Journal of Public Economics 93 (2009) 899–916

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Can good projects succeed in bad communities?

Asim Ijaz Khwaja

Harvard University, Cambridge MA 02138, USA

ARTICLE INFO

Article history:
Received 17 December 2007
Received in revised form 30 September 2008
Accepted 13 February 2009
Available online 5 March 2009

JEL classification:
O12
Z13
D7
H41
D02
D23

Keywords:
Public goods
Development
Collective action
Social capital
Institutional design

ABSTRACT

The lack of "social capital" is frequently given as an explanation for why communities perform poorly. Yet to what extent can project design compensate for these community-specific constraints? I address this question by examining determinants of collective success in a costly problem for developing economies — the upkeep of local public goods. It is often difficult to obtain reliable outcome measures for comparable collective tasks across well-defined communities. In order to address this I conducted detailed surveys of community-maintained infrastructure projects in Northern Pakistan. The findings show that while community-specific constraints do matter, their impact can be mitigated 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. The use of community fixed effects and instrumental variables offers a significant improvement in empirical identification over previous studies. These results provide evidence that appropriate design can enable projects to succeed even in "bad" communities.

1. 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 et al., 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 group-specific factors are important (see Baland and Platteau, 1996a,b; Durlauf and Fafchamps, 2005; Alesina and La Ferrara, 2005 for reviews).

Yet how essential are such group attributes in determining collective success? In particular, can the adverse impact of poor group attributes be compensated for by better project design? The literature on social capital has largely ignored this question. This paper utilizes primary data collected on 132 infrastructure projects from 99 rural communities in Northern Pakistan to show that even communities with low social capital can do well in collective tasks, if these tasks are well-designed. The measure of collective success used is the maintenance level of the projects. While these communities have over 651 potential projects, detailed functional and technical estimates of project condition were obtained for the 132 projects that were of comparable types and whose upkeep was solely the community's responsibility. In a third of the communities — half the project sample — this provides two separate projects, allowing for valuable within-community outcome variation.

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 considered to not 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, and networks present in the group, or its degree of inequality and fragmentation.

This substantial within-community outcome variation allows for better identification of the impact of project-design factors. By only comparing differences in design features between two projects in the same community, one can better isolate the project-design impact. Not doing so would bias estimates if, as is likely, communities choose/receive different design features. For example, if "better" communities choose more challenging design features, then not accounting for this results in a downward bias in the design factor estimates. Including community fixed effects avoids this bias.

E-mail address: asim_ijaz_khwaja@harvard.edu.

0047-2727/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpubeco.2009.02.010

The community fixed effects specification shows that design factors have a large impact: a standard deviation increase in the degree of project *Complexity* (indexed by higher capital/skilled labor requirements and lower experience) leads to a 13 percentage point drop in maintenance (measured on a 0–100 scale). The impact of *Community Participation* in project decisions made during its design/construction, usage, and maintenance stages is not beneficial for all types of decisions: While community participation in *Non-technical Decisions* (e.g. selecting project type, usage rules, etc.) of projects is beneficial, community participation in the *Technical Decisions* that have to be made in the same project (e.g. deciding project design, scale, etc.) is detrimental. The results indicate that a 10 percentage point increase in community participation in non-technical decisions raises maintenance by 6 percentage points. In contrast, a 10 percentage point community participation increase in technical decisions in a project lowers maintenance by 4 percentage points. *Inequality* in the distribution of project returns has a 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 point 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 fixed effects for each of the seven types of infrastructure projects considered[1] and a variety of project controls (such as project age, cost, etc.). This ensures that the comparison is between 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.

These project-design factor estimates are compared to the impact of 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. A hypothetical comparison shows that the best-designed project (in terms of feasible changes in 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 those with the best and worst average performance, i.e., the highest and lowest community fixed effects. This suggests that even a few design factors can mitigate the impact of poor group-attributes.

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 affects 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 well identified. By instrumenting for leadership with attributes of hereditary-leader households, such as whether the household has a healthy male member between 25 and 50 years old, this paper is able to do so. While these results for community-specific factors 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 more positively to both a lower project complexity and the project being 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 one recognizes that greater inequality may translate into less representative and hence less effective community 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, project design is even more important for communities with poorer attributes and lower social capital.

Regarding the generalizability of these results, while the sample is representative of the study region (randomly selected from the universe of communities), a note of caution is that the region itself — containing rural communities in a high altitude setting — may be different in ways that could affect the results. It is nevertheless noteworthy that the sampled region does have comparable community characteristics to rural communities in other developing countries and the estimates on community factors, such as inequality and heterogeneity, are quite comparable to those in previous studies. In addition, while the sample is relatively small, this is not atypical in this literature and small sample asymptotics do not pose a threat. Selecting a smaller set of projects from the over 650 possible ones was a conscious choice. This enabled cleaner comparisons and measurement — both by restricting to comparable projects and ensuring their upkeep was exclusively the community's responsibility — and is borne out in the resulting precision of the estimates.

The broader import of this work is significant: At a time when the literature is concerned with the importance of group-specific factors (such as a group's social capital) in determining group performance, this paper concludes 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 an older literature that emphasizes appropriate design/mechanisms as a means of solving collective problems (Hirschman, 1967; Clarke, 1971; Groves, 1973; Stewart, 1978; Betz et al., 1984; Tendler, 1997; Ostrom, 1990). 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.

The problem of project upkeep examined in the paper is also of substantial social interest, 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, 1996). Existing work in this area has consisted mostly of anecdotal evidence and case studies, and only recently econometric analyses of collective action determinants (Wade, 1987; Lam, 1998; Dayton-Johnson, 2000; Agarwal and Goyal, 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, allowing for cleaner estimates.

The paper also offers novel results for determinants of collective performance. It presents one of the few relatively clean estimates of the effect of leadership on group performance (see Jones and Olken, 2005 for an alternate instrumentation strategy). While the result on community participation not being beneficial in all project decisions is also novel, this is the sole focus of a previous article (Khwaja, 2004) and so is not stressed in this paper. Khwaja (2004) develops a theoretical model to explain how participation in all project decisions may not be beneficial, and then shows how the community

[1] The seven project types are Irrigation Channels, Protective Flood-Works, Pipe/Siphon Irrigation, Lift Irrigation, Micro-Hydel Electricity, Link Roads, and (protective) Boundary Walls.

Fig. 1. Map of Pakistan and Baltistan (inset).

participation results support the model. In contrast, the main focus of this paper is on the relative importance of a range of project-design and community-specific factors.

2. Data and methodology

2.1. Data

The sample consists of 99 rural communities in Baltistan, a state in the Himalayan region of Northern Pakistan (Fig. 1). These communities, either villages or *mohallahs* (sub-villages), were randomly selected from the population of communities obtained through a local NGO. Since the NGO had over 90% coverage in the area, this population is representative of the region. Fig. 2 shows a map of the region, indicating the selected communities (Government of Pakistan, 1998).

While these communities had an average of slightly over 6 public projects each, and very basic information was obtained regarding these (651) projects, detailed data was gathered only for the smaller sub-sample of projects (132) that satisfied the selection criteria — that the project be a comparable (and externally initiated) infrastructure project, and that *only* the sampled community be responsible for its upkeep. A second project satisfying these criteria was found in a third (33) of the sampled communities. Table 1 gives the breakdown of the sample by type of project. Although the cost of imposing these project selection criteria reduced sample size, this is justified by an improvement in measures and estimation strategies.

Since measuring collective success for any task is non-trivial, restricting the task to infrastructure projects simplifies matters considerably as reasonably precise and comparable engineering-based measures of current project state/condition are possible. The restriction to externally funded projects, while not essential, enables one to obtain accurate project cost, age, and characteristics data

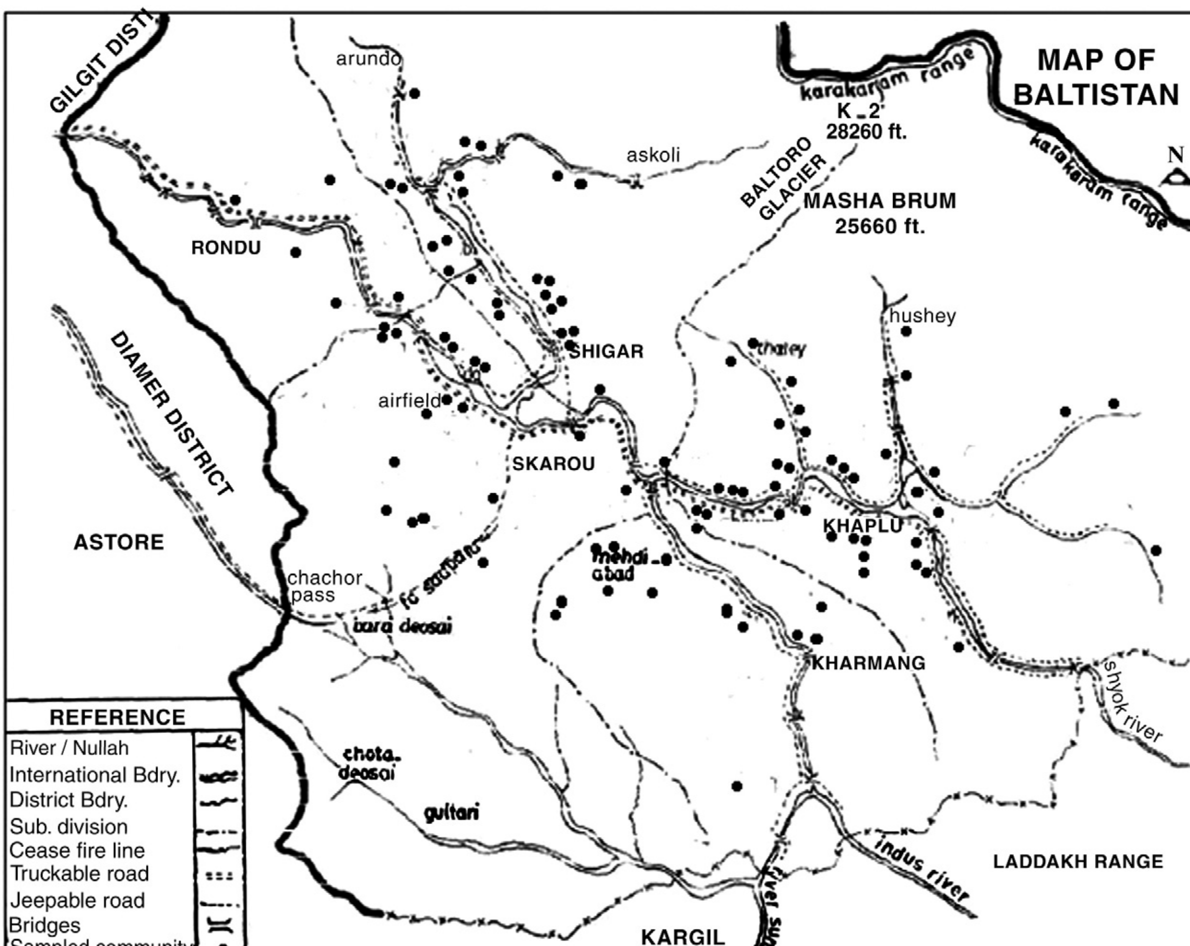

Fig. 2. Map of Baltistan with sampled communities indicated.

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

through these external agencies (government and non-governmental organizations). The projects primarily involved two 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).[2] Both initiate similar projects and only provide technical and financial support during construction. After construction, each organization's role ends and only the community is responsible for maintenance. As a result, it is not only possible to compare similar projects, but the variation in design between the two agencies' approaches can also be exploited.

The second consideration, that the project upkeep be solely the community's responsibility, ensures that the outcome of interest — the project's current state — provides an accurate measure of the community's collective abilities. While other outcomes like project completion/construction success may be confounded by external agencies' initial support, project maintenance was solely the community's responsibility.[3] Care was taken to ensure that in cases where multiple infrastructure projects did exist, they were only sampled provided the *same* group was exclusively responsible for the project's maintenance. This is in contrast to previous studies that rarely sample multiple tasks for a group or take care to ensure the underlying group remains the same for such multiple projects.

There are also several features of the chosen environment 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 due to the remoteness of the region,[4] but harsh weather conditions[5] 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 terrain 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 7000 to 12,000 ft above sea level. Thus, in addition to easily identified and distinct groups, there is significant variation across groups/communities. 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 in order to estimate the degree to which group-specific differences affect collective performance.

2.1.1. Survey instruments

Detailed community, household, and project-level surveys were carried out in the sampled communities. The primary surveys consisted of four separate questionnaires. Trained enumerators administered the first three and a team of engineers undertook the fourth. These surveys were complemented by secondary data sources, particularly regarding project construction cost details.

The first primary survey instrument, a group questionnaire, gathered 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. Finally, 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.

In terms of overall information gathered, these sources provide project state measures (current physical and operational condition, 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. Table 2 presents the descriptive statistics for variables of primary interest. Appendix Table A1 details the construction of the primary measures used.

Given its salience, the construction of the main outcome variable, the current state of maintenance of a project, is detailed here. This outcome is captured by three complementary measures: the project's *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 receiving water), and *Maintenance-work Score* estimates the percentage of required maintenance needs carried out.[6] The latter two, while more subjective, provide useful complements to *Physical Score*. Since these measures are highly

[2] 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 (World Bank, 1990).

[3] 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.

[4] 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.

[5] Temperatures vary from −20 °C in the winter to 40 °C in the summer. Floods, landslides, and avalanches occur frequently, and with very damaging consequences.

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

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)[a]	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[b]	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[c])	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/h)	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

*Both direct and indirect participation (household participated through a representative) are 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.

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

[a] 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.

[b] 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.

[c] There are 8 kanals in one acre (43,560 ft^2).

correlated (correlations ranging from 0.73 to 0.94), and the results are robust to using any one, a combined measure is used. 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.[7]

These scores represent a significant improvement over previous studies that typically rely on self-reported and simple rankings of projects. 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. The correlation between any of these sources is always more than 0.6 for all three scores. It is noteworthy 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.

2.2. Methodology

This section introduces a simple framework to help motivate the distinction between design and group-specific factors and outlines the empirical methodology used.

While the 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.[8] Therefore, distinguishing the impact of group-specific factors from those of task design that vary within a given group, such as the degree of complexity of the task, helps inform us of the relative importance of social capital. Moreover, most of the literature that emphasizes group attributes as determinants of collective success posits that such attributes have a direct/additive effect on outcomes (see Durlauf and Fafchamps, 2005 for a detailed review, also Knack and Keefer, 1997, and Alesina et al., 1999) suggesting the following additive specification:

$$Y_{ij} = \alpha + \mathbf{P}_{ij}\beta + \mathbf{X}_i\gamma + \eta_i + \varepsilon_{ij} \qquad (1)$$

where 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

[7] Annual project returns are also estimated for each project (Khwaja, 1999). 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 and Jan, 2008) with 6–8 members per household).

[8] Putnam (1995) views social capital as the network of social relations that links a group of agents and facilitates coordination and cooperation. The World Bank defines social capital as including, but not limited, to social networks and associated norms. Coleman (1988) considers it to be social structures that facilitate actions (neither within actors nor in physical implements of production). Frank (1993) treats it as the amalgam of formal and informal mechanisms for solving collective action problems. See Durlauf and Fafchamps (2005) for a detailed discussion on the range of definitions. Whether the empirical measures constructed as proxies for social capital are participation in associational activities (Putnam, 1995), or community income inequality and ethnic heterogeneity (Alesina and Le Ferrara, 2000), the point is that they are defined as invariant to the group, i.e. a group-specific attribute.

group-specific factors, η_i represents unobserved group characteristics and ε_{ip} is a general error term.[9]

2.2.1. Improving identification

One of the primary empirical challenges this paper deals with is to isolate the impact of (project/task) design factors. Since it is likely that groups choose or influence design factors, a simple OLS estimate of Eq. (1) would yield biased 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 biased downward (upward).

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

$$Y_{ij} = \alpha_i + \alpha_{j-\text{type}} + \mathbf{P}_{ij}\beta + \varepsilon_{ij} \quad (2)$$

where α_i is the group/community fixed effect[10] and $\alpha_{j\text{-type}}$ is a separate intercept for each project type. 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.[11] By not comparing across communities one is therefore able to deal with the concern that different communities may pick different design factors.

Estimating Eq. (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 identification concerns arising from omitted group factors. Surveys rarely sample multiple collective tasks for a given group. Even when this is done, a closer examination often reveals these projects do not fully overlap in terms of the responsible/beneficiary group. Second, one wants an outcome measure that can be reliably compared across projects. As discussed previously, care was taken to address both these concerns. Groups were carefully delineated and multiple projects only sampled if they shared exactly the same group, and restricting to externally initiated infrastructure projects allowed for comparable outcome measures.

Note that specification (2) also includes separate fixed effects for the type of the collective task in which type is one of seven different 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 type of the collective task/project. 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. Forcing the comparisons to be only *within* a given project type eliminates such concerns. A variety of other project level attributes such as project age, construction costs, etc. are also included as controls. Finally, I run further checks in cases where particular identification concerns are suspected. For example, for community participation in project decisions, halo effects are a potential concern: community members may (falsely) report higher participation in more successful projects. This is directly addressed by using individual level data to show that halo effects are unlikely.

The inclusion of project-type fixed effects also improves (policy) interpretation of the design factor results. In contrast to the type of a project, which is determined more by need, the 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. As these factors can be manipulated while addressing a given need, they 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 a given project type.

2.2.2. Sample size

Using specification (2) to obtain project-design factor estimates does impose a cost in terms of reducing sample size: Only a third (33) of the original community sample and half (66) of the original sampled projects have more than one project in the community that meets the project selection criteria. This raises an empirical concern as to whether the sample drop biases estimates of the design factors.

I address this concern in several ways. First, the data suggests that the sample drop, while systematic (one is more likely to obtain multiple projects of the required types in larger communities), is unlikely to bias the 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 that have multiple projects meeting the selection criteria really only differ in that they are somewhat larger and have roughly 1.5 more projects — factors that can be controlled for and which (it turns out) have little impact on project performance. Along all characteristics such as community inequality and the outcome measures, these communities are not different. Second, I can compare the specification (2) results to specifications without the group fixed effects both in the full and reduced sample. When I do so (Table 3), I find that the design estimates are not affected by change in sample size but rather (as one would expect) by the inclusion of the group fixed effect.[12]

A remaining concern is external validity. Since the 99 communities in the sample are randomly selected from essentially the universe of communities in the area of study, a state in Northern Pakistan, the sample is likely to be representative of this region. A harder concern to address in this or any similar study is whether communities in this setting — rural communities in a high altitude region — differ in a way that would affect generalizing the qualitative results. While it is not possible to directly address this without having similar data from other regions and countries, it is noteworthy that the observed community characteristics in the sample are comparable to rural communities in other developing countries, and these estimates of community factors such as community inequality and heterogeneity are similar to those from studies in other regions.

2.2.3. Comparing the impact of design and group-specific factors

I adopt two strategies to compare the impact of project design and group-specific factors. First, I compare the estimates of design factors to the magnitude of the group fixed effect estimated in specification

[9] A more general specification would also allow for variables to vary over time. However, given this analysis makes use of cross-sectional data, I 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.

[10] 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).

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

[12] One could also ask whether the restriction to specific types of infrastructure projects that selects roughly a fifth (132) of all (651) public projects biases results. While this is hard to rule out without detailed data on all the 651 projects and a way to construct meaningful and comparable outcome measures (these projects include schools, community halls, public graveyards, etc.), this additional concern seems unlikely. The restriction 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.

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)	1316*** (234)	1285*** (231)	1088** (502)
Non-technical decisions com. 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 com. 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	0.93 (0.43 if Community FEs only)	0.94	0.27	0.30	0.93	0.94	0.93
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).

(2). Under the assumption that Eq. (2) is correctly specified, the group fixed effect provides a comparison of the "worst" and 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 having to define or measure them. A disadvantage, in addition to relying on the particular form of specification (2), is that one cannot separately consider the effect of different group factors.[13]

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 the inclusion of a set of group-specific factors. While this approach has the advantage of estimating the impact of group factors separately, to the extent the unobservable group 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. I minimize such concerns by focusing on group factors that are more likely to be exogenous. For example, given almost non-existent land markets in this area of study, land inequality is likely to be based on initial settlement patterns in a community and not be affected by more recent factors such as project design or outcomes. In addition, in estimating the impact of leadership, I use instrumental variable techniques to improve identification.

2.2.4. Interaction effects

Eq. (2) can also be extended by allowing for interaction effects between group and design factors. This is particularly relevant, since if it is found 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 the results will show, design factors matter even more for groups with lower social capital.

Since omitted group factors are an issue as in Eq. (1), interaction effects are estimated in the reduced sample that allows the inclusion of group fixed effects:

$$Y_{ij} = \alpha_i + \alpha_{j-\text{type}} + \mathbf{P}_{ij}\beta + \mathbf{I}_{ij}\delta + \varepsilon_{ij} \quad (3)$$

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 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.

13 To the extent that group factors also influence design, they have both a direct and indirect (through 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 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 the results show, the indirect effects generally work in reducing the overall group effects (since better groups tend to pick or be assigned more challenging designs). Since this paper argues that design factors are as important, I stick to the larger (direct) estimates of the group-factors as this makes the comparison with design factors even more compelling.

3. Results

3.1. 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 *Total Score* (over projects in a given community) 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, they were not surveyed in 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 this sample is not used in the empirical analysis.[14] However, it still serves 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 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.

Figs. A1–A2 in the Appendix 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 the differences in plotted 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.[15] 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 importance (i.e., the number of tasks, or type of tasks matters), the question of interest is how much project "design" factors (rather than just type or number) matter relative to community-specific factors — i.e., whether changes in project design can counter any detrimental effects arising from poor community attributes. The next sections carry out this analysis.

3.2. Project-design factors

Table 3 estimates specification (2) to identify the impact of project-design factors. 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.

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. 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 (see Khwaja, 2006, for a formal treatment of this effect). A failed lift-irrigation project offered a stark illustration: While cash inputs were needed to maintain the lift-pump, cash constraints were an unlikely cause for failure since the community was well off and individuals successfully operated private pumps. Community members were *unwilling*, not unable, to provide cash as they were unsure of how it was spent and disagreed on whether the reported expenditure had been necessary or even taken place. Members believed that projects not requiring regular cash contributions would fare better, even if they needed a cash-equivalent amount of labor.

Project Share Inequality, a measure of how the project returns are distributed among community members, has a U-shaped effect on maintenance: Increasing inequality by 0.1 unit 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., I compare how differences in the distribution of benefits between two different projects in the *same* community impact their maintenance.

Field evidence (see Olson, 1965; Baland and Platteau, 1996a,b, 1998; Bardhan et al., 2005; Khwaja, 2006 for related models) offers potential explanations for the non-linear relationship. The initial detrimental effect arises from free-riding considerations — as a

[14] An added concern 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 it does not suffer from this error.

[15] This is possible if comparisons between communities have less noise as it is averaged out, or it if there is just more noise in outcomes within communities.

member's project share increases, they do not raise their own contribution enough to compensate for the losing members' contribution drop.[16] However, as shares increase further, members with larger shares may be more willing to undertake the maintenance costs by using (outside) paid labor/inputs. In the extreme, the project is effectively privatized and maintenance work (mostly) contracted out. While this may have poor distributional implications for the project, its maintenance improves.

A note of caution in forwarding such an interpretation, however, is that the estimates show the positive marginal impact starts after the Gini becomes greater than 0.14. This is not a particularly high level of inequality and therefore it is hard to think that the project has become privatized at this stage. While it is possible that the project return inequality is an underestimate (see Appendix Table A1), this also suggests that a more general mechanism may be at work. For example, Fritzen (2007) suggests that control by a small number of (diverse) members/elites may be beneficial if it allows for improved investment incentives while retaining representation. Thus the positive effects of increasing inequality may kick in earlier if a small number of people *collectively* gain sufficient shares.

Either way, under these explanations, one would expect to see that members with greater project returns should be contributing more. An examination of several projects from the field indeed revealed this to be the case. Moreover, the data also provides further evidence. While individuals were not asked their actual benefits for each project, they were asked to rank their relative benefits and maintenance costs for each project. This was asked on a 1 (a lot less) to 5 (a lot more) scale. There is a very strong positive relationship between the two and this correlation becomes twice as large for projects with inequality greater then the 0.14. This suggests that indeed those who gain more do contribute more and this relationship gets stronger when the project has a more unequal return distribution.

One of the more surprising results is for *Community Participation* in project decisions. Since the specification uses community fixed effects, the participation effect does not pick up features of the community that 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; Uphoff, 1996) but the results in Column (1) show otherwise: While community participation in *Non-technical Decisions* of a project (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 in 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.[17]

Since this result is the sole focus of a previous shorter paper (Khwaja, 2004), the reader is referred to it for a more detailed explanation and formal model. The explanation, outlined here, draws on the property-rights literature (Hart and Holmstrom, 1987) that posits that, given non-contractible investments, an asset is optimally owned by the party with the more important investment. I extend this to decisions by treating the decision-right as an "asset" and participation in this decision as capturing the relative influence/ownership of the participating agent. For example, consider two decisions for the same project — a non-technical one, such as deciding how to manage the project, and a technical one, such as selecting the appropriate design. The former likely benefits more from the community's investment, while the latter benefits from the external agency's investment (since it has more technical expertise). As such, increased community participation helps in the first decision but can hurt in the second one, since it provides lower incentives for the external agency's investment. The lower incentives arise because the external agency's influence is reduced due to the greater community participation.

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.

External Organization type also has a significant effect, as NGO-initiated (AKRSP) projects are associated with a 23.6 percentage points higher maintenance score 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 a result of initial construction quality: The NGO constructs better projects and so, regardless of the community's maintenance effort, the projects remain in a better state. I offer several reasons why this is less likely, suggesting the result is more consistent with better community collective ability for the NGO-initiated projects. Column (6) 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,[18] 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, one would expect that current physical score is determined primarily by construction quality, and therefore, project *Physical Score* provides a potential measure of initial construction quality. Column (7) 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 therefore suggest that the NGO effect is not as likely to be primarily driven by higher construction quality or a higher funds invested effect.[19] While the data does not allow sharper tests to

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

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

[18] This amount is underestimated for the NGO if it incurs greater, unaccounted for overhead costs. 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.

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

isolate potential explanations, some tentative hypotheses will be offered later in the paper.[20]

I also examine the impact of whether the project was built from scratch (*Project New*) or was a (substantial) *Extension* or *Continuation* of an existing community project. It is not clear ex ante 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 or a greater need that leads to a higher likelihood of being well-maintained. The results show that continuation (extension) projects are associated with a 41.9 percentage point higher maintenance score than new projects (Column (1)). Since project-specific factors such as project type and complexity are controlled for, the result is not driven by new projects being of a different type or more complex. The estimate is also robust to controlling for project need as proxied by community members' rank (1 to 4). Moreover, interaction effects presented later show that the poor performance of new (and complex) projects is less of an issue when the project is initiated by the NGO. This is consistent with the idea that new projects pose more of a challenge by requiring the set up of management systems and rules that anecdotal evidence suggests the NGO is better able to coordinate. However, given data limitations it is hard to isolate a particular explanation.

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 percentage 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. I 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.

3.2.1. 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 required if community fixed effects are to be used. Is the restriction to communities where multiple projects meeting the selection criteria were found likely to bias the design factor estimates? Recall that the data description had already shown that these communities only differ in that they are generally larger in size. Columns (3)–(4) together show that the design factor estimates depend on being able to correct for omitted community-level variables rather than the 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 designs. 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. A generalized Hausman test rejects equality of coefficients between the two estimations with and without community fixed effects (Columns (1) and (3)). This reinforces the importance of using community fixed effects in correctly identifying design factor impacts.

Column (4) then asks the next question — does the sample restriction matter? It runs the same specification as in Column (3) (without community fixed effects) but in the full 132 projects sample. If the sample restriction is 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. the sample size reduction does not seem to matter.

Column (5) next 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 similar results.

Columns (6) and (7) were discussed previously when examining the impact of the external organization. To review, Column (6) adds additional controls 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 an initial external agency-specific condition. Column (7) carries out a similar but more stringent test of accounting for initial conditions in Column (6) by controlling for a project's physical score. It shows that the design factors also have a similar impact on the current functional/operational state of the project *conditional* on its physical/construction quality.[21]

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, the results remain the same even when explicitly allowing for right (or both left and right) censoring and estimating tobit regressions.

3.3. Comparing design to community factors

How large is the impact of design factors compared to community ones? As described in the methodology section, two different comparisons are provided — the first, draws a comparison to the magnitude of community fixed effects and the second, to the effect of specific community factors.

[20] While the local government employs more staff, it pays them lower salaries. This hints at lower incentives in the local government offices. The NGO may also be more aware of the local environment and needs, less prone to corruption, have greater accountability and transparency, attract a more dedicated staff, or elicit greater community participation. The latter is likely since the NGO carries out all project decisions through a village organization that has community representation. In contrast, the local government has no such explicit emphasis and funds are usually disbursed through a representative appointed by local politicians. However, the data is unable to distinguish among 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) I will present results that suggest that NGOs may be better at setting up the management systems required for the newer and complex projects.

[21] It is interesting to note that the coefficients on the design factors are similar to those in Column 1. This suggests that a lot of the variation explained in the main outcome, *Total Score*, is in its functional score component. However, note that this does not mean that any one of the three measures used to construct the main outcome measure is not by themselves affected by these factors. Separate regressions (not shown) using each of the three component scores (physical, functional and maintenance-work) as outcomes shows that these project-design factors all have comparable impacts.

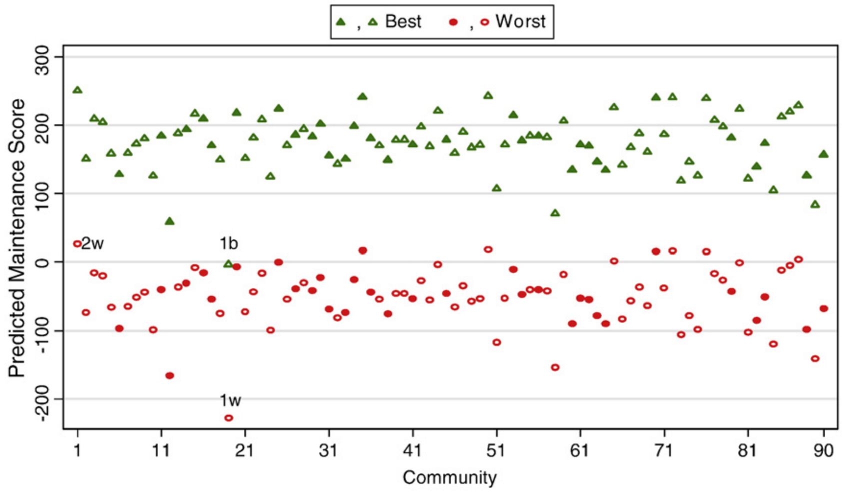

Fig. 3. Predicted outcome (maintenance) for best and worst design project in community. The figure shows predicted maintenance scores for the best and worst designed project in the sample communities. It "forces" each community to have an irrigation project (most common) with a median age of 8 years. 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 the full sample of projects is used — 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. Note that the vertical distance between the best and worst prediction is the magnitude of the community fixed effect, as estimated in the specification in Column (1), Table 3.

3.3.1. Community fixed effects

The estimates of the community fixed effects allow one to compare the impact of the above design factors with the *combined* impact of all observed and unobserved community-specific factors.

Fig. 3 illustrates this by conducting the following thought experiment: If one could change project-design 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 sample* values. The predictions are based on estimates in Column (1) of Table 3 and also include communities with only one project in the 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, points 2w and 1w in the figure) is the *largest* community effect as estimated in the regression in Column (1), Table 3. Provided Eq. (2) is appropriately specified, 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 can then be 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 Fig. 3 shows the *best-designed project in the worst community (1b) does as well as the worst-designed project in the best community (2w)*. A potential caveat is whether it is feasible to "set" design factors to their best and worst values. While this is hard to answer since one lacks the counterfactual, the field context suggested that for the design factors considered this is likely the case — one can solicit the best values of community participation, external agency, etc. In fact, it is for this reason that while the best return inequality value is perfect inequality (Gini = 1), the exercise conducted instead uses perfect equality as the best design value since this is likely to be more desirable. While one may question whether this is feasible, in this context even for irrigation projects where one may have thought land distribution was given a priori, the benefits generated were often for new land which could be distributed equally.

Nevertheless, one can also provide a comparison using 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 Fig. A3 performs a similar hypothetical comparison of best and worst projects. Instead of comparing a hypothetical 8-year-old irrigation project in each community, it uses the actual project type and age in the community and changes the project-design factors as before. Thus communities with multiple projects will have a best and worst predicted score for each of the projects.

These figures illustrate the main observation this paper makes: design factors have a comparable, if not larger, impact relative to 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.[22]

3.3.2. Community factors

I next estimate the impact of particular community-specific factors that have been emphasized in the literature, such as group inequality

[22] 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, Glaeser 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 this case, it could be that it just happens that desirable group attributes in this 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.

Table 4
The impact of community-specific factors on maintenance.

Variables	(1) IV	(2) IV	(3) IV	Variables	(4)	(5)
Land inequality	−238.77* (138.53)	−321.22** (139.11)	−262.05* (141.67)		1st stage (1) and (2)	1st stage (3)
Land inequality squared	301.70 (236.54)	451.62** (221.15)	452.06* (236.55)	Hereditary family 25–50 healthy male?	0.30* (0.18)	
Social heterogeneity	−38.37* (22.95)	−48.99** (19.87)	−39.55* (20.09)	Hereditary family absence (1–3)	−0.32*** (0.12)	
Project leader?	32.31** (16.13)	36.46** (17.25)	30.23* (16.75)	Hereditary family average age	0.02** (0.01)	
Leader quality			42.84** (17.83)			
Number of public projects		−0.09 (1.03)	0.55 (1.03)			
Community size	0.06 (0.06)	−0.01 (0.10)	−0.04 (0.08)	Hereditary family 25–50, educated, present male		0.10*** (0.04)
Travel time	−0.03 (0.03)	−0.12 (0.10)	−0.04 (0.08)	Hereditary family non-farm?		−0.20** (0.08)
Walk time	−0.16 (0.10)	−0.04 (0.05)	−0.03 (0.04)	Ideal leaders in community? (1–4)		0.21*** (0.08)
Total cultivable land		0.001 (0.002)	0.002 (0.002)			
Single cropping zone?		−14.12* (7.20)	−14.60** (6.63)			
Shopkeepers fraction		−63.40* (35.66)	5.29 (42.54)			
Skilled workers fraction		22.07 (18.03)	7.75 (20.65)			
Tertiary education fraction		−8.65* (5.20)	−3.24 (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[a]	Human and physical capital measures[a]			
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 an 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.
[a] 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 Household 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.

and heterogeneity (Alesina and La Ferrara, 2000). While these factors are of independent interest, estimating their impact also allows a relative comparison to the design factors, both individually and (at the end of the section) jointly. Columns (1)–(3) in Table 4 present the results from estimating Eq. (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, and its 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 examining the impact of salient community-specific factors, a note of caution is that these estimates are potentially biased since the 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.[23] Only the community factors of interest are discussed.

[23] 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 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 these estimates.

The literature has emphasized the impact inequality has on collective action, although the sign of the impact is ambiguous with models suggesting that inequality has no effect, (Bergstrom et al., 1986; Warr, 1983), is detrimental (Baland and Platteau, 1998; Alesina and La Ferrara, 2000), or has a positive impact for at least at some parts of the distribution (Olson, 1965; Baland and Platteau, 1998). Khwaja (2006) suggests a quadratic effect of inequality for reasons similar to those posited for the impact of inequality in a project's share. Before presenting the results, it is worth noting that the region underwent land reforms in the 1970s that transferred ownership rights from local rulers to tenants (Dani, 1989; Afridi, 1988). As a result, 79% of farmers own their land and only 2.2% are tenants (Government of Pakistan, 1984). Therefore land-holdings in this primarily agrarian setting provide a reasonable measure of wealth distribution and hence community inequality.

The results show that *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 point *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. Moreover, as is common in 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 likely to be exogenous to project outcomes and therefore the bias in the estimates is low.

The measure of fragmentation emphasized in the literature 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. 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. Since political affiliations are also based on religious and familial associations rather than party platforms, there is little movement across party lines. Thus social fragmentation in a community is also likely to be exogenous to project outcomes and biases arising from reverse causal impacts are unlikely to be significant.

While there are 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, 2005), the empirical literature on this is limited (Jones and Olken, 2005 is a notable exception). Part of the problem arises in drawing causal inferences: 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 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 minimizes this problem by allowing one to instrument for leadership. 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. As noted by MacDonald (1994), "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." 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 the instruments will be community-invariant, I 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[24] 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 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 with leader quality.[25]

The estimation in Table 4 also considers other community-specific factors, some of which are noteworthy. First, note that *Number of Public Projects* does not matter suggesting that there is little concern that community resources are stretched if they undertake 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 literature, it is not size that matters per se but the greater inequality and heterogeneity (more likely to be present in larger groups) that hinders collective action. *Community Remoteness* (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. Columns (2)–(3) also included measures for a community's physical wealth (real estate, mechanized assets, off-farm income, etc.) and infrastructure (access to potable water, electricity, and health facilities) but find these factors have few significant effects.[26]

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 3 and 4) 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 community-factors.[27] This is not surprising

[24] 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.

[25] 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).

[26] 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 in-migrated recently is associated with a one percentage point fall in maintenance. These estimates control for human/physical capital, and project-specific factors.

[27] 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.

Table 5
Community and project-design factor interactions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
					IV	IV	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 and Project type fixed effects, and project-design controls[a]	Community and Project type fixed effects, and project-design controls[a]	Community and project type fixed effects, and project-design controls[a]	Community and project type fixed effects, and project-design controls[a]	Community and project type fixed effects, and project-design controls[a]	Community and project type fixed effects, and project-design controls[a]	Community and project type fixed effects, and project-design controls[a]
R^2	0.95	0.95	0.95	0.95	0.94	0.90	0.94
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).
[a] 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.

given the results depicted in Fig. 3 that compared the impact of the selected project-specific factors to all community-specific factors.

3.4. Heterogeneity in the impact of project design

One may expect that there are interactions between the project-design and community-specific factors. The presence of such interaction effects, while requiring a more nuanced interpretation, adds to the salience of design particularly if design factors are even more important in worse off communities. Given the sample size, a thorough exploration of interaction effects would be asking too much of the data. Nevertheless, Tables 5 and 6 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 the 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. 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 of inequality also fare poorly if the project is new rather than a continuation project. 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)

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 and project type fixed effects, project-design controls[a]	Community and project type fixed effects, project-design controls[a]	Community and project type fixed effects, project-design controls[a]
Adj R^2	0.96	0.94	0.93
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.
[a] 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.

shows *NGO-initiated* projects outperform *Government-initiated* projects in both equal and unequal communities.

However, in the case of community participation, Column (4) shows community inequality *dampens* both the beneficial and detrimental effects of community participation. While a 10 percentage point increase in community participation in non-technical decisions leads to a 7 percentage point increase in project maintenance in an equal community, the same increase has barely any effect (a 1 percentage 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 percentage point drop in project maintenance in equal communities, there is again little effect (1.4 percentage points) in unequal communities though the interaction term is only significant at 18%. These interactions are plausible if one considers that the 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 given more weight than others. Thus any effect of a community's participation, whether positive or negative, is dampened in more unequal communities.[28]

Columns (5)–(7) next explore interactions between leadership and project-design factors. As in Table 4, 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 percentage 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.[29]

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 between project-specific factors. Columns (1) and (2) show that the poorer performance of *Government-initiated* projects is most apparent 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 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 any extension of these projects is able to make use of this existing capital. While these interactions help further the understanding of how these design factors matter, the limited power to detect them means these results are tentative and serve better as suggestions for future lines of enquiry.

4. 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. This is particularly significant at a time when the debate on 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. 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-specific 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.

Acknowledgements

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 thank the co-editor, Thomas Piketty, and two anonymous referees 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.

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

[29] While not significant at conventional levels and not as robust, the interaction of leadership (instrumented) with whether a community is large (above median size) is also positive. This offers further support that leaders matter more where the task is more challenging, whether because of the nature of the project, or simply due to the challenges involved in managing a larger set of contributors.

Appendix A

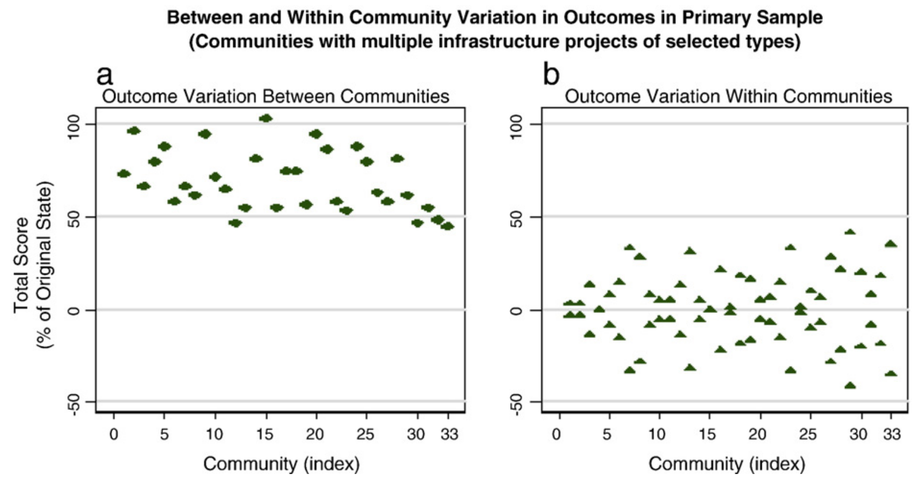

Fig. A1. a and b. Between and within community variation in outcomes in primary sample (communities with multiple infrastructure projects of selected types).

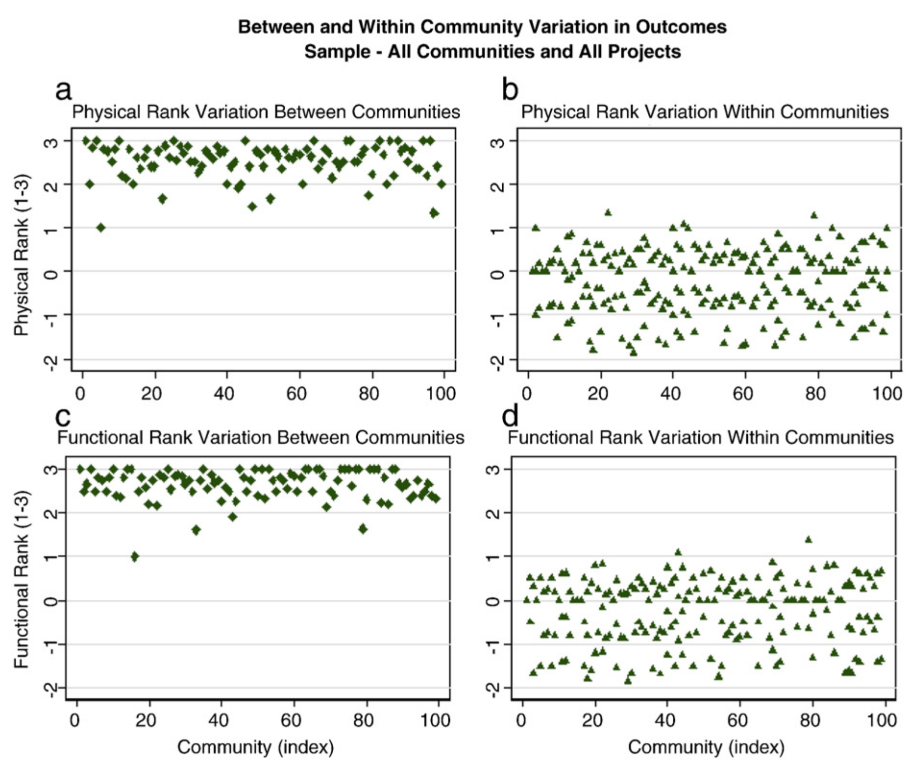

Fig. A2. a–d. The figures show variation in outcomes measures between and within communities and illustrate that there is substantial, if not greater, variation within communities. Fig. A1a–b do so for the primary (multiple-project) sample and outcome measure used in the paper (corresponding to the sample 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 is therefore of high quality. Each point in Fig. A1a represents the average score obtained for projects in a given community — differences across points therefore represent average performance differences across communities. Fig. A1b 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: if community attributes are paramount then the points in Fig. A1b would cluster around a value of 0. A potential concern is that multiple projects meeting the selection criteria are only found in 33 communities. Fig. A2a–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 meet the 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.

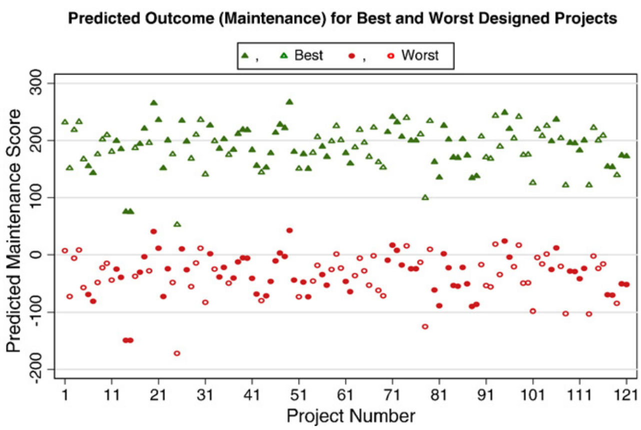

Fig. A3. This figure is similar to Fig. 3 in the main tables and figure, but instead of forcing each community to have the same type and age of project, best and worst outcomes are predicted using the actual type and age of project in each community. As such communities with two projects will have a good and bad design-predicted score for each project (and consecutive project numbers).

Table A1
Variable construction

Outcome variable: Total score (described in main text)	
Selected community factors	
Land equality	
Construction	Constructed in a manner similar to a Gini coefficient except based on grouped (rather than individual) data due to feasibility constraints. To construct the inequality index, 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 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 then be constructed.
Details	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 outliers. Moreover, the problem of misreporting 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.
Social heterogeneity	
Construction	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).
Details	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 of s_k^2) where s_k is the proportion of the k^{th} group in the community. Higher values of the index represent greater heterogeneity.
Project leader	
Construction	Binary variable that indicates whether the project has a leader or not, i.e., an individual selected by the community to manage the project.
Details	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 only exploits community-level attributes (instrument using demographic characteristics of the hereditary leader household), for the purposes of this paper leadership is considered a community-specific factor.
Leader quality	
Construction	Average of the five community individuals' evaluation (good or bad) of the project leader's quality.

Table A1 (*continued*)

Project design factors	
Project share inequality	
Construction	Calculated in a manner analogous to the land inequality gini measure above using "grouped" data. It measures the inequality in observed division of project returns.
Details	For projects where the primary purpose was to irrigate land (the three types of irrigation projects: irrigation channels; pipe/siphon; and lift irrigation), the measure uses the distribution (across households) of land (both old and/or newly developed) in the command area of the project. For flood protection and boundary wall projects, since the primary purpose of these projects was to protect cultivated land (from flood damage or animals), the distribution of the protected land is used. For Micro-Hydels projects, the distribution of the total wattage generated from the project is employed. For Link road projects, pilot surveys revealed the main benefits were the utilization of the road and, in slightly over half the cases, an additional benefit of land appreciation. The measure was therefore constructed using both the distribution of road usage (in terms of monthly trips) and the distribution of any appreciated land.
Project complexity	
Construction	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	
Construction	Dummy variable indicating whether the project is a new project instead of a continuation project, i.e., a project made as an extension to an existing community project.
Details	While extension work for a continuation project could 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.
Government/AKRSP project	
Construction	Dummy variables indicating the type of external organization involved in project construction.
Details	The primary external organization comparison is between the local government and an NGO, the AKRSP. There are a few other semi-governmental external agencies in the sample but too few observations to allow any meaningful comparisons.
Community participation	
Construction	Measured using various decisions that are made during a project's construction, usage, and upkeep. While participation responses were solicited for 20 such decisions, the final measures are grouped into two types for each project: non-technical and technical decisions. Each measure is constructed by averaging over community participation in each individual decision included in the two types. Community participation in an individual decision is measured as the fraction of the five randomly selected community respondents in the individual questionnaire who answered affirmatively to whether their household participated (directly or indirectly, i.e., through a proxy) in the decision.
Details	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

(continued on next page)

Table A1 (*continued*)

Outcome variable: Total score (described in main text)	
Project design factors	
Community participation	
Details	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. A cautionary note is that these participation measures are constructed using only five individual respondents. However, as noted above, each of the two measures is obtained by averaging over reported participation in several different decisions (5 separate decisions for the technical participation measure and 15 different decisions for the non-technical measure). Moreover, it is reassuring that the measures constructed from the individual-level data strongly correlate with measures from corresponding questions asked for each decision in the group level questionnaire regarding villagers' participation in the decision. I find that the participation level as obtained from the individual questionnaire are on average 5.5 times higher when the group questionnaire reports that the decision involved most villagers.

Note: Questionnaires can be accessed at http://ksghome.harvard.edu/~akhwaja/Balt/Allquest.pdf.

References

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

Agarwal, Arun, Goyal, Sanjeev, 2001. Group size and collective action: third party monitoring in common-pool resources. Comparative Political Studies 34 (1), 63–93.

Alesina, Alberto, La Ferrara, Eliana, 2000. Participation in heterogeneous communities. Quarterly Journal of Economics 115 August.

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

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

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

Baland, Jean-Marie, Platteau, Jean-Philippe, 1996. Inequality and collective action in the commons. CRED. University of Namur.

Baland, Jean-Marie, Platteau, Jean-Philippe, 1998. Wealth inequality and efficiency in the commons. Part II: the regulated case. Oxford Economic Papers 50, 1–22.

Bardhan, P., Ghatak, M., Karaivanov, A., 2005. Inequality and collective action (2006) In: Bardhan, P., Bowles, S., Baland, J.M. (Eds.), 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.

Betz, Mathew, McGowan, Patrick, Wigand, Rolf, 1984. Appropriate Technology: Choice and Development. Duke University Press, Durham.

Bowles, S., Gintis, H., 2002. Social capital and community governance. Economic Journal 112 (8).

Clarke, E.H., 1971. Multi-part pricing of public goods. Public Choice 11.

Coleman, James, 1988. Social capital in the creation of human capital. American Journal of Sociology 94, S95–S120 Supplement.

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

Dayton-Johnson, Jeff, 2000. The determinants of collective action on the commons: a model with evidence from Mexico. Journal of Development Economics 62.

Durlauf, S. and Fafchamps, M. (2005). "Social Capital", Handbook of Economic Growth, Chapter 26, pp. 1639–1699.

Easterly, William, Levine, Ross, 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.

Fritzen, Scott, 2007. Can the design of community-driven development reduce the risk of elite capture? Evidence from Indonesia. World Development 35 (8).

Glaeser, E., Laibson, D., Sacerdote, B., 2002. The economic approach to social capital. Economic Journal 112 (8), F437–F458.

Government of Pakistan, 1984. 1981 District Census Report of Baltistan. Government of Pakistan, Islamabad.

Government of Pakistan, 1998. Population and housing census of the northern areas 1998. Census Bulletin-12. Government of Pakistan, Islamabad.

Groves, T., 1973. Incentives in teams. Econometrica 41, 617–631.

Hart, Oliver, Holmstrom, Bengt, 1987. The theory of contracts. In: Bewley, T. (Ed.), Advances in Economic Theory, Fifth World Congress. Cambridge University Press, Cambridge.

Hausman, J., Taylor, W.E., 1981. Panel data and unobservable individual effects. Econometrica 49 (6), 1377–1398 (Nov., 1981).

Hirschman, Albert, 1967. Development Projects Observed. Brookings Institution, Washington.

Isham, Jonathan, Narayan, Deepa, Pritchett, Lant, 1996. Does participation improve performance? Establishing causality with subjective data. World Bank Economic Review 9 (2).

Jones, Ben, Olken, Ben, 2005. Do leaders matter? National leadership and growth since World War II. Quarterly Journal of Economics 120 (3), 835–864.

Khwaja, Asim Ijaz, 1999. Revisiting rural infrastructural investments. Aga Khan Rural Support Programme.

Khwaja, Asim Ijaz, 2004. Is increasing community participation always a good thing. Journal of the European Economic Association 2 (2–3), 427–436 Apr–May.

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

Knack, S., Keefer, P., 1997. Does social capital have an economic payoff? A cross-country investigation. Quarterly Journal of Economics 112 (4), 1251–1288.

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

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R.W., 1997. Trust in large organizations. American Economic Review Papers and Proceedings 87 (2), 333–338.

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.

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. Cambridge University Press, New York.

Parvez, Safdar, Jan, Ehsan-ul-Haq, 1998. Growth, poverty and inequality in the programme area. Working Paper, Aga Khan Rural Support Program.

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

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. Simon and Schuster, New York.

Stewart, Frances, 1978. Technology and Underdevelopment. Macmillan, London.

Tendler, Judith, 1997. Good government in the tropics. Johns Hopkins Studies in Development. Johns Hopkins University Press, London.

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

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

Warr, P.G., 1983. The private provision of a public good is independent of the distribution of income. Economic Letters 13, 207–211.

Whiteman, P.T.S., 1985. Mountain Oases: A Technical Report of Agricultural Studies (1982–1984) in Gilgit District, Northern Areas, Pakistan.

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. The World Bank, Washington, D.C.