting Degradation of Natural Resources. Is There a Role for Rural Communities, Oxford: Clarendon Press.

Baland, Jean-Marie and Jean-Philippe Platteau. (1996). "Inequality and Collective Action in the Commons," CRED, University of Namur.

Baland, Jean-Marie and Jean-Philippe Platteau. (1998). "Wealth Inequality and efficiency in the commons. Part II: The regulated case," *Oxford Economic Papers*, 50, 1-22.

Baland, J.-M, O. Dagnelie and D. Ray. (2007). "Inequality and Inefficiency in Joint Projects," *Economic Journal,*, 117:522, pp. 922-935

Banerjee, Abhijit and Esther Duflo. (2003). "Inequality And Growth: What Can The Data Say," *Journal of Economic Growth*, v8(3,Sep), pg 267-299.

Bardhan, Pranap and Jeff Dayton-Johnson. (2002). "Inequality and Conservation on the Local Commons: A Theoretical Exercise," *Economic Journal*, Vol. 112, pp. 577-602.

Bardhan, Pranap. (2000). "Irrigation and Cooperation: An Empirical Analysis of 48 Irrigation Communities in South India". *Economic Development and Cultural Change.*

Bardhan, P.; M. Ghatak and A. Karaivanov (2005). "Inequality and Collective Action" in P. Bardhan, S. Bowles and J. M. Baland (eds.), (2006) "Economic Inequality, Collective Action, and Environmental Stability", Princeton University Press.

Bergstrom, T., Blume, L., Varian, H. 1986. On the private provision of public goods. Journal of Public Economics 29, 25-49.

Bowles, S. and H. Gintis (2002). "Social Capital and Community Governance". Economic Journal, vol. 112(8).

Clarke, E. H. (1971). "Multi-part Pricing of Public Goods". *Public Choice.* 11: 17-22.

Coleman, James. (1988). “Social Capital in the Creation of Human Capital”. *American Journal of Sociology.* Vol. 94, Supplement S95-S120.

Dani, A. H. (1989). History of the Northern Areas of Pakistan, Islamabad: National Institute of Historical and Cultural Research.

Dayton-Johnson, Jeff. (2000a). “The Determinants of Collective Action on the Commons: A Model with Evidence from Mexico,” *Journal of Development Economics,* 62.

Dayton-Johnson, Jeff. (2000b). “Choosing Rules to Govern the Commons: A Model with Evidence from Mexico,” *Journal of Economic Behavior and Organization,* 42.

Dayton-Johnson, J. and Bardhan, P. (2003). ‘Inequality and the governance of water resources in Mexico and South India’, in (J.M.Baland, P. Bardhan and S. Bowles, eds), Inequality, Cooperation and Environmental Sustainability, New York: Russell Sage.

Deininger, Klaus and Pedro Olinto. (2000). “Asset Distribution, Inequality and Growth,” *Policy Research Working Papers,* 2375, World Bank, Washington, D.C.

Durlauf, S. (2002). ‘On the empirics of social capital’, *Economic Journal*, vol. 112 (November), pp. 459–79.

Durlauf, S and M. Fafchamps. (2005). “Social Capital”, *Handbook of Economic Growth*, Chapter 26, pp. 1639-1699.

Easterly, William, and Ross Levine. (1997). “Africa’s growth tragedy: policies and ethnic divisions,” *Quarterly Journal of Economics,* 112.

Frank, Robert. (1993). “Social Capital and Economic Development,” Workshop on Social Capital and Economic Development, American Academy of Arts and Sciences, Cambridge July 1993.

Glaeser, E.L. and DiPasquale, D. (1999). ‘Incentives and social capital: are homeowners better citizens?’, *Journal of Urban Economics*, vol. 45(2), pp. 354–84.

Glaeser, E., Laibson, D. and Sacerdote, B. (2002). ‘The economic approach to social capital’, *Economic Journal*, vol. 112(8), pp. F437–58.

Glaeser, E., Laibson, D. Scheinkman, J.A. and Soutter, C.L. (2000). ‘Measuring trust’, *Quarterly Journal of Economics*, vol. 65, pp. 622–846.

Government of Pakistan. (1984). “1981 District Census Report of Baltistan,” Islamabad.

Government of Pakistan. (1998). “Population and Housing Census of the Northern Areas 1998,” Census Bulletin-12, Islamabad.

Groves, T. (1973). “Incentives in Teams,” *Econometrica* 41: 617-31.

Hart, Oliver and Bengt Holmstrom. (1987). “The Theory of Contracts,” in T. Bewley, ed., Advances in Economic Theory, Fifth World Congress. Cambridge: Cambridge University Press.

Hausman, J. and W. E. Taylor (1981). “Panel Data and Unobservable Individual Effects”, *Econometrica*, Vol. 49, No. 6. (Nov., 1981), pp. 1377-1398.

Hirschman, Albert. (1967). Development projects observed, Washington: Brookings Institution.

Holmstrom, Bengt. (1982). “Moral Hazard in Teams,” 13 *Bell Journal of Economics* 324-40.

Isham, Jonathan, Deepa Narayan and Lant Pritchett (1996). “Does Participation Improve Performance? Establishing Causality with Subjective Data,” *World Bank Economic Review*, 9(2).

Jones, Ben and Ben Olken (2005). “Do Leaders Matter? National Leadership and Growth since World War II” *Quarterly Journal of Economics 120* (3), pp. 835-864.

Khan, Akhtar Hameed. (1996). Orangi Pilot Project: Reminiscences and Reflections, Karachi: Oxford University Press.

Khan, Shoaib Sultan and Tariq Hussain. (1984). “Principles and Implementation for Small Farmer Development,” First Annual Review, Aga Khan Rural Support Programme.

Khwaja, Asim Ijaz. (1999). “Revisiting Rural Infrastructural Investments,” Aga Khan Rural Support Programme.

Khwaja, Asim Ijaz (2006). “A Theory of Participation and Empowerment.” Mimeo. Kennedy School of Government, Harvard University.

Khwaja, Asim Ijaz (2004). “Is Increasing Community Participation Always a Good Thing?” *Journal of the European Economic Association*. Apr-May 2(2-3):427-36.

Knack, S. and P. Keefer (1997). “Does social capital have an economic payoff? A cross-country investigation”, *Quarterly Journal of Economics*; vol. 112, no. 4, pp. 1251-1288.

Lam, Wai Fung. (1998). Governing Irrigation Systems in Nepal: Institutions, Infrastructure and Collective Action, Oakland, California: ICS Press.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R.W. (1997). "Trust in large organizations." *American Economic Review Papers and Proceedings*, 87(2):333-38.

MacDonald, Kenneth Ian. (1989). “The Impact of Large Landslides on Human Settlements at Hopar, Karakoram Himalayas, Northern Pakistan,” *Annals of Glaciology,* 13, 185-9.

MacDonald, Kenneth Ian. (1994). “The Mediation of Risk: Ecology, Society and Authority in Askole, a Karakoram Mountain Agro-Pastoral Community,” Doctoral Thesis, University of Waterloo, Ontario, Canada.

Miguel, Edward. (2001). “Ethnic Diversity and School Funding in Kenya,” Center for International and Development Economics Research. Paper C01-119.

Narayan, Deepa. (1995). “The Contribution of People’s Participation: Evidence from 121 Rural Water Supply Projects,” *ESD Occasional Paper Series 1.*World Bank, Washington D.C.

Olken, Benjamin. (2005). “Monitoring Corruption: Evidence from a Field Experiment in Indonesia” NBER Working Paper #11753.

Olken, Benjamin. (2006). "Do Television and Radio Destroy Social Capital? Evidence from Indonesian Villages". NBER Working Paper No. W12561.

Olson, Mancur. (1965). "The Logic of Collective Action: Public Goods and the Theory of Groups," *Harvard Economic Studies* No. 124, Harvard University Press, Cambridge.

Ostrom, Elinor. (1990). Governing the Commons: The evolution of Institutions for Collective Action, New York: Cambridge University Press.

Parvez, Safdar and Ehsan-ul-Haq Jan. (1998). "Growth, Poverty and Inequality in the Programme Area," Working paper, Aga Khan Rural Support Program.

Putnam, Robert (with Robert Leonardi and Raffaela Nanetti). (1993). Making Democracy Work. Princeton: Princeton University Press.

Putnam, Robert. (1995). "Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America," *Political Science and Politics*. December, 1995.

Putnam, Robert. (2000). "Bowling Alone: The Collapse and Revival of American Community," New York. Simon and Schuster.

Stewart, Frances. (1978), Technology and underdevelopment, London : Macmillan.

Tang, Shui Yan. (1992). Institutions and Collective Action: Self-Governance in Irrigation, San Francisco : ICS Press.

Tendler, Judith (1993). New lessons from old projects: The workings of rural development in Northeast Brazil, A World Bank Operations Evaluation Study. Washington, D.C.: World Bank Operations Evaluation Department.

Tendler, Judith. (1997). Good government in the tropics, Johns Hopkins Studies in Development. Baltimore and London: Johns Hopkins University Press.

Uphoff, Norman. (1996). Learning from Gal Oya: Possibilities for Participatory Development and Post-Newtonian Social Science, London: London IT Publications.

Wade, Robert. (1987). Village Republics: Economic Conditions for Collective Action in South India, Cambridge, U.K: Cambridge University Press.

Warr, P.G. 1983. The private provision of a public good is independent of the distribution of income. *Economic Letters* 13, 207-11.

Whiteman, P.T.S. (1985). "Mountain Oases: A Technical Report of Agricultural Studies in the Hunza, Ishkoman and Yasin Valleys of Gilgit District," FAO/UNDP.

World Bank (1990). "The Aga Khan Rural Support Program in Pakistan: A Second Interim Evaluation," World Bank Operations and Evaluation Department, Washington, D.C.

World Bank (1996). The World Bank Participation Sourcebook, Washington, D.C.: The World Bank.

Figure 1: Map of Pakistan and Baltistan (inset)

Figure 2: Map of Baltistan with sampled communities indicated

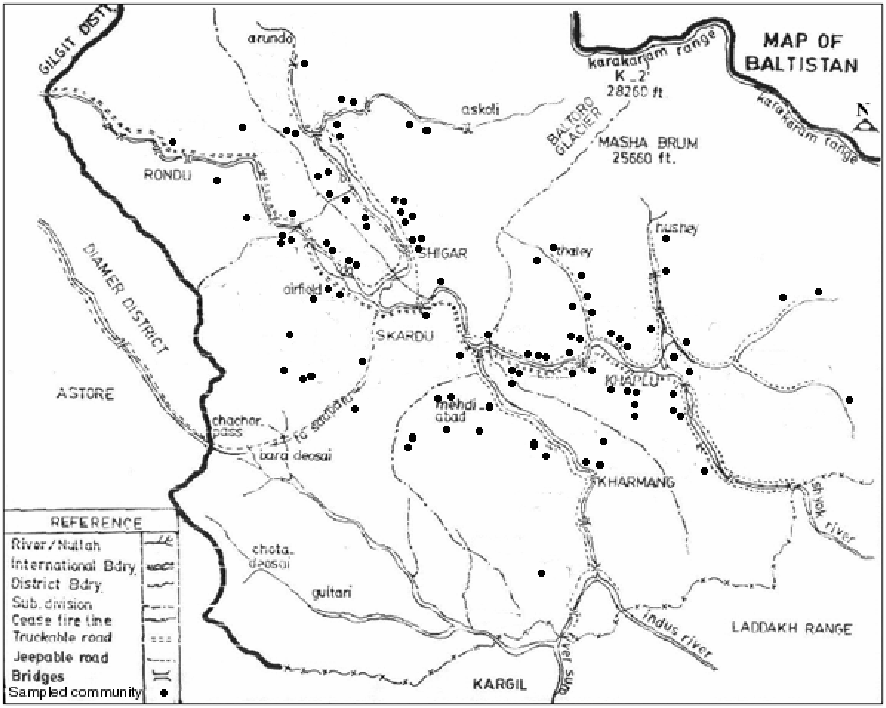


Figure 3: Predicted Outcome (Maintenance) for Best and Worst Design Project in Community

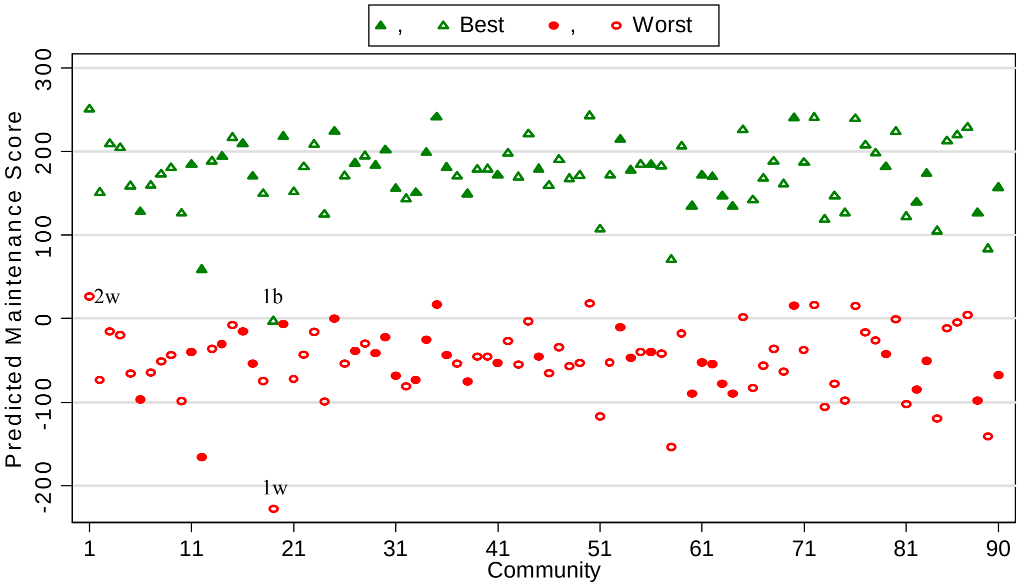


Figure 3 shows predicted maintenance scores for the best and worst designed project in our sample communities. It "forces" each community to have an irrigation project (most common) with a median age of 8 years. It then predicts the outcome (maintenance score) for this project in each community based on the estimates obtained in the community fixed effects regression reported in Table 3 by setting project-design factors at their worst (best) within sample levels (unlike the Table, in the figure we use the full sample of projects – even communities where only one project was sampled – the sample corresponding to the Table is indicated by the filled in geometric shapes). Thus the best (worst) project score in village 1 is obtained by predicting the score under the following values: project complexity index = 0 (3); community participation in non-technical decisions = 100 (0); community participation in technical decisions = 0 (100); project is a continuation (new) project; project is initiated by the NGO (Government); and project share inequality =0 (0.143). Note that the best value for project share inequality given the estimates is the maximum inequality value (1 in the sample). However, the best-case prediction chooses perfect equality since it is likely to be preferred for equity reasons. Doing so gives a lower estimate for the impact of the project-design factors. Thus the vertical distance between the best and worst prediction, is the magnitude of the community fixed effect for the given community.

Table 1: Sampled Projects by Project Type

1. Irrigation Channels	2. Protective Flood-Works	3. Pipe/ Siphon Irrigation	4. Lift Irrigation	5. Micro-Hydel Electricity	6. Link Roads	7. Boundary Walls
34	20	16	6	7	29	20

Table 2: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max	Mean (Multiple Projects sample)
Project Variables:						
Total Score	132	67.6	25.8	6.7	103.3	70.6
Physical score (0-110)[#]	132	74.8	20.5	0	110	76.9
Functional score (0-110)	132	71.1	35.4	0	110	75.3
Maintenance-work score (0-100)	132	57	28	10	100	60
Project Share Inequality	123	0.19	0.15	0	1	0.18
Project New? (1= new project)	132	0.83	.	0	1	0.84
AKRSP Project? (1=AKRSP, 0=LB&RD or Other)	132	0.77	.	0	1	0.53***
Government Project? (1=LB&RD, 0=AKRSP or Other)	132	0.17	.	0	1	0.36***
Project complexity (0-3)	132	1.3	1.1	0	3	1.2
Project leader exists? (1=yes)	132	0.68	.	0	1	0.67
Leader Quality	132	0.69	0.39	0	1	0.70
Community (direct) participation: non-technical decisions	132	30	19	0	78	29**
Community (total) participation: technical decisions	132	45	30	0	100	39
Project Age (years)	132	8.3	4	0[*]	29	8.4
External Funds in Project (000 Rs)	132	139.6	165.1	2.2	1400	153
Community Variables:						
Land Inequality	99	0.28	0.09	0.08	0.52	0.27
Social Heterogeneity	99	0.34	0.13	0.00	0.71	0.35
Community size (Number of households)	99	59	48	13	235	77***
Number of Public Projects	99	6.6	3.2	2	16	8.1***
Travel Time (min) – to capital by jeep	99	163.2	79.4	10	360	174.6
Walk time (min) – to road on foot	99	9	24.8	0	180	13.5
Community cultivatable Land (kanals[$])	99	1334	1409	80	7000	1613
Shopkeepers fraction	99	0.08	0.07	0	0.54	0.09
Skilled workers fraction	99	0.12	0.12	0	0.60	0.12
Basic education (Primary to higher secondary) fraction	99	1.20	0.88	0.03	4.60	1.08
Tertiary education (Graduate/post-graduate) fraction	99	0.08	0.10	0	0.45	0.09
Mean off-farm household income (Rs)	99	1971.7	1478.6	80	8000	2160
Mean real estate value (000 Rs/kanal)	99	136.1	110.3	1	700	138.3
Community Wage (Rs/hour)	99	64	10.7	30	130	65
Households with mechanized assets (no)	99	2.05	3.70	0	30	2.36
Single Cropping zone? (1=yes)	99	0.23	0.42	0	1	0.21
Access to Electricity? (1=yes)	99	0.62	0.49	0	1	0.64
Access to Health facilities? (1=yes)	99	0.47	0.50	0	1	0.55
Access to Potable water? (1=yes)	99	0.47	0.50	0	1	0.48

[#] The index for physical and functional score ranges from 0-110 instead of 0-100 as the score is increased by 10 if the community has made substantial extensions/modifications to the project in an effort to better capture the community's performance. This increase does not affect the results.

[*] Two projects in the sample were completed recently (several months prior to the survey) and are assigned an age of 0. Project scores are not significantly higher for these projects (since they took a couple of years to complete, earlier parts of the project were damaged) and the results are not driven by them. They are retained to reduce small sample biases.

* Both direct and indirect participation (household participated through a representative) is included for technical decisions since both will have a negative effect on maintenance, as they crowd out external organization participation. In the case of non-technical decisions, only direct participation is considered, as indirect participation is not a good measure of maximizing community participation and knowledge. Nevertheless, including or excluding indirect participation in either decision category, does not significantly affect the empirical findings.

[$] There are 8 kanals in one acre (43,560 square feet)

Table 3: The Impact of Project Design Factors on Maintenance

Variables	(1) FE	(2) FE	(3) OLS	(4) OLS– Full Sample	(5) Alt Score FE	(6) FE	(7) Functional Score – FE
Project Complexity	-12.76*** (3.85)	-15.44*** (3.92)	-4.06 (5.15)	-6.13** (2.55)	-12.58*** (3.94)	-14.16*** (4.40)	-14.82* (7.24)
Project Share Inequality	-373.3*** (67.7)	-422.03*** (69.26)	-83.93 (80.28)	-94.1*** (29.8)	-376.1*** (72.2)	-359.7*** (69.4)	-320.31* (155.41)
Project Share Inequality squared	1304*** (225)	1,381.46*** (211.16)	179.8 (154.21)	90*** (30)	1,316*** (234)	1,285*** (231)	1,088** (502)
Non-technical decisions participation	55.43* (28.29)	50.87* (24.24)	31.76 (25.49)	30.41** (15.14)	55.00* (27.67)	53.17 (30.93)	73.11 (48.52)
Technical decisions participation	-38.49* (18.56)	-34.00* (16.68)	-24.97* (14.70)	-17.61** (8.70)	-36.76* (19.12)	-30.46 (20.52)	-46.43 (31.01)
Government project?	-23.63*** (7.95)	-18.18** (8.03)	-17.06* (8.55)	-10.88* (5.76)	-23.41*** (7.89)	-22.44** (9.08)	-23.54 (15.71)
Project New?	-41.92*** (13.67)	-46.77*** (15.06)	-21.11 (12.81)	-22.09*** (4.69)	-39.85** (13.92)	-38.75** (14.62)	-58.54** (22.93)
Project Leader?		13.36 (8.42)					
Project External Funding?						-1.01e-06 (15.2e-06)	
Project Initial design/cost problem?						0.47 (10.08)	
Physical Score							0.42 (0.31)
Controls	Community Fixed Effects, Project Age and type	Community Fixed Effects, Project Age and type	Project Age and type	Project Age and type	Community Fixed Effects, Project Age and type	Community Fixed Effects, Project Age and type	Community Fixed Effects, Project Age and type
R^2	.93 *(0.43 if Community FEs only)*	.94	.27	.30	.93	.94	.93
N	64	64	64	132	64	63	64

Huber-White robust standard errors in parentheses
***, **, *: Significantly different from zero at 1%, 5% and 10% respectively

Columns 1-5 all include community and project-type fixed effects. Column 1 presents the primary regression. Column 2 checks to see whether the results remain similar once the potentially endogenous (due to halo effects) participation measure is excluded. Column 3 checks to see whether the external agency effect remains once the amount of external funds used in project construction and an indicator of the project's initial design/cost being a problem are controlled for. Column 4 uses an alternate outcome measure that employs factor analysis to combine the three underlying physical, functional and maintenance-work scores (rather than their simple average that is used in Column 1). Column 5 uses only functional score as the dependent variable while controlling for physical score to ensure that the results are not just due to initial construction quality. Column 6 runs the same specification but without using community FEs and in the larger sample (also includes communities where only a single project was surveyed).

Table 4: The Impact of Community-Specific Factors on Maintenance

Variables	(1) IV	(2) IV	(3) IV	Variables	(4)	(5)
Land Inequality	-238.77*	-321.22**	-262.05*		1st Stage	1st Stage
	(138.53)	(139.11)	(141.67)		(1) & (2)	(3)
Land Inequality Squared	301.70	451.62**	452.06*			
	(236.54)	(221.15)	(236.55)	Hereditary family	0.30*	
Social Heterogeneity	-38.37*	-48.99**	-39.55*	25-50 healthy male?	(0.18)	
	(22.95)	(19.87)	(20.09)	Hereditary family	-0.32***	
Project Leader?	32.31**	36.46**	30.23*	absence (1-3)	(0.12)	
	(16.13)	(17.25)	(16.75)	Hereditary family	0.02**	
Leader Quality			42.84**	average Age	(0.01)	
			(17.83)			
Number of Public Projects		-0.09	0.55			
		(1.03)	(1.03)			
Community Size	0.06	-0.01	-0.04			
	(0.06)	(0.10)	(0.08)	Hered family 25-50		0.10***
Travel Time	-0.03	-0.12	-0.04	educated,present male		(0.04)
	(0.03)	(0.10)	(0.08)	Hereditary family		-0.20**
Walk Time	-0.16	-0.04	-0.03	non-farm?		(0.08)
	(0.10)	(0.05)	(0.04)	Ideal leaders in		0.21***
Total Cultivable land		0.001	0.002	community? (1-4)		(0.08)
		(0.002)	(0.002)			
Single cropping zone?		-14.12*	-14.60**			
		(7.20)	(6.63)			
Shopkeepers fraction		-63.40*	5.29			
		(35.66)	(42.54)			
Skilled workers fraction		22.07	7.75			
		(18.03)	(20.65)			
Tertiary Education fraction		-8.65*	-3.24			
		(5.20)	(5.60)			
Project-Specific Controls	Age, Type, External Agency	Age, Type. Project variables in Table 3, Column 1	Age, Type. Project variables in Table 3, Column 1			
Community-Specific Controls		Human and Physical Capital measures±	Human and Physical Capital measures±			
Observations	132	132	132		132	132
R-squared	0.17	0.45	0.61		0.12	0.10

Huber-White robust standard errors in parentheses
Disturbance terms clustered at the community level
***, **, *: Significantly different from zero at 1%, 5% and 10% respectively

Columns 1-3 consider 2SLS estimates for the impact of community-specific factors in the full sample. Column 1 considers the base specification with leadership presence instrumented for. Column 2 adds extensive set of project and community-specific controls to the estimation in Column 1. Column 3 adds (and instruments) for leadership quality. Column 4 shows the instruments for leadership presence and Column 5 for leadership quality.

± Human Capital variables include: Shopkeepers fraction, Skilled workers fraction, Basic Education fraction, Tertiary Education fraction, and whether there is a High school in community. Physical Capital Variables include: Mean off-farm hh Income, Mean real estate value (000), Community wage, No. of households that own Mechanical assets, and whether the community has access to Electricity?, to a Health facility?, to Potable Water?

Table 5: Community and Project-Design Factor Interactions

	(1)	(2)	(3)	(4)	(5) IV	(6) IV	(7) IV
Design Factor 1 =	Project New?	Project Complexity	Government project?	Non-technical Participation	Project Complexity	Government project?	Project New?
Design Factor 2 =				Technical Participation			
Variables							
Design Factor 1	-14.73 (16.25)	-8.88*** (2.78)	-33.64*** (8.63)	72.05** (29.57)	-50.87*** (17.72)	-61.92 (41.20)	-6.63 (28.40)
Design Factor 1* Unequal Community	-42.02** (16.23)	-11.5** (4.95)	6.44 (10.89)	-60.94 (41.29)			
Design Factor 2				-50.29** (22.85)			
Design Factor 2* Unequal Community				36.52 (26.00)			
Design Factor 1* Leadership					36.18** (15.44)	65.79 (46.71)	5.21 (14.83)
Controls	Community & Project Type Fixed effects, and Project-design controls [#]	Community & Project Type Fixed effects, and Project-design controls [#]	Community & Project Type Fixed effects, and Project-design controls [#]	Community & Project Type Fixed effects, and Project-design controls [#]	Community & Project Type Fixed effects, and Project-design controls [#]	Community & Project Type Fixed effects, and Project-design controls [#]	Community & Project Type Fixed effects, and Project-design controls [#]
R^2	.95	.95	.95	.95	.94	.90	.94
N	64	64	64	64	64	64	64

Huber-White robust standard errors in parentheses

***, **, *: Significantly different from zero at 1%, 5% and 10% respectively

Columns 1-7 examine interactions between the particular Design factor(s) indicated under the Column heading and the community-specific factor indicated in the rows. All regressions include community-level FEs. Columns 1-4 examines project design factor interactions with community inequality where the community-specific factor is an indicator variable for whether a community has above sample-median inequality or not. Columns 5-7 consider interactions of project design with leadership, where leadership is instrumented by attributes of hereditary community leader households (see Column 4, Table 5 for the instruments used).

[#] Project-design controls are all the remaining project-specific design factors in Column 1, Table 3 not shown in the Column. Thus in the specification in Column 1, whether the project is new or not is the interaction and variable of interest (and therefore shown in the Column), while the project-design controls are project share inequality and its squared, project complexity, whether it was initiated by the government (or NGO) and community participation in non-technical and technical decisions.

Table 6: Design Factor Interactions

	(1)	(2)	(3)
Design Factor 1 =	Government?	Government?	Complexity
Design Factor 2 =	Complexity	Project New?	Project new?
Design Factor 1	8.11 (17.19)	1.81 (17.37)	-24.43 (17.83) 19%
Design Factor 2	0.37 (7.93)	-20.85 (17.68)	-53.81** (23.62)
Design Factor 1* Design Factor 2	-20.19** (8.91)	-30.87 (19.49) 13.5%	13.21 (20.30)
Controls	Community & Project Type Fixed effects, Project-design controls #	Community & Project Type Fixed effects, Project-design controls #	Community & Project Type Fixed effects, Project-design controls #
Adj R^2	.96	.94	.93
N	64	64	64

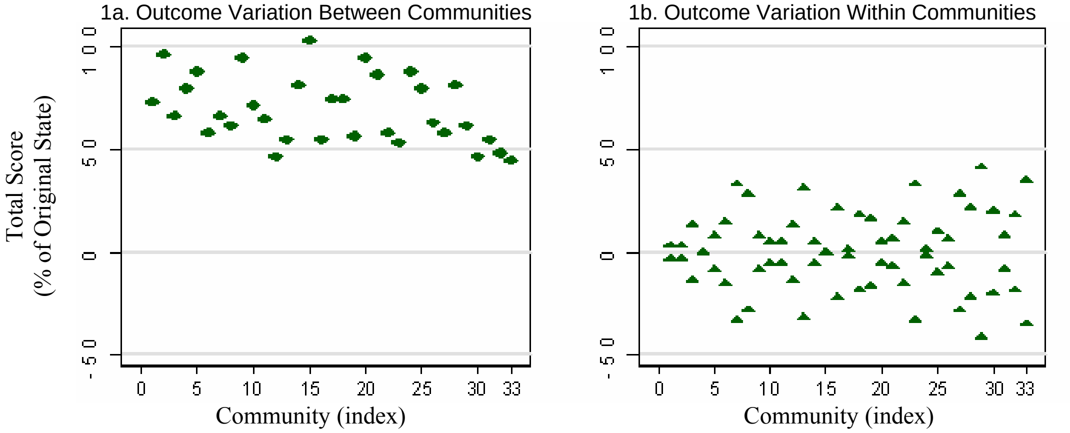
Huber-White robust standard errors in parentheses
***, **, *: Significantly different from zero at 1%, 5% and 10% respectively

Columns 1-3 examine interactions between the Design factors indicated under the Column heading. All regressions include community-level FEs.

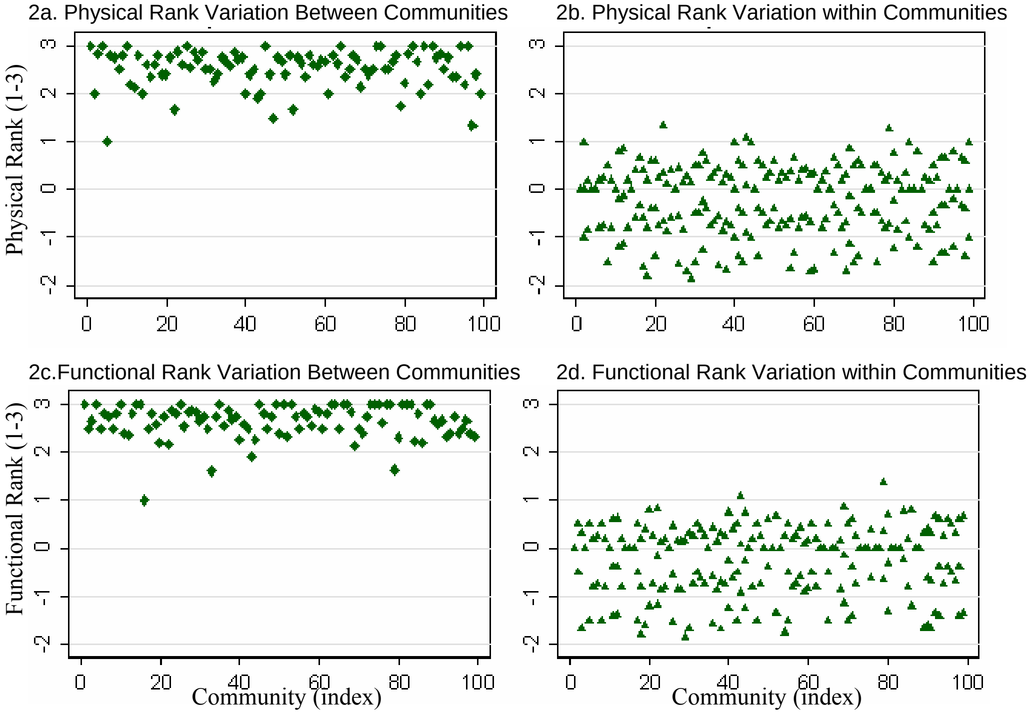
[#] Project-design controls are all the remaining project-specific design factors in Column 1, Table 3 not shown in the Column. Thus in the specification in Column 1, whether the project is government-initiated or not, project complexity and their interaction are the variables of interest (and therefore shown in the Column), while the project-design controls (not shown but included in the regression) are project share inequality and its squared, whether the project was new or not, and community participation in non-technical and technical decisions.

APPENDIX I

Figures 1a-b: Between and Within Community Variation in Outcomes in Primary Sample (communities with multiple infrastructure projects of selected types)

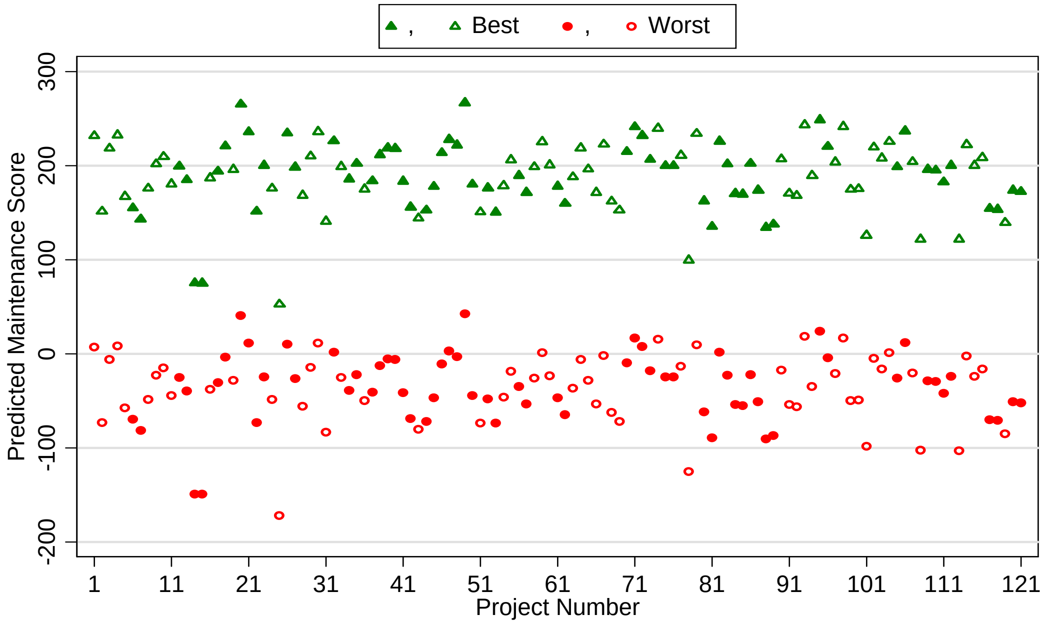

1a. Outcome Variation Between Communities
1b. Outcome Variation Within Communities

Figures 2b-d: Between and Within Community Variation in Outcomes Sample - All Communities and All Projects


2a. Physical Rank Variation Between Communities
2b. Physical Rank Variation within Communities
2c.Functional Rank Variation Between Communities
2d. Functional Rank Variation within Communities

The figures show variation in outcomes measures between and within communities and illustrate that there is substantial, if not greater, variation *within* communities.. Figures 1a-b do so for the primary (multiple-project) sample and outcome measure used in the paper (corresponding to the sample in Table 3). The outcome measure represents the percentage of the project that is in perfect physical, functional and maintenance-work condition and is based on community reports, surveyor and engineer visit to the projects and therefore of high quality. Each point in figure 1a represents the average score obtained for projects in a given community – differences across points therefore represent average performance differences across communities. Figure 1b instead plots outcomes for each project after subtracting the communities' mean outcome. Differences in points for a given community therefore reflect within community outcome variation: Were community attributes paramount then the points in Figure 1b would cluster around a value of 0. A potential concern is that multiple projects meeting our selection criteria are only found in 33 communities. Figures 2a-d address this concern by repeating the same exercise but for *all* (651) public projects in all 99 communities surveyed. Since a lot of these projects do not met our selection criteria, only crude outcome measures are available (self-reported ranks on a 1-3 scale of the project with 1 being a low and 3 a high rank). However, these figures show the same qualitative patterns of large outcome variation within communities.

Figure 3: Predicted Outcome (Maintenance) for Best and Worst designed projects



This figure is similar to Figure 3 in the main tables and figure, but instead of forcing each community to have the same type and age of project, we predict best and worst outcomes using the actual type and age of project in each community. As such communities with two projects will have a good and bad-design predicted score for each project (and consecutive project numbers)