SPNA Endline Report (final draft)

A four-wave impact assessment of a stabilisation programme in North-East Afghanistan

January 2019

ARC GbR. Koehler & Gosztonyi,
Berlin, Germany

ARC
Analysis Research Consulting

Table of Contents

LIST OF ACRONYMS

AHRRAO	Afghan Human Rights Research and Advocacy Organisation
AIHRC	Afghan Independent Human Rights Commission
ALP	Afghan Local Police
ANA	Afghan National Army
ANCOP	Afghan National Civil Order Police
ANP	Afghan National Police
ANSF	Afghan National Security Forces
Arbab	village headman (in this context)
Arbakees	local informal anti-Taliban militias
CDC	Community Development Council
CDD	Community-Driven Development
CLDC	Cluster-Level Development Council
DDA	District Development Assembly
DFG	Deutsche Forschungsgemeinschaft (German Research Foundation)
DRRD	Department of Rural Rehabilitation and Development (the sub-national level unit of the MRRD)
ICRC	International Committee of the Red Cross
IDLG	Independent Department of Local Governance
ISAF	International Security Assistance Force
KfW	Kreditanstalt für Wiederaufbau (development bank)
MoIA	Ministry of Interior Affairs
MRRD	Ministry of Rural Rehabilitation and Development
NABDP	National Area Based Development Programme
NDS	National Directorate of Security
NSP	National Solidarity Programme
NUG	National Unity Government
OSDR	Afghan Organisation for Sustainable Development and Research
PGM	Pro-Government Militia
RC North	Regional Command North
SFB 700	Sonderforschungsbereich 700 (Collaborative Research Center 700)
UNAMA	United Nations Assistance Mission in Afghanistan
US	United States

LIST OF FIGURES

LIST OF MAPS

LIST OF TABLES

EXECUTIVE SUMMARY

BACKGROUND

The purpose of this report is to present the results of the endline survey and to analyse it comprehensively against the baseline data collected in 2010 and 2011 and the follow-up data gathered in 2012 and 2014/15. The report is the fifth part of a longitudinal stabilisation impact assessment of the SPNA programme, covering all original 25 districts in four provinces of the programme as well as the ten new districts added in the current wave.

In order to assess stabilisation effects of SPNA we do the following: (1) We look at what SPNA intends to achieve in terms of stabilisation with the measures it takes. (2) We provide a qualitative assessment of how SPNA works and takes effect via its principle local partner, the District Development Assemblies (DDAs). (3) From there we look at different fields and dimensions of stability and define measurable indicators for stabilisation. (4) We define measurable indicators for most relevant and intended SPNA outcomes, and (5) use statistical methods to trace likely net effects of SPNA measures on different dimensions of stability in the target region. (6) Based on the findings we build recommendations for SPNA regarding where to work for impact under what framework conditions.

A NOTE ON SPNA

SPNA RATIONALE

AKF (Afg)'s proposal to the KfW, defines SPNA's developmental goal as follows:

> "Through the implementation of community prioritized essential development infrastructure, the programme aims to provide an enabling environment for stability by improving access to services and sub-national good governance."

In order to achieve this goal, the programme set itself three main objectives, which remained unchanged throughout the lifetime of the project:

- To **provide basic development infrastructure**, enhancing access to services and livelihood opportunities;
- To **enhance capacity among local governance institutions** for good governance and development planning;
- To **strengthen collaboration between and among institutions** within the sub-national governance system (AKF-A 2010, p. 6; AKF-A 2016, p. 3).

THE DISTRICT DEVELOPMENT ASSEMBLY (DDAS)

SPNA's principle local partner is the District Development Assembly (DDAs), a district-level body composed of representatives of Community Development Councils (CDCs), established from 2003 onwards by the World Bank funded National Solidarity Programme (NSP) and coming under the authority of the Afghan Ministry of Rural Rehabilitation and Development (MRRD). CDCs and their apex body, the DDAs fall under the general developmental heading of community-driven development (CDD) organisations. In the following we will refer to this multi-layered structure of development councils as the "shura structure". On all levels, the component parts of the shura structure (i.e. CDCs,

DDAs) have both developmental as well as local governance tasks. DDAs were established by the National Area Based Development Programme (NABDP) – also funded by the World Bank and falling under the responsibility of the MRRD.

SPNA ACTIVITIES

SPNA was implemented in five phases from 2010-18 by three implementing partners AKF (Afg), the lead organisation, ACTED and Mercy Corp. SPNA had a total budget of € 105 million over the eight years of its existence. In the course of its existence, SPNA concluded a total of 424 main and accompanying infrastructure measures (completed and in warranty; see SPNA Monthly Inventory November 2018). Regarding the sectors covered by the SPNA-funded infrastructure projects, the largest share of the budget was spent on education (65%), followed by energy and water infrastructure (12%), roads and transportation infrastructure (11%), health (9%) and governance (2%). These figures show SPNA to be a **very significant source of development funding in the four North-Eastern target provinces**. In particular towards the end of its existence, i.e. 2016-18, it was one of the only programmes providing funding for larger district level development projects.

In addition, SPNA developed 12 basic and advanced training modules that were delivered to DDAs (see AKF (Afg) Activities for SPNA-KfW, 2017).

Overview of SPNA training modules for DDAs

DDA Level trainings	Advanced Advocacy
	Advanced Gender Mainstreaming and Planning
	Advocacy
	Disaster Risk Reduction (DRR)
	Establishment and Support of Operation and Maintenance Committees
	Participatory Monitoring & Evaluation
	Project Cycle Management
	Advanced Conflict Resolution and Peace Building
	Conflict Resolution on Common Property Resources
	District Planning
	Gender Planning
	Visioning Building

Table 1: SPNA training modules for DDAs

ARC'S ASSESSMENT APPROACH

Before presenting the actual assessment results, we briefly introduce our assessment approach.

DEFINING STABILITY

Taking into consideration the concept of civic stability outlined in SPNA project proposals (AKF (Afg) 2010) and theoretical thought on dynamic stability of social order, we arrived at a working definition of stability comprising four fields:

PHYSICAL SECURITY: stability is defined by low levels of socially unacceptable violence (some forms of violence may be socially accepted and are therefore not detrimental to stability).

GOVERNANCE: stability is defined by functioning governance institutions. The more complex society and its segments get, the more important the reliable and legitimate regulation of collective tasks, issues and conflicts becomes.

ECONOMIC REPRODUCTION AND DEVELOPMENT: Stability of society and its segments is also defined by the ability of those segments to materially sustain themselves; economic reproduction is therefore the third defining aspect of civil stability.

MODERNISATION (ADAPTIVE CHANGE)**:** The three components outlined above require a vital fourth component of stability – the ability to adapt to changing environments via innovation and development. Hence, capacity for adaptive change is an intrinsic aspect of social stability. Under current conditions of internationally driven state-building, the challenge of adaptation predominantly relates to this normative concept of modernisation.

IMPACT MODEL

As a next step we develop an *impact model* depicting how programme measures, infrastructure development and DDA capacity-building (outputs) intend to contribute to stabilisation, defined by our working definition as consisting of four fields (security, governance, economy & development, and modernisation).

Programme measures and (plausible) resultant stabilisation do not take place in a vacuum, but are also powerfully influenced by (a) the overall context in which the measures are carried out, and (b) by the way programme measures are implemented.

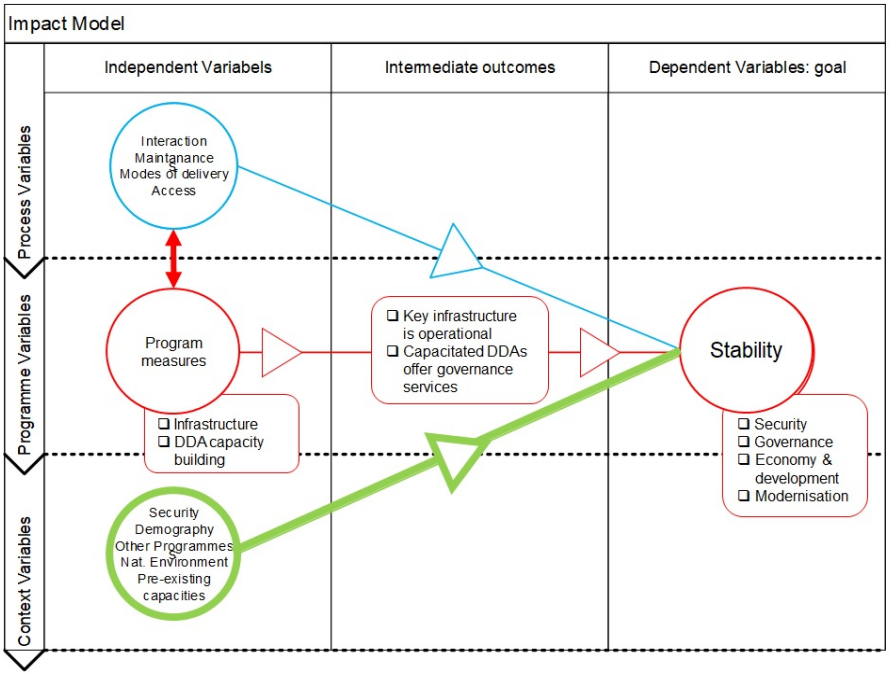


Figure 1: SPNA impact model

ASSESSMENT METHOD

In our assessment methodology and analytical strategy, we build the case for SPNA impact on stability based on three legs:

- Through an **in-depth qualitative research** that analyses how the two components of SPNA – the capacitating of DDAs and providing of infrastructure – unfolded and how they relate to stabilisation over the implementation period. The qualitative analysis flanks the quantitative survey and qualitatively traces causality from SPNA inputs, over DDA activities to DDA interactions with a district's population, leading to changed perceptions that are measured in our (quantitative) perception survey.

- We construct treatment and comparison groups through differences **in the awareness and assessment of the DDA and of SPNA infrastructure** by respondents within one district. We take awareness and assessment as proxy indicators for the immediate outcomes SPNA intends to achieve – more visible and better DDAs and visible and useful infrastructure. The assumption is that higher values for these immediate SPNA outcomes predict higher values across our survey-based stabilisation indicators. Within the same district we expect people that are more exposed to SPNA and have better opinions about SPNA outcomes to also have better views about stability than people less exposed to and with lower opinions about SPNA. We test SPNA's impact hypothesis with regard to stabilisation via **cross-sectional and panel time series analysis**.

- Through **robustness checks on our main statistical models**: firstly, by drawing on specific features of SPNA rollout, and resulting exposure to SPNA (**instrumental variable**). Districts differed substantially in terms of the timing and duration of exposure to SPNA activities due to the step-wise roll-out and termination of the program. We made use of this quasi-random SPNA-exposure to induce exogenous variation into our main explanatory variables (the awareness and assessment of the DDA and of SPNA infrastructure) and thereby improve the identification of causal effects on stabilization outcomes. Secondly, we replaced the survey-based main explanatory variables of DDA visibility and quality with an **expert-coded performance index to trace alternative (perception-independent) channels of SPNA effects**.

FIELDWORK

Concerning the actual implementation of the four survey waves (see table below), together with the Mazar-e Sharif-based organisation, the Afghan Human Rights Research and Advocacy Organisation (AHRRAO), we conducted the baseline in 2010/11 (Wave 1), the first follow-up survey in 2012 (Wave 2), the second follow-up in 2014/15 (Wave 3) and the endline in 2017/18 (Wave 4). In each of the first three survey waves we revisited the same districts, village clusters and villages surveying between 231 and 251 villages and more than 5,000 respondents per wave. In the final wave we revisited the same 25 districts that formed the core of our survey since the baseline and added – upon the request of AKF (Afg) – an additional ten districts. The total endline survey thus comprised 35 districts with 352 communities and 8,458 respondents. The total number of head of household interviews for the four waves thus adds to 23,959.

Overview and key features of the four survey waves

Wave	Survey period	# districts	# communities	# respondents
1	Sep-Dec 10 / May-Aug 11	25	231	4,928
2	Sep-Dec 12	25	251	5,203
3	Nov 14-May 15	25	245	5,370
4	July 17-Apr 18	35	352	8,458

Table 2: Key features of survey waves

RESULTS

SUMMARY OF STABILISATION CONTEXT TRENDS

Wave 1: SPNA was launched in 2010 in the context of the US-led military and civilian surge announced by President Barrack Obama in December 2009. We carried out the *baseline for Cohort I* (the first 15 districts of SPNA) in autumn 2010 amidst intensifying fighting in the survey region, as US and other international troops were deployed to the North and increasingly began to challenge the insurgents in their strongholds. The *baseline for Cohort II* (the second batch of 10 SPNA districts), implemented in spring 2011, already encountered improved security and a generally changed landscape, as the military surge pushed the insurgents back and dismantled their governance structures.

At this time, the DDAs in the survey region had all been set up between 2006 and 2009, implemented their NABDP and were increasingly being taken seriously by development actors (IOs and INGOs) who began to consult them when implementing projects in a district. Aid inflow at this time was very high with our survey villages reporting an average of 1.8 projects having been implemented "in the past two years".

Wave 2: ARC's *first follow-up* stability assessment, implemented in autumn 2012 and early 2013, experienced significant changes on the ground. The security gains achieved by the military surge still largely held, but international military forces progressively took the back seat, allowing the increasingly well-trained and capable ANSF to take the lead.

Turning to the governance context, this wave coincided with a reorganisation of the DDAs of the survey region. Shortly before and during the Wave 2 survey, 18 of the original 25 survey DDAs were reorganised and held renewed DDA elections. A further four DDAs were reorganised and held new elections shortly after the Wave 2 survey, while three districts kept their old structure and old composition without new elections.

Wave 3: We carried out the *second follow-up* in late 2014 and early 2015 in the context of a dramatically transformed political and security context in the whole of Afghanistan and also in the survey region. *One crucial game changer* was the election of President Ashraf Ghani as Afghanistan's new president in 2014. The second game changer was the withdrawal of most foreign forces and the concurrent handover of security responsibility to the ANSF by the end of 2014. The *third key development* shaping the context of this was the escalation of the insurgency and increasing territorial gains by insurgents in the North.

Wave 4: We implemented our expanded *endline* survey between August 2017 and April 2018. The problematic trends noted in the previous wave continued. Taliban and commander-controlled areas expanded, while areas formally governed by the government decreased.

Developments in the field of sub-national governance were also detrimental to DDAs and to SPNA's legacy. In spite of the technocratic approach of the Ghani administration to sub-national governance reform, progress in this field was slow, leaving the DDA in a limbo of uncertainty. Options repeatedly discussed in recent policy papers included elevating DDAs to the status of district councils, integrating them as development committees into – yet to be elected – district councils or eliminating them altogether following future district council elections. A long-awaited sub-national governance policy adopted in April 2017 *de facto* abolished DDAs as official governance bodies and mandated them to be replaced by District Councils – to be elected in the October 2018 elections. Unexpectedly, district council elections were scrapped at the last minute leaving a governance vacuum at the district level.

Immediate outcomes (qualitative results)

As was described, SPNA worked with DDAs to deliver key infrastructure and to improve their capacity to prioritise as well as support the implementation of development projects and to deliver a variety of governance services. As outlined in the impact model above, installed infrastructure funded by block grants and training measures for DDAs are the **output** of SPNA while (a) operational infrastructure and (b) capacitated DDAs are the **intended immediate outcome** of SPNA. These immediate outcomes are then expected to lead to improved stabilisation in the four fields (security, governance, development / economy and modernisation) which in the logic of our model corresponds to the programme's **overarching developmental goal** or **impact**. Relying on our ***qualitative research*** we ask, what were the ***immediate outcomes*** of SPNA measures (outputs).

Immediate Outcome 1: operational key infrastructure

Regarding the first outcome, the operationality of SPNA-built key infrastructure, we asked heads of Community Development Councils (CDCs) and of Cluster Level Development Councils (CLDCs) in the survey districts (the latter were by definition also members of the DDA) to assess all SPNA projects in their districts. These elected councillors have usually participated in the implementation of SPNA projects in their areas and are legitimate representatives of their communities.

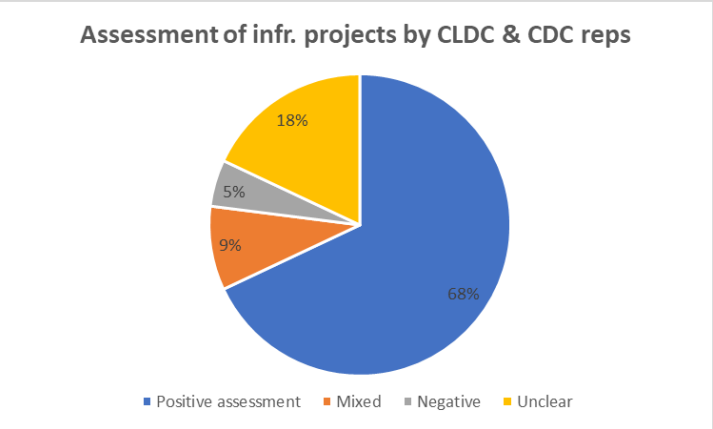


Figure 2: Assessment of SPNA infrastructure projects by CDC and CLDC / DDA representatives

We thus understand their assessments to be expert opinions conveying the position of their respective communities. Responses were then coded as positive, mixed, negative or unclear, if respondents had no concrete knowledge of a project or if the responses were not clear.

The majority of projects (68% or 194 out of a total of 205 measures in the survey districts) were evaluated **positively**. This means that – in the opinion of community representatives – the project had been implemented well and was now fully operational. We coded a much smaller number of projects as **mixed** (9% or 19 out of 205 projects in the survey districts). For example, all cluster and CDC heads interviewed were aware of the Gawhar Bridge in Yamgan. All but two interviewees confirmed that it was in use and operational, but that it was too weak and only small cars could cross it. In another case

(Nawbahar Girls' School in Pul-e Hisar District), the quality of the construction was assessed positively, but respondents criticised that too few female teachers were assigned to the school by the Ministry of Education. The proportion of fully non-operational projects (**negative** assessments) is even smaller (5% or 11 out of 205 projects in the survey districts). In this group we find cancelled or suspended projects or projects which, mostly due to the bad security situation, were out of use, such as the Qala-i Zal-Aqtapa Health Clinic, which reportedly remained incomplete and was used as an army base in this hotly contested district.

In summary, regarding ***Immediate Outcome 1*** (operational infrastructure), based on the projects that our interview partners were aware of, SPNA achieved a fairly high degree of success, as judged by our expert interview partners, the CDC and CLDC/DDA representatives. As a result, the immediate outcome, or in other words, the intermediate step in the impact chain leading from infrastructure outputs to stabilisation impacts, was achieved. It is thus justified to assume that SPNA infrastructure measures could have an impact on perception-based stabilisation indicators higher up in the impact chain.

Immediate Outcome 2: capacitated DDAs

In tracing the effects of SPNA capacity-building measures, we proceed in two steps: firstly we look at the trainings that have been conducted and, secondly, we discuss how DDAs, presumably as a result of trainings, began to perform their tasks.

Trainings: SPNA offered a number of training modules to DDAs. Some of these training measures were (a) directly linked to the developmental tasks of DDAs, such as the modules relating to project prioritisation, project cycle management and participatory monitoring and evaluation, while (b) other training modules capacitated DDAs to perform more general governance functions in their districts (e.g. advocacy, gender planning, disaster risk management and conflict resolution). While not the only training provider to DDAs, SPNA was without doubt the most important, conducting, according to our rough estimate, between half and two thirds of trainings in each wave (a precise calculation is not possible as DDAs we interviewed could not reliably recall which training was provided by which organisation).

Overall, the trainings were well received. Comments, such as the following, were common:

> "At first, none of the members of the DDA knew anything about the work they had to carry out. After going through the trainings [already] mentioned, the members of the DDA now have good administrative abilities, can prioritise projects, implement them and monitor [their implementation]" (interview with Kalafgan DDA head on 15 Dec 2012).

We only noted five critical statements out of a total of 73 DDA interviews (three critical comments during Wave 2 and two during Wave 3). Criticism tended to focus on the value content of certain trainings, e.g. gender, or on the fact that some DDA members were too uneducated to properly comprehend training contents. Based on our assessment that SPNA was by far the main training provider to DDAs in the survey region and on the positive assessment of the trainings by DDAs we interviewed, we conclude that SPNA is likely to have meaningfully contributed to any observed improvements in DDA performance.

DDA performance: In order to systematically analyse DDA activities, we coded the performance of these assemblies across a total of 11 activities associated with the tasking of DDAs. The main source for the coding were our qualitative interviews conducted on the district, cluster and CDC levels. Many

of the activities considered in the scoring also formed part of the SPNA training modules for DDAs (e.g. development related activities, gender, conflict resolution, maintenance, etc.).

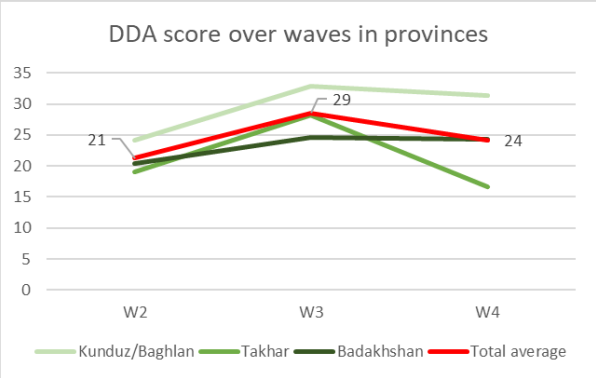


Figure 3: DDA performance scores for three waves and for provinces; the red line depicts the average for the 25 core districts

Looking at the resultant scoring, we find that DDAs, indeed, started to work in line with their mandate and along the lines taught by the training modules. As an example, DDAs took on an ever more active role in development planning, implementation and conflict resolution, and partly even started to internalise and act upon the contents of gender and women's rights related trainings. As such, **DDAs were active in the field of development, maintenance, conflict resolution, security and via their usually good contacts with government institutions intervened on behalf of their constituencies**.

We also find that DDAs varied very strongly in terms of the actual quality of their work and the degree to which they actually fulfilled their mandated tasks. This is visible in the differences in total scores across the different DDAs. Furthermore, we find characteristic differences across the three waves captured by this analysis. Wave 2 starts with relatively high scores for DDAs. Scores further improve somewhat for Wave 3, only to drop sharply for Wave 4. The drop in Wave 4 appears to be partly due the disablement of some DDAs under full Taliban occupation. The more important reason, however, seems to be the phasing out SPNA. DDAs stopped implementing SPNA-related infrastructure and frequently became inactive during Wave 4. This observation provides further **circumstantial evidence that SPNA was indeed instrumental in motivating and capacitating DDAs to perform developmental and governance-related tasks in their districts**.

For the further impact assessment, we thus conclude that DDAs became active and performed their activities to a large extent as a result of SPNA's interventions. **Based on the DDA activities we observed, we are justified in expecting results across the four fields of our working definition of stability: security, governance, economy/development and modernisation.**

SPNA EFFECTS (INFERENTIAL STATISTICAL ANALYSIS)

Using more or less exposure for inference about SPNA effects

Lacking any outright control group of districts, i.e. survey districts that did not participate in SPNA and did not benefit from similar measures funded by other donors, we decided to **gauge the effects of SPNA measures through comparing higher or lower degrees of exposure to different SPNA measures.** The assumption was that persons or communities with more exposure to SPNA, would display SPNA-associated outcomes to a higher degree than communities with less exposure to SPNA. This would allow us to make inferences as to the potential impacts of hard and soft SPNA measures.

With the aim of achieving a higher degree of conceptual robustness, we designed three distinct approaches to model more or alternatively less exposure to SPNA measures. **All three, conceptually rather different modelling procedures, delivered very similar results, making us very confident of having identified genuine SPNA effects.**

Overview of SPNA effects

Table 1 summarises the main results for SPNA's immediate outcomes across the four fields of stabilisation at the time of the endline. We use a simple colour coding indicating statistically significant correlations at the 95% confidence interval or higher. The colour coding shows the direction of the significant relationships – green for positive, red for negative stabilisation effects. We refer to the results of the additional models of the pooled cross-sectional analysis covering all four waves (which is not reflected in the table) when they differ from the results at the endline.

	DDA visibility	DDA assessment	DDA contribution per development sector	SPNA infrastr. visibility	SPNA infrastr. assessment (beneficial for district)	SPNA infrastr. assessment (beneficial for *own* hh)
Security perceptions						
Confidence security provision by the state						
District admin. cares						
Social cohesion in communities						
Government official most powerf'l in district						
Corrupt conflict resol.						
General development assessment						
State contribution to development						
Self-assessed material well-being						
Access to agricult. Land						
Int. dev. as value threat						
Rejection of school curriculum						

Table 3: Main effects for Wave 4 with all controls (35 districts) and with all controls plus baseline control (25 districts)

The results of the three different analytical perspectives can be summarised as follows. **Firstly**, **the two-pronged stabilisation approach implemented by SPNA -** capacitating DDAs and providing infrastructure through a transparent participatory procedure - **worked**. With high likelihood the immediate outcomes of SPNA have had mostly ***positive net effects*** (22 statistically significant relationships out of a total 34) **in** the stabilisation **fields of security, governance, development** and **openness to modernisation**. ***Negative effects*** (12 out of 34) **cluster around** the normative (perception-based) ranking of **who is most powerful** in the district and, **local social cohesion,** proxied by how mobilised the community is to do collective unpaid work (hashar obligations). In the fields of security and modernisation the results are more mixed – confidence in the state as future security provider turned negative in Wave 4 and there were mixed results regarding the reduction of perceptions that international development work is a threat to local values and Islamic norms, as well as regarding the rejection of the public school curriculum.

Secondly, most of the **positive effects increased and became more consistent over time**. In the first follow-up in 2012 we observed no consistent effects across all districts. By 2015 effect patterns had

emerged that consolidated by the time of the endline. Considering effects that are consistent across all three perspectives (Wave 4, all waves and the simplified models), the **capacity-building component** for DDAs seems to **produce most lasting (longer-term) positive results**. The effects of infrastructure projects seem to fade more quickly with time passing (weaker in Wave 4 when most projects had already been completed awhile before). However, positive effects on confidence in the state's capacity to provide security in the future and on value-threat perceptions, manifest in 2015, had vanished by 2017. The same is true for the initially negative effects on local social cohesion and a separate, though fairly strong, negative correlation between the visibility of SPNA infrastructure and perceptions of corruption in dealing with conflicts.

Thirdly, and very importantly, **we record these mostly positive results and trends for SPNA stabilisation effects against the background of overwhelmingly negative general trends in all fields after 2012**. In other words, despite a general increase of instability, a further fragmentation of local governance, an institutional retreat of the state from the hinterland and even some administrative centres, and an increase in the ideological challenge to the international engagement, the net effects of SPNA remain positive and have even consolidated over time. This is remarkable. However, it is also obvious that even consolidated positive net effects of SPNA are far too weak to turn the tide of general disintegration in North Afghanistan. What can be claimed with some confidence is that **without SPNA the situation would have deteriorated even further**; to what extent exactly, we are, however, unable to judge.

Fourthly, the observable effects of **SPNA outcomes** at the time of the endline are **mostly positive, despite years of institutional limbo regarding** the main local counterpart, **the DDAs**. This has been a different – and arguably even more consequential – challenge to SPNA effectiveness than the deteriorating framework conditions on the ground; it has also been more avoidable. To put it bluntly, despite the explicit commitments made in the Tokyo declaration of 2012, the Afghan government has been unwilling or unable to develop, let alone implement, a coherent local governance policy in accordance with the provisions of the constitution. As a consequence, governance below the provincial level has been stuck in limbo. Non-survey data clearly shows that this uncertainty very much affected and weakened the DDAs. They remain dependent in their core functions (development) on SPNA or similar external inputs (Takhar, Badakhshan, few exceptions in well-connected districts).

Fifthly, we find **indications of the functional adaptation of the DDAs**, possibly explaining the positive correlations between and across SPNA outcomes and more informal and traditional power structures of districts. While we find consistent evidence that the SPNA approach fosters the output legitimacy of the state (i.e. the perception that the state administration cares) we do not find similar evidence for SPNA strengthening perceptions that the state is effective and powerful in the districts. The DDA, equipped with a procedure to deliver infrastructure, works **as a transmission belt between the people it represents and any kind of local authority**; i.e. not only **state authority but also informal power holders**.

Sixthly, while the capacity-building effects of SPNA on the DDAs seem to be more lasting than those of the infrastructure component, **DDAs** nonetheless **need the incentive of tapping development funds to remain active.** DDAs of districts in which SPNA concluded before Phase V, were more often inactive, or were even often fully defunct. In contrast, DDAs with ongoing projects – even if only O&M funds – met more regularly, were more active and, while regularly meeting to discuss the implementation of ongoing projects, also fulfilled governance functions not directly linked to the project itself.

SUMMARY AND RECOMMENDATIONS

SUMMARY

So far we have demonstrated that, in line with the programme's theory of change, SPNA achieved its two main intended outcomes to a fairly high degree: delivering operational key infrastructure and capacitated DDAs that engaged until at least Wave 3 more effectively in development planning and implementation, and in the provision of governance services to the populations of their districts. Based on the finding that the intended immediate outcomes were achieved and on the comparison between respondents and communities with greater or lesser degrees of treatment by SPNA, we are also rather confident that we have identified a number of SPNA-associated impacts. These impacts we identified are mostly positive and relate to all four fields of stabilisation (thus registering a fairly comprehensive effect). While these positive stabilisation effects could not turn the tide in the North-East, it nonetheless mitigated the overall negative stabilisation trends in the region.

RECOMMENDATIONS

Stabilisation effects via capacity building of DDAs and the provision of infrastructure projects through them are real and – with a time lag of two to four years – consistent. Based on the findings, their implications for SPNA and the lessons learned that we presented above, we make the following general recommendations regarding achieving, increasing and sustaining the impact of programmes like SPNA or building up SPNA.

Impact requires time

For SPNA it took four years for consistent stabilisation effects to materialise. This has a number of important implications for similar future programmes.

Manage and focus expectations. The political agency behind such stabilisation programmes needs to understand that this is not a "quick impact approach" but a sophisticated approach with a focus on mid-term stabilising effects. During the first two years effects may be messy, local and contradictory.

This brings us to the next point: **be flexible and adaptive** in the first phases. During the first phase we saw that different districts reacted very differently to the SPNA measures they were exposed to. It thus makes sense to tailor the programme more to the specific needs and challenges of individual districts. A competent pre-assessment makes sense. This should be taken into account in districts with low levels of trust and high levels of corruption particularly in the development sphere (in the first follow-up we identified a number of districts in which SPNA infrastructure related negatively to corruption perceptions). Districts with high levels of grievances from pressing demands for *ushr* and infighting militias are a completely different environment. The degree of physical presence of the state varies greatly. All this would need to be taken into account. A one-fits-all approach for this initial phase may not be the best approach. This flexibility should be highlighted already in the initial offer to the donor. Initial flexibility, tailoring programme measures to different local contexts, and factoring in a time lag regarding consistent impact, are operational challenges many development programmes face (and are often unable to address); for stabilisation programmes focussing on district-level governance and infrastructure development under conditions of violently contested state formation, it seems to be a more pressing priority.

The initial disparity of SPNA effects at the time of the first follow-up was the reason that we introduced the District Effect Clusters in the second follow up. The idea was to systematically assess whether differing levels of violence, state penetration and connectedness of clusters of districts explain in a more general way how SPNA produces different effects under different conditions. We did find that **there are** indeed **differences** in how certain SPNA measures work in certain subsets. We also showed that **all three** subsets are relevant as categories for understanding differences in impact that could inform the choices and priorities a stabilisation programme can take. Nevertheless, we also found more consistent effects across all districts from the second follow-up onwards. Hence, a further investigation into what clusters of districts are more receptive towards different stabilisation instruments and strategies would make sense when considering any continuation of the approach after the first dust has settled. Our recommendation would be to further investigate the implications of such effect clusters – in close cooperation with AKF staff and programme designers – in order to be as practical and concrete as possible.

For the initial implementation period of two to three years, however, a generalisation of likely effects under only a few different framework conditions is not promising – here a district-by-district analysis of specific needs, risks and potentials may be more constructive.

A third recommendation is to **clearly communicate the rules of engagement** for the programme to the principle local partner. If those conditions are not met during the first phase, it makes sense to withdraw from that district or to change to a mode of limited engagement. What should not happen is what happened with Warsaj. Here SPNA was implemented without adjustment from 2010 to 2013 despite the fact that neither the DDA nor the shura structure was operational at all. Warsaj may very well be the district with worst examples of (non-violent) dysfunctional Kabul-level patronage. Positions in the government administration in this district have been and still are held by a high-ranking bureaucrat with Jihadi background who originates from the district. He has an astonishing track record of appointing and retaining completely incompetent people in positions of consequence. The head of DDA is a case in point. He showed no interest in his job and the DDA under his leadership was almost fully dysfunctional. Nonetheless, for a timespan of three years it received all trainings and participated in project identification.

Lobby government and its international backers to push for institutionalisation of district-level community representation

As of the finalisation of this report, there is a massive gap in the sub-national governance system on the district level and below. The new sub-national governance policy confirms CDCs as the main village-level governance bodies. However, and in contrast to the 2010 sub-national government policy, by not mentioning DDAs, it basically relegates this – as our analysis shows – crucially important governance body to informality and thus condemns it to a slow death. In its place it foresees the establishment of elected district councils, which should have been elected in October 2018.

However, with the cancellation of the district council elections in September 2018 and the de facto dissolution of DDAs, there is now a huge gap in the sub-national government system in terms connecting and linking local communities to the district-level state. With their demise several key aspects of governance will suffer: the delivery of district-level development infrastructure, maintenance of already existing infrastructure, a structured and effective hierarchy of conflict resolution, an institutionalised channel for communities to communicate their issues to the state – and vice versa, for the state to contact its population.

Given the current grave situation, we strongly suggest (1) using the results of this robust assessment to advocate with the Afghan government to reinstate – if necessary by decree, and if wished only as a temporary mid-term measure – the DDA and the associated cluster structure to act as a de facto district council. (2) to lobby for the design of a citizen-charter type of grand programme for DDAs and clusters, and (3) in the meantime to continue supporting existing, now informal, DDAs from own and other donor funds.

The alternative, i.e. to connect villages directly to the provincial level of government, would risk ignoring one historic and one contemporary challenge to state-formation in Afghanistan. Historically, the political power in Afghanistan has been highly centralised only on paper. In reality, central state power tended to stop in the provincial administrative centres and hardly ever reached the communities in meaningful ways. Subnational governance around the NSP-shura structure had started to change this. The second challenge is that the state is more fragmented and in retreat than at any other time in the past decade. Clipping away the district-level representation of communities and hoping that villages and clusters will link to the state in the provincial centres is not a very likely scenario.

Be conscious about the limitation of controlling all aspects of impact

A programme like SPNA can stabilise not only official patterns of authority but also informal or even Taliban authority. When the state retreats and leaves territorial control, security and other governance functions to local commanders, the communities themselves or even the Taliban, the DDAs and shura structure will work with whoever is in charge. SPNA has no control over the way communities and their representatives are using the capacities it built in order to survive under changing framework conditions. This is not bad per se, but the agencies involved should be conscious of this risk.

Do not compromise the governance focus for quicker and technically better results on the infrastructure component

Do not bypass the principle partner organisation – the DDA – as has been done in some districts. As a stabilisation programme this link has been vital and cannot be replaced by working directly through the administration. Participatory approaches were shown to likely deliver a broad range of positive stabilisation affects.

CONCLUSION

SPNA invested € 105 million over a time span of eight years, with the goal of achieving meaningful stabilisation effects across a total of 52 districts of Northern Afghanistan. The programme set out to achieve this goal by supporting district-level institutions and providing high-impact infrastructure via those capacitated institutions.

Has this goal been achieved?

The task of this assessment was to facilitate the discussion of the above question by supplying evidence in a methodologically sound and transparent way. We did so by assessing what SPNA did over time in both qualitative and quantitative terms, how stability – divided into four distinct subfields

of stabilisation – changed, and how these two observations and measurements of change likely relate to each other.

When we look at what SPNA did, we see a programme that invested systematically in training its key partner at the district level, the DDA, and in providing infrastructure via procedures that for the most part involved the DDA as a key decision-making body.

For the capacity-building component of SPNA, our coded qualitative indicators (DDA functionality index in the 25 survey districts observed by us since the baseline) shows moderate improvements during the time of the most intensive interaction with SPNA (Wave 2 to Wave 3) relating to the preparation and implementation of block grants for infrastructure. The survey-based indicators for the intended immediate SPNA outcomes initially show little change in the visibility of the DDAs (Wave 2 and Wave 3 around 60%) but a strong increase at the time of the endline (Wave 4 with 79%). The two DDA-assessment indicators we used, based on both general and sector-specific survey questions, show a moderate decline between the waves. The development and sustainability of DDAs was clearly influenced by forces outside the control of SPNA, most importantly by the legal and organisational limbo the DDAs faced particularly towards the end of SPNA's lifespan, and their continuing dependence on development programmes' agreeing to work with them.

For the infrastructure component of SPNA, our coded qualitative indicator (operationality of infrastructure projects according to CDC and CLDC guideline interviews) indicate that most are fully (68%) or at least mostly operational (9%). The qualitative assessment also provides a strong indication that projects processed in a participatory manner, via the DDA, are more likely to be coded as fully operational, than those that were not selected via the DDA. This emphasises at the importance of greater ownership through participatory processes. The general assessments of the utility of the infrastructure projects for the district are almost invariably positive across waves (between 2.98 and 2.99 when 3 is the maximum value). Under the given framework conditions, the infrastructure component is a substantive achievement. It implemented infrastructure that is visible, non-divisive, appreciated and (mostly) works according to design.

When we look at how the four fields of stabilisation developed, we see a deterioration of the security context between 2009 and 2011, which, after brief improvement, worsened again between 2013 and 2015. The governance context reflects these changes – formal state rule has been under pressure since the baseline in 2010/11.

Turning to those stabilisation indicators that SPNA may influence (which are different from the general state-related and security forces-related context indicators above) we observe the following: Taken together, with the exception of the local economic situation (household and village) and value-threat perceptions, most of our stabilisation indicators either stall or worsen between baseline and endline assessment.

Does this mean SPNA failed to have a positive effect on stability?

The descriptive statistics only show how SPNA, and the stabilisation indicators it intended to positively influence, trended separately under the given context conditions. However, our statistical models show consistently, after controlling for context factors, that more SPNA exposure predicts positive net effects on most of the stabilisation indicators defined in the four fields of intervention. In other words, without SPNA, the decline in stability would have been worse. We also find that the defining aspects of the SPNA approach to delivering infrastructure – community participation via the shura structure – interacts positively with the stabilisation effects observed. Put differently, participation and the

functionality of the development shura structure boost the effects of SPNA on stabilisation. These are key components of AKF (Afg)'s philosophy of doing development and engaging with beneficiary populations.

Does the impact accounted for justify the investment or meet the expectations in terms of stabilisation results?

This is likely a political question, and the answer depends on the expectations different stakeholders have about what can and should be achieved in terms of development-induced stabilisation. We believe that there are a number of inhibiting factors with more general implications that decreased the potential impact of SPNA on stability – factors that were beyond the control of SPNA and not predictable at the start of SPNA.

INTRODUCTION

TASKING

AKF-A tasked ARC with **conducting an *endline* impact assessment for KfW's Stabilisation Programme for Northeast Afghanistan** (SPNA). The SPNA programme is funded by the German Federal Foreign Office via the KfW Development Bank and implemented by AKF-A (lead implementing organisation), ACTED and Mercy Corp. ARC carried out the endline survey in 35 north-east Afghan districts between July 2017 and April 2018. 25 of the districts surveyed formed part of the original panel surveyed in three previous waves (2010/11 baseline, 2012 first follow-up and 2014/15 second follow-up), while ten districts were added only for the endline wave.

The purpose of this report is to present the results of the endline survey and to analyse them with reference to the baseline data collected in 2010 and 2011 and the follow-up data gathered in 2012 and 2014/15. The report is the fifth part of a longitudinal stabilisation impact assessment of the SPNA programme, covering all the original 25 districts across four provinces of the programme as well as the ten new districts added in the current wave.[1]

In order to assess the stabilisation effects of SPNA we do the following: (1) We look at what the SPNA measures intend to achieve in terms of stabilisation. (2) We provide a qualitative assessment on how SPNA works and how SPNA measures are implemented via the principle local partner, the capacitated District Development Assemblies (DDAs). (3) From there we look at different fields and dimensions of stability and define measurable indicators for stabilisation. (4) We define measurable indicators for most relevant and intended SPNA outcomes, and (5) use statistical methods to trace likely net effects of SPNA measures on different dimensions of stability in the target region. (6) Based on the findings, we build recommendations for SPNA about where to work for better impact under what framework conditions.

STABILISATION PROGRAMME NORTH-EAST AFGHANISTAN (SPNA)

PROGRAMME RATIONALE

The Stability Programme North-East Afghanistan (SPNA) is funded by the German Federal Foreign Office via the KfW Development Bank. It was implemented in five phases from 2010-18 by three implementing partners AKF (Afg), the lead organisation, ACTED and Mercy Corp. SPNA had a total budget of € 105 million over the eight years of its existence. Starting point for the design of SPNA was a governance provision gap identified by AKF (Afg)'s original proposal from March 2010 "between the village level" and higher-level subnational and national institutions (district, provincial and national). While governance at the village level was relatively well serviced by the Community Development Councils (CDCs) established by the National Solidary Programme (NSP), the

> "middle development decision-making bodies, at the district level, are under-resourced and under-supported …, causing gaps in development planning in the sub-national governance system" (AKF-A 2010, p. 4).

[1] The signing of the contract between AKF (Afg) and ARC in May 2017 was preceded by lengthy coordination – facilitated by AKF (Afg)'s MERL Unit – between the key stakeholders of the assessment, in particular the KfW and AKF (Afg) about the goals, scope and expected outcomes of the endline. This coordination between the stakeholder contributed to the success of the endline.

This gap gave rise to frustration, a loss of legitimacy and anti-government sentiments. In order to address this gap, SPNA set out to work with District Development Assemblies (DDAs), a district level body composed of representatives of the already mentioned CDCs. SPNA explicitly builds on the successes of the village level NSP. It intends to replicate the "NSP-approach" on the district level by providing infrastructure development funds combined with trainings to build the capacity of a district-level representative structure. This structure would focus on development while also performing other important governance tasks. The original proposal summarises this approach as follows:

> "Using **infrastructure investment as an entry point for greater community engagement**, the programme incorporates concurrent capacity-building elements for DDAs to ensure **more capable and participatory institutions** as well as more effective infrastructure implementation. In this way, **the programme supports stability by increasing access to development services** through essential infrastructure, **while enabling sub-national governance structures to achieve more relevant planning and effective implementation** on behalf of their constituencies" (p. 3).

SPNA GOAL, OBJECTIVES AND INSTRUMENTS

In line with the programme rationale outlined above, SPNA's logframe defines the programme's goal as follows:

> "Through the implementation of community prioritized essential development infrastructure, the programme aims to provide an enabling environment for stability by improving access to services and sub-national good governance."

In order to achieve this goal, the programme set itself three main objectives, which remained unchanged throughout the lifetime of the project:

- To **provide basic development infrastructure**, enhancing access to services and livelihood opportunities;
- To **enhance capacity among local governance institutions** for good governance and development planning;
- To **strengthen collaboration between and among institutions** within the sub-national governance system (AKF-A 2010, p. 6; AKF-A 2016, p. 3).

SPNA used, in particular, two main instruments to achieve these objectives:

- **Block grants;** SPNA provided block grants worth € 310,000 (later € 300,000), which beneficiary DDAs could use for high-impact infrastructure projects in their districts. In the final stage of the programme, SPNA also offered small-scale infrastructure grants (up to € 30,000) and operations and maintenance (O&M) support. In the course of its existence, SPNA concluded a total of 424 main and accompanying infrastructure measures (completed and in warranty; see SPNA Monthly Inventory November 2018). Regarding the sectors covered by the SPNA-funded infrastructure projects, the largest share of the budget was spent on education (65%), followed by energy and water infrastructure (12%), roads and transportation infrastructure (11%), health (9%) and governance (2%). These figures show SPNA to be a **very significant source of development funding in the four North-Eastern target provinces**. In particular towards the end of its existence, i.e. 2016-

18, it was one of the only programmes providing funding for larger district level development projects.

- **Training and capacity-building measures;** a part of these training measures was (a) directly linked to the developmental tasks of DDAs, such as the modules relating to project prioritisation, project cycle management and participatory monitoring and evaluation, while (b) other training modules capacitated DDAs to perform more general governance functions in their districts (e.g. advocacy, gender planning, disaster risk management and conflict resolution).

SPNA Phases

SPNA was implemented in five phases (AKF-A 2016, p. 4). The different SPNA phases gradually expanded the geographic scope of the programme from initially 30 to 52 districts while also revisiting many of the original participant districts offering their DDAs new block grants and ever more advanced training modules. Moreover, the last phase of SPNA, Phase V, had a clear conceptual focus on final handover and sustainability: it offered small-scale infrastructure grants worth € 30,000 to continue "to engage and strengthen DDAs" and, more importantly, it introduced an operations and maintenance (O&M) component to selected DDAs. This component comprised the setting up of maintenance committees that included DDA and CDC representatives as well as relevant government officials, the training and capacity-building of O&M committee members and the provision of limited funds for O&M activities.

Overview of the five SPNA phases

Phase I	*Phase I (2010-2013)* of the programme involved two sub-phases with two cohorts each. Districts received block grants worth € 310,000 to select high-impact infrastructure projects. Projects were initially identified through consultation with sub-national governance actors, district and provincial governors, and existing Provincial Development Plans (PDPs) and District Development Plans (DDPs). Later, after receiving capacity-building, District Development Assemblies (DDAs) directly prioritised the infrastructure to be built in their districts. During this phase, 30 DDAs received block grants.
Phase II	*Phase II (2012-2014)* of the program expanded into new districts in Baghlan, using the same model of capacity-building for DDAs and complementary € 310,000 block grants for infrastructure construction. During this phase, 35 DDAs received block grants.
Phase |	*Phase III (2013-2015)* involved implementing advanced capacity-building modules for experienced DDAs, selected based on their Phase I performance in Badakhshan, Takhar and Kunduz, paired with additional € 310,000 block grants for new infrastructure construction.
Phase IV	*Phase IV (2014-2016)* focused on expanding target districts within the existing provinces, including delivering advanced capacity-building modules, as well as € 310,000 block grants for infrastructure construction. Furthermore, this phase began to integrate forward-looking components related to *Operations and Maintenance* and *Disaster Risk Reduction*. During this phase, 21 DDAs received block grants.

Phase Very

Phase V (2016-2018) offered 18 block grants of € 300,000 to beneficiary DDAs. In addition, it provided Operations and Maintenance training and support, including capacity-building of user committees and community representatives as well as small-scale infrastructure grants to the value of € 30,000 for essential small infrastructure that continues to engage and strengthen the DDAs, particularly in the districts under volatile security situation. Phase V also included an evaluation to measure SPNA's impact on stability and governance in the project region.

SPNA PHASES IN THE SURVEY DISTRICTS

Participation in SPNA was not uniform across districts. Some districts participated practically continuously in all or almost all SPNA phases (e.g. Jurm District in Badakhshan Province), while others participated in only one or two phases (e.g. Wakhan in Badakhshan). Figure 4 below presents an overview of our survey districts with regard to participation in SPNA phases. Dark grey signifies SPNA in districts that form part of the original 25 districts, for which we have a panel (Waves 1-4), while light grey depicts districts that were only surveyed in Wave 4. The hatched lines depict participation in the O&M component of Phase V.

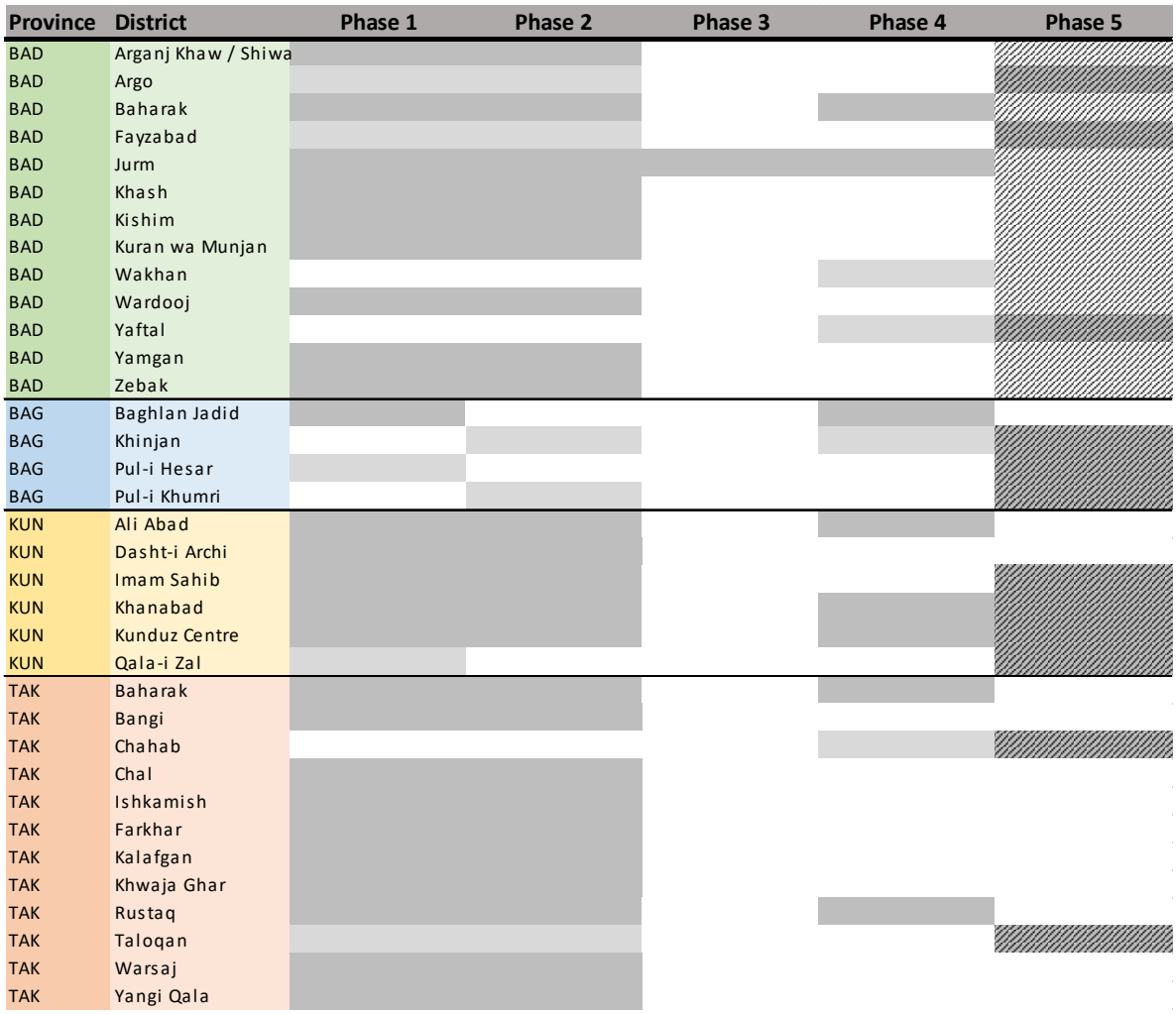


Figure 4: SPNA phases and survey districts

ASSESSMENT STRATEGY, SURVEY APPROACH AND FIELDWORK

So far, we have discussed SPNA's rationale, its objective of supporting stability and its rollout and implementation. In this section we turn to how we designed and implemented a survey to assess SPNA's impact on security. We begin by offering (a) an operationalizable working definition of stability[2], followed by (b) the impact model drawing on SPNA's theory of change. This model conceptually differentiates between independent variables (SPNA measures), dependent variables (stabilisation outcomes derived from our working definition of stability), and control or context variables. (c) Then we turn to the assessment approach and the mixed-method survey, designed to gather data on the aforementioned independent, dependent and control variables. (d) Lastly, we describe the fieldwork for the four survey waves.

DEFINING STABILITY

The key goal SPNA tries to achieve – and which the ARC impact assessment intends to measure – is stability. Stability is a complex concept and there is no generally accepted understanding of what it actually means. As a first step, we therefore developed a working definition of stability. Taking into consideration the concept of civic stability outlined in SPNA project proposals (AKF-A 2010) and theoretical thought on dynamic stability of social order, we arrived a working definition of stability comprising four fields:

PHYSICAL SECURITY: stability is defined by low levels of socially unacceptable violence (some forms of violence may be socially accepted and are therefore not detrimental to stability).

GOVERNANCE: stability is defined by functioning governance institutions. The more complex society and its segments get, the more important the effective and legitimate regulation of collective tasks, issues and conflicts becomes.

ECONOMIC REPRODUCTION AND DEVELOPMENT: Stability of society and its segments is also defined by the ability of those segments to materially sustain themselves; economic reproduction is therefore the third defining aspect of civil stability.

MODERNISATION (ADAPTIVE CHANGE)**:** The three components outlined above require a vital fourth component of stability – the ability to adapt to changing environments via innovation and development. Hence, capacity for adaptive change is an intrinsic aspect of social stability. Under current conditions of internationally driven state-building, the challenge of adaptation predominantly relates to this normative concept of modernisation.

THEORY OF CHANGE AND IMPACT MODEL

Following the presentation of a working concept of stability, we now discuss SPNA's theory of change, derived from AKF (Afg)'s proposal to the KfW and from discussions with AKF staff. SPNA assumes that by strengthening DDAs through financing infrastructure projects and providing capacity-building trainings to them (***outputs***), stability within the target districts can be increased. These activities are assumed to lead to two key ***immediate outcomes***: (1) increased access to key infrastructure and (2) improved DDA capacities regarding development planning and implementation as well as regarding the provision of additional services in the field of local governance (conflict resolution, advocacy,

[2] The concept of stability is not made explicit and is not defined in AKF-A's proposal to the KfW in 2010 or in other subsequent proposals.

disaster risk management, etc.). These outcomes should further lead to the ***aggregated impact*** (goal) of SPNA of comprehensively improving the status quo in the four fields of stability.

As a next step we develop an *impact model* depicting how programme measures, infrastructure development and DDA capacity-building (outputs) intend to contribute to stabilisation, defined by our working definition as consisting of four fields (security, governance, economy & development, and modernisation).

Independent and dependent variables

In our impact assessment *infrastructure outputs* and *capacity-building trainings* for the DDA are, therefore, the principle *independent variables* (also referred to as predictors and explaining variables in statistical models) the impact of which – on stability related dependent variables – we need to measure via quantifiable proxy indicators. We derive the *dependent variables* from the four fields of stabilisation in our working definition of stability: *physical security*, *governance institutions*, *economy and development,* as well as *adaptation and modernisation*. They are the second set of variables we need to account for. We can, however, only statistically establish the impact of SPNA programme measures on stability, once we also consider (in statistical lingo "control for") the likely impact of the *context* and *process variables*. These are thus the third and fourth set of variables.

Context variables

Programme measures and (plausible) resultant stabilisation do not take place in a vacuum but are also powerfully influenced by (a) the overall context in which the measures are carried out, and (b) by the way programme measures are implemented.

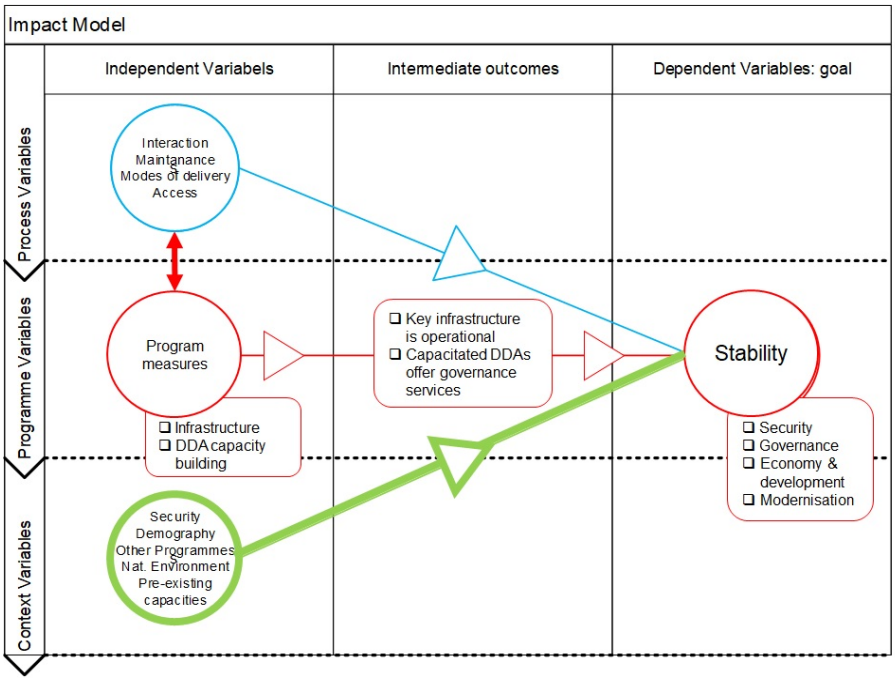


Figure 5: SPNA impact model

The former are usually referred to as *context variables*, and are conceptually independent from the programme but can also influence stability, like the insurgency and counter-insurgency, other development activities, or demographic as well as geographic context conditions. The latter, i.e. the way programme measures are implemented, are referred to as *process or moderator variables* and

are directly linked to SPNA implementation, such as more or less inclusive or participatory modes of delivery, or variation in the prioritised type of the project and in the maintenance of the infrastructure outputs that are intrinsic to the programme itself. We depict the resultant impact model in Figure 5.

As the diagram shows, programme measures are very likely not the most important factors contributing to changes in stability. Instead, the overall security, political etc. context is likely to be the main influence determining stability outcomes from the actual SPNA programme measures and modes of implementation only moderately affect these outcomes. Whether this influence is for the better or worse, and whether one can identify any measurable impact of programme measures on stability at all, are questions the current assessment will tackle.

Assessment Approach

The challenge for causal inference

We derive our assessment approach from the SPNA activities and the impact model (both discussed above). Identifying the likely impact of a programme ideally implies an experimental design in which randomly selected "treated" units are compared to equally randomly selected untreated, "control" units that have not participated in the programme. SPNA's main treated unit is the district. Due to *conceptual* and *financial* reasons, when designing this impact assessment in 2010 in coordination with AKF (Afg), this impact assessment did not comprise untreated (control or comparison) districts. The main conceptual reason is that at the time of SPNA's inception, international development spending was at its height. While no other programme offered a comparable mix of infrastructure grants combined with capacity-building as SPNA did, several other projects worked on the district level, offering large-scale key infrastructure development while often also consulting with DDAs. Moreover, with the intensifying US Surge, development spending was expected to further increase.

Under such circumstances it seemed impossible to find a credible set of comparison districts. Given the low likelihood of identifying a valid comparison group at district level – that would also remain untreated in the foreseeable future – the financial implications of adding 10-15 comparison districts (and thus raising costs by 30-50%) were not justifiable.

In our assessment methodology and analytical strategy, we aim at mitigating this shortcoming principally in four ways:

- We construct treatment and comparison groups through differences **in the awareness and assessment of the DDA and of SPNA infrastructure** by respondents within one district. We take awareness and assessment as proxy indicators for the immediate outcomes SPNA intends to achieve – more visible and better DDAs, and more visible and useful infrastructure. The assumption is that higher values for these immediate SPNA outcomes predict higher values for our survey-based stabilisation indicators. Within the same district we expect that people who are more exposed to SPNA and have better opinions about SPNA outcomes will also have better views about stability than people less exposed and with lower opinions.

- Through specific **features of SPNA rollout and resulting exposure to SPNA.** Districts differed substantially in terms of the timing and duration of exposure to SPNA activities due to the step-wise roll-out and termination of the program. Importantly, the roll-out and termination patterns were largely idiosyncratic: they were driven by administrative considerations rather than by any specific socio-economic or political district characteristics. We made use of this quasi-random SPNA-exposure to induce exogenous variation into our main explanatory

variables (the awareness and assessment of the DDA and of SPNA infrastructure) and thereby improve the identification of causal effects on stabilization outcomes.

- Through an **in-depth qualitative research** that flanks the quantitative survey and qualitatively traces causality from SPNA inputs, over DDA activities to DDA interactions with a district's population leading to changed perceptions that are measured in our (quantitative) perception survey.

Mixed-method approach

We designed a mixed-method approach combining a quantitative survey component with extensive qualitative research to gather data for measuring and investigating the four sets of variables (dependent variables as well as independent variables consisting of context, programme and process variables) identified by our model as necessary for arriving at a net effect of SPNA on stability. The quantitative survey forms the backbone of the assessment, complemented by comprehensive qualitative research focusing on the same districts and communities as surveyed by the quantitative research. The extensive qualitative research serves two main purposes: we used its results for developing quantifiable indicators which we integrated into the statistical analysis (e.g. data pertaining to the village) and, crucially, we also used the qualitative research to interpret and explain our statistical findings within the local context. Lastly, qualitative research also helped us identify behavioural and operational patterns of CDCs, CLDCs and DDAs that could be later tested by statistical means to identify differences within districts with regard to exposure to SPNA measures. For example, our qualitative research identified that not all CDCs cooperated with their CLDCs and that not all CLDCs participated in their DDA. CDCs without such a connection to the DDA are likely to have less exposure SPNA benefits as they lack an organisational link to communicate their development and governance-related issues and problems.

Sampling

In order to ensure a valid representation of survey districts, we designed a survey method whereby in each district we selected five *village (CDC) clusters* and within each target cluster we surveyed two villages (in the case of very small clusters of only two to three communities, we only sampled one village). The clusters were selected based on differences of six criteria: *remoteness*, *size*, *ethnic* and *religious* composition, *access to natural recourses* (rain-fed and irrigated agriculture), and *security*. Where we had insufficient information about a district and the geographic location of communities, we selected survey clusters and communities randomly. We chose this qualitative criteria-based mode of selecting survey communities within districts, because in 2010-11, when we designed the survey, there were no reliable population figures available for districts and sub-districts.[3]

Within villages, we randomly sampled households and within each household, we interviewed the head of household (*rais-e khanawada*). The rationale behind interviewing household heads (and not for instance a stratified sample of different age groups and gender) is the consideration that in rural Afghanistan, households still act as corporate units, the political opinions and stance of which are represented by the head of the household. As a result of the above outlined sampling approach and limitation of reliable demographic data, we can claim statistical representivity in a strict sense only on the community level.

[3] According to our knowledge, even today, the Central Statistics Office (CSO) only publishes population estimates and has not yet carried out a comprehensive census.

Survey methodology

Quantitative component: The quantitative component consisted of a standardised questionnaire of some 50 questions, which remained largely unchanged over the waves. One important addition related to the inclusion for Wave 2, 3 and 4, of our key primary independent (explaining) variable asking for knowledge and assessment of SPNA-delivered infrastructure and of knowledge and assessment of the DDA. Further limited changes concerned issues such as the inclusion of ALP as an armed actor in security-related questions after Wave 1, as ALP was only set up in the survey region from 2011 onwards. The quantitative survey was conducted by teams of enumerators hired and supervised by AHRRAO and trained by ARC in quantitative survey methods, the sampling procedures and security routines.

Filling out the questionnaires was solely paper-based in Waves 1, 2 and 3 and inputted into a database using a double-data entry procedure. In Wave 4 we also applied computer-assisted surveying (CAPI) using the World Bank's Survey Solutions software, wherever we deemed it sufficiently safe for enumerators to use such appliances. In more insecure areas, mostly where the Taliban were present, we continued using PAPI surveying as the Taliban were likely to clamp down on the usage of any appliance with a GPS function.

Qualitative component: The qualitative component was, in many respects, methodologically more complex. Its two key components were the conducting of (a) profiling interviews on the district, cluster and CDC/village level and (b) semi-structured interviews conducted on the district level by trained qualitative researchers. The ***profiling interviews*** provided us with historical, economic, demographic, political, governance and security related data regarding *all* units of analysis, i.e. all survey districts, all survey clusters and all survey CDCs/villages. The profiling interviews were carried out by specially trained enumerator team leaders who travelled and supervised the teams of enumerators. The ***semi-structured interviews*** were carried out by trained qualitative researchers on the district level. Interview partners included district-level actors such as the head of the DDA (or a deputy), the woliswal, chief of police, selected line department heads (health, education, agriculture), huquq department, important district elders, an important district-level religious leader, a trader in the bazaar, etc. All in all, 10-11 such interviews were conducted per district. The length of the interviews ranged from 10 to 20 questions and lasted – depending on circumstances – from 30 to 90 minutes. The profiling and semi-structured interviews were conducted in a native Afghan language (mostly Dari or Pashto) and were subsequently translated to English.

A further key component of the qualitative research were the lengthy ***debriefings*** of the enumerators, profilers and qualitative researchers by the ARC consultants. These debriefings gathered contextual information and observations that was not captured in interviews. Moreover, we also used the debriefings for joint analyses using, among others, conflict analysis tools as developed by ARC (Koehler 2015).

In addition to the profiling, semi-structured interviews and debriefings of AHRRAO researchers, the ARC consultants also conducted its ***own interviews*** with AKF and ACTED staff, DDA representatives, government officials, CDCs and cluster representatives. These interviews were usually mostly unstructured and served to orient the research and help identify new issues and topics that were not yet included as items in the semi-structured interviews of the field teams. Lastly in the course of each wave, we also **debriefed AKF, ACTED** and in Wave 4, **Mercy Corp staff** (engineers and social organisers) on the issues and challenges of project implementation. These interviews were semi-structured in order to allow for comparison over the different waves.

FIELDWORK

The fieldwork was carried out by the Mazar-e Sharif-based Afghan Human Rights Research and Advocacy Organisation (AHRRAO). Fieldwork typically consisted of an approximately five-days-long training, during which we trained quantitative enumerators and qualitative profilers and interviewers. After test runs and in-field coaching, AHRRAO took over the management of the fieldwork, while ARC consultants directed the work remotely. At the end of fieldwork, we gathered the teams for a 5-6-days-long debriefing. (The last debriefing at endline lasted ten days due to the larger number of survey districts – 35 instead of 25 during the previous waves.) In addition to managing the AHRRAO field research, ARC consultants also conducted own field visits and interviewed DDA representatives, members of the administration, project beneficiaries as well as AKF, ACTED and Mercy Corp staff to gain a more detailed understanding of the situation on the ground and understand the challenges of implementation from the perspective of the implementing organisations.

Concerning the actual implementation of the four survey waves (see table below), together with AHRRAO we conducted the baseline (Wave 1) in 2010/11, the first follow-up survey (Wave 2) in 2012, the second follow-up in 2014/15 (Wave 3) and the endline in 2017/18 (Wave 4). In each of the first three survey waves we revisited the same districts, village clusters and villages surveying between 231 and 251 villages and more than 5,000 respondents per wave. In the final wave we revisited the same 25 districts that formed the core of our survey since the baseline and added – upon the request of AKF-A – an additional ten districts. The total endline survey thus comprised 35 districts with 352 communities and 8,458 respondents. The total number of heads of household interviews for the four waves thus adds up to 23,959.

Overview and key features of the four survey waves

Wave	Survey period	# districts	# communities	# respondents
1	Sep-Dec 10 / May-Aug 11	25	231	4,928
2	Sep-Dec 12	25	251	5,203
3	Nov 14-May 15	25	245	5,370
4	July 17-Apr 18	35	352	8,458

Table 4: Key features of survey waves

Security posed a serious challenge to the fieldwork in particular during Waves 3 and 4. In order to cope with the deteriorating security situation, ARC designed a five-step security protocol for field teams to negotiate access to insecure communities. Using this protocol during Wave 3, the survey team managed to gain access to all but six survey villages of Wave 2. During Wave 4 security deteriorated further, moreover the scope of the survey was expanded by ten new districts. However, using the above 5-step protocol, the survey team managed to access all but one survey village of the originally 251 communities surveyed in Wave 2 and cover all ten new survey districts.

ARC security protocol for accessing insecure communities

Step 1	If entry to the village or a cluster is denied or non-advisable contact AKF-A, ACTED or Mercy Corp and discuss, if they can help facilitate access to the location.
Step 2	If this does not work, try to negotiate access through DDA representatives and/or elders; the emergency budget can be used to reimburse an elder or a scout to accompany the teams to the research location.
Step 3	If this does not work, cover the village by proxy; in this case AHRRAO manager, Hayatullah Jawad, will personally coordinate the choice and training of proxy surveyors from the village.

Step 4	If the situation is too dangerous even for proxy surveyors we included the option of repeating the selection strategy taken and recorded in the village cover sheets of 2012 to identify the sample of respondents and ask them for an interview in a safe location (e.g. another village, the district center, or similar).
Step 5	If none of the above works, withdraw and wait for a possible change / improvement of the security situation

Table 5: ARC security procedure

During the last wave (Wave 4), we faced a hitherto unknown problem. In the course of an introductory meeting of our Kunduz team with a Taliban representative from the same province, the aforementioned representative expressed his hope that people (i.e. respondents) would confirm that security had improved markedly in the province. This would prove to foreign NGOs that development work was once again possible – even in the Taliban-held areas of Kunduz. Such meetings with government, militia or Taliban representatives were part of our normal security protocol and served the purpose of explaining the survey and thus avoiding suspicion of undue information gathering. In this case, however, the team interpreted the remark of the Taliban representative as a veiled instruction to make sure respondent security assessments were good. Thereupon the team stopped asking respondents certain security-related questions and ticked the same positive response for all interviewed respondents.

At the end of the survey, we quickly spotted the invariant positive responses in three of the survey districts in Kunduz Province. Faced with this evidence the team admitted having manipulated responses to certain security-related questions. ARC informed AKF (Afg) about the problem and asked for an extension of the schedule of the contract in order to be able to rerun the survey in the affected districts. We received a green light from our AKF (Afg) counterparts and succeeded in repeating the survey in two districts. In a third district, Qala-i Zal, in expectation of an impending government offensive, the Taliban forbade all outside access to the areas under their control. The rerun of the survey in the two Kunduz districts was successfully concluded by mid-July.

DESCRIBING THE CONTEXT

NORTH-EAST AFGHANISTAN

GEOGRAPHY AND ECONOMY

The 35 survey districts lie in North-East Afghanistan in a geographically and ecologically diverse region. Low-lying irrigated plains characterise the northern and north-western section of the area along the border with Tajikistan. These irrigated flatlands support a large population, mainly growing rice where irrigated water is abundant, and wheat where they have to rely on scant and unpredictable rainfall.

The fertile irrigated plains are surrounded in the south by the Hindu Kush (highest peak in the research region Koh-e Bankdaka with 6,812m) and the east by the Pamir (highest peak Koh-e Pamir with 6,320m) mountain ranges. The rugged terrain, though abundant in water, limits possibilities for irrigated agriculture. Instead, livestock, mostly sheep, plays a more important role in the subsistence of households. Nonetheless, this high mountain environment only supports much smaller populations than the fertile plains of Kunduz, northern Baghlan and north-western Takhar.

The population of the study area is still predominantly rural, though lacking any census the size of the rural and urban populations cannot be precisely estimated. The only larger urban centres, Kunduz City (in Kunduz Province) and Pul-i-Khumri (in Baghlan Province) are located in the irrigated plains in the west of the research area.

Industrial production is rudimentary only and despite rich mineral deposits, mining is almost exclusively artisanal, limited to gem stones such as lapis lazuli in Kuran wa Munjan. The illegal cash crops, opium and hashish are produced mainly in the mountainous areas of Badakhshan and Baghlan. In terms of traffic infrastructure, two strategically important paved highways cross the western section of the research region. Highway 1 (also called the Ring Road) passes through Baghlan Province and reaches the Afghan capital Kabul after passing the Salang Tunnel. The second road of strategic significance bifurcates from the Ring Road in Pul-i-Khumri and leads northwards from there to Tajik border at Shir Khan Bandar. A third highway of strategic importance is currently under construction. Leading from Kunduz City via Taloqan and Fayzabad, this road would ascend the Pamir and reach China via the Wakhan Corridor – a route that Marco Polo is also said to have taken in the late 1200s. As of the endline, the road is paved until Baharak District in Badakhshan, while construction work is ongoing on the next section of the road in Taliban-held Wardooj District.

Ethnic and religious composition

The study area is home to a large number of ethnic groups, most importantly Tajiks, Uzbeks, Hazaras and Pashtuns. Smaller ethnic groups include, among others, Turkmens, Balochis, Shughnis and Wakhis. Populations are mostly mixed in the fertile plains, but tend to be mono-ethnic – mostly Tajik and to a lesser degree Uzbek – in the high mountain areas. In terms of religious composition, the area is majority Sunni, with Shia and Ismaili minorities. The latter live mostly in high-mountain districts in Badakhshan (e.g. in the survey districts Zebak, Wakhan, Kuran wa Munjan, Shiwa) Takhar (in the survey district Warsaj) and Baghlan.

Historical overview

Following the US invasion of Afghanistan and fall of the Taliban in 2001, an initial period of increasing peace and stability followed. Starting in 2005, security in this part of Afghanistan was the responsibility of the German-led Regional Command North of the International Security Assistance Force (ISAF). In contrast to the South and East of Afghanistan, ISAF insurgency activity was low and ISAF played a more passive role focused on presence and civil military cooperation during those first years (Chiari 2014).

The tipping point in the North occurred in 2007–2008 when the Taliban, after re-establishing themselves in the south, used contacts to former (mainly Pashtun) supporters in the north to systematically re-launch the insurgency in this region (Koehler and Gosztonyi 2014). Between 2009 and 2011, the insurgency managed to destabilize significant parts of the north and some areas even came under full insurgent governance (Giustozzi and Reuter 2011). During this period, the Taliban also succeeded in making important inroads into non-Pashtun parts of the local population.

The country-wide successes of the insurgency and the prospect of losing the war in Afghanistan compelled the US government to increase its military and development resources (McChrystal 2009). The US troop surge, which began in 2009, had significantly disrupted insurgent structures by 2011 and managed to take back large areas that had been entirely lost for the government until late 2010. The military surge was accompanied by a civilian surge of both US (countrywide) and German (North-Afghanistan) development aid, focusing on capacity-building and infrastructure to strengthen Afghan governance capabilities (BMZ 2010, 2014).

Prior to this surge, military security concerns rarely informed aid-allocation strategies of the main development donors in the North-East. Before the insurgency gained momentum after 2007, insecurity was simply no overriding concern for development work and the strategies of state and non-state agencies were explicitly oriented towards development impact. With the partial deterioration of security, development efforts initially stalled in some of the most dangerous areas or continued via local proxies. Development work financed by the German Ministry of Development (BMZ) started to focus explicitly on more secure areas while the military and the German Federal Foreign Office kept covering more insecure areas. The US surge added to this balance with the USAID-financed SIKA programme that focused on strategically important communities in insecure areas. Thus, while the security context influenced strategies of individual donors, it did not determine people's overall access to development aid.

SUB-NATIONAL GOVERNANCE ON THE DISTRICT LEVEL AND BELOW

DISTRICT AND SUB-DISTRICT LEVEL GOVERNMENT

This section provides a brief introduction to some crucial elements of the local system of governance in the survey region. As in most parts of Afghanistan, sub-national governance in the north-east is highly complex, combining institutions of the state administration and influential informal societal organizations (Murtazashvili 2016). These different levels and forms of governance sometimes co-exist, sometimes complement each other, and sometimes compete. In the period under investigation, there were 34 provinces in Afghanistan, each of which was headed by a provincial governor (*wali*) appointed by a government agency directly responsible to President Karzai and now to President Ghani (Nixon 2008).

Provinces are subdivided into districts, the lowest functioning tier of constitutionally recognized state administration in Afghanistan. Districts are headed by the district manager (the *woliswal*), a political appointee. Even though its power is formally limited to coordination, in practice the institution is very powerful on the local level. The woliswal is often described as a "gatekeeper" who controls access to service delivery and to the higher levels of government (Saltmarshe/Medhi 2011). Woliswals and their administration also play a critical role "as the face of government with which most people come into contact, and their interest in and ability to help people greatly influence people's attitudes toward the government as a whole" (The Asia Foundation 2007: 6). Districts also have their own police department, prosecutor, and district court, as well as departments of the line ministries such as education, health and agriculture. As of the writing of this report, constitutionally mandated district councils have not been set up.

The next official administrative unit below the level of the district is the village. As administrative units mentioned in the constitution and thus far further defined only by government policy papers, villages are not functional (Lamb 2012). Instead, the governance space outside of the centrally appointed district-level state representatives (woliswal, police, the court and the line departments) is filled with unofficial and more or less formalised local institutions. (We omit the discussion of municipalities, which, even though increasingly functional in recent years, do not play a major role in our assessment.) Most importantly the three-tiered *shura structure* bridging the gap between communities and the district-level state (more about this structure in the next section) and informal institutions such as traditional village *shuras* or *jirgas* (regular or issue-related councils), as well as various forms of more or less traditional local authority like *maliks*, *arbabs* or *qariya dars* (village headman) and *bozorgan* (community elders).

Where insurgents have wrested control from the government (like in parts of Baghlan, Kunduz and Badakhshan), they replaced the state administration with their own leaner administration usually comprising of a Taliban woliswal, judges, fighters providing security and increasingly also liaison officers with NGOs who are also responsible for development in their district.

In cases of Taliban takeover, some government departments, most importantly education, health and to some extent agriculture, are allowed – and even requested – to continue providing services, but their administrative support structure then works in exile in districts under state control. The Taliban then take over the monitoring of service provision, e.g. the attendance of teachers or medical staff (July 2018 debriefing in Mazar-e Sharif). CDCs tend to continue their work under the supervision of a Taliban representative on the community level, but the higher levels of the structure, CLDCs and the DDA, are mostly defunct in districts under Taliban control.

THE SHURA STRUCTURE

Given the importance of the shura structure for sub-national governance in general and for this evaluation in particular, we discuss it in more detail in this section. By the term "shura structure," we refer to the three-tiered structure of development shuras as established under the Ministry of Rural Rehabilitation and Development (MRRD). The core building blocks of this structure are the CDCs first established in 2003 by the National Solidarity Programme (NSP). The NSP was rolled out in three phases, reaching nationwide coverage by 2009-10. CDCs represent communities comprising between 25 and 300 families. The CDC representatives are elected by secret ballot by all male and female members of their respective community. Contrary to the nationwide elections for the presidency, Wolesi Jirga, or PCs, which are held on the same day throughout the entire country, CDC elections are a lengthy individual process preceded by extensive community mobilisation, which is organised by an NGO facilitating partner, contracted by the MRRD to facilitate the implementation of the NSP in a given district. As a result of this extensive community involvement, CDC elections are not simultaneously held on the same day in a district, but rather consecutively.

The core function of CDCs centres on prioritising, facilitating, monitoring, and partly implementing development projects in their communities. Early on, however, CDCs were intended to serve an additional goal: namely to become "effective institutions for local governance and social-economic development" (MRRD, National Solidarity Programme Phase Three, p. 11). Thus, in addition to their development-related tasks, CDCs also resolve conflicts, organise hashar (unpaid communal work), and represent the community to the outside world. At least in northeast Afghanistan from 2007 onwards, CDCs became the main community-level governance institutions in the study area (Figure 6). Very likely with the disbursement of a second round of NSP block grants to CDCs from 2013 onwards, the importance of the CDC head increased dramatically until 2015, only to drop once again to pre-2013 levels. Concurrent with this increase, there was first a drop (until 2015) and then an increase (in 2017) in the perceived importance of the main alternative contender for authority in the villages: elders. At present a new World Bank-funded and MRRD-led programme focuses on the community level and offers new block grants to re-elected CDCs, the Citizen Charter. However, at the time of the survey only very few communities have entered the programme already, though some places have heard about it and were looking forward to it.

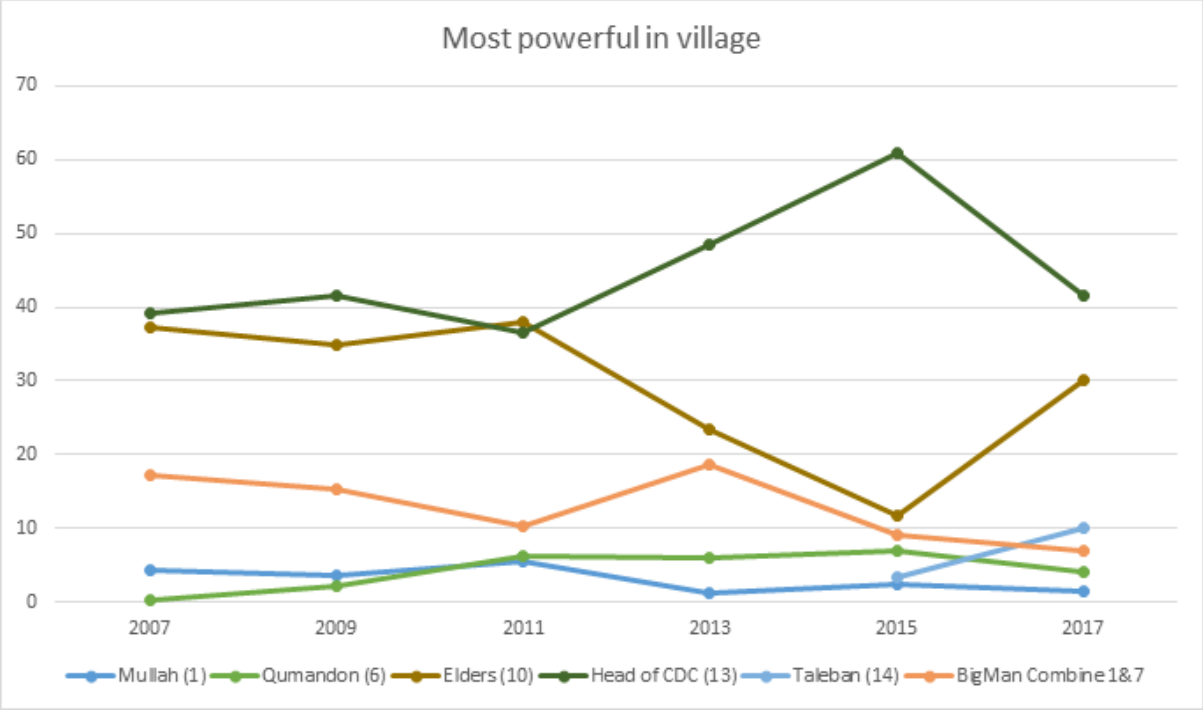


Figure 6: Respondent perceptions regarding the most powerful person in the village. Since the mid-2000s, the head of the shura (CDC) occupies a prominent position as an important village-level authority. Source: Six successive surveys conducted by Jan Koehler in four districts of Kunduz and Takhar provinces; the survey was conducted every two years in the same villages, but did not interview the same respondents. It is thus not a classic panel survey (Freie Universität Berlin, SFB 700 C9)

Building on the CDC structure, two additional institutional innovations have been implemented since 2007 by MRRD. Within the framework of the National Area-Based Development Programme (NABDP), CDCs were gathered into geographically defined Cluster-Level Development Councils (CLDCs) comprising representatives of the member CDCs. CLDCs in turn send representatives to District Development Assemblies (DDAs), the main beneficiary institution of SPNA, composed of the representatives of all CLDCs within a district.

The tasks of CLDCS and DDAs are similar to those of CDCs, but at a higher level (see GIROA, "Sub-National Governance Policy"). They thus prioritise, lobby for, and facilitate development projects, organise hashar, support the maintenance of projects, resolve conflicts, and frequently represent the interests of the community towards the state. The shura structure is hence the only inclusive, locally elected and rooted governance institution of rural areas, in an otherwise highly centralised state. As already mentioned, the main interface of the shura structure with the centralised state administration occurs at the district level.

Of the three tiers of the shura structure, the lowest level (i.e., CDCs) is the most strongly institutionalised and functional, followed by the somewhat less institutionalised DDAs. In recent years, both bodies have received substantial funds and capacity-building support from national and international actors. CLDCs are the least institutionalised level of the shura structure, and they are even absent in some districts (e.g., Warsaj District in Takhar Province, Arganj Khaw / Shiwa in Badakhshan Province and Charkint District in Balkh Province).

AFGHAN SUB-NATIONAL GOVERNMENT POLICY AND THE DDA

A remark is necessary regarding the legal status of the DDA, the key institutional partner of SPNA, and of CLDCs and CDCs, the structures upon which DDAs are built. With regard to sub-district governance, the Constitution is the key legal document. Art. 138-140 of the Constitution requires the establishment of elected district and village councils. So far, however, neither district, nor village council elections have been held and these constitutionally mandated bodies have thus not been established.

Turning to DDAs and CDCs, they are formal institutions of local governance, but they are not the equivalent of the constitutionally mandated district and village councils. They are formal, because their tasks and structure are laid down in decrees and policy documents: for CDCs, a presidential bylaw signed in November 2006 and ratified as per MRRD's Rules and Regulations; for DDAs, the National Area-Based Development Programme (NABDP) implemented by the MRRD with support from the United Nations Development Programme (UNDP 2009). In contrast to CDCs and DDAs, CLDCs have no formal basis. From MRRD's perspective, they were only established to facilitate the election of DDAs. In many areas, however, CLDCs continue to function even after the DDA elections, and some NGOs even support them, as it is convenient to have a functional governance and developmental entity above the CDC, but below the district level. They can thus be described as an informal organisation.

While CDCs and DDAs are not the equivalent of Village and District Councils – in the absence of these constitutionally mandated councils – CDCs and DDAs can perform their functions. IDLG's Subnational Governance Policy (SNGP) states in this respect that:

> "CDCs are presently accountable to the people who elect them. CDCs are also accountable to the Ministry of Rural Rehabilitation and Development who provides them block grants. CDCs will be accountable to people while they perform the future role of Village Councils" (Islamic Republic of Afghanistan/IDLG 2010: 18).

The Subnational Governance Policy gave similar temporary powers to DDAs:

> "DDAs will maintain their current roles and responsibilities until the constitutionally mandated District Councils are elected in 2010 [sic!]. DDAs will perform the planning function of District Councils till constitutionally-mandated District Councils come into existence" (ibid. p. 17).

There have been no plans to hold village-council elections and with the recent setting up of the Citizen Charter Programme and the adoption of a related National Priority Programme (NPP), the survival of CDCs as the main representative village governance body is all but guaranteed. In contrast, the status and future role of DDAs remains in limbo. Holding district-council elections was repeatedly discussed among Afghan policy makers, in particular since the establishment of the NUG in late 2014 (see section above).

The reason for the insistence on district-council elections is that one of the key points of the agreement that founded the NUG, was the commitment to change the Constitution. According to the Afghan Constitution, the only body that can change the Constitution is a Constitutional Loya Jirga (grand assembly). However, convening such a Constitutional Loya Jirga first requires the establishment of District Councils, as article 110 of the Afghan constitution specifies that "presidents of provincial as well as district councils" participate. Moreover, district councillors must also elect 32 members of the Meshrano Jirga (the upper house of Parliament) who, too, are key members of a would-be Constitutional Loya Jirga. DDAs as only indirectly elected bodies cannot substitute for District Councils when calling a Constitutional Loya Jirga.

While district-council elections have been repeatedly postponed (most recently they were supposed to be held on 20 October 2018 together with the Provincial Council Elections, see AAN, 26 Sep 2018), the role of DDAs following the establishment of district councils remained undefined. Options ranged from abolishing them altogether, to integrating them into district councils as a development committee. With the closure of NABDP in 2015 (UNDP 2009), funding and official attention by the MRRD to DDAs also came to an end.

As of now, district-council elections have not been held and it is doubtful whether they actually will be held at all in the foreseeable future. At the same time, the DDA's status is also unclear. A 16-page document from May 2017 titled "Citizen-Centered Governance. A Roadmap for Subnational Reform" is currently being circulated as the new sub-national governance policy. The document lacks the usual insignia of official strategies (logo of the government, information about the publishing ministry or government agency, etc.), but has allegedly been adopted by the Cabinet and the Parliament, though we could find no confirmation of this on electronic news media. Importantly, this document makes no mention of DDAs – only of CDCs as the main "representational body at the village level" (p. 10) – and firmly expects the establishment of district councils.

Should this document, indeed, be an official policy paper abrogating the DDA-related provisions of the 2010 SNGP, this would constitute a highly problematic situation. With district-council elections scrapped, DDAs degraded to mere informal organisations and the DDA councils themselves in decline due to a lack of funding and official attention, there would be a complete absence of institutional linkages between the village level and the lowest level of government administration at the district level.

MAIN DEVELOPMENTS IN THE SURVEY REGION DURING THE SURVEY PERIOD

WAVE 1 (2010-11)

SPNA was launched in 2010 in the context of the US-led military and civilian surge announced by President Barrack Obama in December 2009.[4] The rationale behind the surge was to turn the tide in the protracted conflict between the Afghan Government and the Taliban-led insurgency, which had increasingly gained momentum and by mid-2009 appeared to win the upper hand against the government. The aim of the military component of the surge was to push back the insurgents and create space for the civilian surge to win the "hearts and minds" of the population by providing development, and to strengthen Afghan government structures and prepare them for taking over responsibility for the country, once foreign troops were to withdraw in 2014. Several Western countries participated in both the military and civilian component of the surge. Germany contributed to the military component by raising its troop levels in the north, and giving them a more assertive combat role, and to the civilian component with its *Entwicklungsoffensive (development push) Afghanistan*.

We carried out the **baseline for Cohort I** (the first 15 districts of SPNA) in autumn 2010 amidst intensifying fighting in the survey region, as US and other international troops were deployed to the North and increasingly began to challenge the insurgents in their strongholds. The **baseline for**

[4] The US military surge began 2009 and peaked late 2010 / early 2011. It temporarily deployed an additional 33,000 US troops to Afghanistan raising total US troop number to 101,000 (total NATO forces reached 140,000). As part of the surge, US military presence in Regional Command North was also significantly increased. The withdrawal of US troops began still in 2011 with the last surge troops leaving Afghanistan in 2012.

Cohort II (the second batch of 10 SPNA districts), implemented in spring 2011, already encountered improved security and a generally changed landscape, as the military surge pushed the insurgents back and dismantled their governance structures. The dominant and highly visible actors of the baseline period were, however, international military forces and international development organisations, largely overshadowing the Afghan National Security Forces (ANSF) and Afghan government representatives in importance.

In terms of the governance and developmental context, the shura structure was fully up and running at baseline (Wave 1). Most CDCs had either just completed or were busy implementing their NSP 1 projects. The DDAs in the survey region had all been set up between 2006 and 2009, implemented their NABDP and were increasingly being taken seriously by development actors (IOs and INGOs) who began to consult them when implementing projects in a district. Aid inflow at this time was very high, with our survey villages reporting an average of 1.8 projects having been implemented "in the past two years".

WAVE 2 (2012)

ARC's **follow-up 1** stability assessment, implemented in autumn 2012 and early 2013, witnessed major changes on the ground. The security gains achieved by the military surge still largely held, but international military forces progressively took the back seat, allowing the increasingly well-trained and capable ANSF to take the lead. The phased handover of security responsibilities to the Afghan Government, called *Inteqal*, affected all 25 SPNA target districts of the stability assessment (see NATO 2013) and became thus the dominant context for the follow-up assessment. In this phase Afghan security and government actors increasingly came to the fore, while international security and development actors became less visible and *de facto* also less able to shape developments on the ground.

Turning to the governance and development context, as a result of the conclusion of NSP 1, village-level project counts dropped to 0.8 per survey community (from 1.8 during Wave 1). District-level aid projects continued, however, on a large scale. In particular USAID's Stability in Key Areas (SIKA) has to be mentioned. SIKA concentrated on strategically important districts of Kunduz and Baghlan. In spite of noteworthy similarities to SPNA (SIKA concentrated on the district and had a fund character), it differed from SPNA in that it did not cooperate with DDAs, but established its own *ad hoc* shuras to prioritise and disburse aid in insecure areas.

The DDA structure also experienced a noteworthy change in this period. Shortly before or during the Wave 2 survey, 18 of the original 25 survey DDAs were reorganised and held renewed DDA elections. A further four DDAs were reorganised and held new elections shortly after the Wave 2 survey, while three districts kept their old structure and old composition without new elections.

WAVE 3 (2014-15)

We carried out **follow-up 2** in late 2014 and early 2015 in the context of a dramatically transformed political and security context in the whole of Afghanistan and also in the survey region. *One crucial game changer* was the presidential election in spring and summer 2014. The first round of elections held in April reduced the field of contenders to two leading candidates, Ashraf Ghani, a former World Bank official and Afghan Minister of Finance between 2002 and 2004, and Dr. Abdullah Abdullah, former chief of staff and Ministry of Defence spokesperson of the Rabbani Government in 1992 and main challenger of President Hamid Karzai in the 2009 elections. The second-round run-off in June 2014 resulted in a win for Ghani, but was contested by Abdullah Abdullah with both sides accusing the other of election fraud on a massive scale. The standoff following the announcement of preliminary

results early July brought Afghanistan to the brink of ethno-political conflict and could only be overcome by repeated US mediation. The power-sharing agreement brokered with US support in September 2014 led to the establishment of the National Unity Government (NUG) with Ashraf Ghani as President and Dr Abdullah Abdullah as Chief Executive Officer. The compromise successfully averted a violent escalation, but the three-months-long stalemate, and resultant lack of leadership and uncertainty, also offered the insurgents a chance to intensify their campaign, leading to the first significant deterioration of security in the North following the success of the surge.

The *second crucial game changer* in Afghanistan was the conclusion of *Inteqal* (transition): the withdrawal of most foreign forces and the concurrent handover of security responsibility to the ANSF by the end of 2014. The remaining foreign forces under the new NATO Mission *Operation Resolute Support* were to provide training, advice and assistance to the ANSF with only a limited counter-terrorist and combat mandate. The handover of military responsibilities to the ANSF was accompanied by a similar incremental assumption of political, administrative and development responsibilities by Afghan government actors and a concurrently diminishing role of, and decrease in funds for, international civilian (diplomatic and development) actors. The conclusion of *Inteqal* meant for the North that, in the months following the signing of the power-sharing agreement, the ANSF had to confront the onslaught of the insurgents without any substantial international military support – most importantly without the support of a capable air force. Contrary to some expectations, the ANSF did not break under the pressure. It put up a tough fight, but it soon became evident that it was badly overstretched and suffering losses that were likely unsustainable in the long run.

A *third key development* shaping the context of follow-up 2, is the escalation of the insurgency and increasing territorial gains by insurgents in the North. This resurgence was apparently part of a well-planned Taliban offensive to destabilise areas hitherto considered relatively stable (s. Giustozzi/Mangal 2015). It was, however, also facilitated by the months of power vacuum following the inconclusive run-off elections in June 2014 and the conclusion of *Inteqal* in December of the same year.

These game changers had a profound impact on the governance and development context in the survey region. For one, we noted a resurgence of former jihadi commanders as key military, political and governance actors on the local level. In the years following the collapse of the Taliban regime in 2001 and the establishment of the Karzai Government, jihadi commanders and their men, the main actors during the anti-Soviet jihad and the subsequent civil war, started to lose power and influence. This trend was first halted in 2009 with the initial post-intervention resurgence of the Taliban, when former jihadi networks were called upon by government security structures, and crucially by the US military, to set up militias in insurgency affected areas. The improvement of security following the surge temporarily slowed the re-empowerment of former jihadi structures, but with the renewed deterioration of security in 2014 this once again picked up pace, and became an important trend needing to be considered when interpreting the results of Wave 3.

A further important development on the local level, unrelated to the game changers mentioned above, concerns the rollout of the second round of NSP, which coincided with our Wave 3 survey. As part of the process, a second round of CDC elections were held and the newly elected CDCs were busy planning and implementing projects out of their allocated block grants. Our survey thus noted 150 newly formed CDCs (60.9%) out of a total 246 survey communities and project counts rose once again: to 1.2 per community as compared to 0.8 in the previous wave.

WAVE 4 (2017-18)

We implemented our expanded **endline** survey – now comprising 35 districts instead of 25 in the previous waves – between August 2017 and April 2018. The problematic trends noted in the previous wave continued. Taliban and commander-controlled areas expanded, while areas formally governed by the government decreased. For the first time we noted significant violent conflicts between drug traffickers for trafficking routes and settling scores among each other, suggesting that the state was losing its capacity to informally arbitrate in conflicts among different illegal groups. International interest in Afghanistan also continued to wane. The limited increase and more aggressive mandate of US troops, announced by US president Donald Trump in August 2017, was hardly perceivable in the North-East during survey implementation.

We also noted substantial changes with regard to governance and development. The technocratic approach of President Ashraf Ghani to state-building, governance and development as laid out in his book with Clare Lockhart (Ghani/Lockhart 2008) could now clearly be felt on the ground. Appointments by the Central Government tended to be merit based, not following the previous pattern of allocating posts according to clientelistic networks. While this supposedly led to more competent officials being appointed, it came at a high price. The strong apex patrons who were faulted for lot of the bad governance outcomes on the ground, appeared significantly weakened during this survey wave. Their decline loosened their control of informal armed pro-government militias who, as a result, began to act increasingly independently. The already noted increase in violent conflicts between different drug trafficking groups might also be linked to the eclipse of patronage as an instrument of informal state control and state penetration of the provinces and districts.

In spite of the technocratic approach of the Ghani administration to governance and development, progress in this field was only slow, leaving the shura structure and its top body, the DDA, in a limbo of uncertainty. Options repeatedly discussed in recent policy papers included elevating DDAs to the status of district councils, integrating them as development committees into – yet to be elected – district councils or eliminating them altogether following future district-council elections. The diminishing international interest in Afghanistan was clearly visible through reduced aid on the district and village level (we noted only 0.9 projects per survey community). The only positive change with regard to district-level governance and development during Wave 4 was the commencement of the Citizen Charter National Priority Programme, which in the survey region was by and large conceived as a continuation of the previous NSP 1 and 2. However, in our survey we only identified 23 survey communities with newly elected Citizen Charter CDCs. The impact of this laudable programme on governance perceptions and service delivery is thus likely to be very limited.

IMMEDIATE OUTCOMES: QUALITATIVE EVIDENCE

As was described, SPNA worked with DDAs, the apex council of the shura structure, to deliver key infrastructure and to improve their capacity to prioritise as well as support the implementation of development projects and to deliver a variety of governance services. As outlined in the impact model above, installed infrastructure funded by block grants and training measures for DDAs are the **output** of SPNA while visible and widely appreciated operational infrastructure and capacitated DDAs the **intended immediate outcome** of SPNA. These immediate outcomes are then expected to lead to improved stabilisation across the four fields (security, governance, development / economy and modernisation) which in the logic of our model corresponds with the programme's **impact** or overarching developmental goal.

This **section relies on our *qualitative research*** and **focuses on the activities of the DDA that define SPNA's immediate outcomes**. In the following, we will first briefly discuss the legal status of the shura structure and, in particular, the DDA as the only grassroots elected governance body on the sub-district level. Then we consider the operationality of infrastructure provided by SPNA as reported to us by respondents in survey districts. Subsequently, we turn to a detailed description and analysis of the DDAs we surveyed. Then we ask how sustainable DDAs are and how the end of SPNA has affected the DDAs in the survey. Lastly, we summarise our results regarding whether findings on the outcome level justify any expectations that SPNA activities might be linked to stabilisation impacts as laid out in our working definition.

IMMEDIATE OUTCOME 1: OPERATIONAL KEY INFRASTRUCTURE

INFRASTRUCTURE BUDGET AND ACTIVITIES

The total infrastructure budget of SPNA over the eight years of its existence was, according to the most recent inventory (November 2018), USD 66 million. The relatively largest share of this budget went to Badakhshan (34.2% of SPNA budget; estimated population: 904,800), followed by Takhar (27.9% of SPNA budget; estimated population: 933,700), Kunduz (18.9% of SPNA budget; estimated population: 953,800) and Baghlan (19.1% of SPNA budget; estimated population: 863,700).

In the course of its existence, SPNA concluded, or is about to conclude, 427 main and accompanying infrastructure measures (completed and in warranty; see SPNA Monthly Inventory November 2018). Concerning the sectors[5] covered by these funded infrastructure projects, the largest share of the budget was spent for education (65%), followed by energy and water infrastructure (12%), roads and transportation infrastructure (11%), health (9%) and governance (2%). These figures show SPNA as a **very important source of development funding across the four North-Eastern target provinces**. In particular towards the end of its existence, i.e. 2016-18, it was one of the only programmes providing funding for larger district-level development projects.

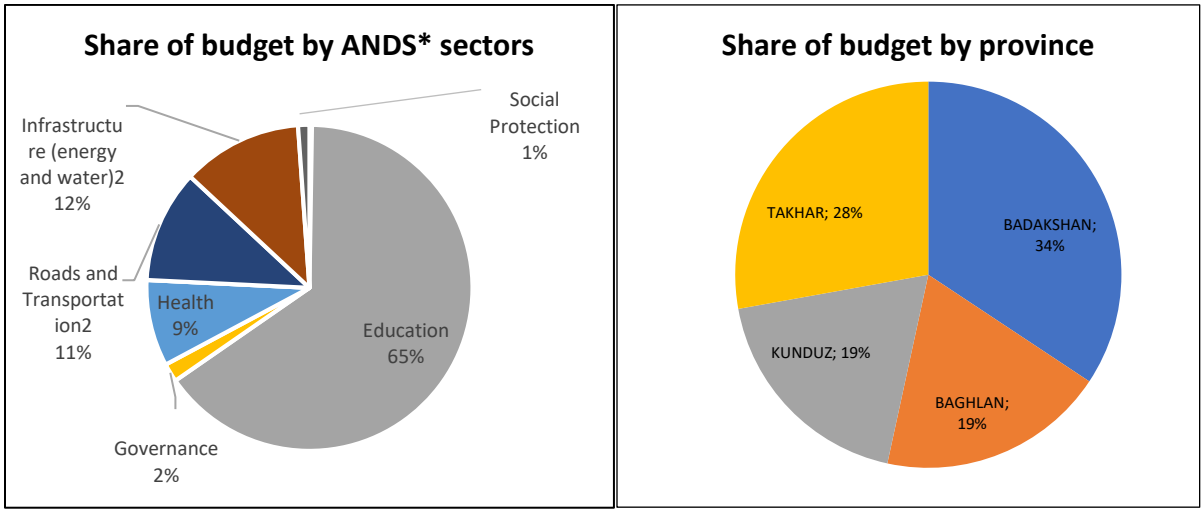


Figure 7: SPNA infrastructure budget by sector and by province (source: SPNA Inventory, May 2017)

[5] The definition of the sectors is based on the Afghan National Development Strategy (ANDS) from 2008.

MODES OF DELIVERY: SELECTION PROCESS

A central feature of SPNA's rationale was the participatory selection of projects by the DDA from the block grant allocated to a district. It is also at the interface of the programme's soft and hard measures. According to the original project proposal of AKF (Afg) to the KfW, in the initial phases of SPNA in a district, projects were supposed to be selected in a non-participatory way, as DDAs were not yet fully capacitated. However, in the later phases, once DDA members had received trainings in district planning and vision building, project selection was to be participatory. Regarding the participation in project prioritisation and modes of deliberation see also the section on "*DDA meetings and female participation*".

Given the central importance of participatory project selection for the logic of SPNA, we coded, based on the SPNA Inventory available to us (May 2017), the mode of project selection for all SPNA-block grant-funded projects in the 35 ARC-surveyed districts, i.e. whether a project was selected in participatory manner or not. Out of the identified 203 infrastructure projects implemented in the 35 survey districts,[6] 162 (80%) were selected in a participatory manner while 39 (19%) were not (for two further projects we could not clearly determine the mode of selection). In the case of the 39 projects that were selected without the involvement of DDAs, it was usually the line departments, or sometimes provincial-level government bodies, that made the suggestions to the implementers for the use of the block grants.

However, a phase-wise breakdown of DDA involvement, based on information provided in the SPNA inventory (May 2017) shows that the situation is not that clear cut. During the initial phases (Phases 1 and 2) non-participation ranged between 21% and 23%, while during the latter phases (Phases 3 and 4) it was still between 6% and 13%. (We disregard Phase 3 as there was only one project implemented in our survey districts in this phase.)

Geographically non-participatory approaches concentrate in Baghlan Province and within Takhar and Badakhshan to Fayzabad (seven out of nine projects without participatory selection selected) and Taloqan Districts (three out of 11 projects without participatory selection). The case of Kunduz Centre clearly shows that even in a provincial centre, participatory project selection is possible. Here an active and competent DDA participated in the selection of 13 out 17 projects (three projects were selected in a non-participatory fashion, while one could not be coded).

[6] For the coding and analysis, we subsumed main and accompanying projects as one single project. As an example, we counted the "construction of Arzbigi mixed school" (KDZ-2010-007) and its accompanying project "GI sheet for latrine, locks, oil painting" (KDZ-2010-007A), as just one project.

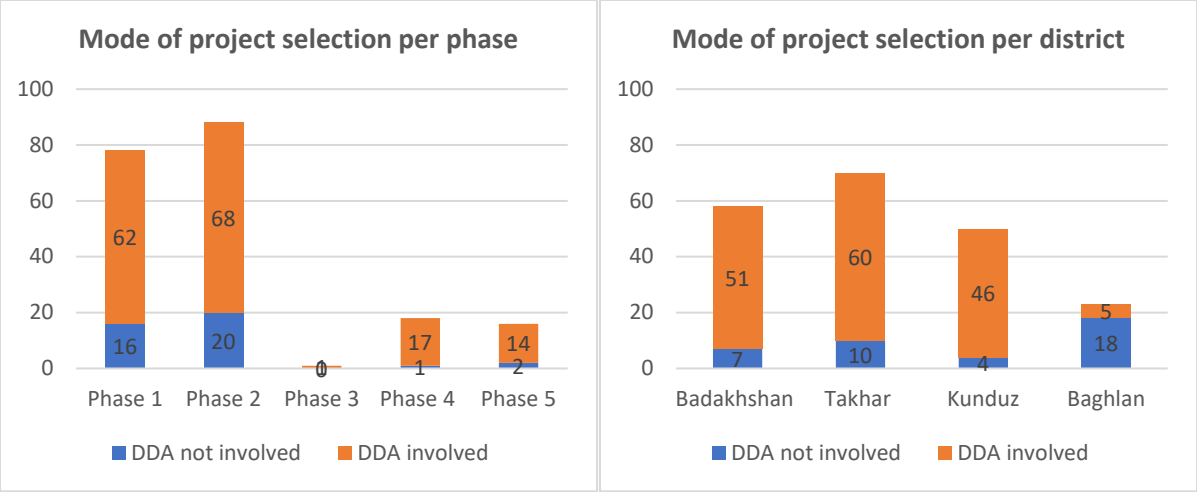


Figure 8: Mode of project selection in the 35 survey districts per SPNA phase and per district

According to our interviews with AKF (Afg)'s Baghlan staff, the provincial office preferred this mode of project selection for districts with larger towns, where DDAs were allegedly not so relevant. They explained that in central districts, projects should be prioritised that are of benefit for the whole province. Hence, they worked via the Provincial Council and the Wali's office. The head of DDA in Pul-i Khumri explicitly took issue with this approach and even refused to meet our team for an interview, because he was upset with what he saw as a lack of cooperation with AKF (Afg). However, looking at the projects in Baghlan that were selected in a non-participatory manner, we find 14 schools (all of which are of local relevance only), three healthcare facilities (of which two might be of broader significance: a public health training centre and a 20-bed hospital for the treatment of drug addicts; both located in the provincial capital Pul-i Khumri) and one road (seemingly also of local importance only).

In Fayzabad, the provincial capital of Badakhshan, the DDA was largely dysfunctional and under strong political influence. This likely explains the selection of projects such as a provincial library, line department office buildings (e.g. health and economy) or university facilities. Outside of Fayzabad, participatory project selection was, however, comprehensively followed in Badakhshan. The situation is similar for Taloqan, the capital of Takhar Province. The implementing organisations had reservations regarding independence of the Taloqan DDA, and thus preferred to implement projects in this capital district in a non-participatory manner. Outside of Taloqan, however, participatory selection was the norm.

Operationality of SPNA infrastructure projects in survey districts

Regarding the first outcome, the operationality of SPNA-built key infrastructure, we asked CDC and CLDC representatives to assess all SPNA projects in their districts. Responses were then coded as positive, mixed, negative or unclear, if respondents had no concrete knowledge of a project, or if the response was not clear. A positive coding signifies that interview partners evaluated a project as fully successful and operational. A mixed coding was used if interview partners generally assessed a project as operational, but mentioned smaller problems (e.g. insufficient number of teachers in an SPNA constructed school) or if statements were contradictory. Negative coding was used if a project was not operational, or if complaints and negative assessments dominated.

As *Figure 4* shows, the majority of projects (68% or 194 out of a total of 205 measures) were evaluated positively. As an example, all five cluster heads interviewed and all ten interviewed CDC representatives were aware of the Jurm Clinic (Project Code: BDK-2011-001) and evaluated it positively. As Hajji Haq Nazar, the head of Chungha Cluster put it: "It is very beneficial for the district. It is operational and all patients from this district go there" (interview on 24 Sep 2017).

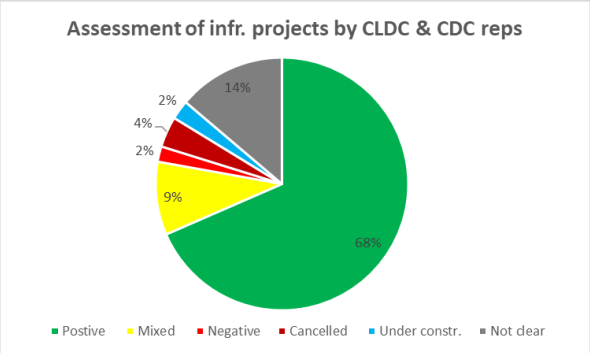


Figure 9: Assessment of SPNA infrastructure projects by CLDC and CDC representatives

We coded a much smaller number of projects as mixed (9% or 19 out of 205 projects). As an example, all cluster and CDC heads interviewed were aware of the Gawhar Bridge (Project Code: BDK-2011-005) and all but two confirmed that it was in use and operational, but that it was too weak and only small cars could pass it: "The contracting company didn't build this bridge properly and it is very small and narrow and [even] small cars can hardly use it but it is good for walking" (interview with the head of Gawhar CDC on 3 Feb 2018). Not all critical remarks are so severe. In the case of the Nawbahar Girls' School (Pul-e Hisar District, Baghlan), some respondents criticised the *tashkil* that only allocated two female teachers to this school: "It is really good and efficient project for girls. This project made girls' education a lot easier, but they have a problem regarding their tashkil. They have only two female teachers, all other teachers are male" (interview with the Head of Ahanagaran CDC on 20 October 2017). Thus, while this project was generally evaluated as positive, the tashkil nonetheless gave rise to some criticism, diminishing to a certain extent the quality of this SPNA infrastructure output in the eyes of the users.

The proportion of fully non-operational projects is even smaller (5% or 11 out of 205 projects). In this group we find cancelled or suspended projects, or projects which mostly due to the bad security situation were out of use, such as the Qalai Zal-Aqtapa Health Clinic (KDZ-2016-004), which reportedly remained incomplete and was used as an army based in this hotly contested district. The situation is similar in Dasht-i Archi, where the Teacher Training Centre (KDZ-2011-027) was damaged by fighting and was, as of the time of the survey, used as an office of the district administration and not as intended as a TTC.

In a next step we examined how project operationality related to developmental sectors, to see, whether certain sectors were more prone to failures regarding operationality than others. For this exercise, we used developmental sectors derived from our questionnaire, instead of the ANDS sectors introduced above (see *Figure 7*). The two categorisations differ somewhat from each other. However, since for the statistical analysis presented later on in the report, we use the categories of the questionnaire, at this point we also use these same categories. This eases the integration of the qualitative assessments of SPNA projects with respondent assessments of improvements in these same sectors.

The questionnaire asks for six categories. However, since in the 35 survey districts SPNA conducted only very few projects, categories such as electricity or drinking water were collapsed to a residual "other" category. *Figure 10* shows the breakdown of project operationality as per development sector derived from the questionnaire. The figure shows the largest number of SPNA-prioritised projects in the field of education, followed by roads and bridges, agriculture and healthcare. Looking at the percentage breakdown of operationality as reported to us by CLDC / DDA and CDC representatives,

we note that most positive results are reported for “roads and bridges” and “education” (schooling), while most negative results, as relates to operationality, concern “agriculture” and “healthcare”. More specifically, respondents evaluated only 35% of all projects in the field of agriculture (6 out of 17 projects) as fully operational and a further 6% received mixed assessments (1). In contrast 4 (24%) projects were cancelled and 6 (36%) further projects were not sufficiently well known to our respondents, to give a codable assessment to us.

The four cancelled projects are irrigation canals in Yamgan, Farkhar, Chal and Khanabad. Two of the projects were cancelled after the feasibility study showed them to be too expensive (Abi Dara water supply in Farkhar and Zanborak Canal in Chal; both located in Takhar Province). A further project, Zard Kamar siphone in Khanabad (Kunduz) was one of the few projects that were cancelled due to security reasons, while the construction of an irrigation canal and siphon in Gharmi (Yamgan, Badakhshan Province) had to be terminated, allegedly due to the refusal of a commander to let the canal pass through his land. The comparatively high number of cancelled irrigation projects suggests that measures in this field might be particularly difficult. The relatively high number of “unclear” codings in the field of agriculture, i.e. projects about which our CLDC / DDA or CDC respondents could not give detailed assessments, is also somewhat puzzling. One of these projects was an agricultural warehouse, whereas the other five were irrigation canals. A possible hint as to the reason for the low level of knowledge might be that three out of these six uncodable projects were selected in a non-participatory fashion.

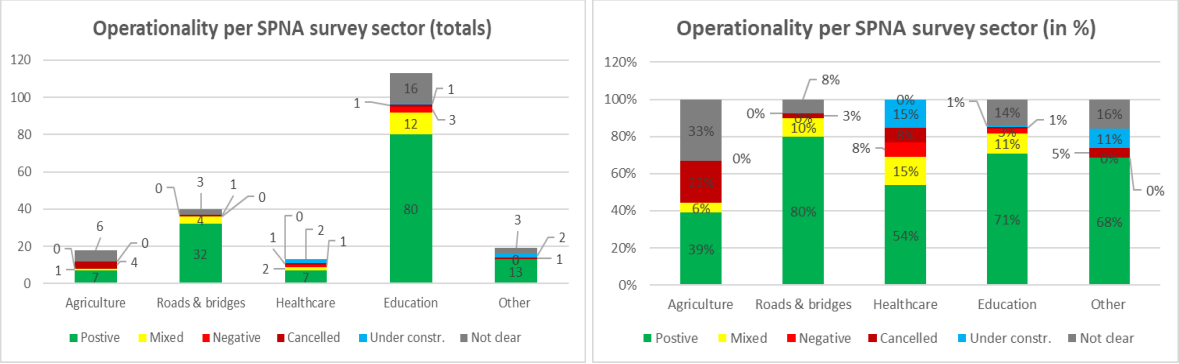


Figure 10: Operationality per sectors using the categories asked for in the survey questionnaire; the figure on the *left* shows total project numbers within each sector while the one of the right depicts operationality as a percentage of total assessments within a sector.

Turning to the health sector, one project received a negative, and two projects a mixed, operationality assessment; a further project had to be cancelled. The negative assessment relates to the Jukhan Basic Health Clinic in Yamgan District (Badakhshan). While the construction of the building appears to be of good quality, the district was taken over by the Taliban soon after the completion of works. As a result, the clinic had not become operational and served for a time as the headquarters of the Taliban. Reportedly, the Taliban vacated the building in 2018 and the Ministry of Health apparently also made moves to allocate staff to the clinic. However, the doctor allocated to the clinic was rejected by the Taliban, who claimed that his medical diploma was faked (debriefing of AKF (Afg) staff on 3 July 2018).

Complaints about the 50-bed hospital in Rustaq (Takhar Province) concern the lack of sufficient medical staff, while regarding the Samandan Clinic in Pul-e Hesar, interviewed representatives criticised the lack of medicine, claiming that it was sold by doctors on the black market. Both projects received a “mixed” coding. Lastly, the Aqtapa Clinic in contested Qala-e Zal District (Kunduz Province) was described by representatives interviewed as cancelled or indefinitely postponed due to the occupation of the site by an ANA camp, while the SPNA inventory listed the project as planned (May

2017). The number of projects within the health sector is, with 13 cases, relatively low. Therefore, it is difficult to assess whether the relatively low assessments of operationality are idiosyncratic or whether they relate to structural issues within this sector. The complaints related to the allocations of insufficient numbers of qualified staff and the black-market sale of medicine appear to be structural issues, whereas a lacking operationality of completed facilities to the armed conflict, seem to be just bad luck. They could nonetheless explain, to some extent, the relatively low assessments associated with the healthcare later in the statistical analysis.

Selection process and operationality

Lastly, we check how the mode of project selection, i.e. participatory or not, influenced assessments of operationality by CLDC / DDA and CDC representatives we interviewed. Of the 203 projects coded for the survey area, we noted 162 projects selected in a participatory fashion and 39 in a non-participatory way. Comparing the two categories of projects, we note a lower number of projects assessed as fully operational: 72% of assessments of projects selected with local participation are positive, whereas only 54% of projects received a positive assessment out of the group of projects selected without local participation.

Concurrently, the percentage of "mixed" assessments is higher among projects selected without local participation, than among projects selected with local participation (23% as compared to only 6% among the latter). The four non-operational projects selected with local participation (none among projects selected without local participation) are all out of use due to the conflict: the aforementioned clinic in Taliban-controlled Yamgan, a high school in similarly Taliban-controlled Wardooj, a teacher training centre (TTC) close to the frontline with the Taliban in Qala-e Zal District, which has been taken over by the ALP. A similar TTC building in Dasht-i Archi was damaged in fighting and taken over by the district administration. A frustrated CLDC head voiced his anger as follows:

> "Now it is the district administration office [in the TTC building] and it doesn't have any students. … Now it is district administration and the building has no quality and importance [for the district]. The budget was US$ 1.5 million. It should have been spent for building schools in other areas. It is a completely failed project. I don't have good idea of it" (Head of Wazir Ahmad CLDC, interview on 23 August 2017).

A further feature differentiates the assessments of projects selected via local participation from those selected in non-participatory ways. Our key interview partners for the assessment of operationality, CLDC / DDA and CDC representatives, were somewhat less well informed about the projects selected without local participation than about those not selected via participatory procedures (only 12% "not clear" respondents as compared to 21% such respondents). This is not particularly surprising, as elected representatives likely participated in the identification of participatory-selection projects and not in that of projects selected without local participation. They had thus more detailed knowledge about participatory-selection projects.

Given the relatively small number of projects selected without local participation, and their partial concentration in central districts (Taloqan, Pul-i Khumri, Fayzabad), we should not overemphasise these results. Nonetheless, it gives a first indication that participatory project selection might work from a stabilisation perspective in a different way than non-participatory selection procedures. In the statistics section we will further investigate whether differences in project selection are also associated with differences in values for stabilisation indicators.

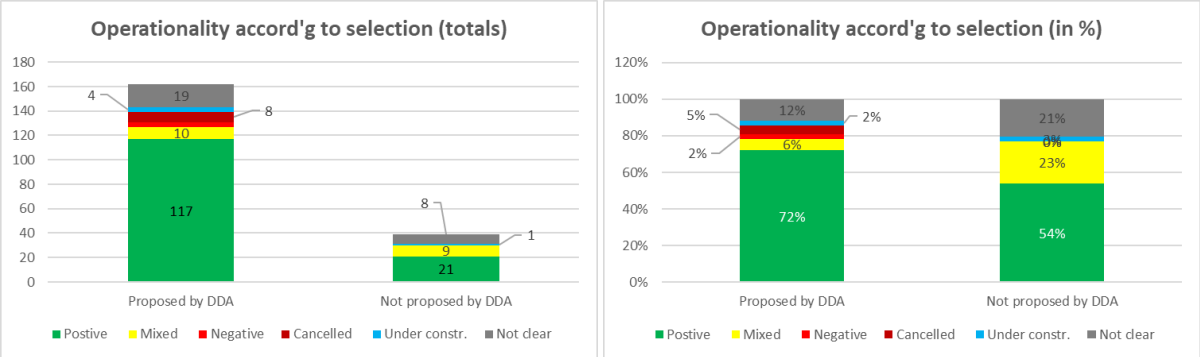


Figure 11: Operationality of SPNA infrastructure projects according to selection mode totals (left) and percentages (right)

IMMEDIATE OUTCOME 2: CAPACITATED DDAs FACILITATING GOVERNANCE

DDA TRAINING AND CAPACITY-BUILDING

SPNA offered a number of training modules to DDAs. A part of these training measures was (a) directly linked to the developmental tasks of DDAs, such as the modules relating to project prioritisation, project cycle management and participatory monitoring and evaluation, while (b) other training modules capacitated DDAs to perform more general governance functions in their districts (e.g. advocacy, gender planning, disaster risk management and conflict resolution). *Table 3* below provides a full overview of trainings and capacity-building measures provided by SPNA to beneficiary DDAs.

Overview of SPNA training modules for DDAs

DDA Level trainings	Advanced Advocacy
	Advanced Gender Mainstreaming and Planning
	Advocacy
	Disaster Risk Reduction (DRR)
	Establishment and Support of Operation and Maintenance Committees
	Participatory Monitoring & Evaluation

Table 6: SPNA training modules

It was, however, not only SPNA, but also a number of further organisations and programmes that provided training to DDAs. The most important was the MRRD's NABDP programme that had actually set up the DDAs. Other organisations and programmes, however, also offered trainings, such as the Afghan Independent Human Rights Commission (AIHRC)[7] or GIZ.

Using our interviews with DDA representatives we can code training providers per district from 2012 until the endline (Waves 2, 3 and 4). In our interviews with DDA representatives we asked whether the DDA had received trainings in the past two years. We coded responses according to whether SPNA was the only training provider, whether the DDA received trainings from SPNA as well as from other organisations or whether only another organisation had provided training. In a number of cases our interview partners did not provide sufficient information with regard to the training provider. These cases were coded as "unclear", though it is almost definite that at least some of these trainings were provided by SPNA.

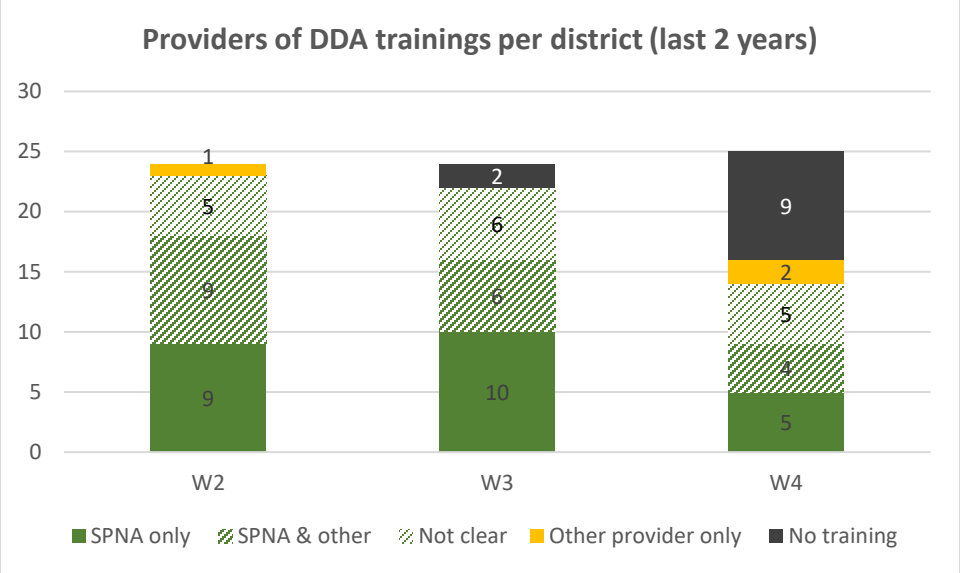


Figure 12: Training providers by Wave

The overview confirms the importance of SPNA as the main training provider in the survey region. In Waves 2 and 3, SPNA was in about two-thirds of districts (76% of districts for Wave 2 and 67% for Wave 3) either the sole provider of DDA trainings or it provided such trainings together with other organisations or programmes. This proportion decreased during Wave 4, but only because a relatively large number of districts had received no trainings whatsoever. Assessing the volume of SPNA-provided trainings as a proportion of all trainings provided to DDAs is difficult. However, a very rough

[7] On topics such as child marriage, or the practice of "*baad*", the exchange of women between households to resolve blood feuds.

estimate suggests than in any wave, SPNA provided anywhere between half to two-thirds or more of all trainings and capacity-building measures to DDAs.[8]

Satisfaction with trainings – not just with SPNA, but with all trainings – was generally very high. DDA respondents typically mentioned how the capacity-building had enabled them to understand and perform their tasks. As the head of the Kalafgan DDA put it:

> "At first, none of the members of the DDA knew anything about the work they had to carry out. After going through the [already] mentioned trainings, the members of the DDA now have good administrative abilities, can prioritise projects, implement them and monitor [their implementation]" (interview on 15 Dec 2012).

The Imam Sahib DDA head voiced similar sentiments:

> "The training workshops were very useful and helpful for our daily activities. Now we have information regarding how to do supervision, prioritization of projects, solving disputes, planning and so on. It has helped our works and our daily activities a lot" (interview on 15 Oct 2012).

We only noted five critical statements out of a total of 73 DDA interviews (three critical comments during Wave 2 and two during Wave 3). Criticism tended to focus on the value content of certain trainings, e.g. gender, or on the fact that some DDA members were too uneducated to properly comprehend training contents. Here are two examples:

> "I do not think these workshops are useful because there is no literacy, knowledge and awareness in the society and people are not ready to accept and enforce these trainings. It is a traditional society" (interview with Deputy Head of Aliabad DDA on 6 Oct 2012).

> "The people were not aware of the rules and regulation of the DDA and the people were also not aware of how to live properly in the society and the CDCs/CLDCs and DDA did not understand their jobs properly. People received information about forced marriage and the people understand the rights of the wife and the husband. However, illiterate individuals have been selected in our CDCs/DDA and they have not learned as much as we expected them to learn" (Farkhar Head of DDA, interview on 5 January 2015).

Moreover, a number of DDAs directly linked their perceived increased in importance in their districts to capacity-building. Here a quote from the Chal DDA to illustrate the point:

> "The DDA has become important. Now the awareness of the DDA members has increased and the members know their responsibilities and their jobs. The DDA has learned a lot about conflict resolution and the prioritisation and implementation of development projects. … There are positive changes in the DDA now, because members of DDA know their job … because now everyone thinks about development" (Interview on 22 Dec 2014).

[8] This rough calculation is based on the following strongly simplifying assumptions: (1) all districts received the same amount of trainings, and (2) in districts that received trainings from both, SPNA and other providers, about half of the trainings were provided by SPNA. To calculate the minimum we simply added the number of districts with "SPNA only" and 50% of the number of districts with "SPNA & other". To calculate the maximum possible proportion of SPNA-provided training as a proportion of all trainings, we further added the number of districts coded as "not clear", assuming that these trainings have been provided by SPNA.

In conclusion, while not the only training provider to DDAs, SPNA was by far the main programme capacitating DDAs. Moreover, the trainings offered to DDAs were generally well received, improved the capacity of DDAs and were then put to use to fulfil the tasks of this body. Therefore, while we cannot isolate trainings provided by SNPA from those provided by other organisations, given the pre-eminence of SPNA as the *main* training provider (proportion ranging from half to two-thirds of trainings in each wave), it is safe to assume that any improvement of DDA capacity is to a significant extent the result of SPNA. This confirms a first link in the causal chain leading from SPNA measures (outputs) to immediate outcomes.

Nonetheless, while we can unequivocally estimate the impact of infrastructure provided by SPNA, as we can precisely identify and geographically locate all such SPNA measures, we cannot fully isolate the impact of the SPNA's soft component from the contribution of other training providers. Therefore, in this regard we can only establish a determining *contribution*.

DDA STRUCTURE

DDAs were set up by the MRRD with the facilitation of an international facilitating partner (FP) like AKF (Afg), ACTED or Mercy Corp. Despite the supposedly uniform structure of DDAs, we noted substantial variation in structure both over time and regionally. According to our interviews, the DDAs of the survey region were all established between 2007 and 2010.

Originally, in order to elect DDAs, districts were subdivided into clusters of CDCs that were roughly proportional to the size of the district. Large districts had more and larger clusters than smaller ones. The number of originally established clusters in the survey region thus ranges from between six to eight in the smallest districts such as Khash or Zebak (with populations reaching 36,000 and 8,667 respectively) to 17 in the largest district of Baghlan-i Jadid with an estimated population of 418,000. Every Cluster Level Development Council (CLDC) sent one male and one female representative to the DDA. In practice, female representatives originally rarely participated. As a result, DDA meetings were mostly meetings of male CLDC representatives with each DDA member being the sole representative of its cluster.

From 2011 onwards, the MRRD (once again supported by FPs) organised a new round of elections, for which they restructured both the cluster system of districts and the structure of the DDA itself. According to the template, the old clusters were dissolved and sometimes 10, sometimes 11 new clusters were established – thus disregarding the actual size of a district. Following the new structure, each cluster now sent two male and one female representative to the DDA. The new DDAs were thus composed of 30-33 members, depending on whether the district happened to have 10 or 11 clusters. Restructuring and new elections were, however, not comprehensively implemented: three of the original 25 survey districts had no second round of elections: Imam Sahib, Ishkamish and Yangi Qala (on Warsaj we have conflicting information whether there were any elections held at all). We are not aware of the reason for omitting these districts from restructuring.

However, we noted not only differences between the structure of DDAs (new versus old structure) and the quality and legitimacy of their elections, but also regarding how their component parts, the CLDCs functioned. While CLDCs have not explicitly benefitted from and were not targeted by SPNA, they are nonetheless essential components of DDAs. As the head of Aliabad DDA, Hajji Haikal put it: "The CLDCs are members of the DDA and if the clusters don't work, the DDA cannot work either" (interview on 14 Feb 2015). Several DDA members voiced similar sentiments.

Clusters have a key two-way transmission role between the village level and the DDA, which its turn interacts with the official government administration at the district level, i.e. the woliswal, the police and the line departments. In this system, CLDCs "forward the problems of the villages [to the DDA], monitor the projects, facilitate the implementation of projects and other issues" (interview with Head of Cluster 5, Aliabad on 16 Feb 2015). Importantly, this view does not just represent the self-conception of members of the shura structure, but is also shared by district-level government representatives, above all by woliswals, as several interviews over three survey waves testify. Thus when asked in interviews where people in villages should turn to if they have a problem, woliswals typically referred to the shura structure.

> "First people should submit their request to the CDC and then the CDC submits the proposal to the CLDC and then [it should be] confirmed by DDA and then it is taken to the woliswali and through woliswali it is referred to the relevant institutions" (interview with woliswal, Baharak Badakhshan, on 27 January 2015).

There are, however, important regional differences with regard to how the CLDCs, the component parts of the DDAs, work. Since Wave 2 we systematically ask cluster-level respondents about their meetings with member CDCs. In Kunduz and Baghlan CLDCs generally considered all component CDCs as members, inviting them to regular large meetings. These meetings often involved 10-15, occasionally even 20 participants. In contrast, Takhar CLDCs usually only considered the elected cluster representatives, i.e. head, deputy and treasurer, as members. Here a quote from Takhar, to exemplify this understanding of cluster membership:

> "The three members who have been introduced to the cluster council have a regular cooperation, but the rest of the councils [CDCs] who don't have membership don't cooperate. For example, when there is distribution of wheat, oil or other…, we gather them, but otherwise they don't cooperate" Rustaq, Head of Cluster 4, W2.

Meetings were accordingly smaller and more informal – often only taking place in the wake of DDA meetings. In some extreme cases, the CLDC head considered himself the only valid representative of the cluster and avoided having CLDC meetings altogether. Badakhshan appears to be located somewhere inbetween inclusive Kunduz and Baghlan and "village elitist" Takhar. On occasion CLDCs elected more management members than originally foreseen or invited more CDC members to their meetings than just the three key members.

We are aware of at least one conflict in Badakhshan between DRRD representatives and CLDCs in Kuran wa Munjan (Badakhshan) regarding cluster membership. Reportedly, DRRD representatives visited the district and instructed the CLDCs to decrease their members to the size foreseen by the official structure (head, deputy and treasurer), hence some CDCs have no membership in their cluster. This caused tensions and people complained about it as some CDCs were now well-represented, while others not (debriefing 2013, Kuran wa Munjan, W2, 3 March 2013).

These differences in cluster organisation are not trivial. Since CLDCs represent the key institutional channel for communities to communicate their problems upwards towards the state administration, a higher or lower degree of inclusiveness will likely have profound effects on how broad the governance-related impact of DDAs is likely to be. Where clusters are more inclusive, CDCs are likely to have better chances in making their voice heard. In contrast, where clusters are more exclusive, many CDCs are likely to have no access to higher levels and are thus less likely to benefit from DDA activities – and ultimately from SPNA.

Since the differences appear on the provincial level, it is unlikely that the degree of inclusiveness is dictated by one of the SPNA FPs, i.e. AKF-A, ACTED or Mercy Corp, as more than one FP is usually active within one province. Rather, it seems to have been the approach of the provincial DRRDs to the tasking of clusters that explains these differences. The conflict we noted for Kuran wa Munjan appears to confirm this impression.

DDA CAPACITY

We coded DDA performance regarding a total of 11 activities associated with the tasking of DDAs. The main source for the coding were our qualitative interviews conducted on the district, cluster and CDC levels. Occasionally, we also consulted additional qualitative material, such as debriefing notes and interviews with SPNA staff, if our qualitative interviews were incomplete or inconclusive on certain aspects. Many of the activities considered in the scoring also formed part of SPNA-training modules for DDAs (e.g. development related activities, gender, conflict resolution, maintenance, etc.).

1. The quality of elections
2. Female participation W2-4
3. Frequency of DDA meetings W2-4
4. Participation in development work W2-4
5. Participation and nature of conflict resolution W2-4
6. Organisation of hashar W2-4
7. Quality of cooperation with the government W2-4
8. Quality of cooperation with the shura structure W2-4
9. Mediating towards the Taliban W2-4
10. Performance of additional governance tasks W3-4 (Wave 2 rather sketchy only)
11. Performance of maintenance tasks W3-4

The 11 activities were expert rated on a score ranging from 1 to 4 (with 1 being the lowest and 4 the highest score). The scores were then simply added up without weighting. The resulting ranking shows quite strong differences between DDAs with the lowest scoring DDA, Warsaj, reaching merely a score of 41, while the top scorer, Aliabad, having 109. We note also important differences in performance between waves within individual DDAs. The most extreme case is remote Kuran wa Munjan, which performed normally during Wave 2 and even quite well during Wave 4 but received only a very low score for Wave 3. The reason is that during 2014/15 the dominant pro-government commander of the district, expelled the district administration; simultaneously he enabled the entire DDA to go on the Hajj. The district was thus temporarily without a district administration and without a DDA.

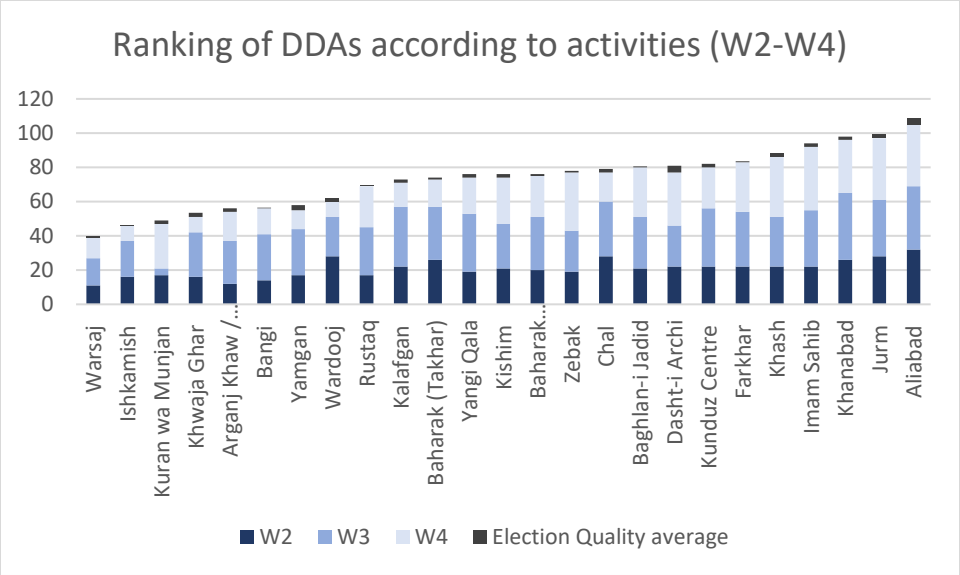


Figure 13: DDA functionality index

Looking at performance according to waves and provinces (we group Kunduz and Baghlan together for this analysis), shows Wave 3 as the peak. After Wave 3 DDA performance dropped sharply in Wave 4 ending close to the level of Wave 2. Disaggregating performance according provinces, shows, however, that most of the drop in performance calculated for the whole sample, actually is driven by Takhar Province where average DDA score dropped from 28 to merely 17 (a drop of 46%). The drop was much less pronounced in Kunduz (from 33 to 31) and Badakhshan (from 25 to 24). The performance of Badakhshan is particularly remarkable as two DDAs of the province, Wardooj and Yamgan, have been disabled by full Taliban occupation of their district, obviously strongly and negatively affecting the provincial average.

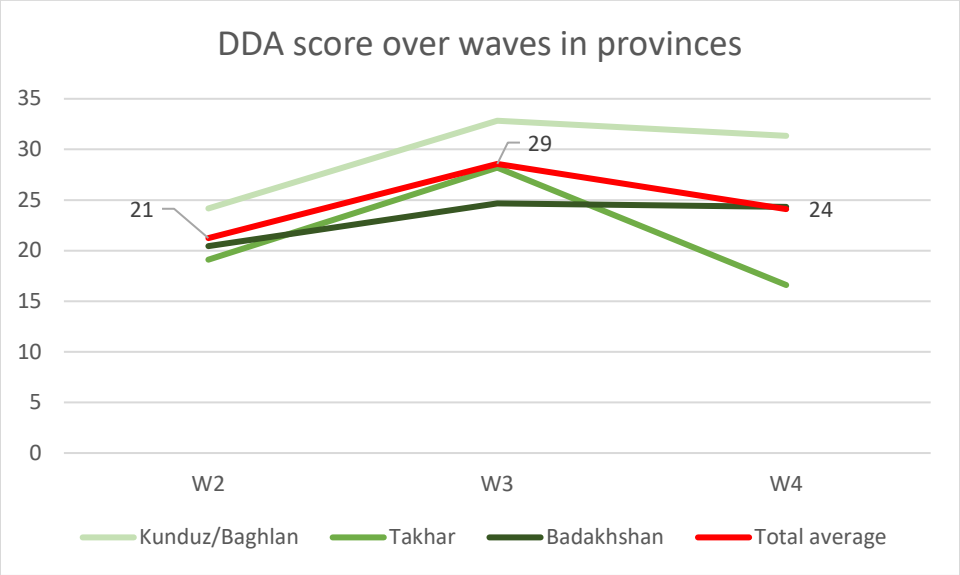


Figure 14: DDA functionality index across Waves

In the following, we will discuss how the individual components of DDA-capacity scoring evolved in the course of the survey.

DDA elections

DDA elections are a key determinant of later DDA capacity. Elections are a crucial feature of the entire shura structure and thus of the procedure defined for establishing DDAs (elected indirectly by cluster representatives). There are two advantages usually associated with elections, which are likely also valid for DDAs. First, elections that follow due procedure convey legitimacy, in the eyes of the electorate, on those elected (see e.g. Luhmann 1983). Second, elections, as opposed to appointments by seniors, make office holders accountable to the electorate and not to superiors. According to theory, they are thus more responsive to the needs of the electorate, while appointed officials seek to please their superiors (for the benefits of decentralisation see e.g. Faguet 2004). Both mechanisms suggest a higher degree of efficacy and accountability of DDAs, whose who were selected as a result of faire elections.

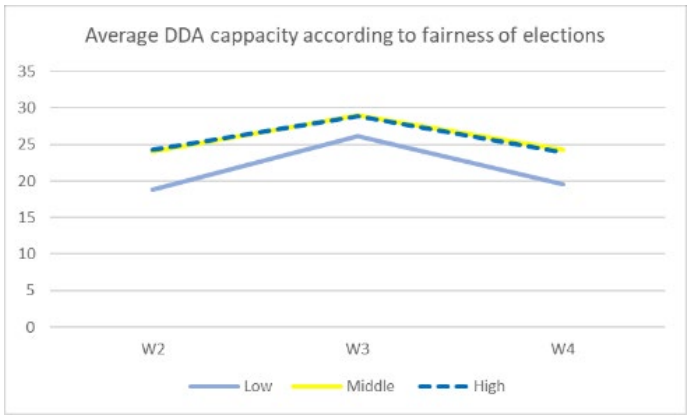


Figure 15: DDA functionality & elections

We have only limited information on the conduct of the first DDA elections between 2007 and 2010. Nonetheless, in at least four cases, local power brokers are said to have interfered with the elections (in Ishkamish, Warsaj, Khwaja Ghar and to some extent in Yamgan), while in three cases our interview partners insisted that elections were fair (Aliabad, Dasht-i Archi and Yangi Qala). In contrast, our knowledge of the procedural fairness of the second round of DDA elections is fairly comprehensive. Of those DDAs that had underwent a restructuring and had renewed elections, we noted allegations of serious interference in five and lesser interference in further three cases, often seriously undermining the legitimacy of the effected DDA and hampering the effectiveness of its work. In order to check whether the fairness of DDA elections had any impact on DDA performance, we grouped DDAs into three sets: those DDAs where the last elections were seriously interfered with, those that were allegedly slightly interfered with and those that had fair elections. We then compared to total average capacity scoring of these three groups of DDAs.

The chart summarises results. It is quite obvious that DDAs with manipulated elections (light blue line) perform in other measures of capacity (the scale on the y-axis: conflict resolution, hashar, cooperation with the government, involvement in development work, etc.) worse than DDAs with clean or only slightly manipulated elections (the dashed blue and yellow lines). In fact, there is no difference between the averages of the latter two groups of districts. Our qualitative knowledge suggests that causal relationship is as follows: a disinterested and incompetent person is elected through manipulation as the head of the DDA. This person is then negligent in the performance of DDA tasks and this negligence manifests itself in low DDA performance (low scores). A prime example is Warsaj, where the DDA was not elected but appointed by the woliswal the protégé of the main political patron of the district patron, the powerful former head of the Independent Civil Service Communication for Merit Based Appointments, at the helm of the DDA. A prime example is Warsaj, where the head of the DDA was not elected but appointed by the woliswal, thus putting a protégé of the main district patron, then holding a powerful position in the bureaucracy at the central state level, at the helm of the DDA. This person showed no interest whatsoever in his job as a DDA head. As a result, the DDA withered under his leadership, ending up as the worst-performing DDA in our sample.

DDA meetings and female participation

A key indicator of DDA activity are DDA meetings. Following their establishment, DDAs began to have meetings – usually to discuss development related problems, plan new projects, follow the implementation of ongoing works, resolve conflicts, organise hashar and to debate the security situation, the quality of public services and similar. According to our interviews, decisions at these meetings are consensual and in some cases by voting, e.g. when prioritising projects. Whatever benefits DDAs provide to their districts, it is likely that they are decided on and coordinated at these meetings.

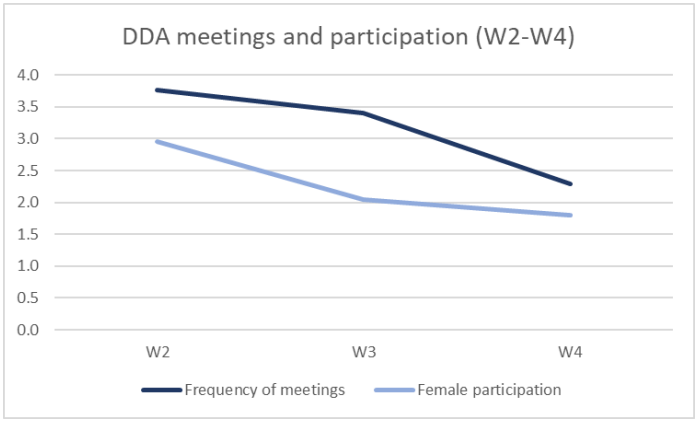


Figure 16: Meetings

A note is necessary regarding the decision-making process. Consensual decision-making has a long tradition in Afghanistan (see e.g. Newell 1972) and was mentioned by qualitative studies of CDCs as the predominant form of decision-making for example during project selection (see e.g. Nixon 2008, p. 31-32). Consensual decision-making does not imply unanimity of opinions or lack of controversy at the outset. Rather it implies a time-consuming process of "repeated meetings to achieve consensus" (ibid. 32). Our finding of consensual decision-making for DDAs thus fits well with ethnographic accounts of traditional communal decision-making in Afghanistan and with qualitative findings on the functioning of lower levels of shura structure, i.e. of CDCs. In contrast, the less time-consuming – but by definition more controversial – process of voting, is usually considered a "Western innovation". Consensual decision-making procedures are so pervasive that even for project selection within the framework of SPNA, several DDAs conducted informal consensus-seeking exercises prior to voting, which was only held once consensus had been reached and which were then unanimous (debriefing of AKF staff on 15 March 2015).

From Wave 2 onwards, we have systematically asked DDAs how often they held meetings and whether female DDA members participated or not. Regarding the frequency of meetings, we gave highest scores for frequent regular meetings (4) and lowest for no meetings in past 12 months (1). Female participation we coded by estimating the number of women usually participating in DDA meetings and then relating this figure to the total number of elected female DDA members. Participation rates below 25% were coded 1, above 75% were coded as 4.

As our chart for the three survey waves shows, the **frequency of meetings** dropped slightly from Wave 2 to 3 and then steeply for Wave 4. Initially, all but three DDAs (Ishkamish, Warsaj and Rustaq) held frequent regular meetings. By Wave 3, eight DDAs held only irregular meetings, two of the original low performers (Rustaq improved its performance in this regard) while six "new" DDAs also became laxer about their meetings. These DDAs are all among the lower performing assemblies of the survey. By Wave 4, nine DDAs stopped having meetings altogether with many of them stating that their DDA was defunct, while six others only held irregular meetings. The reasons for the drop in the frequency of meetings are manifold. Of the 15 defunct or mostly inactive DDAs, nine had finished their last SPNA projects and were not involved in any Phase V activities either. They thus had no compelling incentive

to keep holding regular meetings. The Baharak (Takhar) DDA made the link between lacking projects and the weak performance of the DDA explicit:

> "We have two kinds of meetings: monthly and emergency. We haven't conducted our monthly meetings for more than one year *due to lack of development projects* [our italics], but emergency meetings we have conducted twice when organizations [i.e. INGOs] assembled the members of DDA. … We haven't been financially supported by any source ever since the beginning of national unity government and we don't have any budget either. The reason for the weakness of the council [the DDA] is this problem" (Baharak, Takhar, interview with head of DDA on 2 Oct 2017).

Conflict also seems to play a role. Dasht-i Archi and Wardooj were hotly contested between the government and insurgents[9], making any serious work very difficult. In Kuran wa Munjan, the dominant (pro-government) commander of the district, temporarily expelled practically all government staff from the district and allowed the entire DDA to extract sufficient lapis lazuli from the famous mines of the district to collectively go on the Hajj. As a result, the entire DDA temporarily stopped its work in 2014-15.

In two further cases, however, it was the Taliban occupation that led to a discontinuation of meetings (Wardooj from 2014-15 and Yamgan from 2016 onwards). The Taliban, while supporting the CDCs, seem to be suspicious about the district-wide authority of the DDA and seems to discourage its functioning in districts wholly or mostly under its control (e.g. in Qala-e Zal, Wardooj from 2016 onwards, Yamgan from 2017 onwards). In other cases, however, the Taliban either cooperate with the DDA (e.g. in Khanabad) or take it over (Wardooj from 2012-15, to practically disband it in 2016-17), possibly depending on the degree of control they exercise in a district. Aside of full Taliban control over a district, bad security alone, in the sense of frequent security incidents, does not seem to effect DDA meetings. As such, several DDAs in hotly contested districts continue having regular meetings (e.g. Khanabad, Baghlan-i Jadid, Aliabad, Imam Sahib, Jurm and Baharak-Badakhshan).

Female participation also dropped dramatically between Wave 2 and 4. The main reason for this drop is clearly security. In terms of security, Wave 2 was the most secure. Security was worse during Wave 1 and also deteriorated after Wave 2 (Waves 3 was second worst and in Wave 4 improved security is of a different, less state-associated and less stable quality). As an example, in Jurm (Badakhshan) women stopped participating in DDA meetings during 2014/15 when fighting was particularly intense, and the district centre was at risk of falling to the insurgents. As the Jurm DDA explained:

> "Yes, female representatives used to show up in the meetings before the insecurities. In the past we even took the female representatives several times to the DRRD at the provincial level; and the money for their transportation was also paid by DDA. We used to take them to the provincial level for communication purposes and we had a lot of success with their visits and got several projects approved" (Jurm, Head of DDA, interview on 28 February 2015).

Later, as fighting calmed down in 2016/17, women once again began to participate in DDA meetings, albeit only from government-controlled areas (CLDCs in particular under foreign mujahedin control,

[9] Dasht-i Archi was violently contested for the whole period of the SPNA implementation, i.e. from 2010-2018. Wardooj, in contrast, was relatively stable during baseline (2010), but fell to the Taliban in 2012 (Wave 2), was partly retaken by the Government in 2014-15 and fell once again to the insurgents in 2016 and remained under insurgent control during the endline (2017-18).

cannot participate in the meetings of the DDA). Once Taliban territorial control expands, impediments to female participation increase, as a statement from Yamgan shows.

> "Female members can't participate in the meetings due to insecurity. None of them have been present so far. We had the female members elected as per the requests of the NGOs, just in order to have the female members in our assembly. But none of them have been present in our meetings due to fear from the Taliban; neither are we interested in their presence" (Yamgan, head of DDA, interview on 23 Feb 2015).

Turning to the quality of women's participation in the DDA, our data is insufficient for a systematic assessment and scoring. We nonetheless tentatively discern two contradictory trends. In areas under more or less firm government control, female participation seems to have frequently improved, albeit it frequently remained narrowly restricted to projects that are of direct interest for women (e.g. the monitoring of income generation projects for women). A female DDA member from Kunduz explained:

> [My tasks included] "participating in workshops and seminars of the Kunduz DDA but important issues had been done by male members … including the selection of projects and I didn't have any role regarding that. … Even though I'm an educated person, when I shared my opinions, male members didn't take notice – just as with other uneducated women in the assembly.

Asked about changing attitudes of male DDA members[10], she added:

> "male members [increasingly] agreed with the participation of women in the Kunduz DDA but personally, due to insecurity, I couldn't go to villages to have contact with women, listen to their problems and convey them to the DDA" (interview on 28 December 2019).

This assessment was broadly confirmed in 2012 in an interview with the head of the Kunduz DDA:

> "in the past they [women] were just symbolically members. Now it [women's participation] has increased. It is fifty percent better than in the past. For example, we have tailoring projects for women funded by UNHABITAT, another project is about poultry farming for women; there is also a carpet weaving project for women and men. So when we go there to monitor these projects we also take the female M&E person" (interview on 26 September 2012).

Other women, such as the deputy head of the Aliabad DDA, are more positive about their participation.

> "I work as a representative of women in a CDC. I collect women's problems in the villages and covey them to DDA. For example, I suggested literacy courses, hen farms and small agricultural projects for women to the DDA but they were not executed due to the lack of funds. … The meetings of Ali Abad DDA are provided based on needs and I participate in all of them. [At the meetings] I totally share my opinions and the participants listen to it. [For example] I also had an active role regarding the selection of a drinking water project in the district for digging water wells.

She added regarding changing attitudes:

[10] The precise question was phrased as follows: "Do you think that – maybe as a result of trainings – the male DDA members became more sympathetic towards women's problems and participation in councils?"

> "Yes, exactly [male DDA members became more sympathetic towards women's issues]. They appreciate the participation of women. Male members believe that in the past, before the Karzai Government, they would have been opposed to the participation of women in social activities but today they agree with it and consider it a requirement" (interview on 27 December 2019).

A female DDA from Baghlan-i Jadid, echoed this upbeat assessment:

> "I share my opinions and they are mostly accepted. However, other female DDA members are ashamed of talking amongst men, but there are three other women who are also always active in the discussions of the DDA. For example, I presented the issue of health awareness and legal affairs for women and men in the remote villages of the district".

This female interview partner also sensed an attitude change among male DDA members as a result of trainings:

> "Yes, they [male DDA members] are [more accepting of women]. The acceptance of male members regarding the participation of women and their problems in the DDA is much better in Baghlan-i Jadid District than in other districts. In other districts, women don't have any activity and role in society in general and in [DDA] councils in particular".

The active participation of women in Baghlan-i Jadid's DDA was also confirmed by the Baghlan AKF team debriefed by one of the authors of this report on 1 July 2018. In contrast, in areas under Taliban control – or strongly threatened by the Taliban, this progress is being steadily rolled back. The Taliban are very explicit about their ideological rejection of women's participation in the public sphere. In many areas, the Taliban are even opposed to merely nominally fulfilling the requirement of the Citizen Charter for female representation in the newly elected councils (e.g. in parts of Baghlan-i Jadid and Pul-i-Khumri under Taliban control or in Yamgan or Wardooj, which is fully under the control of the insurgents). It is thus less insecurity as such that limits women's participation, but more the ideological opposition of the Taliban.

Having said that, there is nonetheless variation regarding Taliban attitudes towards women's participation in public life. For example, in most districts of Kunduz the local Taliban leadership agreed to the election of female members of the new Citizen Charter CDCs – as long as these were close family members of elected CDC members. Their objection was concerned less about women executing public functions, and more about unaccompanied women mingling with unrelated men. In contrast, the Taliban leadership in Baghlan-i Jadid and Pul-i-Khumri flatly rejected any female participation in the yet-to-be-elected CDCs on ideological grounds opposing female presence outside of the household. The Taliban reportedly stated that they were "fighting to prevent women from going out" and that it was "against the sharia" and that, therefore, they did not want it (debriefing on 27 April 2018). To the dismay of local communities, the Taliban leadership was thus willing to block the implementation of the Citizen Charter which required the election of female CDC members. In these Taliban held areas of Baghlan-i Jadid girls' schools are also closed, while in many parts of Taliban-held Kunduz several girls' schools continue to function (e.g. Qala-e Zal; debriefing on AKF Kunduz staff on 1 July 2018).

Participation in development work

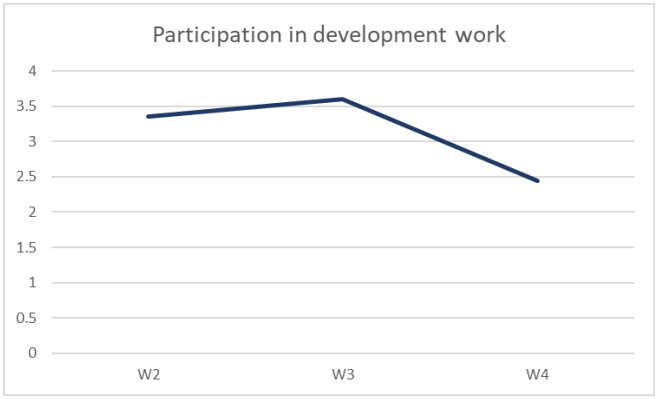


Figure 17: Participation

A key task of DDAs is to participate, prioritise, approve, facilitate the implementation, agree the handover and conduct the maintenance of development projects in a district. This core task was supported by a number of SPNA training modules, such as District Planning, Visioning Building, Establishment and Support of Operation and Maintenance Committees, Participatory Monitoring & Evaluation, Project Cycle Management. In order to investigate DDAs involvement in district-level development activities, from Wave 2 onwards we asked all DDAs whether implementers of development programmes comprehensively contacted and consulted with them and subsequently worked with the DDA. Already at Wave 2, DDAs were generally involved in most development activities in their districts. Importantly, the role of the DDA with regard to development was generally acknowledged by woliswals and other government officials. Here two quotes to illustrate woliswals' understanding and acknowledgment of the DDA's role.

> "When an organization wants to implement a project in district, they should first, either refer to the district governor and or the DDA; then we prioritize the project and seek the consultation of the DDA regarding people's needs … . The projects are implemented by the DDA and our role is more in ensuring security. Mahmood Haidary is the DDA head. He is an educated and active man, he knows his authority and responsibilities very well. He also … cooperates very well with woliswal (Kalafgan woliswal, interview on 4 Jan 2015).

> "The DDA prioritizes the development projects and woliswal confirms it. The only difference is the prioritization of the development projects. The woliswal himself makes the decisions in political, criminal and administrative cases; and the DDA has nothing to do with these issues. It is only the development sector and sometimes social affairs that DDA has a more important role" (Khanabad woliswal, interview on 18 Feb 2015).

Subsequent to Wave 2, DDAs' engagement in development increased further (Wave 3) only to drop markedly in Wave 4. At this time a number of DDAs – mostly those without ongoing SPNA or other projects – stopped functioning. They explained their lacking involvement in development activities by referring to the complete lack of projects.

Conflict resolution

Conflict resolution is a further key task of elders in general and of DDAs in particular. Within the framework of SPNA, DDAs received repeated trainings on the subject (e.g. Conflict Resolution on Common property resource and Advanced Conflict Resolution and Peace Building). While conflict resolution and mediation is common task of Afghan elders, the trainings seem to introduced some new elements to traditional conflict resolution, such the writing down of agreements and their signing by participants and witnesses as well as the archiving of such agreements in DDA books (e.g. interview with Kunduz DDA head on 21 September 2012).

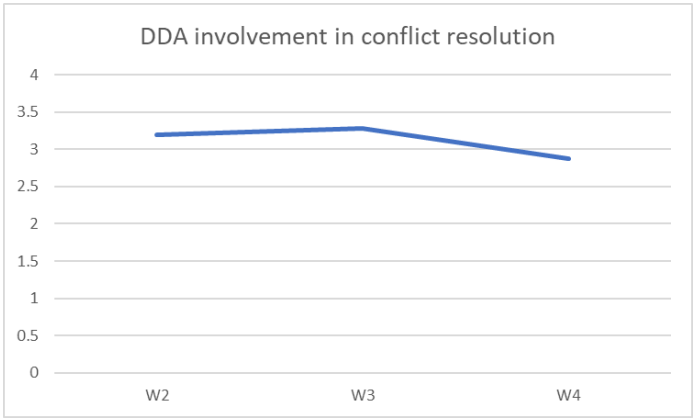


Figure 18: Conflict resolution

Our qualitative research clearly shows that DDAs are engaged in the resolution of a very large number of smaller routine cases as well as nearly all development-related conflicts. There is usually a clear hierarchy of how such cases are dealt with. As conflicts emerge, people usually try resolve conflicts locally under the mediation of the local CDC and of local elders. When this fails, the case is usually submitted to a higher level of the shura structure, mostly to CLDCs – and if their attempts to resolve the conflict also fail, that is when the DDA gets involved. This hierarchy of conflict resolution is acknowledged and supported by state officials, as for example the statement of the Jurm chief of police shows:

> "The local [CDC] shura tries to resolve the conflicts inside the village because it is a small issue among the people. If a conflict is not solved by [CDC] shura, the case will be referred to the head of shura in the centre of the district [i.e. DDA]. The shuras have a positive role in solving conflicts in the area. The shuras do a good work in resolving conflicts and we consider it very lucky that they are so active" (Jurm, Chief of police, interview on 25 September 2017).

More serious cases, murder or grave injuries, are often immediately referred to the DDA (as well as to relevant government officials, most importantly the woliswal and the chief of police) before they further escalate. For example, an interview with the deputy head of the Pul-i-Khumri DDA was once broken off because the DDA was called to mediate urgently following the shootout between two Pashtun household in which five people were killed (field diary, 17 March 2014).

Probably as a result of their gender-related trainings, DDAs occasionally also get involved if e.g. underage girls are being forced to marry. Here a case from Kalafgan:

> "There was another conflict over the selling of a young girl [for marriage]. An addicted person in Chinar Village had sold his daughter for AFN 130,000 to another person from Kunduz. The DDA got involved in the case as soon as it received information about the issue. With the help of the security department, the DDA gathered the involved parties and admonished and advised the girl's father not to sell his daughter for marriage. The father promised that he wouldn't marry his daughter before she was eighteen years old; the buyer [the would-be husband] also regretted and promised to withdraw from his action. Both sides were satisfied" (Kalafgan, head of DDA, interview on 5 January 2015).

In this case and a number of other similar cases we have identified over the years, the DDA represents a more "modern" understanding of traditions, and gender relations. Here three further examples:

- The Chal DDA intervened on behalf of a woman who ran away from her husband (reasons unknown). The DDA supported the right of the woman to divorce her husband and even forced him to pay her money (district debriefing of AKF staff on 23 September 2012).
- In Zebak a Sunni boy and an Ismaili girl married – against the will of their parents – at the district court. The DDA confronted a large crowd protesting against the marriage and argued that the marriage was rightful. The intervention was successful in preventing harm to the couple but could not prevent their expulsion from the district (debriefing of the qualitative survey team on 24 February 2013).
- In Bangi the DDA successfully intervened to prevent the marriage of an orphaned 8-year-old girl to her 8-year-old cousin (debriefing of the qualitative survey team on 26 February 2013).

That DDAs take the gender-related trainings seriously is also suggested by the statement of the head of DDA of Farkhar District, a district lying adjacent to Kalafgan, where one of the recounted events took place:

> "They [the trainings] have been completely beneficial for people's awareness. We got information about "*baad"* [the custom when a killer's family forces one of their daughters to marry one member of the murdered household as a compensation for the killing] and other inappropriate issues like underage brides and women's inheritance rights" (interview on 27 August 2017).

In many districts, DDAs are not just involved in the resolution of interpersonal problems, but also attempt to mediate in a number of very significant and large conflicts that threaten the stability of the entire district. Typical examples relate to conflicts between commanders and armed militia groups. Here an example from remote Kuran wa Munjan District (Badakhshan).

> "The district settles all the main conflict with the help of the DDA. For instance: a conflict broke between irresponsible armed groups lead by [dominant commander of the district; *NAME OMMITTED*] on the one hand and [two lesser commanders; *NAMES OMMITTED*] on the other. It was with the mediation of members of DDA and the authorities of district administration [woliswali]. This is one example of our activities on the level of the district" (Kuran wa Munjan, Head of DDA, interview on 19 April 2018).

In order to grade DDA performance, we coded DDA involvement in conflict resolution with mediation in serious district-level conflicts receiving a 4 and no involvement whatsoever receiving a 1. The chart summarising the scoring shows a high level of DDA involvement for Wave 2, which further increases for Wave 3 and then drops off during Wave 4 as a number of DDAs became defunct or at least reduced their activity. A further reason for the drop in Wave 4 is that with the deteriorating security situation, conflicts with insurgents, powerful militia groups became so serious that DDAs and district elders gave up on attempting to mediate.

Cooperation with the government

Strengthening collaboration between and among institutions within the sub-national governance system is a key objective of SPNA. In line with this objective, DDAs have been trained and encouraged to cooperate with district-level government officials, most importantly the woliswal, the chief of police and with line departments. Relevant training modules in this respect are Linkage with Government and Civil Society, Visioning Building as well as Advocacy and Advanced Advocacy. Most DDAs to a very high degree fulfilled the hopes they would foster good relations with government officials. Many DDA counterparts interviewed in fact used the metaphor of a "bridge" to describe the role of the assembly: "The DDA is like a bridge between the people and the government and organisations [NGOs; development organisations]" (interview with Aliabad DDA representatives, interview on 18 September 2014).

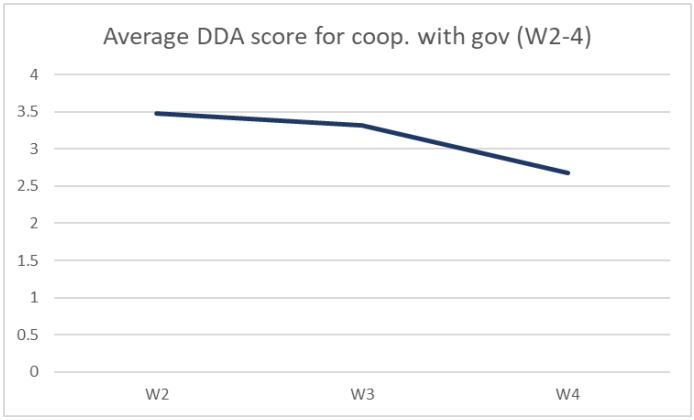


Figure 19: Cooperation with government

This cooperation is very broad, affecting development issues, public services (such as education, health and agriculture), maintenance issues, issues related to district security, conflict resolution and advocacy activities. Woliswals and line-department heads occasionally participate in DDA meetings and vice versa, top DDA representatives are often invited to the meetings of the woliswal or other line departments. Regular cooperation with line departments, for example, includes participation in the distribution of seeds or other agricultural inputs (department of agriculture), support for vaccination campaigns (department of health) or finding accommodation for teachers in remote areas (education department). In all these cases it is the pyramidal structure of the shura structure covering the whole district that makes line departments' interested in cooperating with the DDA. As point of interest, it mostly only the Huquq, a government mediation department under the Ministry of Justice, that is often critical about the DDA – most probably as it sees it as competition in resolving civil conflicts.

The shura structure is, however, also important as a channel of two-way communication. On the one hand, the police or the woliswal frequently use the DDA communicate key messages to the population or even if they want to apprehend or simply talk to a certain person. As the Khwaja Ghar chief of police explained:

> "Yes, we fully cooperate with the DDA. Through the DDA and the CDCs we arrest the criminals and the CDCs introduce the criminals to the district authority. However, according to the law, our fields of work are separate [of the police and the DDA]. Our effort is to maintain security and the work of the DDA/CDCs … is mostly on the field of the development of the district" (chief of police, Khwaja Ghar, 8 February 2015).

On the other hand, the shura structure can also transmit concerns of the population to government officials. How efficient this structure is in terms of mobilisation and information transmission was repeatedly explained to us by our DDA interview partners, e.g. DDA members in Yaftal:

> "Every four-five villages are joined in a cluster. The members of a cluster meet regularly – informally, for example for weddings and other events, and formally for CLDC meetings. So if

there is a DDA meeting, the cluster transfers the information to the heads of its member CDCs. The CDC head then conveys the information to his community. People also ask him. He mostly does this after the evening prayer in the mosque. If there is an important news, in maximum two days, the whole district knows it" (interview on 6 July 2018).

Similarly to the previous DDA activities, we also coded DDA cooperation with the government on a scale of one to four (with 4 representing comprehensive and good cooperation and 1 not cooperation at all). Results show an almost flat curve between Waves 2 and 3 and a sharp drop at Wave 4. The minimal drop in average score at Wave 3 is mostly due to the rather idiosyncratic developments in Kuran wa Munjan in late 2014. As was mentioned before, the dominant commander of this remote, but – due to its lapis lazuli mines – important district, temporarily expelled government officials from the district and facilitated the entire DDA to go on Hajj. Besides that, low cooperation with government institutions during Wave 2 and 3 is mostly due to temporary or longer lasting, often personal, conflicts between the head of the DDA and the woliswal or other key government official.

The sharp drop from Wave 3 to 4 is, however, the result of the *de facto* collapse of a number of DDAs, partly probably as a result of the termination of SPNA and the lack of other district-level development programmes (Ishkamish, Warsaj and Khwaja Ghar) and partly as a result of Taliban occupation (Wardooj and Yamgan).

Cooperation with the shura structure

The structure and functioning of the shura structure was described in detail in previous sections. The DDA, as the apex body of this structure has a key role in making the structure work and actually linking the communal level organisations (CDCs) up with the government. We coded the cooperation with the shura structure using our guideline interviews with DDAs and with cluster heads, who are actually themselves members of the DDA. As previously, we graded DDAs cooperation with its component CLDCs on a scale of one to four (4=comprehensive and good cooperation, 1=no cooperation at all).

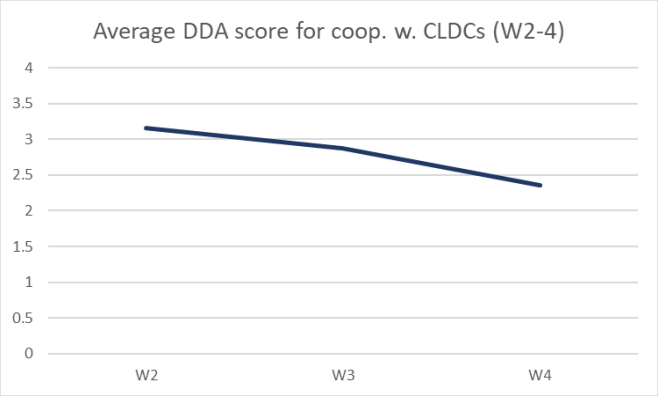


Figure 20: Cooperation with CLDCs

Our results show two districts with fully dysfunctional CLDC structure throughout all waves: Warsaj and Arganj Khaw / Shiwa. The dysfunctionality of the shura structure was comprehensively confirmed through different qualitative research methods and was mainly the result of a fraudulently elected DDA in Warsaj. Concerning the Arganj Khaw / Shiwa, the reasons for the dysfunctionality of the cluster structure are not clear, though low capacities and the eventual Taliban takeover in this very remote district likely have contributed. Aside of that, it was in particular the deteriorating security situation that contributed to dropping DDA cooperation with clusters – affecting in particular districts in Badakhshan. The head of the Baharak DDA, a former commander and mayor of Baharak Town, re-activated himself as a militia commander to fight the Taliban and completely neglected his duties as a DDA head, while his deputy did not have the standing to fully replace him. In conflict-torn Wardooj the head of the DDA sided ever more openly with the Taliban, making him untrustworthy as a neutral

arbiter or go-between for the government. Lastly, the aforementioned temporary absence of the Kuran wa Munjan DDA (the entire assembly went on Hajj) also contributed to lower scores in this province.

Wave 4 saw a continuation of the trend for deteriorating DDA cooperation with the shura structure. The absence of district-level projects (mostly in Takhar) and the de facto ban on DDA activity in Taliban-controlled districts in Badakhshan appear to explain negative results in this wave.

The Insurgents and the DDA

Mediating between the government and the Taliban or representing the interests of the population towards the Taliban is another key security-relevant role many DDAs take upon themselves. There appear to be no directly relevant training modules for this form of mediation apart from the two conflict resolution-related components: Advanced Conflict Resolution and Peace Building Conflict Resolution on Common property resource. It has been common in wartime for elders to try to negotiate with warring parties on behalf of their communities. In this case, however, the DDA represents a formally constituted institution with bottom-up grassroots legitimation, a district-wide reach and training in mediation that takes this task upon itself.

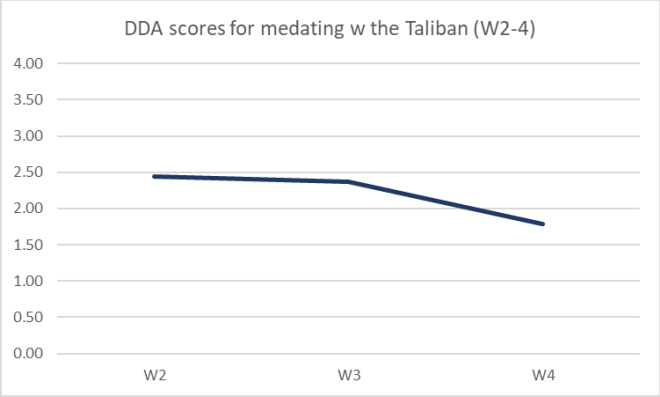


Figure 21: Mediation with Taliban

Some DDAs did indeed achieve remarkable feats regarding mediation with the Taliban. As an example, the Jurm DDA negotiated repeatedly with the local Taliban leadership, for example, to allow for the construction of the SPNA-funded Zair-e Qala suspension bridge in Jurm. The bridge spans the Jurm River, is on the frontline between the Taliban and the government and is essential for several communities living under Taliban control on the eastern banks of the river to reach the Jurm district centre. In order for the project to go ahead, the DDA had to negotiate a limited ceasefire around the project site from both the government troops and the Taliban in 2016. In order to prevent damage to the construction, once the bridge was handed over mid-2017, the DDA negotiated a renewed commitment from both parties not engage in fighting a larger radius around the bridge. As of the time of writing, this ceasefire has held.

We noted similar accounts from several other districts (e.g. Khanabad and Kishim). Negotiations with the Taliban are, however, risky as the Zebak DDA found out:

> "People of this district [Zebak] don't have the ability to negotiate with the Taliban. Recently nine policemen of this district were captured by the Taliban. The councils [DDA and cluster representatives] and we [elders who are not members of the shura structure] went to the Taliban and the Taliban appreciated the request of the people and released those nine policemen. However, the security department [police] and people who were the enemies of the released policemen spread rumours among people of this district that our district had some secret deals with the Taliban and this caused the councils and the elders to be worried. Elders don't [anymore] have the courage to mediate between the Taliban and the government because of the disunity of government" (Zebak, Elder, interview on 10 November 2017).

This latter case shows, how tricky and sensitive such negotiations are. Contacts with the respective other conflict party can lead to allegations of collaboration and partiality and thus put the participating DDA members at risk. This is likely one of the main reasons why DDAs seem to be less and less inclined to mediate between the government and the insurgents. As the insurgents gain ground and the conflict escalates, mediation on their own initiative becomes simply too sensitive and too risky for many DDAs.

It is also possible that – in some areas at least – the Taliban changed its attitude regarding mediation by the DDA. The Khanabad DDA, for example, previously often mediated between the Taliban and the people – and on occasion the government. In this last wave, however, such mediation attempts appear to have stopped. A respected elder of the district explained:

> "No, the DDA and other shuras do not help in such cases [where people have problems with the Taliban] because nobody can stand against the Taliban. They do not listen to anyone" (interview on 8 August 2017).

In contrast, this same DDA continues negotiating if there are conflicts between local militia commanders and the population, suggesting that it is not the DDA that became less active, but the attitude of the Taliban that changed.

> "Yes, the [DDA] shura can also resolve the problems between the commanders and the people. There were some conflicts between the people and local commanders a few months ago. The DDA could resolve it fairly" (ibid).

Moreover, mediation towards the Taliban is usually only possible with local Taliban. Foreign jihadists reject such mediation.

> "The Taliban were behaving in a good way until about a year ago. They used to respect the elders and old men. But now the foreign Taliban and Taliban from other provinces are staying with them and they won't allow anyone to go to them. They say, 'you are the subjects of a traitorous government'" (Yamgan, Elder / community leader, interview on 23 February 2015).

We coded DDA engagement in mediation with the Taliban as before (4=important mediation, 1=no mediation). However, in the course of the waves, several districts had no meaningful Taliban presence: 16 in Wave 2, nine in Wave 3 and only six in Wave 4. The districts without Taliban presence were coded as non-applicable and left out when calculating the average score for a wave. The chart of DDA mediation with the Taliban shows a flat curve between Wave 2 and 3 and a rather sharp drop from Wave 3 to 4. The drop is puzzling as Taliban presence has increased in this period from 14 of the 25 survey districts to 19 (or from 56% to 76%).

Considering geographic patterns of mediation vis-à -vis lack of mediation by DDAs does not produce conclusive results. Mediation is practically non-existent in Takhar. However, Taliban presence is also less prevalent in this province: only 4 out of 10 Takhar districts had significant insurgent presence during Wave 4. In contrast, all other survey districts of all other provinces of the core sample were comprehensively affected during the last survey wave. Moreover, Takhar is also the province with the highest number of dysfunctional DDAs where SPNA implementation had comprehensively stopped after Wave 3. Therefore, the main reason for the absence of DDA mediation vis-à-vis the Taliban seems to be not the local context, but the absence of the Taliban coupled with a less functional DDAs – partly as a result of the discontinuation of SPNA.

The reasons for this surprising drop in the badly needed DDA activity in this respect is likely twofold: the declining functionality of the DDAs and the escalating conflict situation, which makes mediation by the DDA ever more difficult. Moreover, the Taliban seem to be willing to accept the DDA and shura-structure representatives and intermediaries between them and the government in contested districts. However, in districts fully under their control, they tend to de-activate the DDA and CLDCs, only maintaining the CDCs as local-level governance institutions. Even these bodies are monitored by what appear to be local Taliban commissars. We can only speculate about the reasons of the Taliban for not tolerating shura-structure bodies above the level of CDCs. The most likely rationale is that – as a revolutionary movement – the Taliban mistrust any body with a broader clout that is not fully under its control and that could become the leader of potential districtwide dissatisfaction with Taliban rule.

Hashar

The organisation of hashar, unpaid communal work, is a further important task of DDAs. Hashar is a common feature of Afghan (as well as of Central and South Asian) communities traditionally being organised by communal elders and enforcing participation through social control within the face-to-face group of the village. Organising larger-scale hashar, even though often essential (e.g. for the maintenance of roads, district level irrigation canals, disaster relief), poses a problem for this type of face-to-face social control-based mobilisation because the population to be mobilised is larger than the face-to-face group. Therefore, social control is not likely to effectively function. In the recent past, district-level commanders occasionally enforced participation in communal work that went beyond the limits of the village. Nowadays, the DDA and its component CLDCs have become the probably dominant actors organising large-scale hashar. The organisation of hashar was not discussed in any single training module, but was brought up in a number of different modules. Relevant training modules in this regard seem to be District Planning, Establishment and Support of Operation and Maintenance Committees and Disaster Risk Reduction (DRR). To illustrate the capacity of DDAs to organise hashar, here are two examples, one of which shows a hashar organised in cooperation with a government body, the other, a hashar that was based on the sole initiative of the DDA:

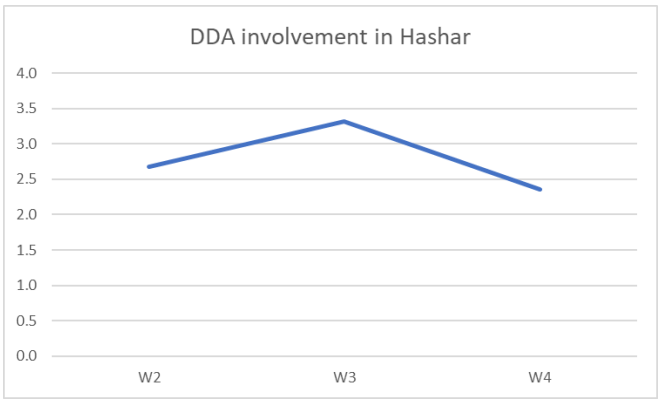


Figure 22: Hashar involvement

> "We do make hashar on the cluster level always. However, one of the biggest hashars was the cleaning of canals about two years ago. The municipality [Imam Sahib] called for the cleaning of the canals of the town and some 500 people participated. Even in the SPNA capacity-building it was in one of our training modules that we should participate in hashar organisation, so we then called on all CDCs to participate. So, the cleaning was initiated by the municipality and the DDA and information about the hashar was disseminated through the shuras [i.e. the shura structure]. I was asked by the municipality if I could provide some manpower. I said 'yes, I have almost 200 CDCs, I can call all of them" (interview with Imam Sahib DDA on 16 September 2012).

> "From here in Shar Shar the road was bad. It is a 15km-long road. We did it. We cleaned it. Ten villages participated in it, some 500-600 people, maybe even a 1,000. It took five days of work. Everybody participated, but it was not mandatory. For Muslims, if we do something for the community, god will be happy. It was an initiative of the DDA. We communicated with the woliswali, then through the clusters we contacted the CDCs. They announced it in the mosques" (interview with the Kalafgan DDA on 14 September 2012).

Frequently, hashar-type activities are essential for disaster relief. Here an example from the high-mountain district of Yamgan:

> "We cooperate in maintaining security for development projects, in project identification and in conflict resolution. More important than anything else: a few days ago, five people were killed and seven others injured in a snow avalanche in Kalafzaar village. We organised people and managed to pull out two dead bodies out of the piles of snow. Three others are still missing. We have been mobilizing and gathering people for hashar. We told from nearby villages to pick up their shovels and tools and help search for the dead bodies. We had sent letters to the development departments including the provincial level to help and support their households. WFP has come forward with offering food supplies which will distribute to them" (Yamgan, DDA, interview on 23 Feb 2015).

Lastly, DDAs do not just organise hashar where people voluntarily provide their labour, but also other collective efforts in support of public services and development in their districts. As the head of the Department of Education of Arganj Khaw / Shiwa explained:

> "Yes, we do cooperate with the DDA and the DDA has cooperation with us. For example, in this year we collected [with the shura] money from the families of students in Wuran Shar and Deh Payan-e Wuran Shar schools for repairing some damages in these schools (Arganj Khaw / Shiwa, Head of Education, interview on 21 August 2017).

Our usual coding shows an increase in hashar organisation by the DDA from Wave 2 to 3 and a drop thereafter to a level below the original Wave 2. The drop is driven in particular by the DDAs that had stopped activities, such as e.g. Bangi, Ishkamish, Chal, Warsaj, Kalafgan, Wardooj or Yamgan. The latter two lie in Badakhshan and are fully occupied by the Taliban, while the former five are located in Takhar, where SPNA implementation had mostly been discontinued.

Maintenance

Maintenance is a key responsibility of DDAs, with two training modules dedicated specifically at enhancing it: DDA Project Maintenance and Sustainability and Establishment and Support of Operation and Maintenance Committees. With the looming phasing out of SPNA, attention to maintenance further increased with Phase V. This phase contained a specific O&M component, with limited funds, organisational support to districts to set up O&M committees and a training module for DDAs and the government

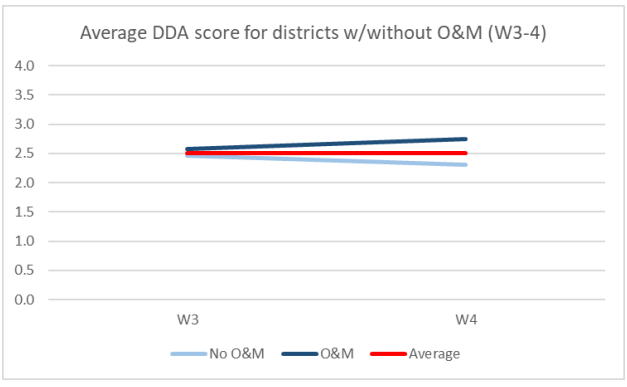


Figure 23: Operation & Maintenance

and non-government members of the O&M committees. Contrary to previous training activities, the O&M component only affected a select set of SPNA districts (12 out of our 25 core survey districts).

The role of the DDA with respect to maintenance ranges from setting up and facilitating complex systems of maintenance, and responsibility for maintenance, via merely morally encouraging maintenance activities, right up to virtually ignoring the problem and leaving it to the state, to CDCs or to nobody. Structured and well-organised maintenance systems – before the implementation of Phase V – relied more strongly on the shura structure with little formal involvement of government structures in monitoring. While sufficient as a provisional measure, relieving the state of all responsibility for maintenance appears to be an unsustainable approach in the mid- to long-term and is also undesirable from a state-building perspective. In the cases we observed, it was usually the monitoring committee or the project-cycle management committee of the DDA that was responsible for maintenance. In some cases, competent DDAs even set up a special maintenance committee. The Kunduz DDA's maintenance system is an example for such a systematic approach:

> "Yes, the DDA council has a maintenance and monitoring committee. They monitor the projects after completion. If a project has a problem, they inform the related sector [line department]. In addition, the heads of CDCs are also responsible to maintain and monitor the development projects in their villages" Kunduz DDA, interview on 2 Apr 2015.

In contrast, Rustaq (Takhar Province) is an example for a less structured - apparently lacking a specialised maintenance committee – less proactive approach that mostly just encourages "people" to participate in maintenance:

> "The DDA is involved [in maintenance] and the DDA performs capacity-building of the people. We encourage the people to maintain the infrastructure projects. We repeatedly discuss in our meetings about the maintenance of the development projects. We motivate the people through CDCs and mosques to maintain completed development projects and they should protect development projects from being damaged" (Rustaq DDA, interview on 30 Dec 2014).

All in all, during Wave 3 we observed 11 districts with more sophisticated and active maintenance setups, a further ten with just rudimentary structures in place and four with mechanisms whatsoever to ensure to maintenance. The most sophisticated and well thought-out maintenance mechanisms were located in the districts of Aliabad, Kunduz Centre, Khanabad, Kalafgan and Chal. These DDAs are among the general top-scorers of Wave 3 and of our assessment in general. In contrast, three of the four districts with no maintenance structures at all, were among the least active and least competent DDAs of Wave 3: Warsaj (the lowest performing DDA of the entire sample), Bangi (located in a troubled and conflicted district and also one of the all-time low-scorers) and Kuran wa Munjan (with the entire DDA temporarily absent during Wave 3).

With Phase V's O&M component things changed. The O&M committees set up under Phase V are independent of the DDA, though six of its 11 or 12 members (there are some differences between the districts) are DDA members. These O&M committees are headed by the executive manager of the district. The O&M committee establishes sub-committees for each SPNA project. As such Phase V O&M is first and foremost a structure for the maintenance of the SPNA projects. While not their main task, O&M committees also consider the maintenance situation of non-SPNA projects. Should the O&M teams establish a maintenance issue, they usually inform the maintenance committee who might contact the relevant line department or an NGO for funds or support or might organise a hashar for maintenance work or collect money for small repairs.

Additional governance tasks

Finally, a number of DDAs also perform a whole set of additional tasks and governance functions that benefit the populations of their districts. All in all, they make life more bearable for people in Afghanistan's poor and conflict-affected districts. These activities are partly at least derived from the Advocacy and Advanced Advocacy modules. For Wave 4 we noted 14 DDAs that had taken up themselves to carry out additional governance activities, while 11 DDAs had not done so. This section only discusses those DDAs that had initiated such additional activities and thus only contains positive examples.

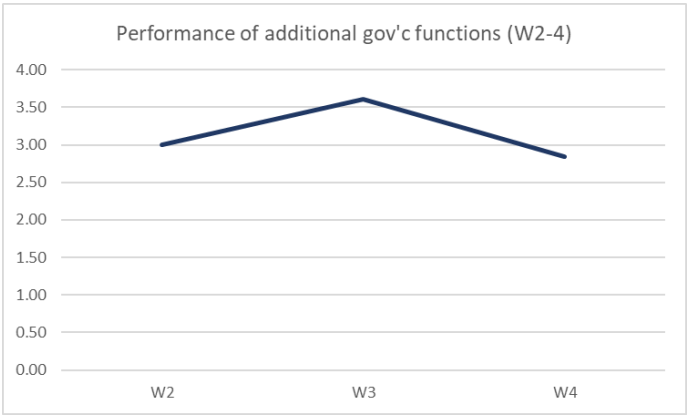


Figure 24: Additional functions

These additional DDA activities include, among others, the field of security:

> "There was an active conflict in our district between the Taliban and the army. The army didn't allow medicaments and health workers to pass to the Taliban area. The CDCs [on the Taliban side] complained to the DDA. We organised a meeting with the woliswal and the woliswal managed to convince the army commander to allow the medicine through to the population on the Taliban side" (Jurm DDA, interview on 6 July 2018).

> "We represent the people. One of our main problem were the arbakees. All the robbers and criminals joined the arbakee forces creating problems for people. We discussed with the CDCs about how to solve this problem and took the issue to the woliswali. The woliswal took this issue to the provincial level. Now all the arbakees were integrated into the ALP which is under government control" (Aliabad DDA, interview on 14 Feb 2015).

Interventions of the DDA also support the delivery public service:

> "The ways are very long for teachers to reach their schools. Therefore, CDCs and CLDCs have the responsibility to find horses for teachers, in case they cannot reach their schools due to long distances. These teachers are women" (Zebak, Treasurer of DDA, interview on 3 Dec 2017).

> "When teachers' salaries aren't paid, we ask the education department about the reason. When the police don't behave well with people, we go together with the effected people to security commander to complain and to solve it" (Baghlan-i Jadid, head of DDA, interview on 11 August 2017).

In other cases, the DDA advocates for a better treatment of the population by government departments:

> "We represent the problems of people towards the government. For example, there were many complaints about the hospital of the district which has 50 beds. The doctors and nurses

> of that hospital did not treat people politely. We advised them to be politer, but they did not change their behaviour. So, we referred the issue to the local government [woliswali] to resolve the problem. Finally, they are now changing their behaviour" (Imam Sahib, DDA interview, W4, interview on 5 Aug 2017).

Several DDAs also get involved in cases of abuse of power by the police or the army:

> "About ten minutes ago the police had arrested a guy who had criticized the security commander on Facebook. I have to go and see the commander and talk to him about this, once you finish your interview with me" (Baharak DDA, W3, 27 January 2015).

Support for the poor in times of need is a further field of activity:

> "About four years ago, there was a very heavy winter. It snowed for forty-five days and the poor people did not have food to eat. We could ask the merchants and traders of the bazaar for help and managed to get rice, wheat, oil, tea and basic supplies and distributed it among the poor. We actually helped about 500 needy families at that time" (Aliabad DDA, interview on 14 Feb 2015).

Lastly, the DDA and the shura structure in general also play an important role in confirming the identity of individuals applying for personal documents and certificates.

Our coding for these additional tasks shows the same pattern as for previously examined DDA activities. The DDA score rises from Wave 2 to 3 and drops again rather sharply for Wave 4. The reason is, as regarding other activities, the reduced activity of many DDAs.

How lasting are DDAs?

During the previous discussions, we repeatedly spoke about inactive DDAs and tentatively suggested that the termination of SPNA could be a key driver of DDA inactivity. In this final qualitative section, we check how long DDAs in a given district were active in implementing infrastructure projects. This perspective differs somewhat from an approach that only looks at phases of implementation, as project implementation in some districts was a rather lengthy procedure, while in others, it went quite fast. The advantages of looking at project implementation, as opposed to just SPNA-phases, is that project implementation on occasion lasted much longer than the actual project phase. While project implementation lasted, the DDA had an incentive to meet regularly. These regular meetings – while mainly focusing on project implementation – also provided opportunities for DDA members to discuss other urgent governance issues, such as important district-wide hashar activities, conflict resolution, or relevant campaigning or advocacy activities. Once the implementation finished, the incentive for regular meetings ended, and opportunities for discussing these "additional" governance tasks decrease.

Figure 20 below shows the length of project implementation – from the start of the first SPNA-funded project to the end of the last project. Project implementation in the core 25 districts is depicted in dark grey while implementation in the ten new districts is light grey. The dashed vertical lines depict the approximate timing of the different survey waves.

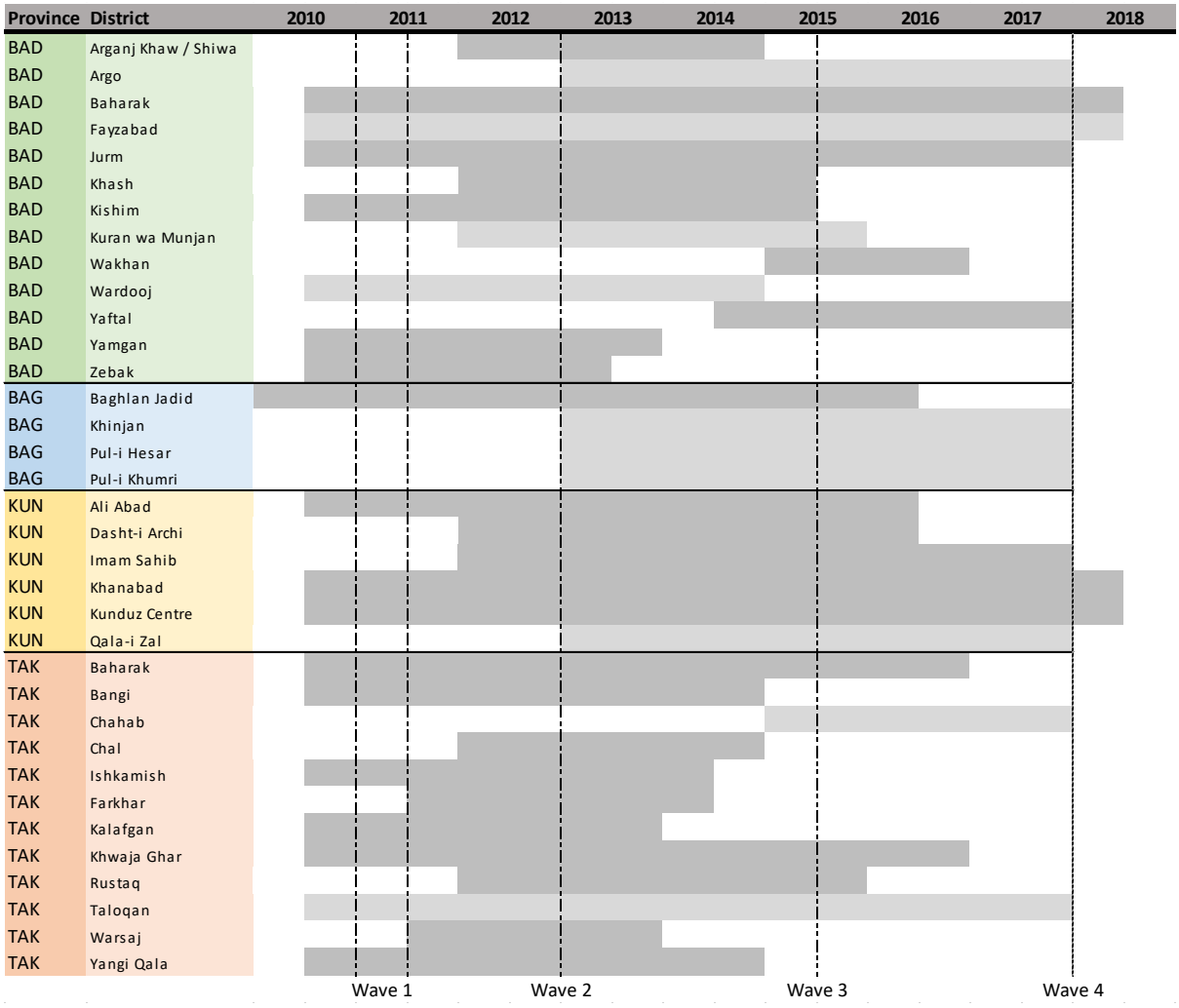


Figure 25: SPNA project implementations

We argue that project implementation provides a rationale for DDAs to meet regularly. During these meetings, DDAs do not just discuss issues related to project implementation, but also decide about other governance-related tasks. An ongoing project thus keeps the DDA active. *Figure 20* above shows considerable variance with regard to when SPNA project implementation terminates. In a second step we look at how DDA inactivity relates to the status of implementation in a district. For this analysis we also consider the implementation of the O&M component of SPNA (see *Chart 21*). The chart fairly unequivocally shows more defunct DDAs where SPNA had ended and more active DDAs where either full project implementation or at least the O&M is ongoing. These findings thus tentatively confirm a link between SPNA implementation and the degree of activity of DDAs.

The link between the lack of projects and the reduction or complete suspension of meetings was partly confirmed by two DDA interview partners, e.g. Ishkamish and Baharak (Takhar). The lack of projects is, however, not the only reason for failing to meet. Insecurity was also mentioned by several DDA respondents (e.g. Bangi, Yangi Qala), the lack of initiative and the incompetence of the DDA head (Khwaja Ghar) or Taliban occupation (from the Munjan section of Kuran wa Munjan under the Taliban) and the de facto disbandment of the DDA by the Taliban (Wardooj, Yamgan). In the case of Yamgan and Wardooj, the lack of projects might partly also be the result of Taliban occupation, which also led to the *de facto* disbandment of the DDAs.

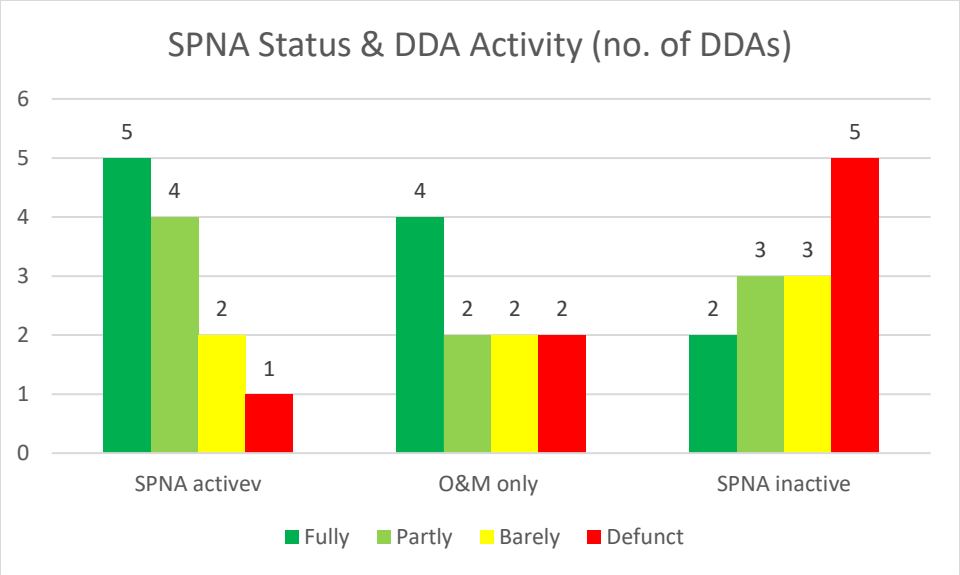


Figure 26: SPNA status and DDA activity

At the end of this section regarding *Immediate Outcome 2* (capacitated DDAs) we sum up our findings. We observe that SPNA was the largest provider of capacity-building in the survey region and that the trainings provided to DDAs were well appreciated and assessed as useful by these DDAs. Therefore, the capacity increase of DDAs is, to a large extent, very likely the result of SPNA's interventions.

Examining DDAs' actual activities, we find that these councils, indeed, started to work in line with their MRRD-provided mandate and, as sporadic evidence suggests, along the lines taught by the training modules. For example, DDAs began to internalise and act upon the contents of women's rights related trainings. As such, **DDAs were active in the field of development, maintenance, conflict resolution, security and via their usually good contacts with government institution intervened on behalf of their constituencies**.

We also find that DDAs varied very strongly in terms of the actual quality of their work and the degree to which they actually fulfilled their mandated tasks. This is visible in the differences in total scores for the different DDAs. Furthermore, we find characteristic differences between the three waves captured by this analysis. Wave 2 starts with relatively high scores for DDAs. Scores further improve somewhat for Wave 3, only to drop sharply for Wave 4. The drop in Wave 4 appears to be partly the disablement of some DDAs under full Taliban occupation. The more important reason, however, seems to be the phasing out of SPNA. DDAs that stopped implementing SPNA-related infrastructure, frequently became inactive during Wave 4. This observation provides further **circumstantial evidence that SPNA was indeed instrumental in motivating and capacitating DDAs to perform developmental and governance-related tasks in their districts**.

For the further impact assessment, we thus conclude that DDAs became active and performed their activities to a large extent as a result of SPNA's interventions. **Based on DDA activities observed, we are justified in expecting to see results across the four fields of our working definition of stability: security, governance, economy/development and modernisation.**

QUANTITATIVE INDICATORS AND MAIN DYNAMICS

In this section we present the patterns and dynamics of three different kinds of indicators.

First, we present indicators pertaining to the four fields of stability (security, governance, economy & development, and modernisation) at the time of the endline assessment. We start the discussion of each field by presenting non-perception-based data on the wider context of stability in that field (among them incident-based security trends, shifts in territorial control over time, general development activities, size of district bazaars).

We *then* look at survey-based indicators for each field of stabilisation. These indicators are the dependent variables of our statistical models, i.e. we want to know if SPNA has had the intended impact on those variables. We use the full dataset of 35 districts assessed in 2017-18 and compare the indicators across the four provinces covered by the assessment. *At the end* of this section we present a trend analysis, for which we limit our sample to the 25 districts covered since the baseline in 2010/11. We discuss (a) how each indicator changed between the baseline, (b) the two follow-ups and (c) the endline assessment, again differentiating trends by province.

Following the discussion of indicators reflecting stabilisation across the four fields, we turn to indicators pertaining to SPNA: our independent predictor variables for SPNA outcomes in terms of its capacity-building as well as infrastructure component. Again, we first look at the state of SPNA outcomes at the time of the endline, then at the trends since the baseline.

INDICATORS / DEPENDENT VARIABLES AND MAIN DYNAMICS IN THE FOUR FIELDS OF STABILISATION

In this section we describe the **main trends** pertaining to the four fields of stabilisation (security, governance, economy/development and adaptation/modernisation) as outlined in the **working definition of stability** (see section *Defining Stability* above). For each field we first (a) offer an impact hypothesis (see textboxes below), then (b) describe the main trends via "**context indicators"** and subsequently (c) turn to the discussion of what can be termed as "**stabilisation indicators"** – i.e. indicators that SPNA seeks to positively influence. The context indicators are mostly non-survey-based (and thus non-perception-based), while the outcome indicators are exclusively survey-based. In our statistical models presented in later sections (sections *Summary of results of the* DESCRIPTIVE STATISTICS

The descriptive statistics presented above provide evidence of how the different dimensions of stabilisation trended since SPNA took off, how immediate SPNA outcomes developed over time and how the modes of delivery of SPNA evolved. The most important observations with implications for SPNA's impact on stability are summarised below.

Firstly, "objective" (incident-based) **security** improves towards the endline, but the quality of security changes. In contrast to the earlier improvement observed in 2012, when state and international forces managed to push pack the armed insurgency, the number of incidents in 2017 dropped as a result of prior gains of the insurgency and government inactivity. The result is a stalemate. Under these conditions, subjective security perceptions and trust in the state to provide security in the future on average across the sample deteriorated further (for differences between provinces see detailed discussion above).

Secondly, we note a deterioration across most stabilisation indicators in the field of **governance**. In particular, the indicators for state presence, state legitimacy and official state power stagnate or worsen. The decline towards the endline assessment reflects a trend towards state retreat and government fragmentation.

Thirdly, subjective indicators of general **development** activities decline towards the endline. Our "objective" (project count at village level) indicator trends vary widely across the provinces. In comparison to the baseline, they have declined in Kunduz, stagnated in Takhar, but increased in Baghlan and particularly in Badakhshan. Compared to the baseline, the district, village and household **economy** indicators have either increased (cars, self-assessed material well-being, size of district bazaars), or, after an initial deterioration, recovered to baseline level (access to land). Compared to the baseline values mechanisation of agriculture (tractors) stagnate on average.

Fourthly, in the field of **adaptation/modernisation** we find a decrease in the ratio of public schools as compared to religious schools. Perceptions of international development aid as a value threat improved towards the endline (less threat).

Fifthly, for **immediate SPNA outcomes** we note an increase in DDA visibility towards the endline, while the assessment of the DDAs' contribution to district development declined. Regarding the infrastructure component, visibility of projects declined slightly over the last years of the SPNA lifespan, while the evaluation of the utility for the district remained very high (close to 99% positive responses) but declined slightly for the household utility question.

Lastly, at the time of the endline, most infrastructure projects financed by SPNA were being **maintained** to some degree (reportedly between 80% in Takhar and 100% in Baghlan). Most projects had been identified in a **participatory** manner, i.e. with the involvement of the DDA (above 80% in all provinces but Baghlan, which is the exception here with just under 20% participatory project identification). In terms of the functionality of the **CDC-shura structure**, upon which DDAs are founded, Kunduz, Baghlan and Badakhshan score between 70% and 80% on the functionality index; here Takhar is the exception with a score just below 50%.

Thus far we looked at these trends for stability, SPNA outcomes and modes of delivery separately. It is the task of the following analytical section to enquire as to how the different dimensions of stability, SPNA outcomes and modes of delivery relate to each other. In more formal terms, we will use statistical tests to verify the hypothesis that immediate SPNA outcomes and modes of delivery have a significant effect on the different dimensions of stability identified in the models.

Analysis: Stabilising effects of SPNA), the context indicators will be treated as **control variables**, while the outcome indicators will be our **dependent variables**.

The introduction of the context indicators concludes with a critical discussion of the implications of the concept for SPNA's theory of change. Here, we also highlight any relevant changes in the meaning and relevancy of the indicator across the observation period of eight years of SPNA implementation.

At the end of each of the discussion of each field of stabilisation, we relate patterns observed and trends in the descriptive statistics of context and stabilisation indicators back to the SPNA impact hypothesis of each field.

SECURITY

SPNA's impact hypothesis: We assume that the SPNA programme components infrastructure and capacity-building reduce subjective fear levels and increase confidence in future security provision by the state because they increase visibility of, cooperation with and standing of the official institutions in the eyes of the communities.

Context indicators: incident counts

Incident counts, gathered from a number of different sources (see textbox below for a critical discussion of the data sources) constitute our first context indicator in the field of security. Incidents are coded from the perspective of the affected populations as *negative* (attacks, explosions, assassinations), *neutral* (e.g. security operations) or *positive* (arrests, surrenders, release of hostages) and are disaggregated according to province. Two features, in particular, are noteworthy with regard to security incident trends. First, Kunduz has experienced by far the most violence in recent years (total negative incidents peaking at 531 in 2015), followed by Baghlan (322 negative incidents in 2015), Takhar (190 negative incidents in 2010) and Badakhshan (170 negative incidents in 2015).

The second remarkable feature relates to the two peaks in violence. The first at the time of the initial escalation of hostilities in 2010 (Kunduz, Baghlan, Takhar) and 2011 (Badakhshan). This first climax is followed by a substantial de-escalation in 2011-12 in the first three provinces and with a time lag in Badakhshan, too. From 2013-15 we note a renewed escalation of violence in the wake of the withdrawal of foreign troops and the fall of Kunduz, which peaked in 2015. Following 2015, we see a renewed de-escalation, albeit under a dramatically changed equilibrium. While the drop in violence in 2011-12 was a result of the US surge and signalled an expansion of government control, the de-escalation of 2016-17 is characterised by a withdrawal of the state and the expansion of insurgent control. (For more on the political, security and governance trends during the survey period, see the context section "*Main developments in the survey region during the survey period*").

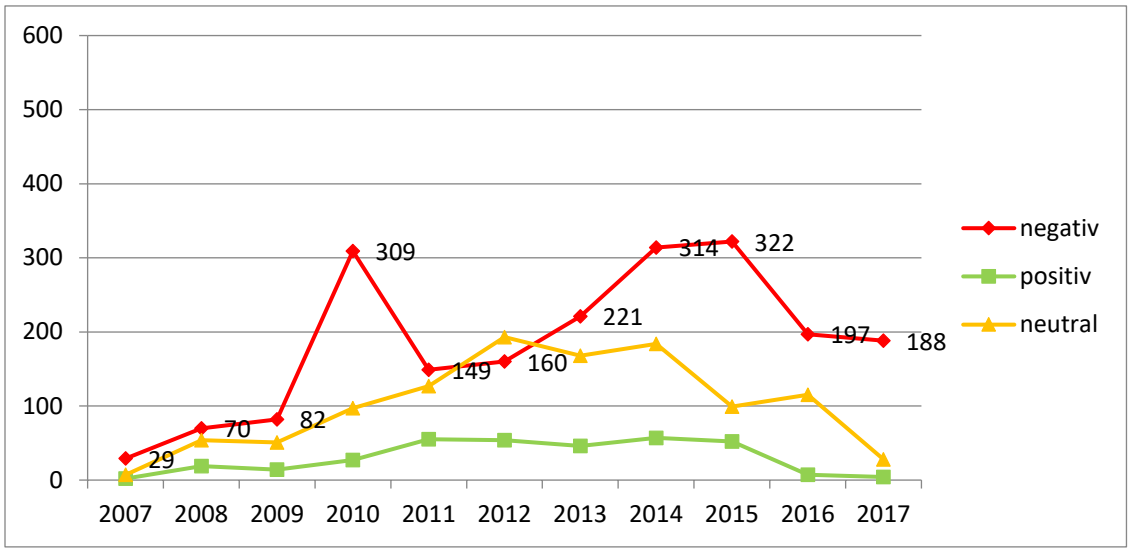


Figure 27: Baghlan security incidents

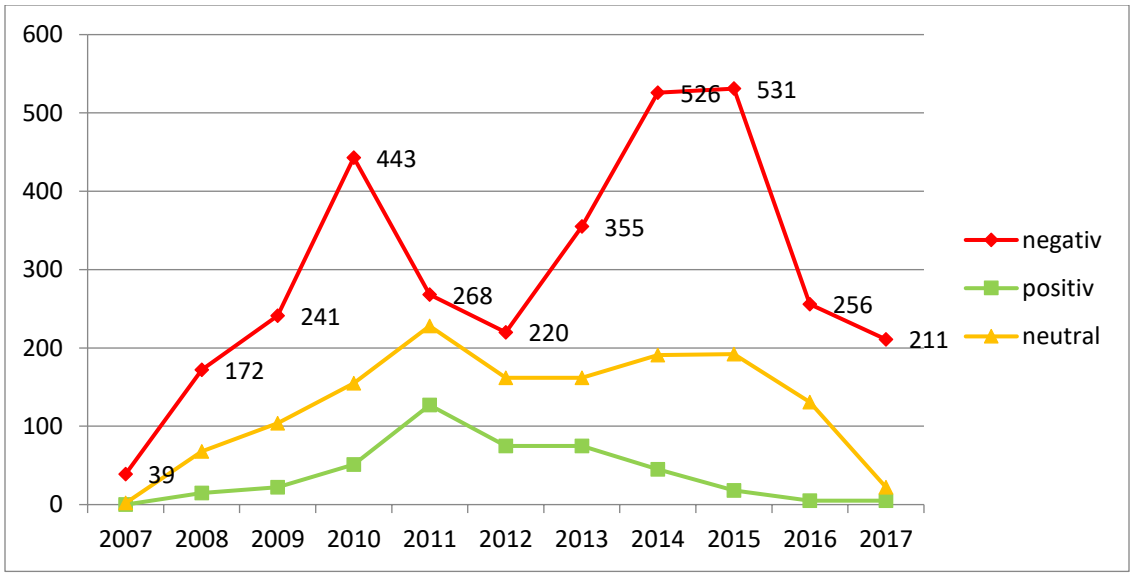


Figure 28: Kunduz security incidents

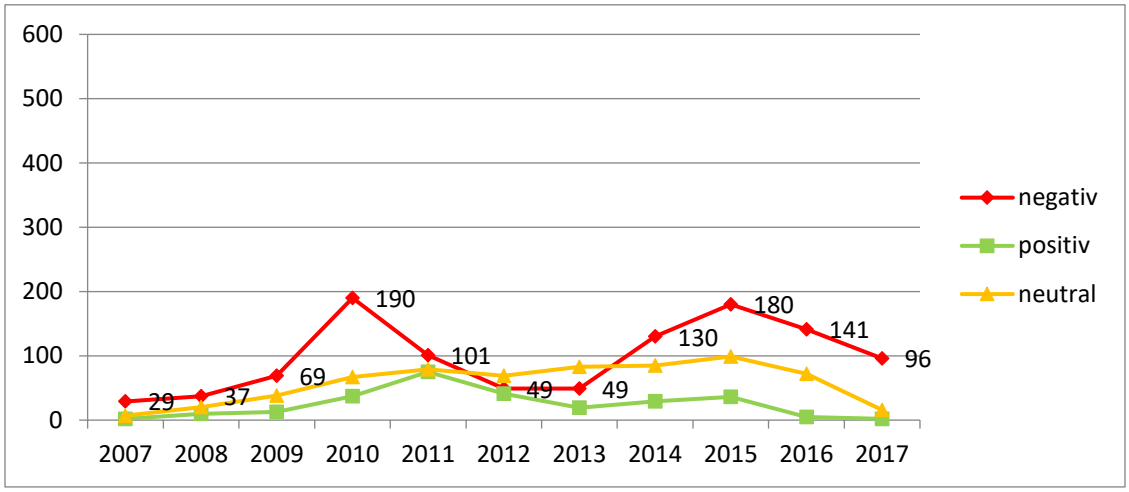


Figure 29: Takhar security incidents

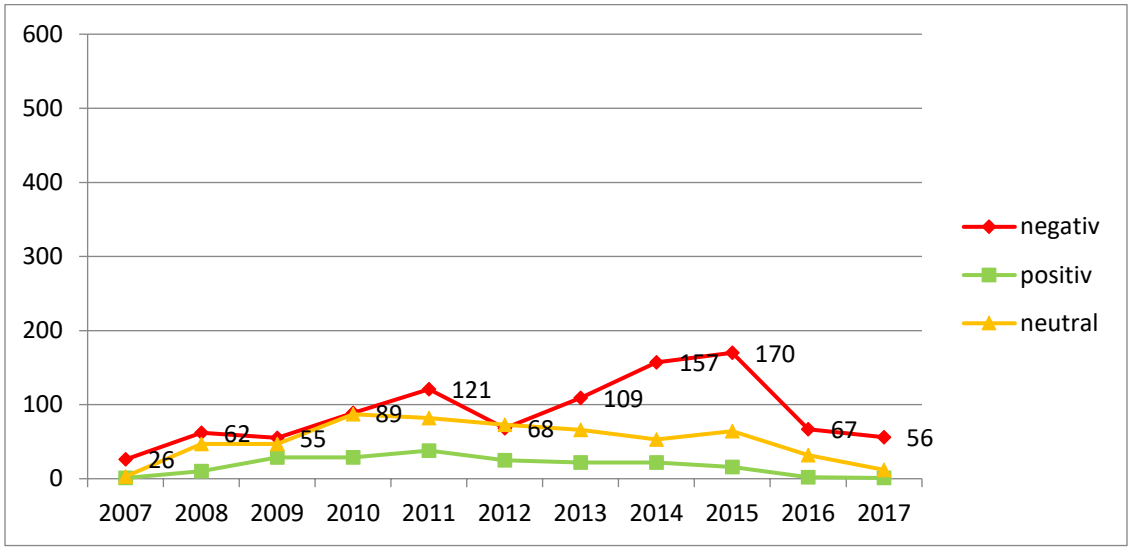


Figure 30: Badakhshan security incidents

Critical discussion of the indicator: Georeferenced security incidents are a widely used, but problematic approximation of spatial physical insecurity. They are problematic because of a range of issues:

- Definition, types of incidents. Different agencies consider different kinds of incidents relevant, hence incident records of UN agencies, the foreign military, Afghan security agencies or development organisations and NGOs are usually quite different from each other. Put simply, international security forces tend to highlight insurgency-related incidents, UN agencies political, human rights-related and incidents involving civilians and risk management units of NGOs and development agencies focus on those incidents that may affect development work, including criminal incidents.
- Measurement bias. The "sensors" recording incidents that are of relevance to the different interested agencies are usually not spread evenly across the area of observation. This may lead to an exaggeration of incident numbers in areas with high agency presence (typically well-connected administrative centres and the lines of communications between them) and underreporting of incidents occurring in remoter and less-connected areas. This is less problematic for insurgency-related or political incidents which tend to cluster around strategic locations; but common violence, clashes between communities or acts of local oppressive powerbrokers may remain unreported.
- Extent, quality, accuracy. A simple count of georeferenced incidents is only a rough approximation of local physical security. A targeted IED attack against foreign forces has, from the perspective of the villages located close by, different implications than prolonged security operations, an indiscriminate suicide attack or the targeted killings of local government employees by the Taliban. The geography of insecurity and quality of insecurity for the local population is different in each case. Furthermore, georeferenced incident lists are often not very accurate below the level of districts. We noted a tendency with various incident lists to place unidentified locations either in the geographical or administrative centre of a district or province.
- Data consistency. Agencies collecting incident data may introduce changes in the method of data collection, and they may add or lose data sources over time. Such changes can affect the data structure in a way that is not related to real changes on the ground, and may, hence, be misinterpreted both in qualitative assessments as well as in statistical modelling.

For this impact assessment, we were granted access to IMMAP's comprehensive incident database. IMMAP uses a range of sources that were consistent for most of the observation period. However, in 2014 one important source stopped providing data and we noted a data gap between 1 July 2014 and 30 September 2015, when IMMAP had reorganised their assessment routine with a replacement source. As we have done in the past, we compared IMMAP data for our target districts with other sources available to us (A/INSO, RMO, open sources, declassified military sources) for data accuracy and general trend confidence. We supplemented the missing data with those other sources. While this approach is not ideal from the viewpoint of data consistency, we are nonetheless confident that it captures the general trend in incident-based physical security in a credible and reliable way. Having said this, we cannot exclude that the drop of incidents in 2016 may be partially at least due to switching from a key international to a national data source on the side of IMMAP.

For statistical modelling in the analytical section (*Analysis: Stabilising effects of SPNA*) below, we use negative incidents occurring in the immediate vicinity of our assessment villages (3 km radius) over a period of two years; we want to be confident that the incidents are relevant to the communities

assessed. The risk of measurement bias in this part of the analysis is mitigated by the fact that in the cross-sectional analysis we are controlling for districts, i.e. we are only comparing differences between respondents of one and the same district. Since in most cases measurement bias would affect whole districts (remote or well-connected) this bias would be minimal. In the panel analysis we are comparing changes between villages and districts over time, so we would not expect measurement bias to affect our results for this control variable.

Stabilisation indicators

Against this backdrop of objective changes, we select two indicators reflecting subjective perceptions of security in the districts where SPNA has been active: (1) the mean of responses to the question of how respondents assess the security of their household and of the district on a 4-point Likert scale from "not secure at all" to "very secure"; (2) Respondents' agreement with the statement that the Afghan state will be strong enough to deal with the security issues after the departure of foreign forces (question only asked since Wave 2 – when the planned departure became evident).

Subjective household and district security: In Wave 4, a majority of households in Baghlan and Badakhshan assessed the local security situation as very or rather negative, while responses in Kunduz and Takhar are much more positive. Subjective security does not necessarily reflect objective security or security in absolute terms. Looking at the incident graphs, Kunduz province has been consistently the most insecure province in our sample. However, while "objective" security assessed via incidents has dropped strongly (i.e. become more 'secure') across provinces since the recent peak of 2015, this drop is steepest in Kunduz. For security perceptions in that province, this is especially relevant after the shock Taliban takeover of the city twice, first in 2015 and then again in 2016. This in part explains how Kunduz turned from being the province with the worst subjective security scores from Wave 1 to Wave 3, to scoring nearly on par with the consistently subjectively least insecure province Takhar.

However, we believe that there are two more forces at work explaining why Kunduz trends different from the other three provinces in terms of subjective security. Both relate to the increasing control exercised by the Taliban and to the way the state has reacted to this. Before Kunduz fell to the Taliban in September 2015, the province had been the epicentre of insurgency as well as counter insurgency (COIN) activities in the North. Part of the COIN strategy after 2009 re-introduced militias and empowered local commander structures that had never fully vanished in Kunduz. In some districts – Khanabad, Kunduz, Dashti Archi and to a lesser extent also Aliabad – the population experienced these militias more as arbitrary forces of insecurity than as security providers. The fall of Kunduz drastically showed the weakness, incompetence and lack of local backing of/for the militias. The Taliban managed to push the militias back from many areas they used to control and established a more reliable, if oppressive local order. In 2016/17 the increasing territorial control of the hinterland by the Taliban has not been challenged in any meaningful way by the state or by the morally discredited and organisationally broken militias. The debriefing of our field teams confirmed that in Kunduz the frontlines are, for the time being, settled and there are few areas actively contested.

The second reason relates to the specific ethnic makeup of Kunduz as an ethnically diverse province with a high share of Pashtun residents. Across the waves, Pashtuns are more likely to be afraid of militias and much less likely to be afraid of the Taliban. Hence, the pushback of militias and increase of Taliban control adds up to a positive security perception within this section of the population.

Finally, the debriefing produced some evidence that in Taliban-controlled areas we have a stronger-than-usual expected response bias with regard to some questions that people may have seen as sensitive from the perspective of the Taliban rulers. Security was among those questions likely affected (see above, *Fieldwork*).

In Baghlan, a province with a similarly high share of Pashtuns but otherwise less ethnically diverse than Kunduz, the standoff between Taliban, militias and the state did not stabilise in a way similar to Kunduz. This is particularly the case for the two central and contested districts of Baghlani Jadid and Pul-i-Khumri. Additionally, the district of Pul-e Hesar is almost paralysed by internal, non-insurgency related intercommunal conflicts (feuding that escalated beyond social control mechanisms). Hence, security perceptions there remain very low.

Badakhshan and Takhar used to have high levels of objective as well as subjective security (Wave 1). Security in Badakhshan, however, has been deteriorating over the waves while – with the exception of the northern districts along the Tajik border – the downward trend for Takhar remained less pronounced. Badakhshan insecurity relates to inroads made by local and later also external Taliban in districts either disappointed with (e.g. Wardooj) or distant from (e.g. Yamgan) the provincial government. In contrast, insecurity in Takhar is less related to the civil war-like conflict between insurgents and the state, and more to criminal issues, ethnic tensions and, most importantly, the long-standing absence of functioning state institutions and a history of locally powerful commanders.

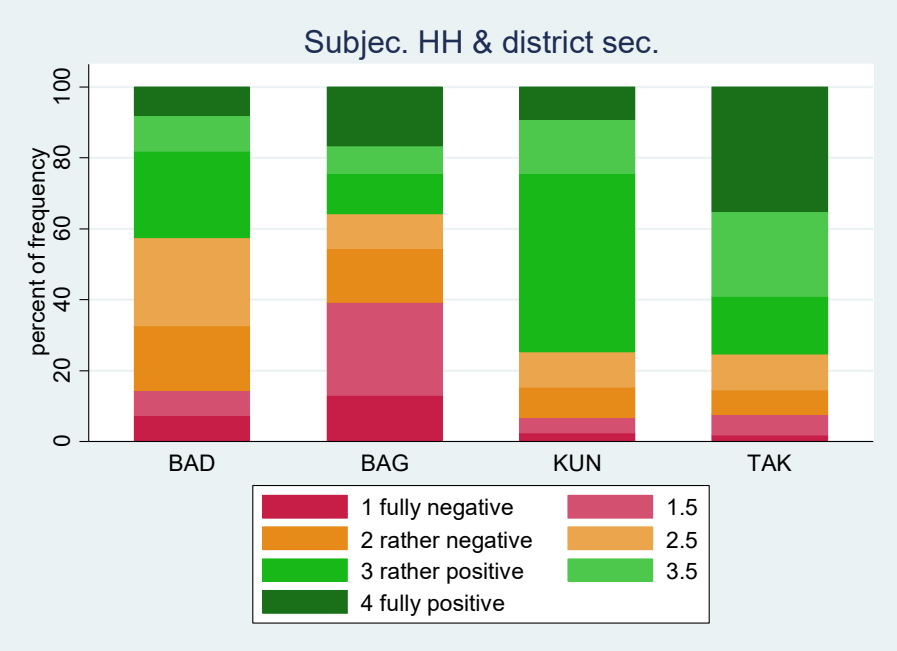
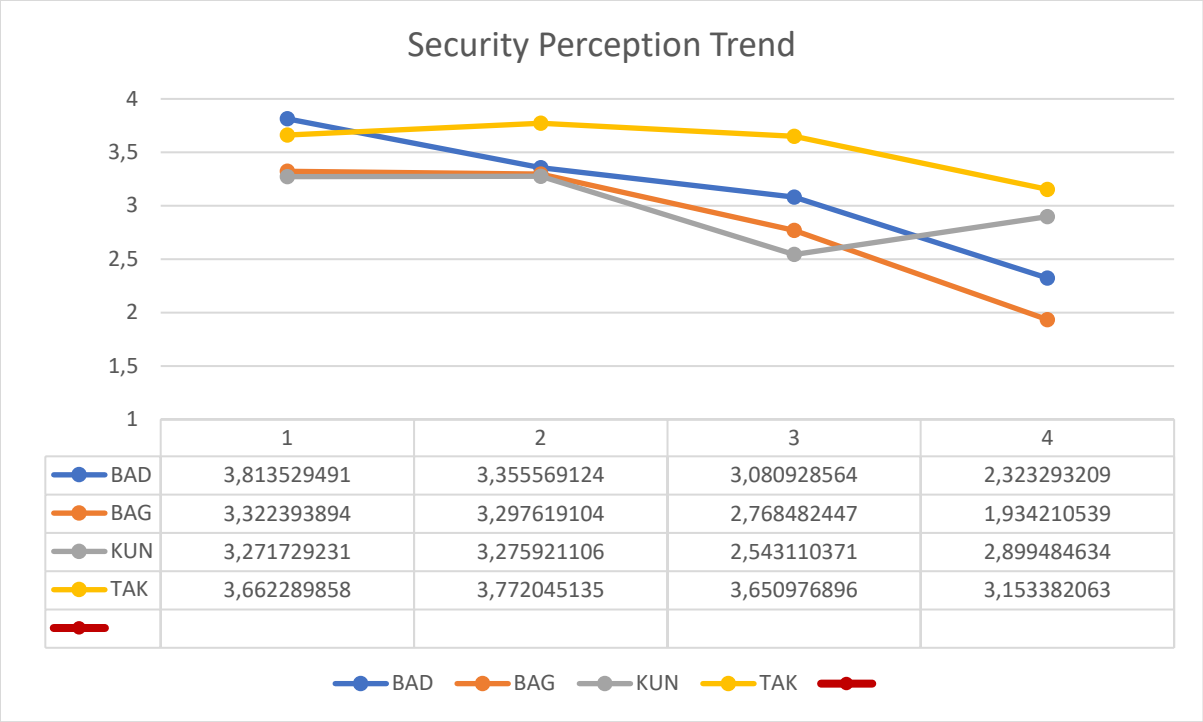


Figure 31: Wave 4, 35 districts – subjective household and district security compared between provinces



	1	2	3	4
BAD	3,813529491	3,355569124	3,080928564	2,323293209
BAG	3,322393894	3,297619104	2,768482447	1,934210539
KUN	3,271729231	3,275921106	2,543110371	2,899484634
TAK	3,662289858	3,772045135	3,650976896	3,153382063

Figure 32: All waves, 25 districts – subjective household and district security trend; 1 is worst, 4 is best score

Security perspective; perceptions of state capacity: After Wave 2, we asked respondents to evaluate a number of statements regarding the future security of the districts. One statement says that the Afghan state will be strong enough to deal with security issues arising after foreign forces will have left. Surprisingly, we register consistently high positive responses in Kunduz, where the state has been violently contested, and Takhar, where the state has been institutionally especially weak. Baghlan and particularly the rather remote Badakhshan are much more sceptical in Wave 4 and both register a steep decline in confidence in the state from Wave 3 to Wave 4.

We believe that three separate dynamics explain the findings. First, our past research has shown that under current conditions, confidence in the state is influenced most strongly by the simple absence or presence of the state – i.e. by factors like remoteness or state retreat (many districts in Badakhshan, increasingly also the case in Takhar in Wave 4). The difference between Takhar and Badakhshan is that in Takhar, local commander domination, while weakening the state as an institution, is locally none the less associated with the government (this is characteristic in districts like Rustaq, Khwaja Ghar, Farkhar). The commanders are the government in the eyes of the people. In Badakhshan the view of commanders seems to be different in this regard – people seem to expect more civilian rule.

Second, a state engaging the insurgents – even if not very successfully so – is perceived as active. This was the case for Kunduz until 2016. It is much less the case in Baghlan or Badakhshan (for a systematic assessment state-society relations in the four provinces see Juan et al. under review).

Lastly, the question links the assessment of the state to the departure of foreign forces. In Kunduz there was, and still is, widespread hope that this departure will actually facilitate a political peace process between the Taliban and the Government that would pacify the province and possibly the country (results of a systematic assessment of the discussion of future perspectives after foreign forces withdrawal in guideline interviews conducted 2012-2017, partially published in Koehler 2014).

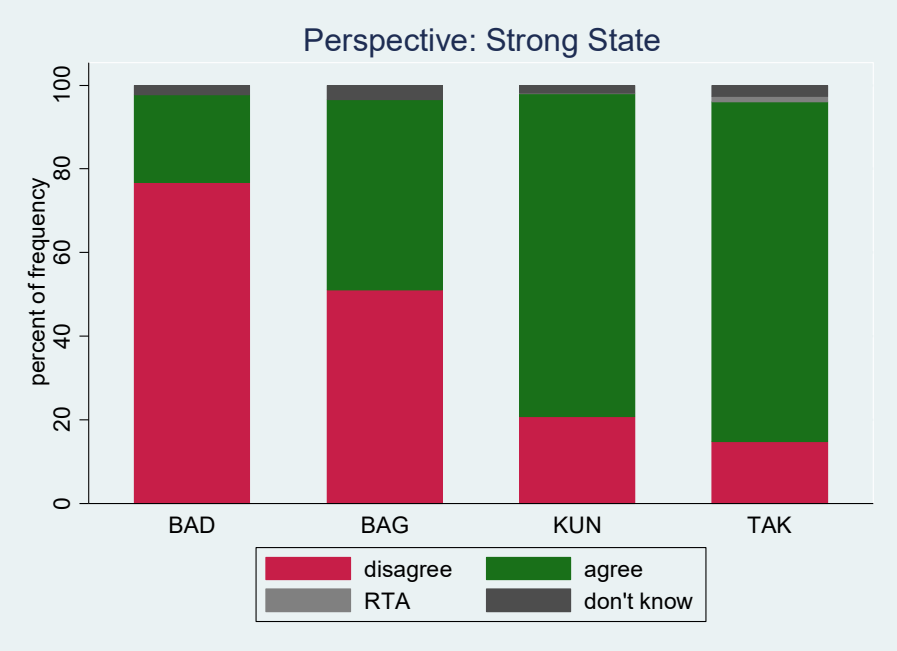


Figure 33: Wave 4, 35 districts - perception of the state as strong enough to deal with security issues without foreign forces

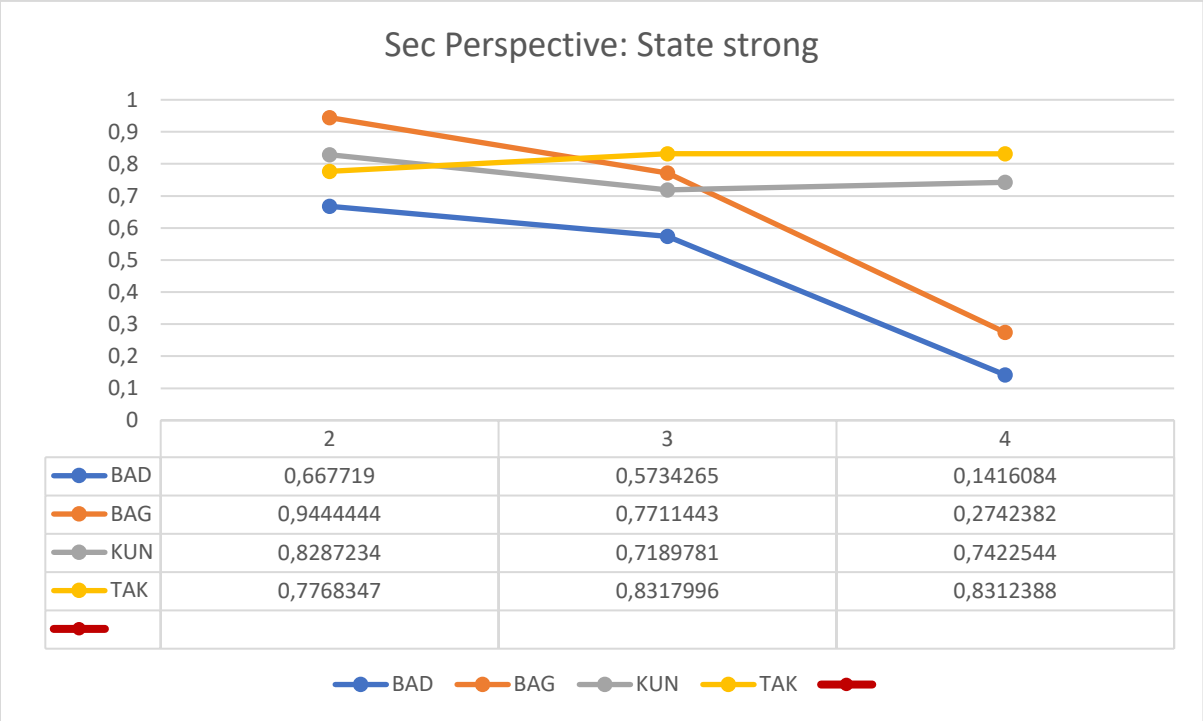


Figure 34: Waves 2-4, 25 districts - perception of the state as strong to deal with security without foreign forces; 0 is worst, 1 is best score

Implications for SPNA impact hypothesis

SPNA, as a stabilisation programme working with the toolbox of capacity and infrastructure development, intends to have a positive direct effect on security perceptions. Under conditions of an

ongoing insurgency and contested statehood, the effect on objective physical security would be indirect and marginal. Local community disputes are an exception, but they rarely make it into the insurgency-dominated incident records used in this report to represent physical security. Hence, the conceptual difference between objective measurements of physical security via incident density, and subjective village and district security, is important for SPNA's theory of change in relations to security-related stabilisation.

Both aspects of security are related to each other –and it is plausible to assume that a higher density of security incidents increases, at least to some extent, subjective perceptions of insecurity. But already the trend analysis presented above shows that objective and subjective security are conceptually distinct from each other – incidents may be decreasing while insecurity perceptions increase, and more violent provinces may show higher levels of confidence in future state security provision than less violent (but more remote) provinces. A Principle Component Analysis (PCA) of variables representing the different dimensions of the stabilisation field "security" show that in Wave 2 and 3, the relationship between incident density and security perceptions was actually positive (higher incidents and higher security perceptions trended together); this changed only in Wave 4. The reason seems to be that in Wave 2, the fruits of (temporary) pacification at the time of the survey had been achieved during the previous year through successful, but intense security operations (the incident density indicator counts the number of incidents around an assessment village in a three km radius in the year before and the year of the survey). In Wave 3, after the end of ISAF, many people viewed the struggle of the state against the insurgents as a sign of government activity and resolve (confirmed in many of the guideline interviews). This relationship had changed by the time of the endline, when incident density and security perceptions are not loading on the same component anymore (see annex).

To better understand the effect SPNA has on the security dimension of stabilisation, we want to know if capacity-building of local governance institutions and district-level infrastructure projects improve security perceptions and trust in the state as a security provider, irrespective of the "objective" physical security environment over which SPNA has no influence. To do so, in the analytical section, we will control for the objective aspect of physical security that we consider conceptually independent from the impact of SPNA measures (see *impact hypothesis* above).

GOVERNANCE

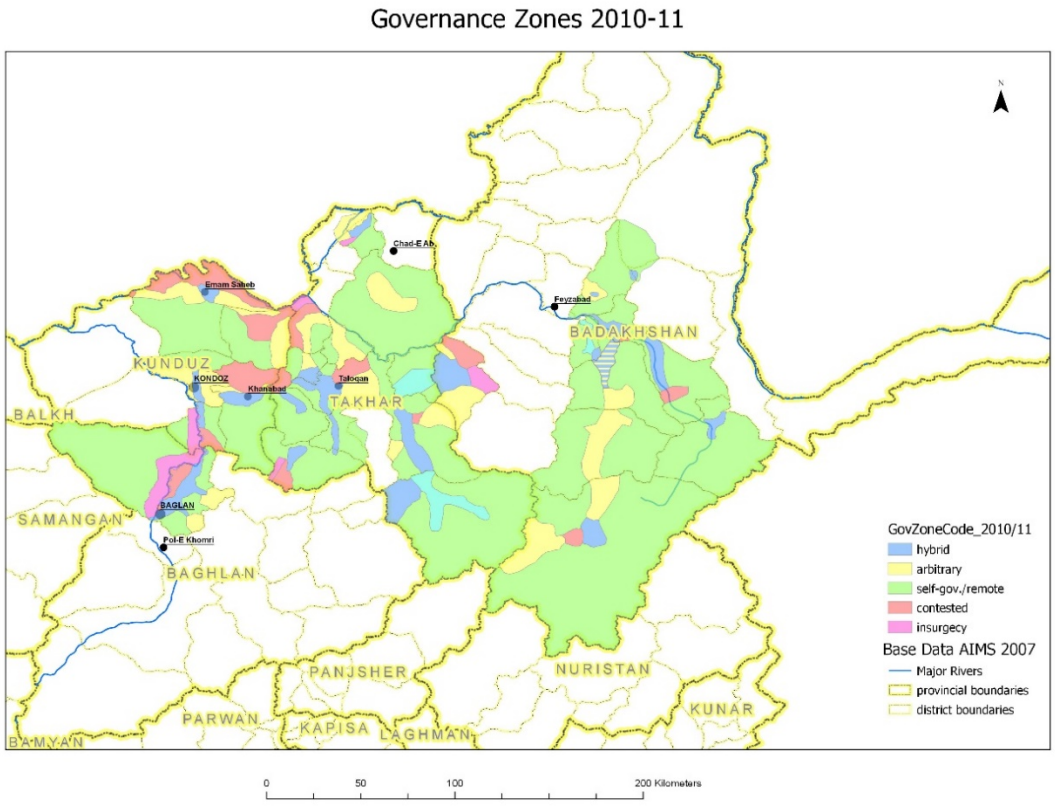
SPNA's impact hypothesis: We assume that the SPNA programme components improve legitimacy and effectiveness of both district-level as well as community-level governance via capacity building, introduction and application of procedures, training "on the job" and by connecting communities with district-level institutions.

Context variable: governance zones

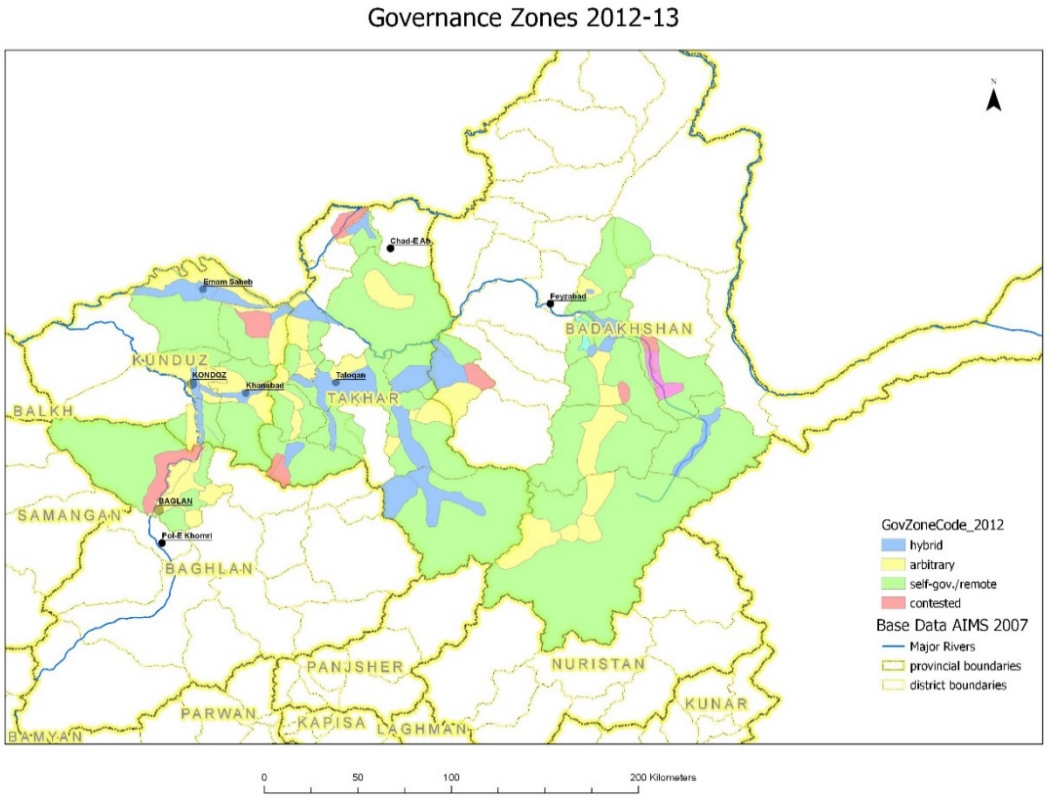
Governance, in the wider sense used in this report, refers to institutionalised ways by which society regulates issues of collective concern. This regulation can take the form of state authority; it can be exercised by (friendly or hostile) alternatives to state authority, like militias or the Taliban; and it can be achieved by societal institutions, like local councils or traditional authorities. In the survey region, governance functions are the result of often fluid arrangements between different sources of regulative authority, which vary according to whether they are institutionalised (persistent rules) or alternatively comprise ad hoc power relations. Despite the fluidity between these arrangements, the models of governance that have emerged over the course of the past two decades are rather stable –

while the geographical extent of the different governance zones is changing. We distinguish between areas in which governance is dominated by the state (governance by government) or a mix of state authority and the authority of local strongmen in state positions (hybrid governance); the rule of informal or semi-official commanders and their militias (arbitrary/commander rule); remote areas where communities mostly look after themselves (remote/self-governance); areas in which the right and authority to govern is violently contested, mostly between insurgents and the state (contested governance); and lastly areas governed by the insurgents.

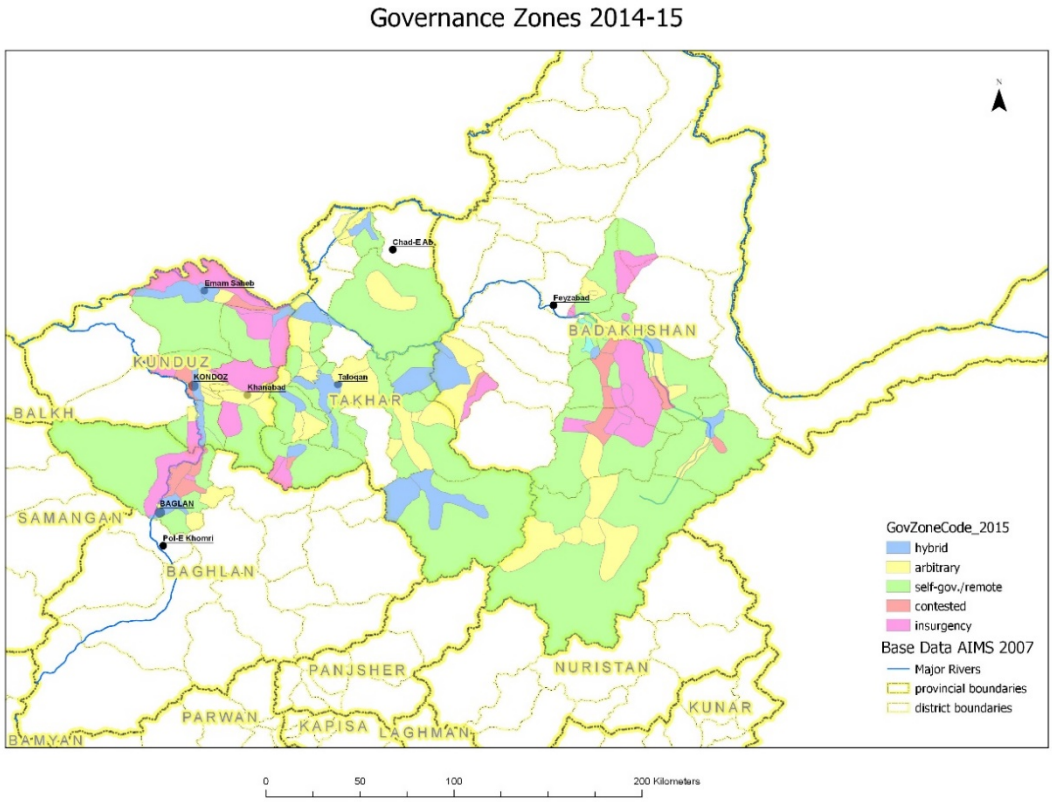
We coded the resulting governance patchworks for each assessment wave. The following maps show the changes between the waves.



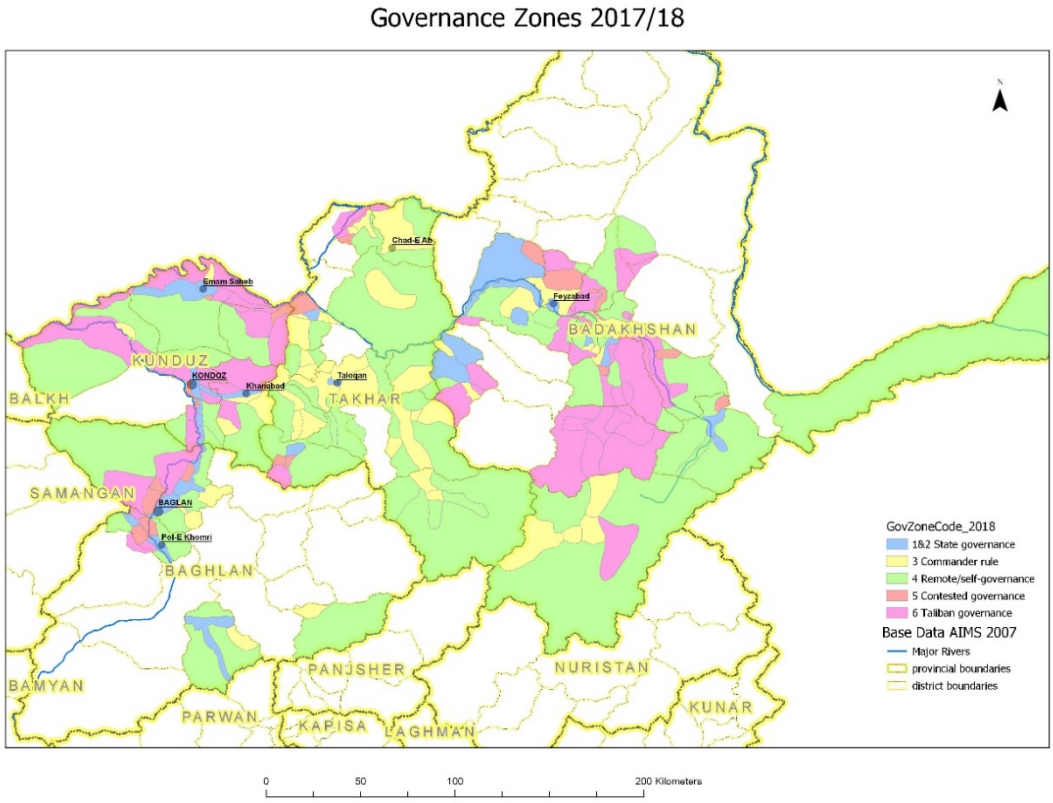
Map 1: Governance Zones 2010-11



Map 2: Governance Zones 2012-13



Map 3: Governance Zones 2014-15



Map 4: Governance Zones 2017-18

Between 2010 and 2012, local governance was still formed by a decade of government rule via patronage, shaped by the relationship between Hamid Karzai and top political patrons of the former anti-Taliban Northern Alliance (state and hybrid governance, light and dark blue on the maps). Locally the former *jihadi tanzeem*s (organisations/parties) and their commanders gained renewed influence, partly because of the reestablishment of local militias (commander/arbitrary rule, yellow on the maps). Hence, civilian governance came under pressure not only from the armed insurgency, but also from the re-emergence of the *qomandan* system. Taliban territorial control and governance provision (insurgency governance, purple on the maps) retreated but was not replaced by sustainable state rule. Rather, it changed to contested areas (red) or to areas ruled by local commanders and their militias (yellow).

The 2014 elections and subsequent stalemate between President-elect Ashraf Ghani and CEO-elect Abdullah Abdullah paralysed decision-making even on urgent issues of sub-national governance and froze the local governance reform process. The DDAs and *woliswals* (district managers) were unsure about how district-level representation of the communities would be organised in the future and how the future responsibilities of the different institution would look like. The state loses more ground, this time to a resurgent Taliban insurgency that extends its reach, even exceeding control over territory held in 2010/11. For the first time, the Taliban make wider and more sustained gains also in Badakhshan that go beyond their strategical stronghold in Wardooj district.

From 2015 to 2018, the state retreated even further and was, for the most part, confined to the administrative centres it still controls and the main lines of communication connecting those. This retreat of state governance was not only a consequence of increasing insurgency control in the hinterland of the provincial centres, but also reflected the inability of the state to reign in increasingly

autonomous local strongmen (mostly in Takhar). In some districts the Taliban gained credit for targeting the arbitrary rule of commanders (e.g. Khanabad in Kunduz, according to field interviews conducted there with representatives of the DDA in 2017/18; also debriefing notes July 2018). The Taliban made strategically important further gains mostly in Badakhshan and Kunduz (the takeovers of Kunduz in 2015 and 2016 falls into this period). However, during the time of the survey in 2017/18 active fighting between the government and the Taliban had decreased across survey provinces but most noticeably in Kunduz (see the drop in incidents discussed in the previous section "*Security*"). Local residents as well as our research teams were able to pass the invisible frontlines of territorial control without problems. Some governance services are provided as the result of an informal agreement between the Government and the Taliban (schooling, hospitals, some development and even village-level elections in some Taliban areas). However, it seems that the Taliban are able to project their power more effectively into government-controlled areas than vice versa. This ranges from enforcing times when mobile phone networks can operate (or are switched off), threats and application of violence against government officials, over interference with elections to enforcing decisions taken by Taliban courts even in areas outside their control. We also find that the Taliban are increasing their penetration of remote, formerly self-governed areas (Yamgan, remote sections of Baghlan-i Jadid) while the state as well as state-associated militias are decreasing their presence in such areas (e.g. Warsaj, Imam Sahib).

Put bluntly, thus far the Taliban are able and willing to substantially reduce everyday security in government-controlled areas, while the government for the most part refrains from doing the same in Taliban-controlled areas. The Taliban bring the war to the government, while the government allows the Taliban to keep their separate peace. There are exceptions to this rule – like the airstrike targeting a graduation ceremony at a madressa in Dashti Archi, Kunduz, on 2 April 2018 – but during the time of data collection for this endline report in the survey districts, such events remained the exception.

Critical discussion of the indicator: The mapping of governance zones implies strict borders between uniform territorial zones; it does not capture transitions or patchwork zones. The territorial coding is a compromise between oversimplification (we code sub-districts rather than districts or whole provinces) and over-differentiation (we are not coding individual settlements or sub-villages). This means, however, that in some cases local differentiation is lost – as one of our researchers from Kunduz provinces remarked during the debriefing: In some areas neighbouring settlements, even neighbouring quarters of some larger settlements, should be coded in different ways. This level of differentiation seemed not practical for our purposes, not least since the teams did not cover all settlements of a district in similar depth; hence, we focussed on dominant patterns of control over the different geographical and demographic zones of the assessed district. This approach provides a good indication of the dominant governance pattern in a valley or sub-district.

Stabilisation indicators: state responsiveness, power perceptions, conflict resolution and hashar

We select four governance indicators that reflect different dimensions of governance-related stability affected by SPNA activities and outcomes. We look (a) at the perceived responsiveness of the district administration (*woliswali* cares for the village) as a proxy for political trust and legitimacy at local level; (b) at the question of who is the most powerful person in the district, where we consider official representatives of the executive as an indication of strengthening state presence and informal gunmen or Taliban as signs of weakening state governance (with two intermediate categories); (c) at

the question of how often people feel that corrupt means are used to win in conflicts and disputes in the district: money (venality), ties (nepotism), and force. Finally, (d) we also look at the readiness of the village community to fulfil their hashar obligations (unpaid community work) as a proxy for social cohesion and collective action, assuming that hashar remained the main and socially accepted form of mobilisation for unpaid collective action.

Responsiveness of the state: A key question reflecting levels of political trust in Afghanistan is whether local communities view state authority as caring for their concerns. The question refers to the lowest level of state representation, i.e. the district administration (woliswali). For most people in rural Afghanistan, the district administration and its head, the woliswal, are the face of the state. The woliswali deals with most administrative issues of concern for the village communities (identity cards, birth certificates, salaries of most local state employees, supervision of public service provision such as education, healthcare) and is the first point of contact in case of conflict, complaints or criminal offenses. The perception of a state that is willing and able to care for the population has been low and is an issue throughout Afghanistan's recent history. It refers not only to the question of output legitimacy, i.e. the acquiescence of citizens with government rule in return for services and public goods, but also a question of being cared for like a citizen and not just as a subject.

In Wave 4 the distribution of more pessimistic and more optimistic responses is similar between provinces, and overall, negative opinions (never, rarely) range between 70% and 80%. There is, however, a marked difference in the distribution of fully negative responses (never): here Baghlan scores highest and Badakhshan lowest.

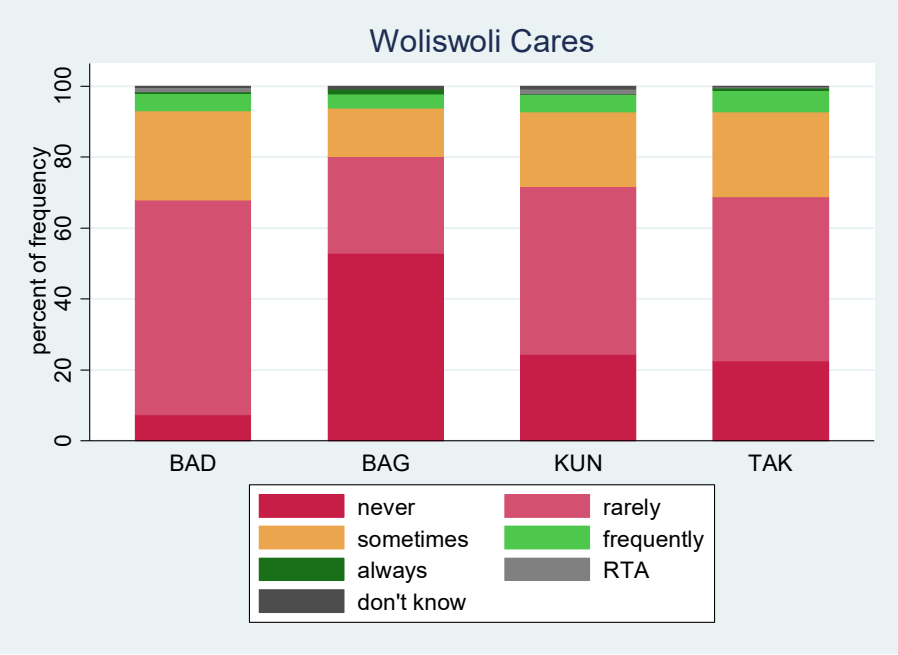
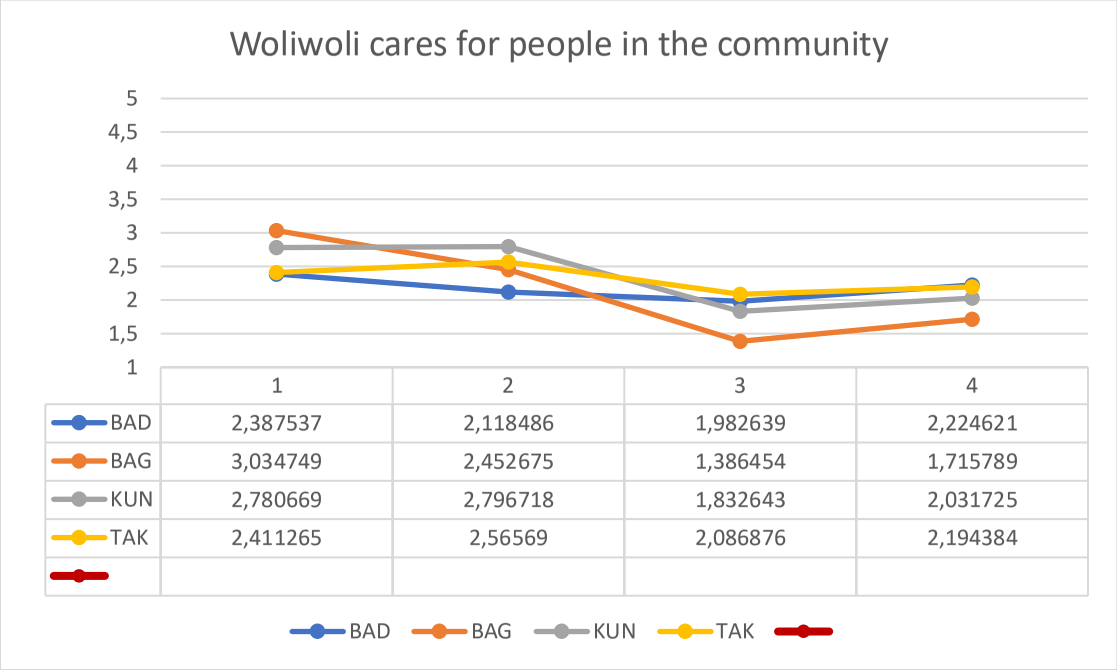


Figure 35: Wave 4, 35 districts - does the district administration care?

In terms of trends, this indicator remained relatively stable between Waves 1 and 2 but had deteriorated in all provinces but Badakhshan by Wave 3 (from an average response score of 2.5, i.e. between rarely and sometimes to below 2, i.e. rarely). Negative changes in Baghlan are most pronounced but we need to consider, again, that in the trend analysis this province is represented

only by the rather volatile district of Baghlan-i Jadid. There are only minor changes between Wave 3 and 4 but those are positive across provinces.



	1	2	3	4
BAD	2,387537	2,118486	1,982639	2,224621
BAG	3,034749	2,452675	1,386454	1,715789
KUN	2,780669	2,796718	1,832643	2,031725
TAK	2,411265	2,56569	2,086876	2,194384

Figure 36: All waves, 25 districts - does the district administration care? 1 is worst, 5 is best score

Most powerful person in the district: The perception of state authority as responsive and caring is one important dimension of the stability state-provided governance. It relates to the output legitimacy of the state. A second important dimension is the actual authority of the state. Hence, we are asking who is considered the most powerful actor in the district, and code official representatives of the executive as best option (woliswal, police chief, head of persecution department) and informal commanders, gunmen or insurgents as the worst option. Inbetween we also code for informal modern powerholders (elected representatives of the shura structure or in the provincial council or parliament) as second best option and informal traditional authorities (maliks, arbabs, mullahs, elders) as second worst option. Note that in the survey we cannot exclude that a respondent considers an official most powerful because that official is, at the same time, a former jihadi commander and still wields informal power. However, it is up to the respondent to decide between the formal (e.g. woliswal) and informal (here: jihadi commander) label for the most powerful actor in the district.

For Wave 4 we find that state officials are considered as most powerful in our target districts of Badakhshan and Baghlan by more than 60% of respondents. In Kunduz and Takhar this share is much smaller and the share of modern unofficial powerholders (elected representatives of councils / parliament) is higher. In Takhar, provincial and national-level deputies are mostly responsible for this, in Kunduz, the heads of DDAs also figure prominently, being associated with both state as well as with Taliban representatives in that province. The power of the Taliban in Kunduz is likely underreported for Wave 4 and this most likely relates to fear of negative consequences for areas under de-facto Taliban control.

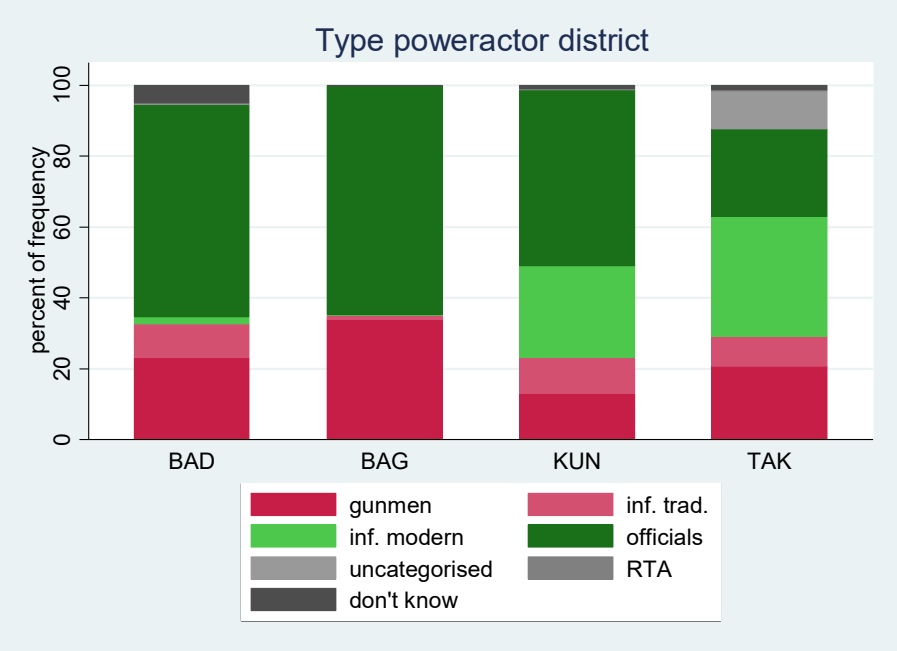
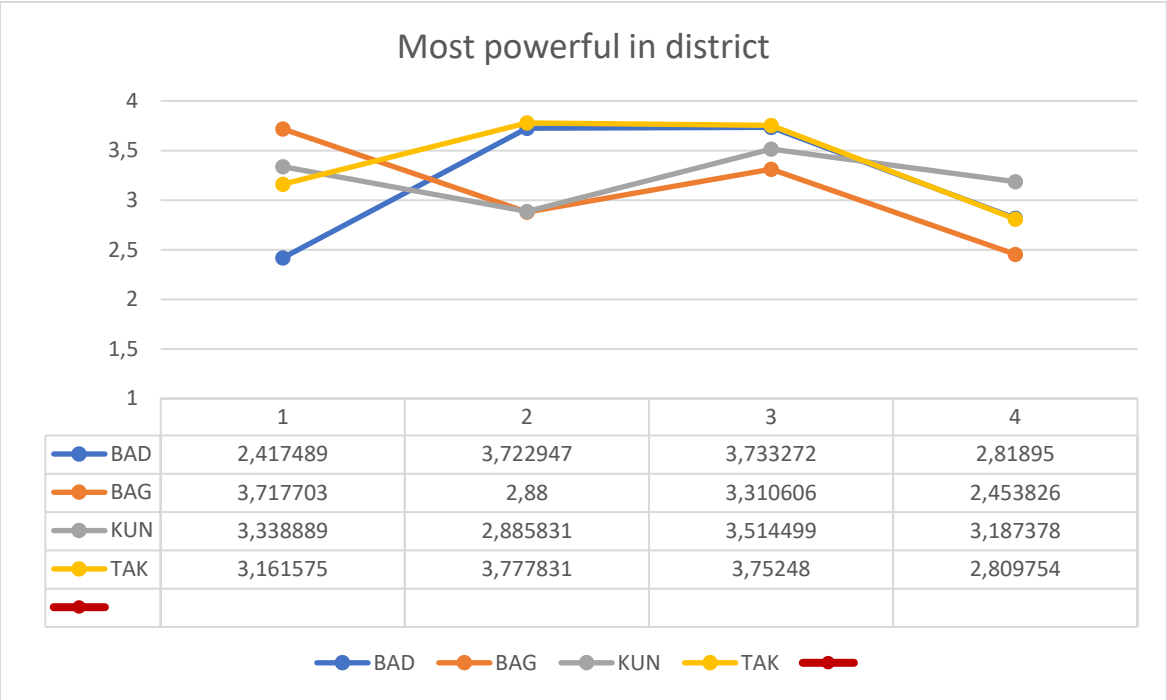


Figure 37: Wave 4, 35 districts - type of most powerful actor in the district

Looking at the trends across waves, we start out with a high spread of average scores for the different categories of powerholders between provinces with Badakhshan at the lower end and Baghlan (i.e. Baghlan-i Jadid) at the upper end of the scale. Up until Wave 3, Baghlan and Kunduz trend together, with a marked deterioration of official powerholders' influence. Badakhshan and Takhar move in the opposite direction, with official powerholders gaining influence until Wave 3. In Wave 4 all provinces report a similar decline in official authority (least pronounced in Kunduz).



	1	2	3	4
BAD	2,417489	3,722947	3,733272	2,81895
BAG	3,717703	2,88	3,310606	2,453826
KUN	3,338889	2,885831	3,514499	3,187378
TAK	3,161575	3,777831	3,75248	2,809754

Figure 38: All waves, 25 districts – index over types of most powerful actors mentioned for each district (4=formal state authority; 3=elected members of official councils, 2=traditional authorities, 1=informal armed leaders)

Corrupt means of dealing with conflict: Sustainable stability depends on the ability of political and social institutions to deal with conflicts of interest in a reliable, predictable and legitimate way. Conflict resolution is also one of the modules that SPNA provided to DDAs as part of their capacity-building component. Hence, we are interested in how fair or corrupt respondents perceive conflict-processing in their districts. We ask how often corrupt means (money, connections, force) are applied in dealing with conflicts and computed a score across the responses via a statistical procedure called Latent Class Analysis (LCA). The three corrupt means correlate strongly, resulting in a score that is easy to interpret. Note that this indicator does not deal with standard definitions of corruption that usually refer to misuse of public office for personal gain. Rather, it refers to the way societal and state actors in a district are dealing with arising conflicts. Based on earlier qualitative research on conflict processing in Afghanistan (Koehler 2015) we assume that venality, nepotism and violent self-help are widespread, but widely considered illegitimate or at least less legitimate than formal as well as informal procedures (courts of law, arbitration, shuras, jirgas).

In Wave 4 the respondents in all provinces are similarly likely to belong to the two bad classes defined by people thinking that corrupt means are always or sometimes applied (between 70% and 80% of respondents). However, the most negative class shows more variation. Twice as many people in Badakhshan and Takhar than in Kunduz or Baghlan feel that corruption is always part of dealing with conflicts in their district.

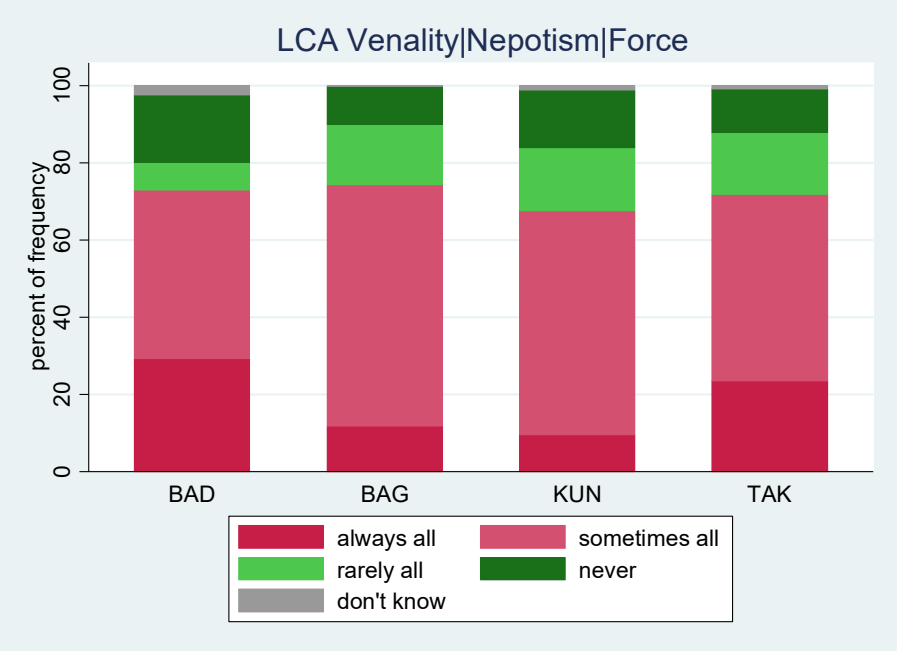
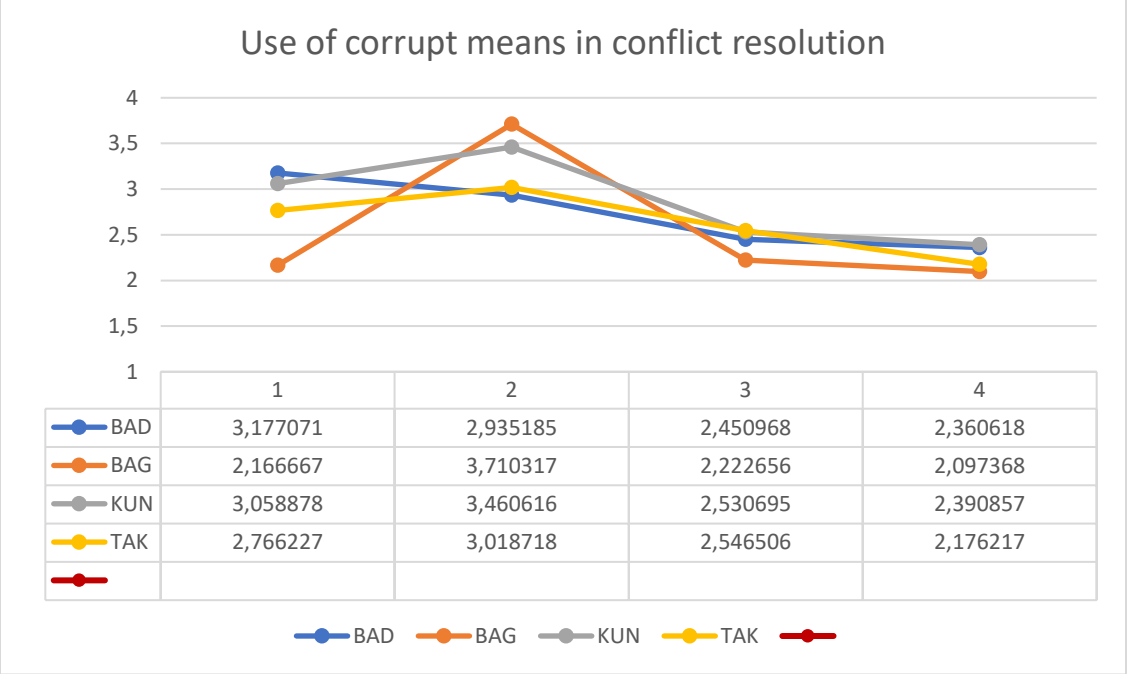


Figure 39: W4 35 districts – corrupt means of conflict processing

The trend across the four waves is defined by two tendencies. First, since Wave 2 the general direction for all provinces is negative. Second, Wave 1 and 2 register a broader spread between provinces which narrows in the last two waves. Baghlan (Baghlan-i Jadid) shows the strongest changes between 2010 and 2015, which are determined by the dynamics around a highly divisive state capture by local jihadi commanders, which was backed by then Vice President Marshal Fahim until his death in 2014. For a

brief period in 2012, this local criminal network of commanders came under direct political pressure (after one of them ambushed and killed Afghan special forces soldiers trained by the US) and temporarily resigned, leading to short-lived improvements of governance perceptions during the Wave 2 survey.



	1	2	3	4
BAD	3,177071	2,935185	2,450968	2,360618
BAG	2,166667	3,710317	2,222656	2,097368
KUN	3,058878	3,460616	2,530695	2,390857
TAK	2,766227	3,018718	2,546506	2,176217

Figure 40: All waves, 25 districts – corrupt means of conflict processing; 1 is worst, 4 is best score

Social cohesion / collective action: The state is a principle provider of governance functions for the public good. However, many governance functions are provided by society itself, drawing on traditional institutions and authority (Murtazashvili 2016). The institution of *hashar*, i.e. unpaid community work, is an important institution in many Islamic societies. Especially in the absence of state-provided public services, hashar is often the only way to mobilise larger groups for the collective good, which is of particular importance in case of natural disaster, but also for vital regular tasks like maintaining irrigation infrastructure or local roads. We ask in how far households uphold their hashar obligations.

In Wave 4, Kunduz stands out, with over 40% reporting that households are not interested in participating in hashar. While hashar generally has stronger institutional roots in more rural settings, the stark decline of this institution in Kunduz is surprising. Over the three prior waves Kunduz scored similar to Takhar and only drops in Wave 4. While the Taliban are interested in enforcing some governance functions – like security, justice, but also school and hospital attendance of staff – they do not seem to affect compliance with hashar at community level in any positive way. A possible explanation is that the Taliban are hostile to any kind of community mobilisation that is beyond their control. Badakhshan is the only province where hashar compliance is increasing across the waves.

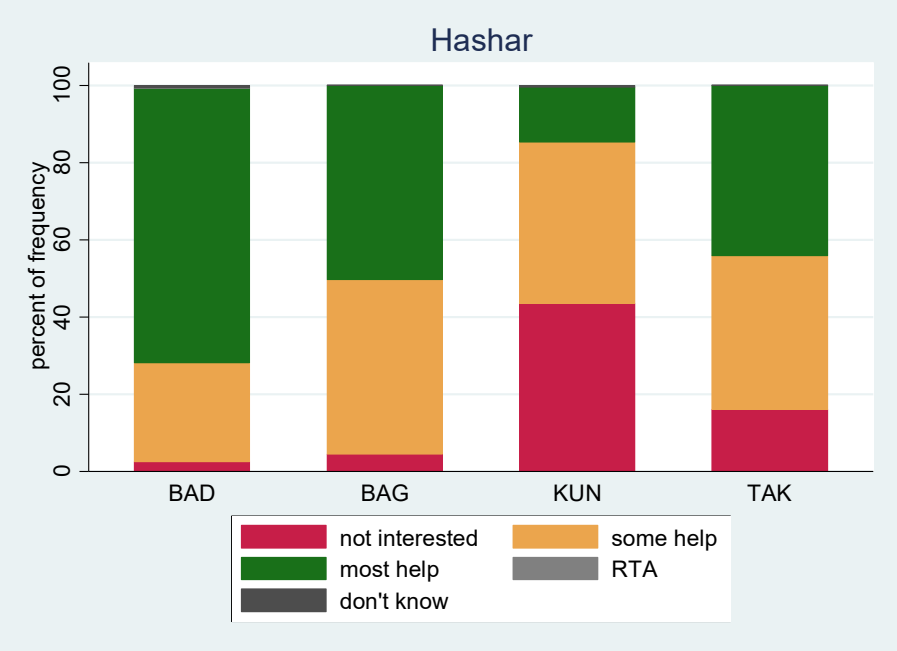
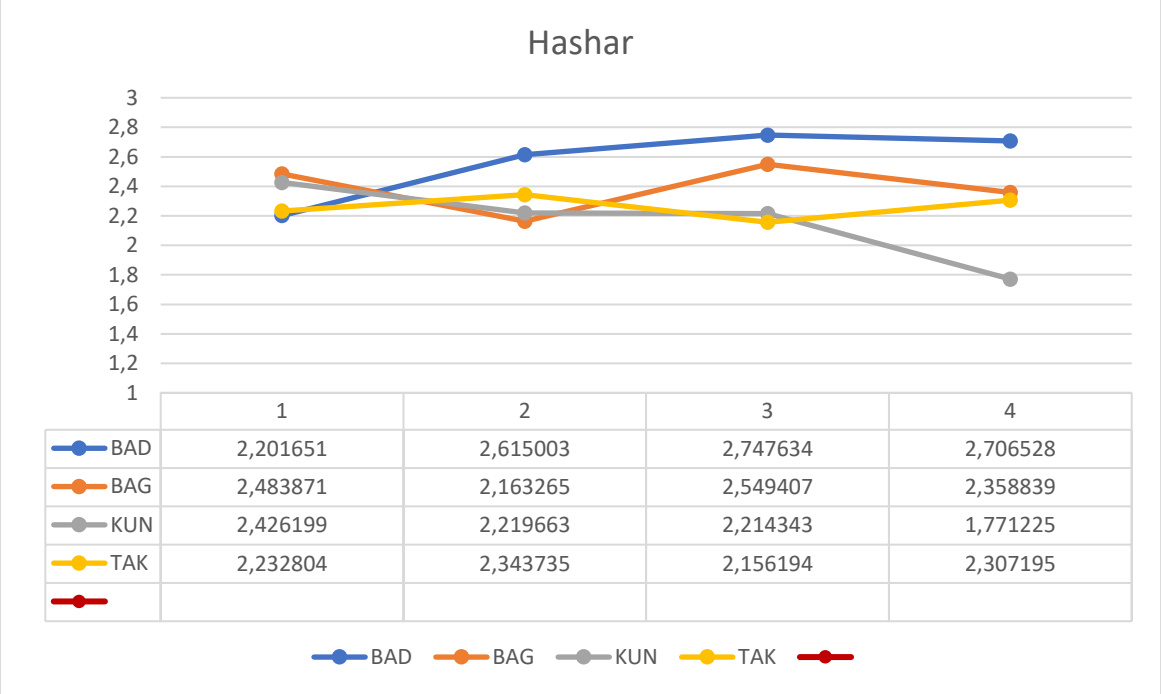


Figure 41: W4 35 districts – compliance with Hashar obligations



	1	2	3	4
BAD	2,201651	2,615003	2,747634	2,706528
BAG	2,483871	2,163265	2,549407	2,358839
KUN	2,426199	2,219663	2,214343	1,771225
TAK	2,232804	2,343735	2,156194	2,307195

Figure 42: All waves, 25 districts – compliance with hashar obligations; 1 is worst, 3 is best score

Implications for SPNA impact hypothesis

In terms of stabilisation expectations, SPNA has a strong focus on improving local, i.e. district-level governance. The indicators capture four important dimensions of governance intentionally targeted in SPNA's theory of change: the **legitimacy** of and trust in local state institutions, i.e. the perception

of the state as caring for the needs of the community; the **effectiveness** of local state authority indicated by the perception of who is the most powerful actor in the district; the degree to which dealing with conflicts in the district is **fair** or affected by corrupt practices; finally, the ability of local communities to **mobilise** for collective action to solve specific governance challenges on their own.

On all four dimensions of governance, the trend from baseline to endline indicates either stagnation or is negative (with the exception of hashar mobilisation in Badakhshan). This trend corresponds to the deterioration of territorial rule by official state institutions and the spread of informal commander or Taliban rule in many parts of the SPNA survey area. Since SPNA has no influence over framework conditions like territorial control and the question of who rules where, we need to separate SPNA's impact on the perception of legitimate, effective, and fair governance from the effect of a general decline in state presence and state rule as a context condition. Hence, in the analytical statistical models we will control for those general context conditions not driven by SPNA (i.e. gov zones).

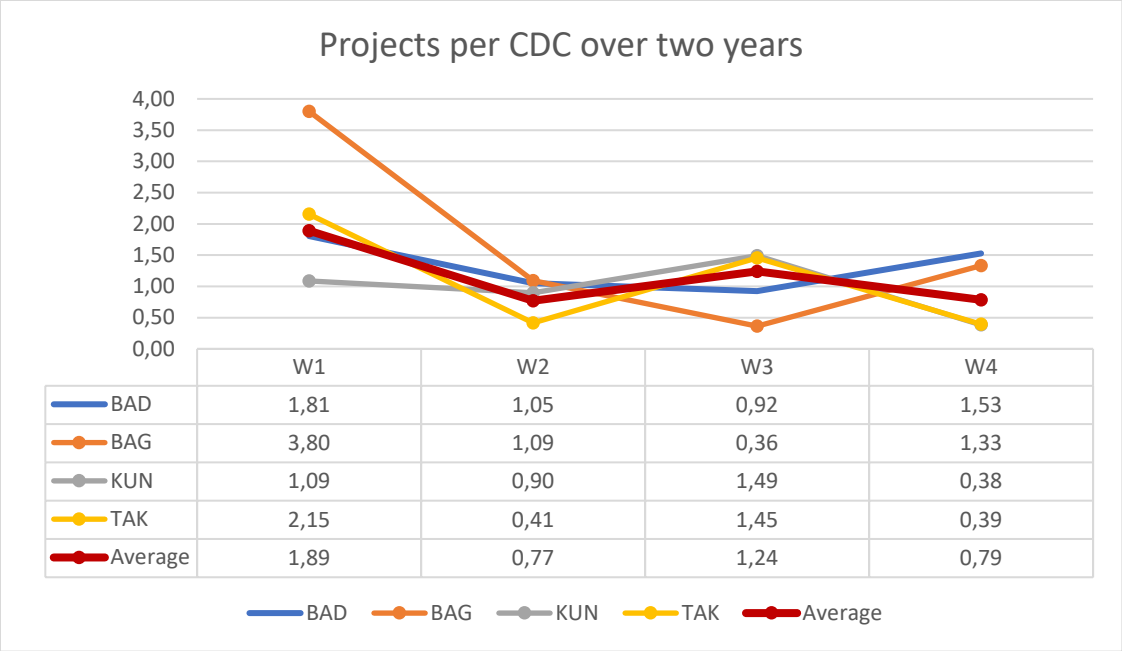
Economy / Development

SPNA's impact hypothesis: We assume that the SPNA programme components improve economic development through infrastructure and other (third-party) development measures that are prioritised by the DDA according to local needs, and increasingly based on smart choices of competent local state and societal institutions.

Context indicators: number of development projects, cars & tractors, bazaar size

In order to capture the development and economic context in which SPNA has been implemented, we identified three context indicators that account for the two distinct dimensions of this field of SPNA stabilisation, i.e. **development** and the **local economy**.

The most important trends in terms of development have been the initial **intensification of development initiatives** alongside the military surge and COIN strategy beginning in 2010 and ending by the time of the Wave 3 survey in 2014/15. Over this period both the German and the US government increased their spending significantly, and put a greater focus on rural development as well as capacity-building and training of subnational governance institutions. Many of those programmes were implemented above village and even district levels. The question as to how far local communities were directly involved in or exposed to development measures is an important general indication for stability. We see stark differences between the provinces and different dynamics between the waves. Between the first two waves we note little change, with Takhar being a negative outlier in terms of reported project implementation (less than half compared to the three other provinces). Takhar catches up in wave 3 with the best performer of that wave, Kunduz. Both provinces record a steep fall back to the lowest levels of projects received (0.4 per village) in Wave 4. Badakhshan remains a stable mid-level performer in wave 1-3 and increases average projects received markedly in Wave 4. Baghlan (which for the reduced sample is always only one district, Baghlan-i Jadid) is relatively stable between 1.2 and 1.4 projects per village in wave 1, 2 and 4. It is the only province showing a steep drop in Wave 3.



	W1	W2	W3	W4
BAD	1,81	1,05	0,92	1,53
BAG	3,80	1,09	0,36	1,33
KUN	1,09	0,90	1,49	0,38
TAK	2,15	0,41	1,45	0,39
Average	1,89	0,77	1,24	0,79

Figure 43: All waves, 25 districts – number of projects implemented over 2 past years per CDC

These results, based on our village profiles, coincide with the information we gathered during the debriefing sessions from our team members. Their written notes and field impressions confirmed that wider development efforts in Kunduz and Takhar had all but stopped at the time of the Wave 4 assessment. The situation in Badakhshan and in Baghlan (including the three additional districts added only in Wave 4) was more diverse in terms of ongoing development activities.

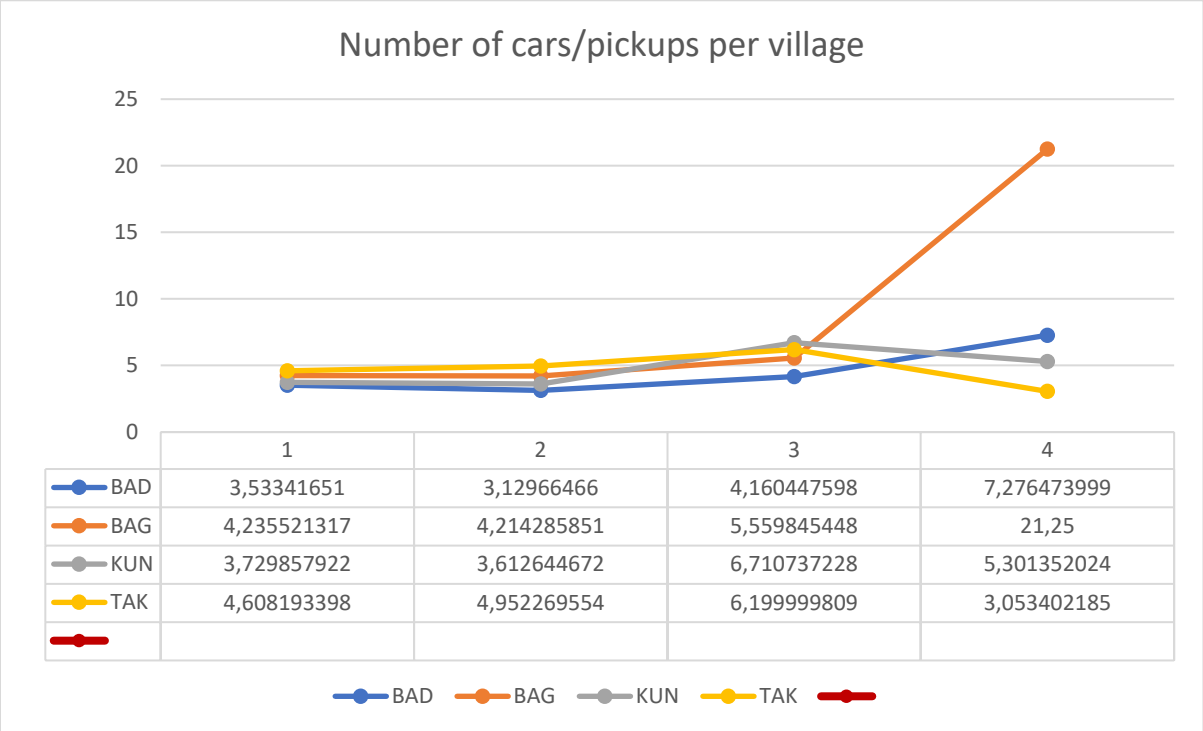
Critical discussion of the indicator: this indicator captures ongoing development activities external to SPNA at village/community level. Assessing the number and kind of projects implemented in a village over a period of two years as reported by CDC representatives generally provides reliable information. The simple project count we use for statistical purposes is, however, only a rough proxy measurement for local development activities. Nonetheless, it does provide a reliable approximation of general trends (high, medium and low levels of activity over a period of two years).

In terms of the **"real" economy** (i.e. the local economy existing independently of development projects and measures), different local opportunities and dynamics play a role. Already for wave 3, we noticed and reported that important aspects of the village-level rural economy as well as the household economy seem to be detached from the dynamics of the internationally backed development sector (Koehler et al. 2015). This observation is consistent across waves (see annex). A decline or increase in our objective as well as subjective measurements for development did not predict a similar decline or increase in our measurements for the "real" economy. Indicators of the "real" economy that seem to be detached from development include the size of the district bazaars, the number of cars and tractors per village as well as subjective indicators like self-assessed material well-being, self-reported land-ownership, or the possession of mobile phones per household.

The same was true for the relationship between security and the "real" rural economy. We found that until 2015, the specific insecurity of the assessment area did not visibly and negatively affect local economic activities. In parts of the research area this may relate to locally dominant illicit and informal economic activities, like informal mining or the opium poppy as well as cannabis economy – activities

that are independent of development efforts and may even thrive under conditions of latent insecurity. Revenues from cash-generating activities were increasingly invested in construction businesses, particularly in the rapidly growing urban centres.

However, since early 2015 (Wave 3) the trend seems to be more contradictory and fragmented. It seems to continue for Badakhshan (judging by the car index as well as self-reported material well-being) but not for Kunduz and Takhar, where the indices are declining. In terms of our indicator for mechanised agriculture – tractors per village – we note a decrease in Kunduz (which had the strongest increase before Wave 3), stagnation in Takhar, recovery in Baghlan-i Jadid and a slight decrease in Badakhshan. The size of bazaars is declining across most districts (see annex for this district-level indicator), but for different reasons: Iranian sanctions against work migration, drought, a retreat of development staff and contractors as well as insecurity are quoted as reasons by district authorities interviewed. The jump in car ownership in Baghlan is driven by four villages in the only district assessed there since 2010 (Baghlan-i Jadid). We double-checked with village representatives and they confirmed this to be the case – however, in three cases the increase of cars relates to the setup of transportation businesses – a car dealer and a taxi/logistics business.



	1	2	3	4
BAD	3,53341651	3,12966466	4,160447598	7,276473999
BAG	4,235521317	4,214285851	5,559845448	21,25
KUN	3,729857922	3,612644672	6,710737228	5,301352024
TAK	4,608193398	4,952269554	6,199999809	3,053402185

Figure 44: All waves, 25 districts – number of cars or pick-up trucks owned in the village

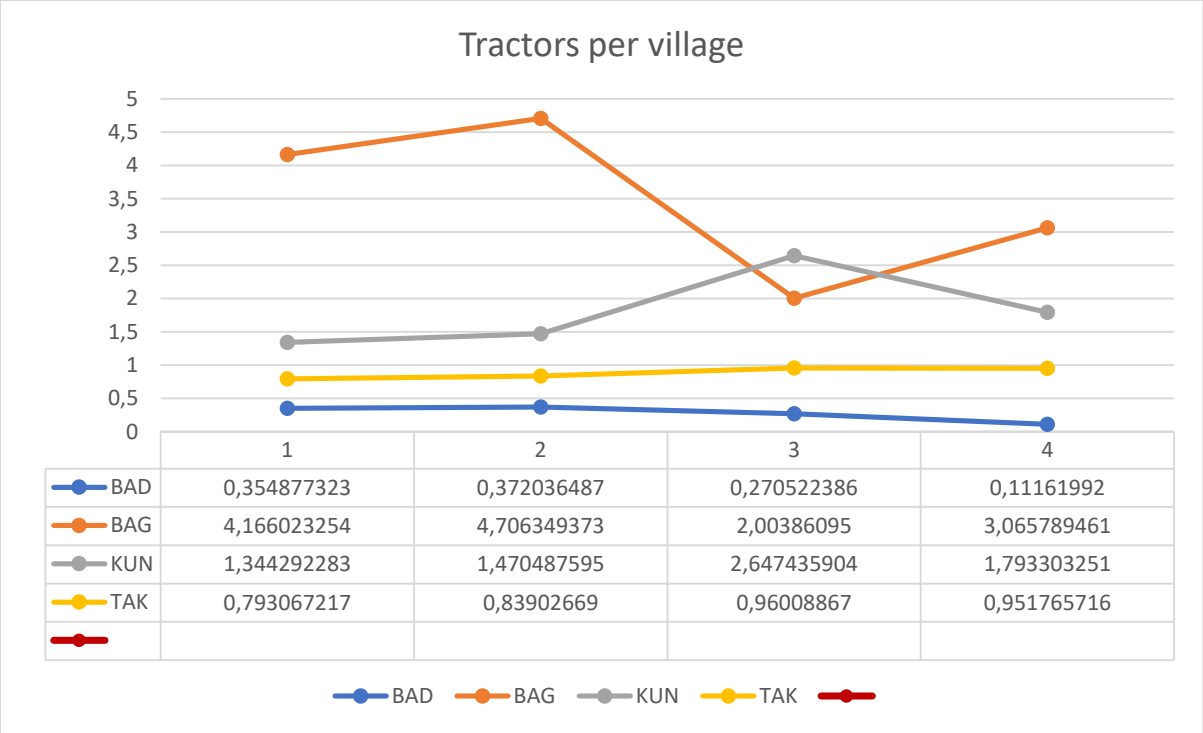


Figure 45: All waves, 25 districts – number of tractors owned in the village

Critical discussion of the indicators: We use the car index per village to reflect the general economic prosperity of a village and the tractor index for the degree of mechanisation of the agricultural sector. Past research has shown that only the surplus household cash income of richer households is invested in cars (often from the illicit or informal economy). The appearance of car pools and taxi services in rural areas seems to be a recent development. Tractors are also usually private investments. All of this seems to be sufficiently independent from immediate SPNA activities to serve as indicators for the wider external context of economic development.

Stabilisation indicators: satisfaction with development, material well-being, access to land

As in 2015, we statistically identify four **distinct dimensions of development and economic progress** that the SPNA programme may influence according to its theory of change: the first two dimensions relate to development programmes (general satisfaction and specific satisfaction with the government associated development provision). These seem to be largely independent from indicators linked to the second dimension, the rural agricultural subsistence economy. Our indicators depicting these two different dimensions continue to show different development trends for both. The high degree of independence of the local subsistence economy in rural areas from development programmes may have two reasons: first, development programmes target a different level of economic activities (not subsistence agriculture, or informal trade); second, the drivers of the local economy are insulated from changes induced by development, as is the case with the illicit economy.

General satisfaction with development: The indicator is built from a question that asks for satisfaction with the development contribution of different actors (INGOs, the government, the community, a

private patron) across seven sectors (drinking water, agricultural production, roads, job creation, electricity, healthcare and schooling). This indicator uses only the highest mark (maximum is 4 for fully agreeing with the statement that development actor X contributed positively to developments in sector Y over the past two years) per respondent with any sector and any development providers. The idea is that if a respondent is happy with at least one sector or actor, he can nonetheless be classified as being exposed to positive development even if nothing much happens in other sectors or by other actors. In other words: the graphs here represent a maximum positive result possible.

Against this backdrop, results in Wave 4 for Takhar are extremely poor (close to 70% of respondents only reach worst or bad results), followed by Baghlan, and then Kunduz with above 40% negative top results). Badakhshan is the positive outlier here, with nearly 90% positive results for development in at least one sector by at least one development actor.

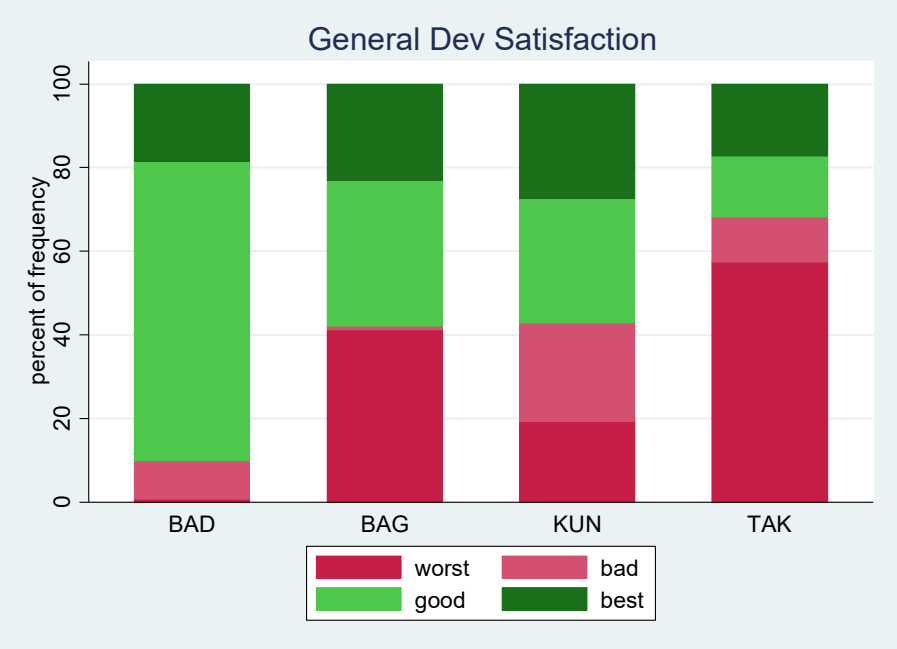
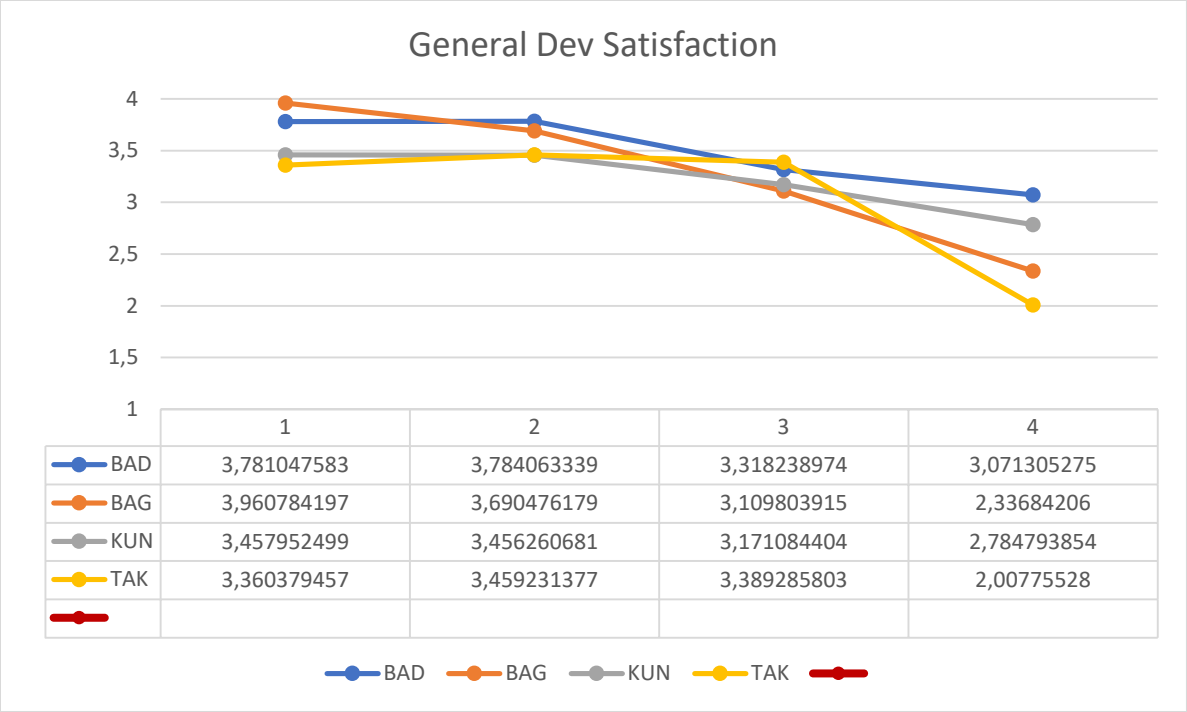


Figure 46: Wave 4, 35 districts - general maximum satisfaction with development in any sector by any actor

The trend across waves for general development satisfaction is negative for all provinces clearly reflecting dropping development funds throughout the survey region (and in general in the country). We register the steepest fall for Takhar between Wave 3 and 4. This is consistent with our non-survey-based development indicators – in Takhar most development work that made sense from a community perspective came to a standstill. Some districts, like Farkhar, received, in absolute terms, more projects than before, but these were mostly gifts of parliamentarians with access to development funds to their local clients and did not follow any developmental logic. According to the debriefing of AKF staff in Taloqan in July 2018 this led to a situation in which it became close to impossible to mobilise communities for any kind of local contribution or participation in development projects. SPNA's small project funds and the O&M component have been negatively affected by this situation.



	1	2	3	4
BAD	3,781047583	3,784063339	3,318238974	3,071305275
BAG	3,960784197	3,690476179	3,109803915	2,33684206
KUN	3,457952499	3,456260681	3,171084404	2,784793854
TAK	3,360379457	3,459231377	3,389285803	2,00775528

Figure 47: All waves, 25 districts - general maximum satisfaction with development

Satisfaction with development associated with the government: SPNA intended to foster stability by improving meaningful cooperation on development between societal institutions representing communities (the DDA-CDC structure) and district-level state bodies. Hence, we are specifically interested in understanding how the state is viewed as a development provider or enabler. Here we are looking at the top mark the state receives in any of the seven sectors we asked for. In terms of positive development associated with the state Takhar and Baghlan fare nearly equally bad, both reaching the worst score (1 fully disagree) as best result in over 60% of observations. In Kunduz, worst scores still stand at around 30% while in Badakhshan less than 10% of respondents fully disagreed with the statement that the government contributed positively to developments in any of the seven sectors.

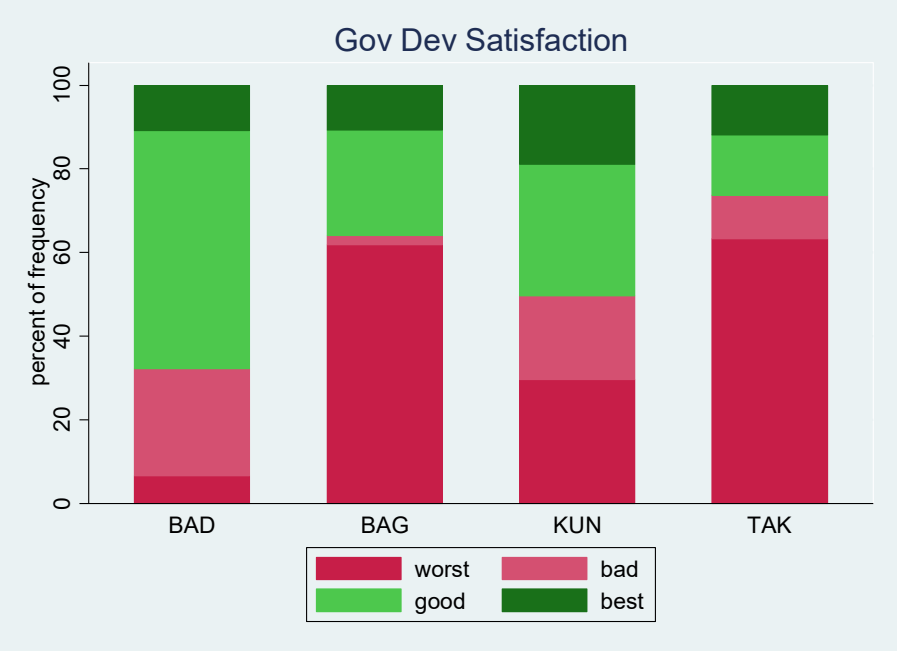
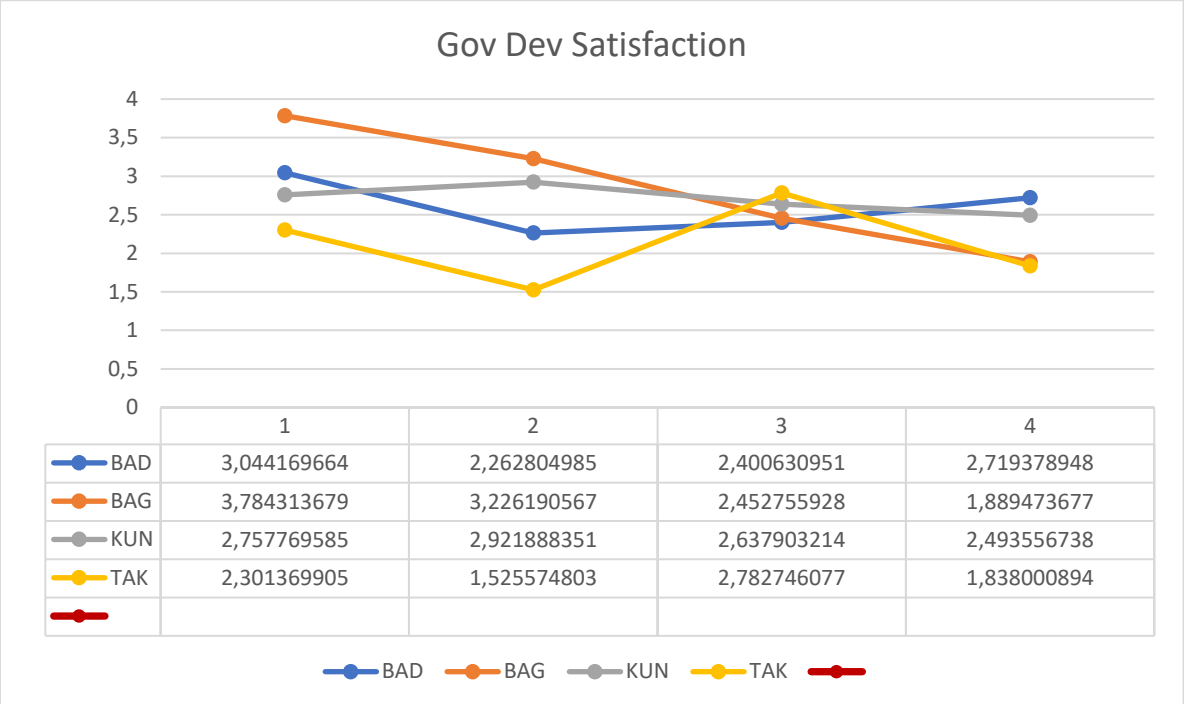


Figure 48: Wave 4, 35 districts - maximum satisfaction with state associated development

The trend over four waves shows a steady decline for Baghlan, and relatively stable medium values for Kunduz. Badakhshan perceptions initially declined (Wave 1 to Wave 2) and steadily recovered since then. Takhar has been the lowest performer on this indicator in all waves except for Wave 3.



	1	2	3	4
BAD	3,044169664	2,262804985	2,400630951	2,719378948
BAG	3,784313679	3,226190567	2,452755928	1,889473677
KUN	2,757769585	2,921888351	2,637903214	2,493556738
TAK	2,301369905	1,525574803	2,782746077	1,838000894

Figure 49: All waves, 25 districts - maximum satisfaction with state associated development; 1 is worst, 4 is best score

Self-reported material well-being of the household: Reducing poverty and improving the material well-being of households is one important dimension of economic sustainability and social stability.

One well-founded reasoning for a positive relationship between increasing material well-being and stability are rising opportunity costs that members of the household would incur if they engaged in destructive behaviour (like the insurgency). The self-assessment of the material status and opportunities of the household has implications for some key decisions relevant for local stability, like work migration or engaging in violent strategies of income generation (insurgency, criminal activities). The participatory and locally informed development approach offered by SPNA may add to poverty reduction at household level, e.g. if it provides economically potentially relevant infrastructure such as roads or bridges, even though this level is not the immediate target of SPNA measures.

In Wave 4 we find that the vast majority of households in our survey districts of Baghlan (82%) and Takhar (73%) consider themselves either very poor or poor. In Kunduz (44%) and Badakhshan (40%) these values are much lower.

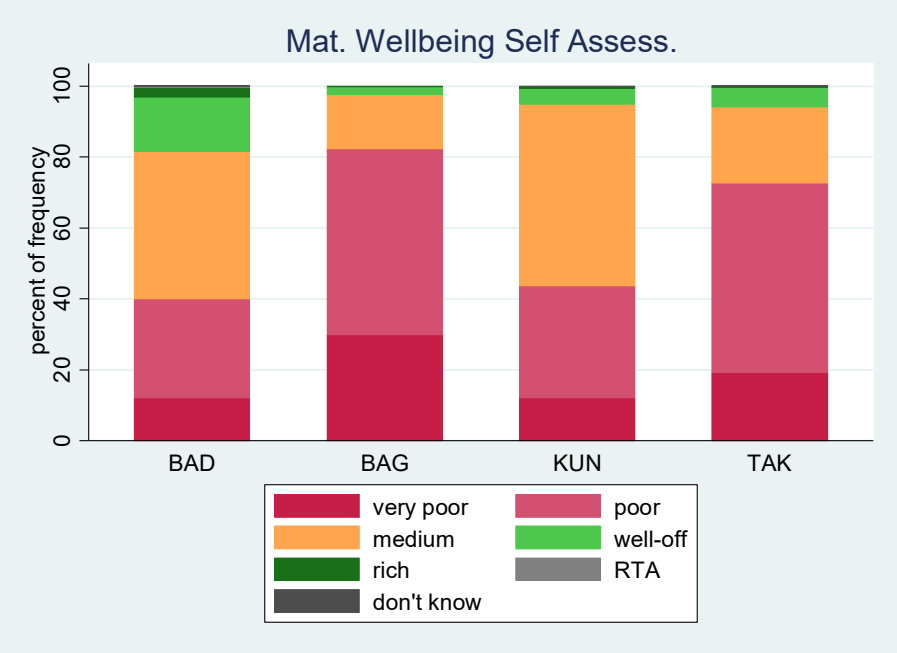


Figure 50: Wave 4, 35 districts – self assessed material wellbeing of the household

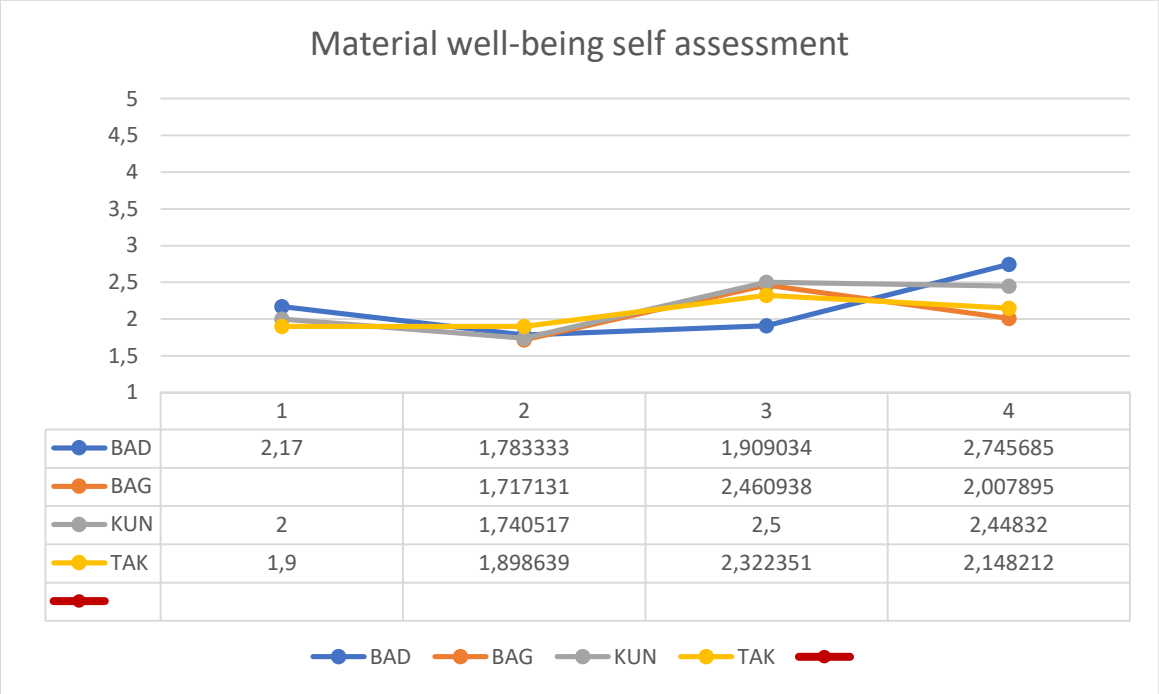
Past research into risks of falling back into civil wars has shown that it is not absolute poverty that increases the risk of renewed violence but rather negative trends, i.e. a deteriorating economy (Collier/Sambanis 2005). While this research looked mostly at macroeconomic data, the underlying causal dynamics explaining this finding relate to a large extent to decision-making at household level. Issues such as decreasing opportunity costs, loss of faith in the state to protect, or the experience of a comparatively small personal economic loss, translate into a stronger political motivation than the prospect of gaining something one does not have ("prospect theory") and are as relevant for household decision-making as they are for larger corporate groups in a society. Hence, we believe that the dynamics of subjective material well-being are particularly important for the stability SPNA hopes to contribute to.

In Wave 1, provinces have roughly similar average levels of material well-being (average value of "2" corresponding to the response option "poor") with little changes in Wave 2[11]. Kunduz, Takhar and

[11] Note that for Wave 1 we only considered the 10 districts of the second cohort for this indicator because of a specific enumerator mistake for this question in the first wave

Baghlan continue to trend together between Wave 2 and 4, however, with a widening spread in Wave 4. Badakhshan shows a different trend – worse for Wave 3 and then with a steep increase in Wave 4. This corresponds to the non-survey-based indicators reported above, and most likely relates to the positive changes in the informal as well as illicit economy in this province (trade in hashish, opium and gemstones).

What is also apparent is that our impression reported in earlier waves is continuing in Wave 4 – important dimensions of the "real" economy are independent of the international presence and the development sector backed by it (see above, development vs. "real" economy).



	1	2	3	4
BAD	2,17	1,783333	1,909034	2,745685
BAG		1,717131	2,460938	2,007895
KUN	2	1,740517	2,5	2,44832
TAK	1,9	1,898639	2,322351	2,148212

Figure 51: All waves, 25 districts – self-assessed material well-being of the household; 1 is worst, 5 is best score

Access to land: We ask respondents how they would describe their household – as landless, as cultivator, landowner or landlord. We use the self-description "landless" as a proxy for poverty; in rural Afghanistan access to land on household level remains the principle resource depicting social and material status. No access to land is a strong indication for poverty and a precarious subsistence.

In Wave 4 Badakhshan has the lowest rate of landless households (11%) and Takhar at 21% the highest rate.

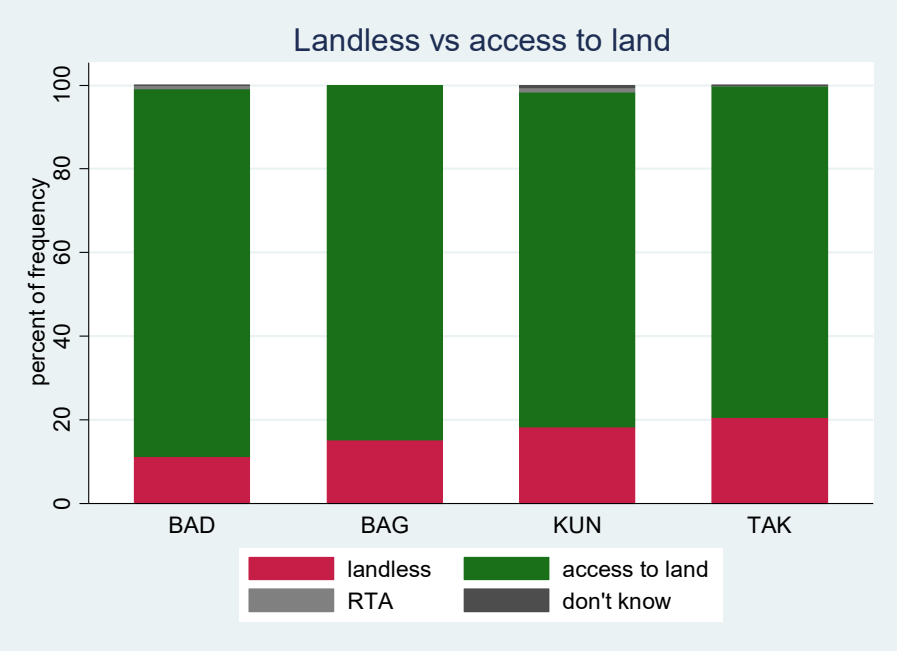
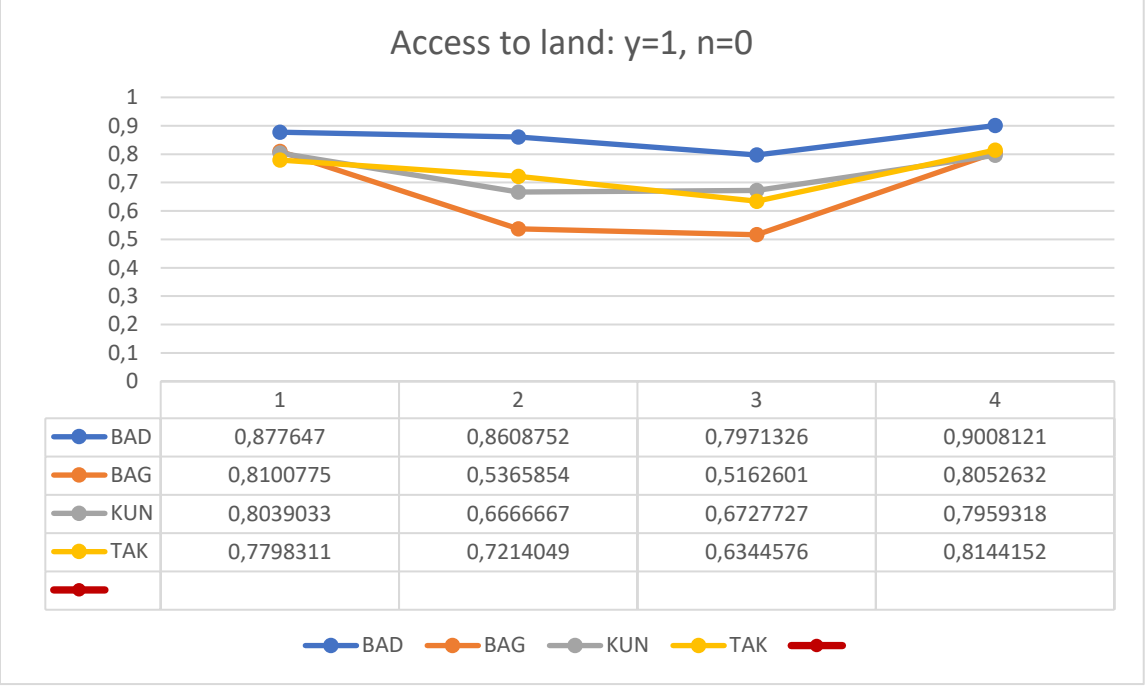


Figure 52: Wave 4, 35 districts – household self-characterisation by access to land

In terms of trends Takhar, Baghlan and Kunduz start at roughly similar levels in Wave 1 and then trend together, first with an average decline of reported access to land by 10% (Kunduz, Takhar) to 25% (Baghlan) in Wave 3 and then recovering to the baseline level in Wave 4. Badakhshan reported consistently higher levels, but also initially declines by nearly 8 percentage points only to recover to recover in Wave 4.



	1	2	3	4
BAD	0,877647	0,8608752	0,7971326	0,9008121
BAG	0,8100775	0,5365854	0,5162601	0,8052632
KUN	0,8039033	0,6666667	0,6727727	0,7959318
TAK	0,7798311	0,7214049	0,6344576	0,8144152

Figure 53: All waves, 25 districts - household self-characterisation by access to land; 1 is best, 0 is worst score

Implications for SPNA impact hypothesis

In contrast to more mainstream development programmes, SPNA development goals are no end in themselves, but are viewed in the context of the wider stabilisation effort. Economic development does not always and necessarily foster district-level stabilisation. Local economic development often has stabilising effects, but it can be divisive, it may be driven by criminal economic activities and destabilising entrepreneurs (as is the case with parts of the drug economy, the illicit gemstone business, or cross-border smuggling of illicit goods). SPNA believes it exercises a stabilising impact on local economic development, by facilitating development via capacitated DDAs applying informed, fair and transparent procedures. Hence, we focus on how development progress is perceived, and how the state is perceived as a development enabler.

We separate the development-related dimension of the local economy from the two household-level self-reported indicators of material well-being (i.e. self-assessment of wealth and self-reported access to land). We believe that successful and stabilising SPNA implementation is likely to have a positive effect on household poverty levels and fairer access to agricultural land. The latter is a key indicator for the level of subsistence of households in rural Afghanistan; no access to land of a household is a strong indication of dire poverty.

Economic development is a key dimension of SPNA-intended stabilisation. Positive assessments of development have been declining between baseline and endline, while the attribution of positive development to the government remained more even in most provinces. The indicators for material well-being and poverty stagnated or improved with improvements being most pronounced in Badakhshan. To assess the contribution of SPNA to these trends we need to analyse them in the context of other factors influencing perceptions of development and material well-being. In the analytical part we will focus on the effects SPNA has on positive perceptions of development and the material well-being of households, while controlling for village-level development initiatives and general economic trends that are not directly linked to SPNA.

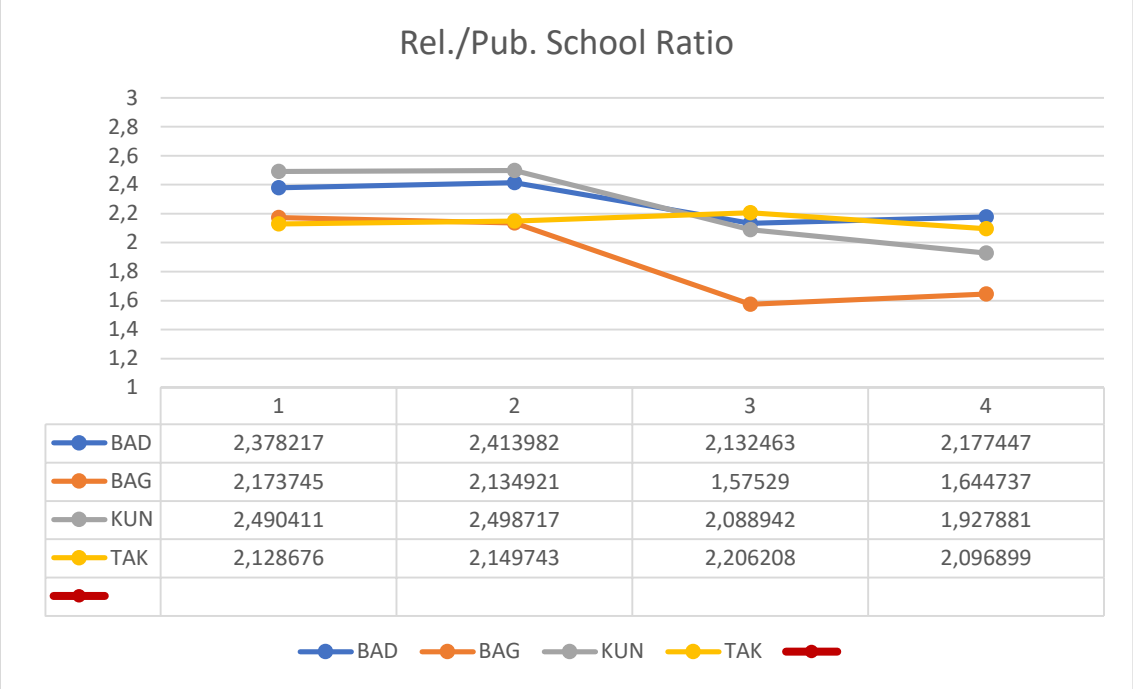
ADAPTATIONS/MODERNISATION

SPNA's impact hypothesis: We assume that the SPNA programme components improve openness and competence to adapt to new "modern" technologies, information, as well as organisational forms. This is achieved via participatory approaches to development in which the DDA is an increasingly legitimate and competent two-way "transmission belt" for development initiatives between the state and NGOs on the one side and communities on the other side.

Context indicators: public school/religious school ratio

The **capacity** and **willingness** of affected communities to **accept change** and **adapt** to new framework conditions is an important aspect of development interventions. This is not only a technical challenge, but also an ideological issue, particularly in Afghanistan where the **normative interpretation of change** is communicated via different traditional and modern media and where development interventions themselves are part of the opposing normative discourses. Key participants in these discourses are different insurgency groups, different parts of the government, and different traditional and modern parts of civil society as well as the development agencies themselves. This contest of interpretation has been intensifying over the past eight years (we noted an increase of more general ideological interpretations of the international presence in the guideline interviews of 2012 when discussing benefits and risks of the international military and development presence in the region).

We use the ratio between (official) public schools and (mostly informal) religious schools (*madrassas*) per village community as a simple indication of modernisation that is independent of subjective survey responses. We coded this ratio as (1) for villages with more religious schools than public schools, (2) equal, (3) more public than religious schools.



	1	2	3	4
BAD	2,378217	2,413982	2,132463	2,177447
BAG	2,173745	2,134921	1,57529	1,644737
KUN	2,490411	2,498717	2,088942	1,927881
TAK	2,128676	2,149743	2,206208	2,096899

Figure 54: All waves, 25 districts – coded ratio of public vs. religious schools per village; 1 is even

With the exception of Takhar we observe a decline in the public vs religious school ratio to the detriment of secular schooling between Wave 2 and 3. This decline is steepest in Baghlan, a low performer already at baseline level. Takhar remains stable and, hence, in relative terms moves from lowest performer in Wave 1 to second best in Wave 4. The decline is most pronounced in districts that are at least partially under Taliban control. Data we received from district education departments also confirm that by the time of the endline, the school enrolment ratio for girls is highest in Badakhshan (close to 50% of total enrolment), followed by Takhar (about 45%). Kunduz and Baghlan are below 40% (see annex for details).

Critical discussion of the indicator: The indicator corresponds to recent reports about a general trend of declining public school attendance after years of substantial increase for both boys and girls (Ministry of Education Islamic Republic of Afghanistan et al. 2018). The changing balance between religious and public schools at village level, particularly in areas affected by the Taliban, does not necessarily mean that fewer children attend schools altogether but it does indicate a turn towards limited religious education and conservative religious values.

Stabilisation indicators: value threat and disagreement with curriculum

We selected two indicators to illustrate attitudes related to modernisation: the perception of development interventions as a threat to local values and Islamic norms; and disagreement with what the children learn at public schools.

Development as value threat: We ask the respondents to tell us how much they agree with the value statement that development aid is a threat to local values and Islamic norms, while it may bring material benefits for the community. We want to capture to what extent people are open to development induced change. For the impact analysis further below, we focus only on the most negative state – fully agree - since rather agree may simply indicate a more differentiated view on the quality of development implementation (even few development enthusiasts would dispute that poorly implemented development aid may be a threat to values and norms). But since the value statement itself is rather uncompromising, fully agree would indicate a more general rejection of development-induced change.

In Wave 4 Badakhshan (2% fully agree) shows the lowest level of value-threat perceptions, Takhar the highest (17%). In Badakhshan we have unusually high numbers of invalid responses (don't know) mostly driven by districts (Yamgan and Wardooj) or parts of districts (Kishim and Jurm) under Taliban control (40% of respondents in those areas chose don't know, and nearly 30% in contested areas as compared to the survey average of 13%). Hence, the positive impression might have a bias.

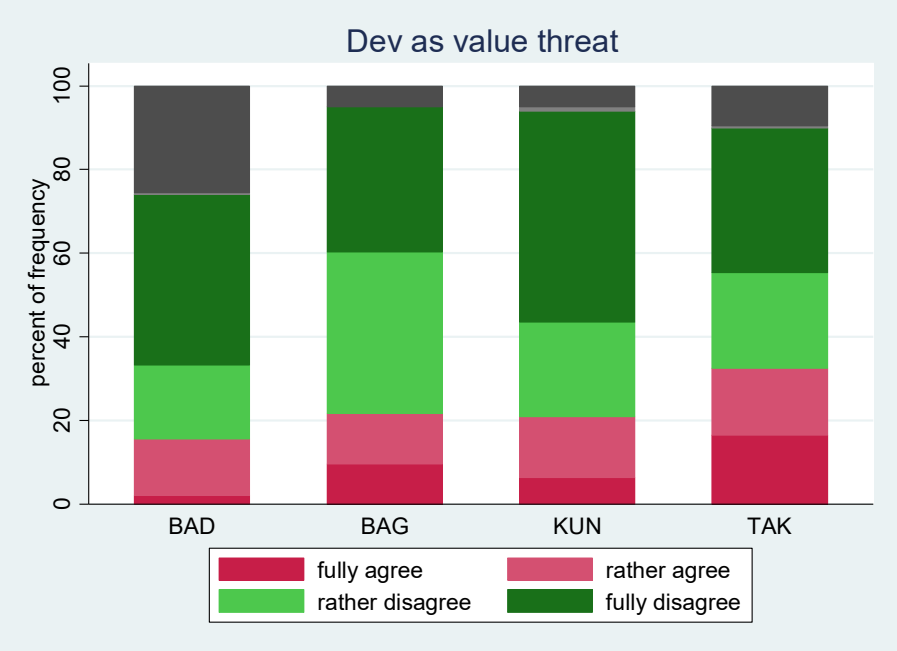
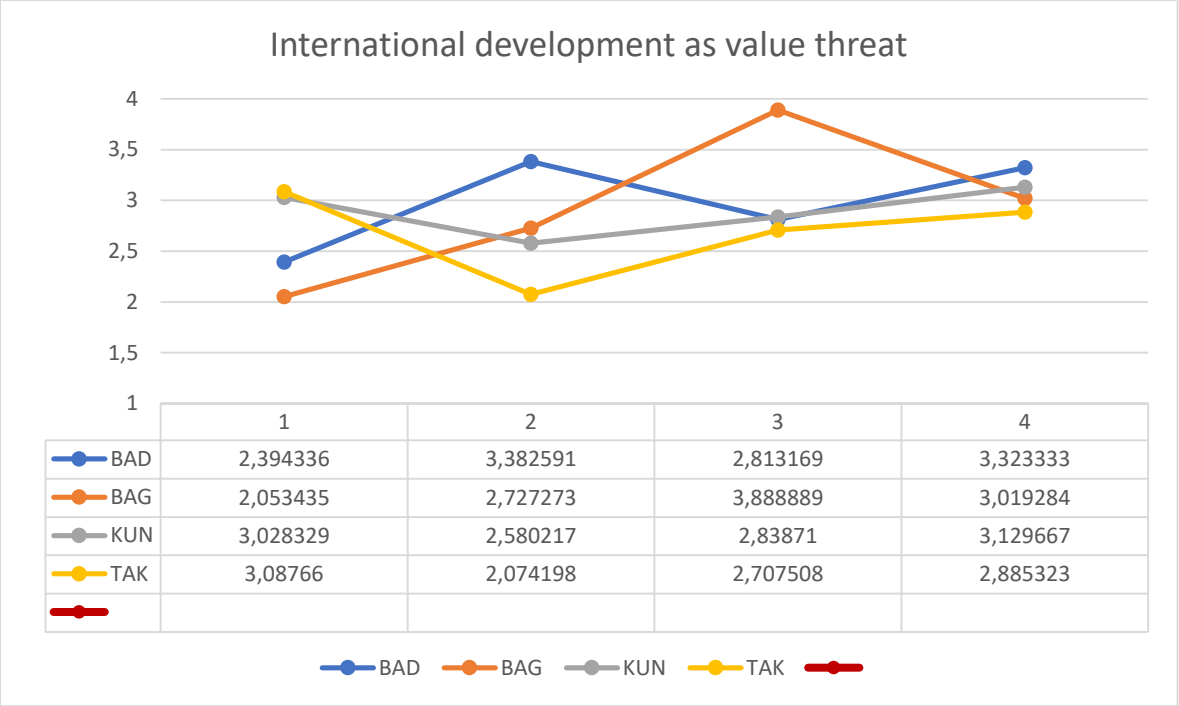


Figure 55: Wave 4, 35 districts – development aid perceived as threat to Islamic norms and local values

The trends across waves for the value threat indicator are inconclusive. Takhar and Kunduz trend together with an initial increase in value threat perceptions between Wave 1 and 2 and a steady decline since then. Badakhshan first shows an improvement from high threat levels in Wave 1, which worsens between Wave 2 and 3 only to improve again in Wave 4. Baghlan (i.e. Baghlan-i Jadid) starts off with the worst average values for Wave 1, increases to top scores in Wave 3 and drops again in Wave 4. Average results across all valid (i.e. don't know and refused to answer are excluded) response options 1-4 are converging for Wave 4 in the four provinces when we consider only our 25 core survey districts.



	1	2	3	4
BAD	2,394336	3,382591	2,813169	3,323333
BAG	2,053435	2,727273	3,888889	3,019284
KUN	3,028329	2,580217	2,83871	3,129667
TAK	3,08766	2,074198	2,707508	2,885323

Figure 56: All waves, 25 districts – development aid perceived as threat to Islamic norms and local values; 1 is worst, 4 is best score

Disagreement with school curriculum: We use the answer "I do not agree with what the children are taught at [public] school" to the question asking the head of household to evaluate how much the children learned at school as indicator for openness to modern education.

In Wave 4 general disagreement with the state school curriculum is low (between 4% in Badakhshan and 7% in Takhar), except for Baghlan, which scores a staggering 24%.

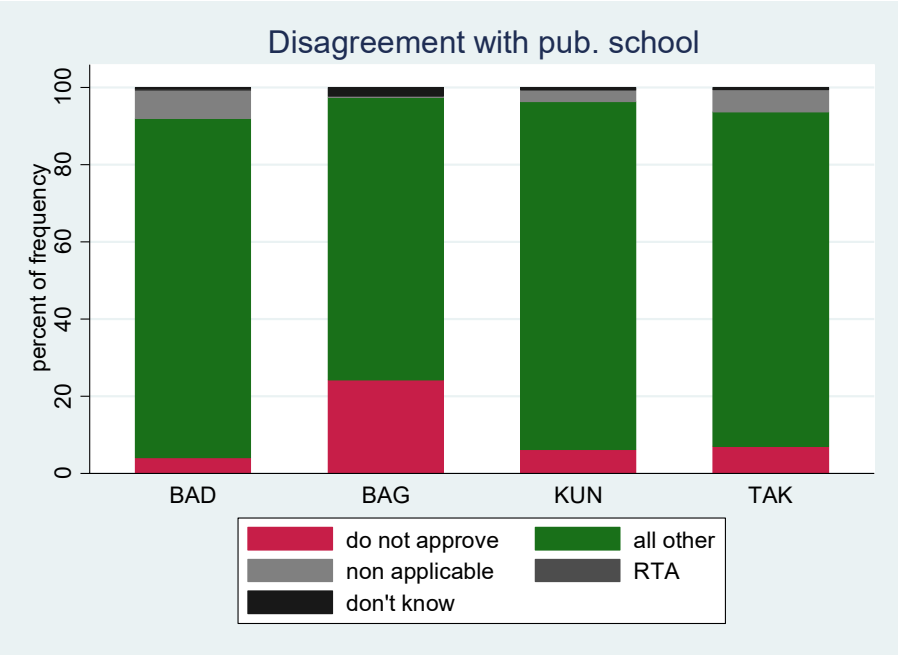
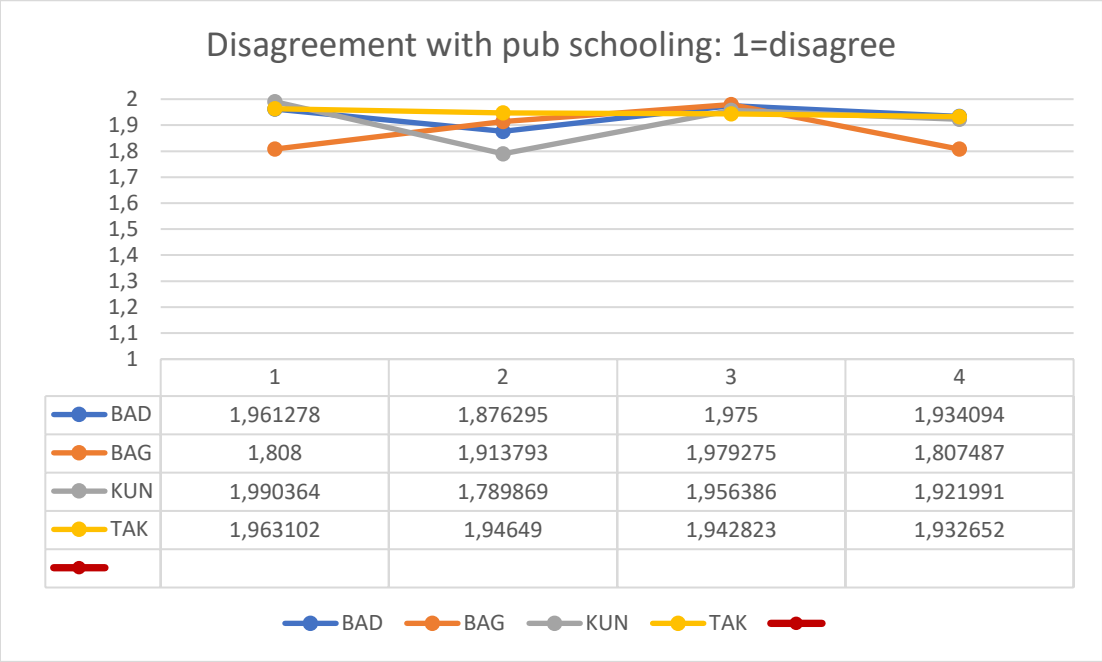


Figure 57: Wave 4, 35 districts – disagreement with public school curriculum

In terms of trends, we find the highest spread between average responses in Wave 2 and the lowest level of negative opinions about the public school curriculum in Wave 3. Baghlan in Wave 1 and 4 and Kunduz in Wave 2 mark the low points.



	1	2	3	4
BAD	1,961278	1,876295	1,975	1,934094
BAG	1,808	1,913793	1,979275	1,807487
KUN	1,990364	1,789869	1,956386	1,921991
TAK	1,963102	1,94649	1,942823	1,932652

Figure 58: All waves, 25 districts – disagreement with public school curriculum; 1 is worst, 2 is best score

Implications for SPNA impact hypothesis

Since the beginning of the international intervention in Afghanistan, the population in the target region of SPNA has been exposed to rapid changes that have, to a large extent, been induced by international development initiatives. Two district governors told us in separate interviews during the baseline research for SPNA in similar words: "My district saw more development in the past 10 years than in the 100 years before this". Both our surveys and the guideline interviews show that development is in general highly appreciated. This relates not only to the technical aspect of development – like improved access to services via better infrastructure – but also to some of the ideologically more contested consequences of development, like girls' access to schooling, off-farm work opportunities for women or a more secular public school curriculum (see annex for the trends of these indicators since the SPNA baseline).

At the same time SPNA has been implemented under conditions of increasing religious conservativism in many districts covered. Religious conservativism is not only driven by the Taliban, but also by former Jihadi commanders and representatives of the local ulema. It connects increasingly to popular discourses via social media that reach beyond the rural confines of access to information (access to smart phones and television is on the rise). In the guideline interviews conducted since the baseline, we see that the moral impact of increased access to information, mobility and some aspects of modernisation are critically discussed (a recurrent issue is loss of traditional control over young women due to mobile phones, more liberal dress codes, bad outside influence on manners and obedience of young people in the household). The guideline interviews also show an increase in more general critical discussion of international (i.e. Western) intentions and ulterior motives in their

presence in Afghanistan. We further notice a decline (from very high levels) in the aforementioned indicators for openness to girls' school education and off-farm labour towards the endline.

The SPNA approach is special in that it processes development projects through an (indirectly) elected body of community representatives. The implementing agencies are in regular exchange with this body through training and capacity-building measures. This opens a line of communication into communities that has the potential to accompany (foreign funded) development-induced change in more credible, meaningful and reciprocal ways than would be the case for less participatory approaches.

For the subsequent statistical modelling, we chose two variables that indicate a more radical rejection of development-induced modernisation: full agreement with the statement that international development is threatening local values and Islamic norms, and a general rejection of what is taught in public schools. In the analytical section we will test if more exposure to immediate SPNA outcomes tends to soften these points of view.

Note that we are using the term "modernisation" and "modernising development" here with an intentional bias, referring to the internationally-backed state-building intervention that started after the fall of the Taliban regime. As of the time of writing, both the Afghan Government and its international backers for the most part still follow this Weberian blueprint of modern national states, despite the practical challenges it encounters. In fact, the ideological commitment of a vision of a state adhering to the principles of "good governance" is under the Ghani Government much stronger than under the Karzai Administration.

INDICATORS FOR SPNA'S IMMEDIATE OUTCOMES

SPNA has two main outputs – active and capacitated DDAs and visible and operational infrastructure. According to SPNA's theory of change it is these immediate programme outputs that are effecting stabilisation positively.

Following the discussion of context and outcome indicators in the four stabilisation fields, we now turn to perception indicators directly related to SPNA. As was discussed, SPNA outputs divide into two principle components – capacity-building for DDAs and the provision of funds and expertise for the implementation of infrastructure development projects identified by the DDAs as most relevant for the district. We identified three survey-based indicators for immediate intended outcomes for each component.

SPNA capacity-building component	**SPNA infrastructure component**
DDA visibility Do the respondents know about the DDA?	*Infrastructure visibility* Do the respondents know about the SPNA infrastructure projects implemented in their district?
General assessment of the DDA	*District level assessment of the infrastructure*

Is the DDA good, neutral or bad for the development of the district?	Is the [infrastructure project XYZ] of benefit for the development of the district?
Sector specific assessment of the DDA Has the DDA contributed to improving [sector 1-7]? The sectors are: drinking water, agricultural production, roads, jobs, electricity, schooling, medical services.	*Household level assessment of the infrastructure* Is the [infrastructure project XYZ] of benefit for the household?

Table 4: Overview of survey-based indicators of SPNA performance

Regarding (a) the **DDA capacity-building component,** the three perception-based indicators are the following: (i) the visibility of the DDA, (ii) the general assessment of the DDA regarding its developmental role, and (iii) the contribution of the DDA to seven development sectors. Concerning (b) the **SPNA infrastructure component**, the respective perception-based indicators are (i) the visibility of SPNA-provided infrastructure, (ii) the assessment of SPNA-provided infrastructure on the district level, and (iii) the assessment of SPNA-provided infrastructure on the household level. Note that survey responses to the above SPNA project indicators were only included in Wave 2 in the questionnaire. A timeline of this indicator is thus only available for Waves 2-4.

In our statistical models, these SPNA project indicators will serve as predictor variables of the programme. Using variance in the perception of SPNA project outcomes (knowledge of and evaluation of project measures) is one way to overcome the already discussed challenge of a lacking control group of districts that were not treated by SPNA (see *The challenge for causal inference*, above).

CAPACITY BUILDING COMPONENT

The principle partner of the SPNA approach is the DDA. It consists of community representatives elected in CDC elections and seconded, via the intermediate body of the CLDC, to the DDA. Capacitating the DDA is the key to achieving all three programme objectives: (1) to provide basic development infrastructure, enhancing access to services and livelihood opportunities; (2) to enhance capacity among local governance institutions for good governance and development planning; (3) to strengthen collaboration between and among institutions within the sub-national governance system.

Visibility of the DDA

In Wave 4 DDA visibility is high in Kunduz (87%), Baghlan (84%) and Badakhshan (77%). In Takhar the situation is different: here, after seven years of SPNA implementation with a focus on the DDA, only 54% of respondents have heard about the DDA.

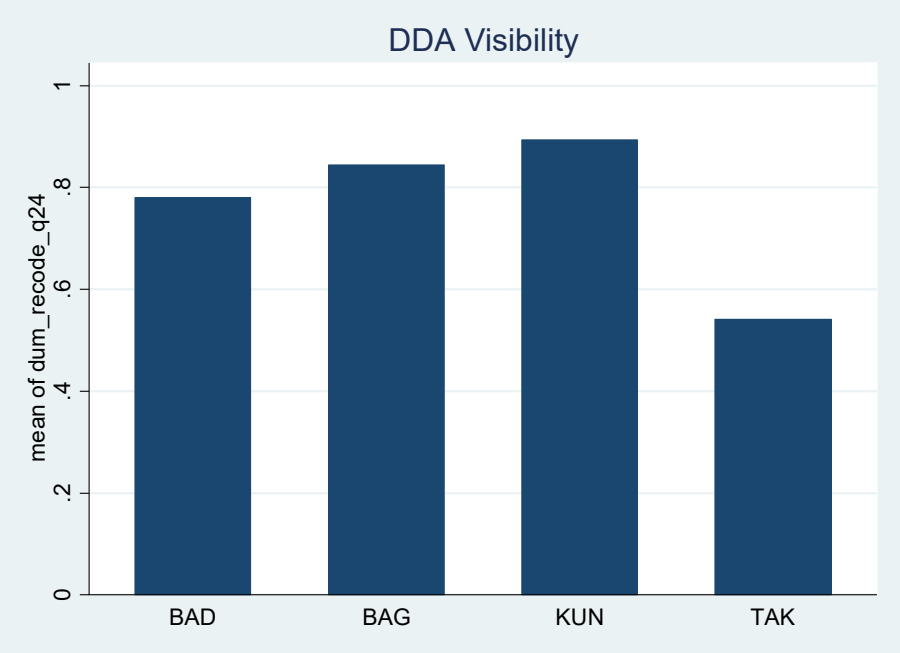


Figure 59: Wave 4, 35 districts – visibility of the DDA

The trends across waves show that DDA visibility was very low (43%) already in Wave 2 when we first asked the SPNA specific questions. Hence, in terms of the increase in visibility Takhar performs not worse than the other provinces. Kunduz, Takhar and Baghlan trend together, with a slight decrease in visibility between Wave 2 and 3 and a strong increase from Wave 3 to 4.

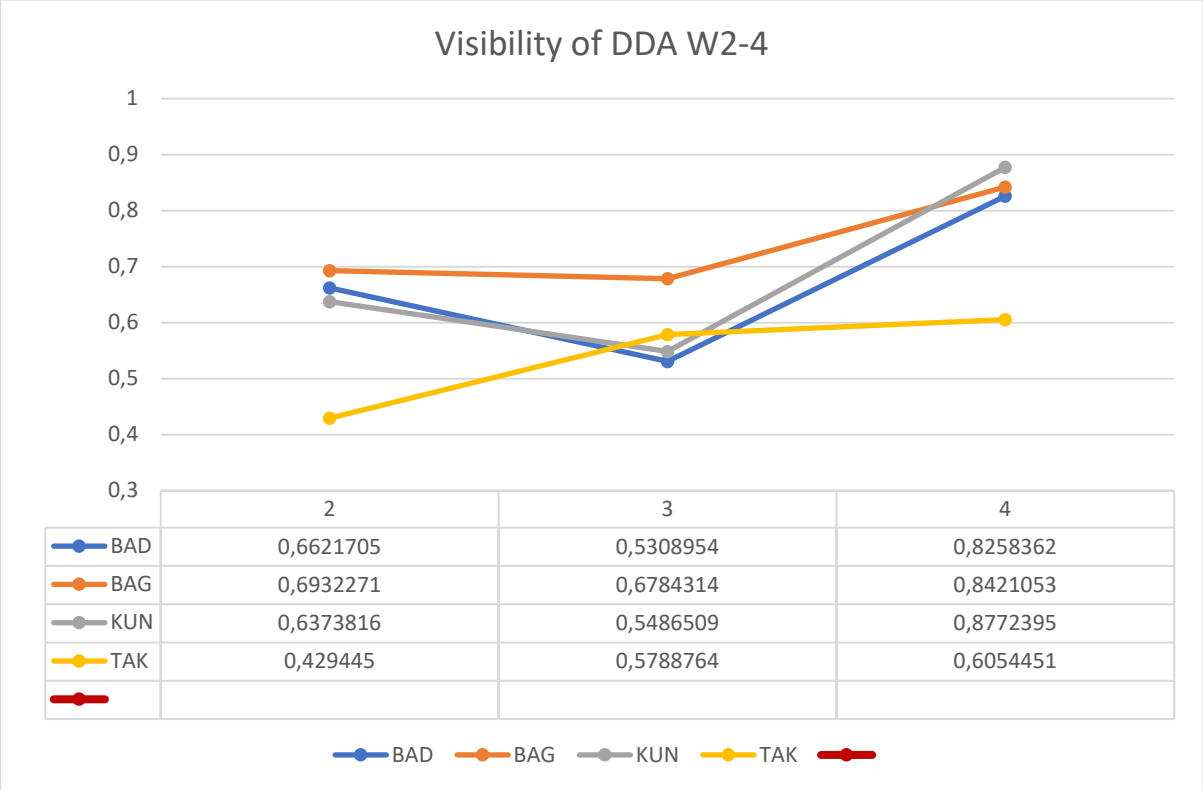


Figure 60: Wave 2-4 - visibility of the DDA

General assessment of the DDA's role for development of the district

While knowledge about the DDA's reached its highest scores in Wave 4, the general assessment of the DDA's contribution to development of the districts by those who had heard about them declined. In Badakhshan only 23% of respondents felt that the DDA contributed positively, and in Takhar 36%. The scores are decidedly better in Baghlan (68%) and best in Kunduz (82%).

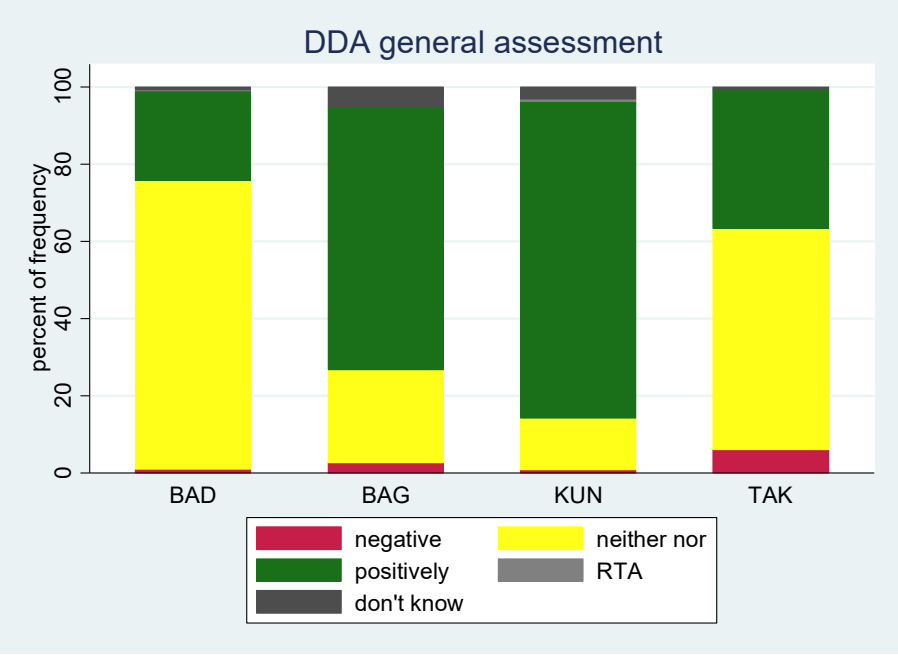


Figure 61: Wave 4, 35 districts – general assessment of the DDA impact on district development

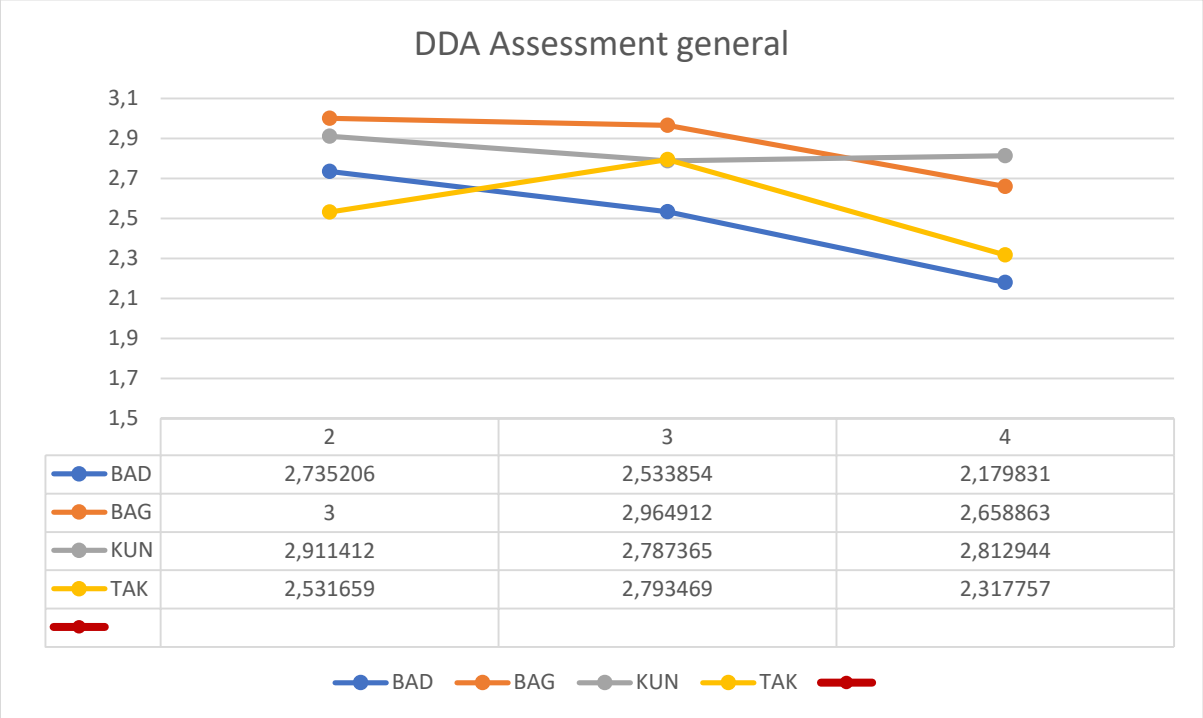


Figure 62: Wave 2-4 – general assessment of the DDA impact on district development; 1 is worst and 3 is best score

Sector-specific assessment of DDA's role in development

The assessment of the DDA's influence on district-level development is much more conservative when respondents are asked about its contribution to seven specific sectors of rural development. Takhar (83% fully disagree that the DDA contributed to any of the 7 sectors) and Badakhshan (74%) are most sceptical, followed by Baghlan (59%) and Kunduz (41%).

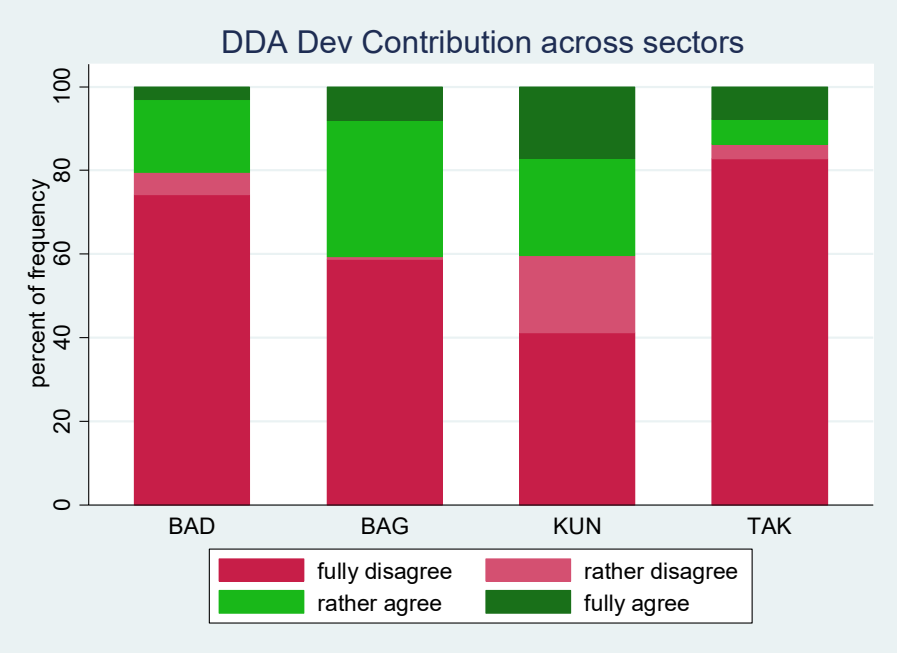


Figure 63: Wave 4, 35 districts – assessment of the DDA by sector

Across the waves, this rather negative assessment has been stable on average; however, the spread was greatest in Wave 2 (with positive assessments dominating for Kunduz and Baghlan, very negative for Takhar and Badakhshan on average close to option 2 "rather disagree"). In Wave 3, the assessments converge between provinces while Wave 4 sees slight improvements for Kunduz and Baghlan and a negative direction for Badakhshan and, strongest, for Takhar. Compared to Wave 2 the overall trend at the Endline is moderately negative.

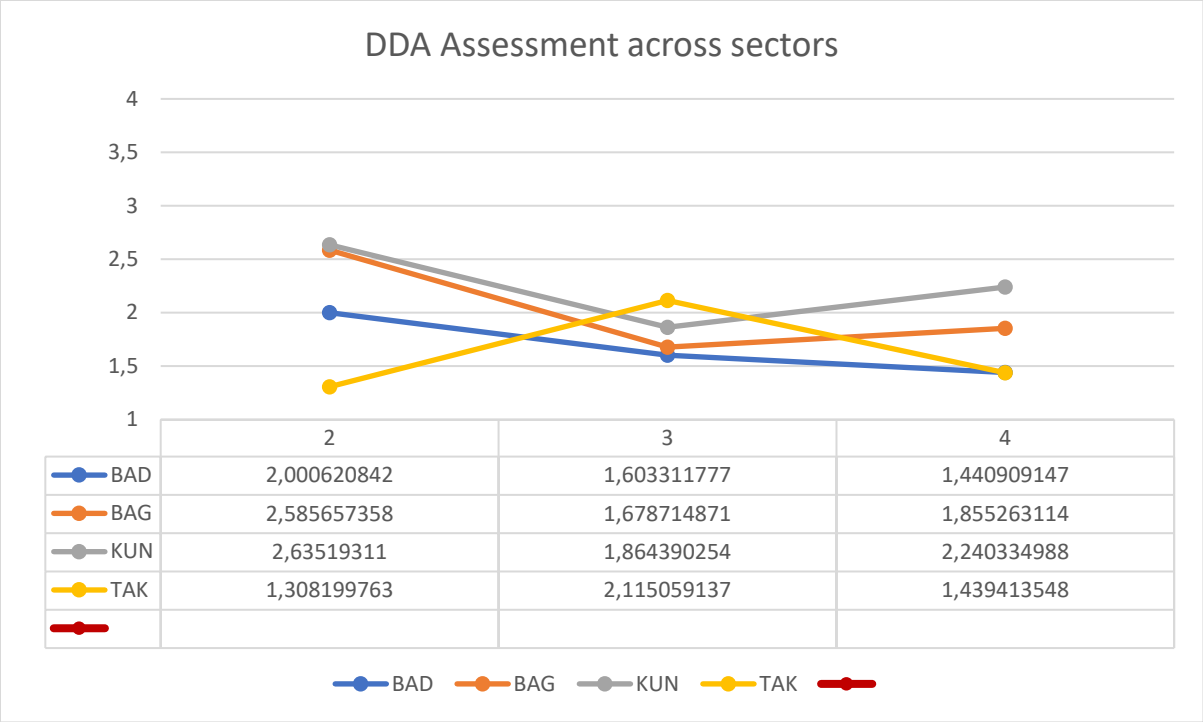


Figure 64: Wave 2-4 – assessment of the DDA by sector; 1 is worst and 4 is best score

Resume for the capacity building indicators

In Kunduz the DDA developed into a rather powerful organisation crossing the border between Taliban and state governance; they are still involved with development issues but also fulfil other, more political, governance functions. DDA representatives are often respected elders of different ethnic communities who also draw on more traditional sources of prestige and social standing. The Taliban seem to take them seriously and work with them. On occasion, Taliban representatives for NGO/development work are embedded with the DDA.

In Takhar, the DDA shura structure seems to be mostly defunct itself and development work has all but stopped after its peak in Wave 3. This is clearly linked to the fact that SPNA implementation in the core survey districts all but stopped in this province after Wave 3, and in several cases even before (see Figure 4). In Badakhshan, DDAs are mostly operational (with the notable exceptions of the DDAs in Taliban-controlled Wardooj and Yamgan) but here, too, SPNA infrastructure project implementation stopped around 2014-15 in six out of the nine original survey districts. DDAs continued to implement SPNA's O&M component as well as other AKF (Afg)-facilitated smaller-scale activities (such as the Health Action Plan for Afghanistan, HAPA, or the Market Development Programme, MDP), but have, for the most part, stopped implementing large-scale and visible infrastructure development projects. Thus, as of Wave 4, they are more of a governance and political than a visible development actor.

In Baghlan, the four SPNA districts assessed perform very differently from one other. In Baghlan-i Jadid the DDA had long been captured by divisive and violent group of local commanders with central backing, though this has changed in the last two waves. Nonetheless, the track record of the head of DDA as a committed development actor is tainted. In Pul-i Khumri, the SPNA implementing partner

even consciously sidelined the DDA for reasons the consultants did not fully comprehend (the head of the DDA complained about this fact; the decision was confirmed during the debriefing of SPNA staff involved in implementation). In Khinjan the DDA seems to be highly active and, according to the SPNA implementing partner AKF, comprises almost a show case institution for district development. The reasons given relate to successful coordination with the district administration and police (the head of police was quoted as jokingly demanding to stop further capacitation of the DDA because no work will be left for the police) and to successful development project acquisition. Some members left the DDA to establish their own commercial construction company. Pul-e Hesar, finally, is a very special district that is totally dominated by traditional local feuds that, while accepted, paralyze any kind of collective decision-making, planning or strategic sustainable development.

INFRASTRUCTURE COMPONENT

Successfully implementing infrastructure that is relevant for the development of the wider district is a complex task for the district administration as well as local self-governing bodies like the DDA. Introducing and enhancing a procedure of identifying, prioritising, and selecting relevant projects, monitoring the construction, handing over the object to the responsible line ministries and beneficiaries, and maintaining the infrastructure has been the backbone of the SPNA approach.

The success of this component integrates with and depends on the first component, i.e. capacitated DDAs that understand and fulfil their mandate and are able and willing to cooperate with other state, societal and commercial (contractors) development actors in their districts. In consequence, at baseline, during the first phase of project identification, the DDAs were not yet ready to take full ownership of the procedure. Hence, in Phase I, many projects were de facto suggested and chosen by the implementing organisation, i.e. AKF (Afg), ACTED or Mercy Corp. This changed in Phase II, after the first rounds of training had been accomplished.[12]

Up until 2012, most projects identified fell into the sector of education (schools and teacher training colleges). Thereafter, projects became more diverse, in part owing to the advice of the implementing partners. At the time of the endline in late 2017, projects in the sector of education had been implemented in 30 districts, road infrastructure in 20 districts, health infrastructure in 11 districts, and one in agriculture (irrigation). In addition, the provincial centres received social and governmental infrastructure projects (government buildings, public library).

Owing to the inclusive and participatory procedure for project identification and implementation we define project visibility within the district and the assessment of the projects as useful for the development of the district as important indicators for immediate SPNA achievements.

Visibility

In Wave 4, with the exception of Badakhshan less than 50% of respondents had heard about SPNA projects already implemented in their districts. At 30% it is lowest in Kunduz. For Badakhshan we find the highest visibility at 65%.

[12] Note that this was the implementation theory; the two phases are only a rough orientation for the degree of participation of the DDA during the first block grant delivery; we used the reports, journals and debriefings with social mobilisers, field engineers and trainers to fine tune our coding of more and less participatory project identification (see moderating variables below).

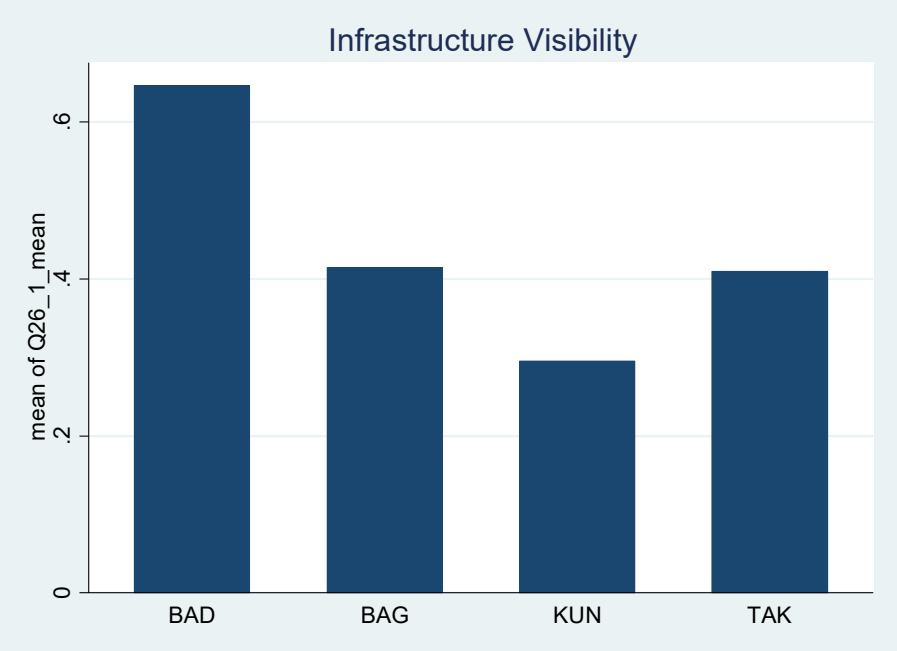


Figure 65: Wave 4, 35 districts – infrastructure visibility

In Badakhshan, SPNA has been more visible than in the three other provinces across all waves. However, project visibility declines across all four provinces from Wave 3 to Wave 4, but least so in Badakhshan. This decline is not surprising if we consider that that the trend analysis here is based on our 25 original baseline SPNA districts. Of those, only three continued to implement infrastructure projects in Phase V of SPNA implementation. It seems that projects implemented in earlier phases slowly fade from public awareness (which is not necessarily the case since we are simply asking if people are aware of a specific infrastructure item – as school, a road, a hospital; if the infrastructure remains in use, awareness could also rise or stay stable, as seems to be the case for Badakhshan).

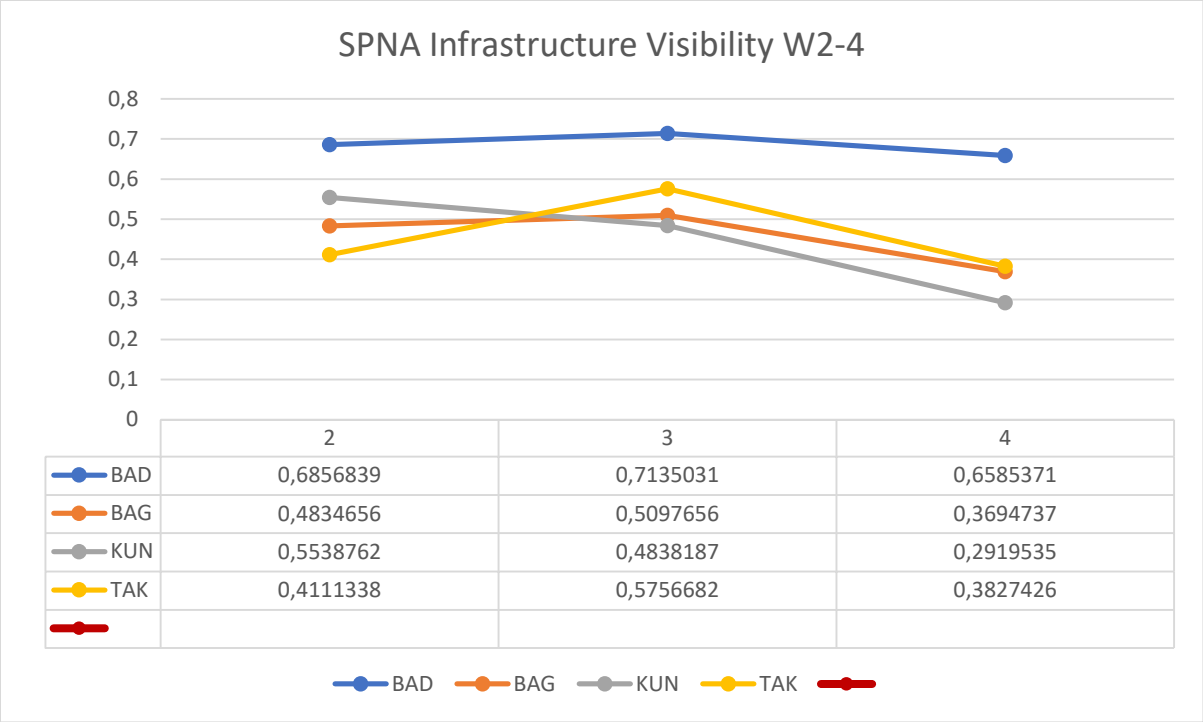


Figure 66: Wave 2-4 – infrastructure visibility

Infrastructure assessment, district level

Of those who are aware of the projects, a vast majority (above 98%) thinks that the projects have been beneficial for the development of the district. Hardly anyone reports any negative impact. While this is positive for SPNA as a whole, suggesting a balanced and needs-oriented selection of projects, it is a problem for the statistical analysis, which needs variance for its predictors to work.

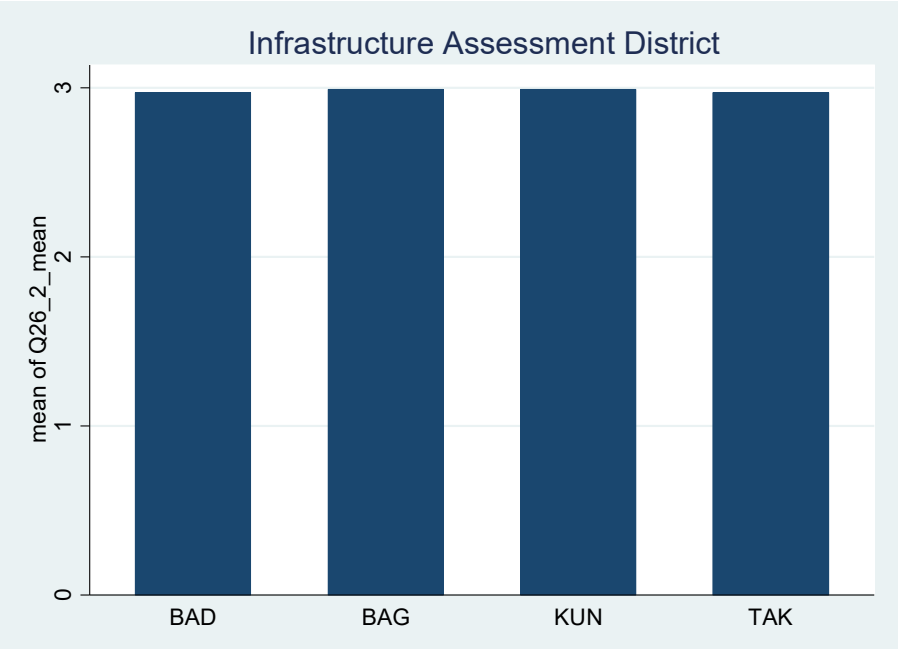


Figure 67: Wave 4, 35 districts - infrastructure assessment district level

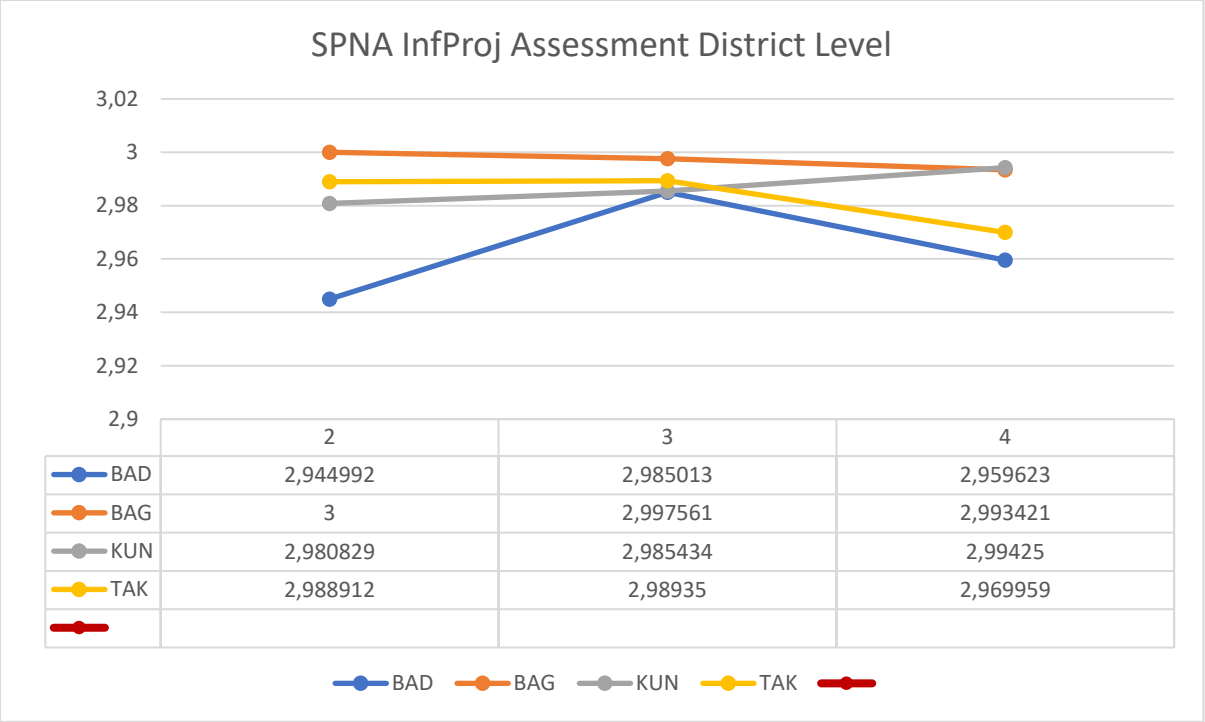


Figure 68: Wave 2-4 – infrastructure assessment district level; 1 is worst and 3 is best score

Infrastructure assessment at the household level

The assessment of SPNA infrastructure projects on the household level, meaning whether a household assessed that it benefitted from an SPNA-funded infrastructure project, shows more variation but is still very positive. It is best for Kunduz (scoring 2.7 on a scale form 1 negative to 3 positive). In comparison, the other provinces reach an average of 2.4.

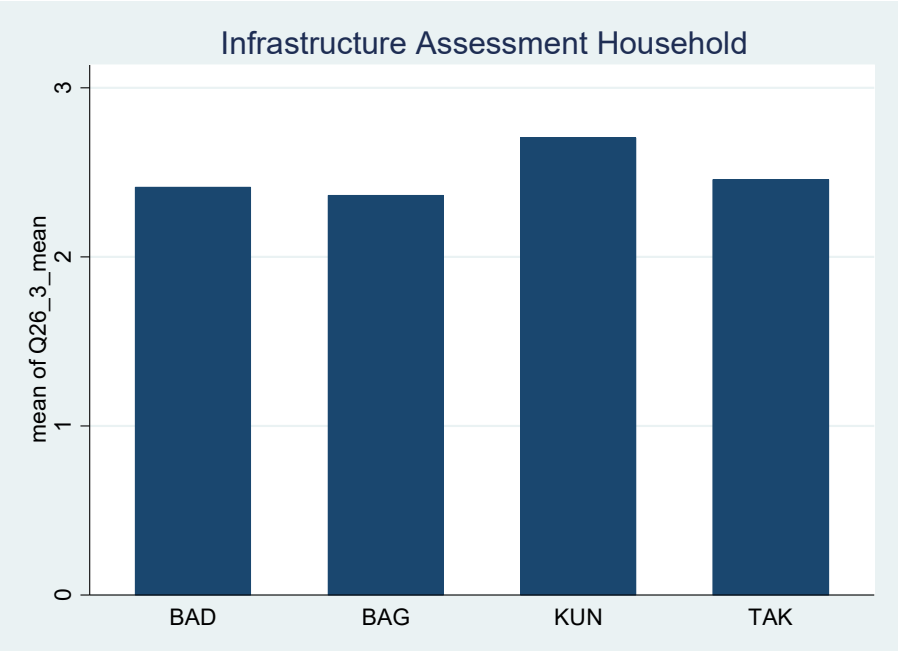
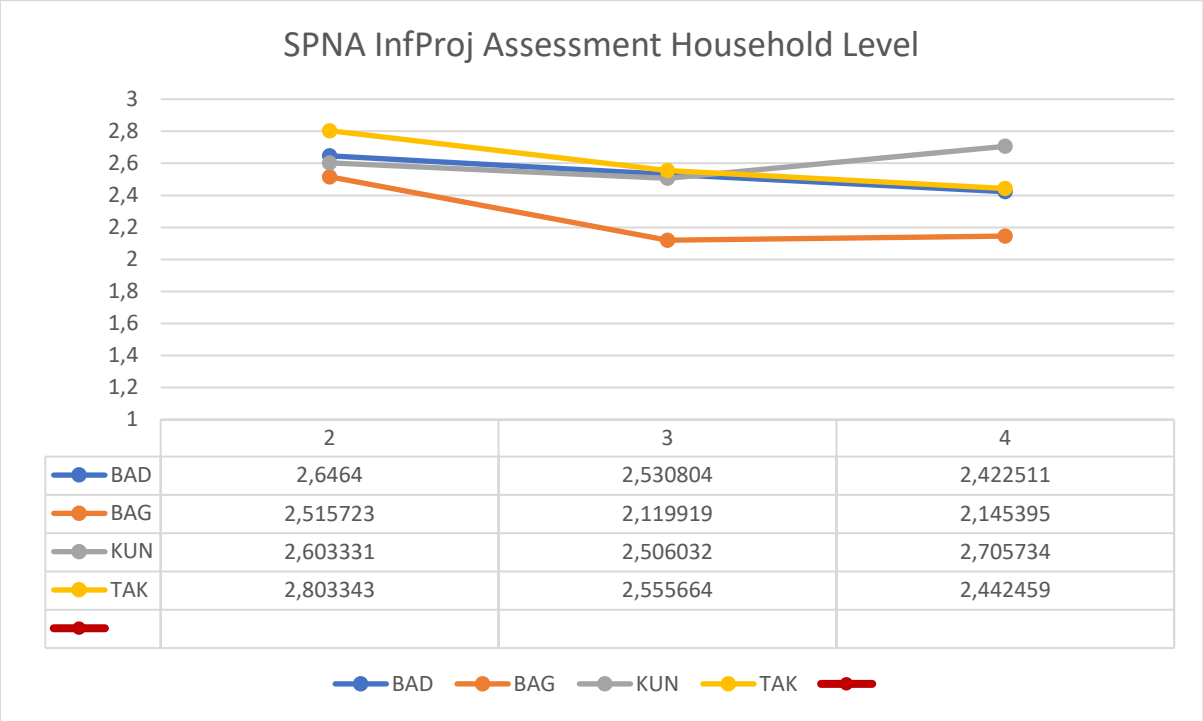


Figure 69: Wave 4, 35 districts - infrastructure assessment household level

The trend for household-level assessment of SPNA usefulness remains relatively stable over the waves with a slight decline from Wave 2 to Wave 3 (steepest decline in Baghlan) and an increase in Kunduz in Wave 4.



	2	3	4
BAD	2,6464	2,530804	2,422511
BAG	2,515723	2,119919	2,145395
KUN	2,603331	2,506032	2,705734
TAK	2,803343	2,555664	2,442459

Figure 70: Wave 2-4 – infrastructure assessment household level; 1 is worst and 3 is best score

Summary of the infrastructure indicators

The infrastructure component of SPNA is visible to between 30% and 70% of the population surveyed. Knowledge about projects seems to decline with the time passing after handover (until 2015 above 50% of respondents were aware of the projects in each province, in 2017 this number declined to below 40% in all provinces but Badakhshan).

Opinions about the infrastructure projects have been overwhelmingly positive, both in their effect for the district as well as for the individual households surveyed. Based on the interviews conducted with a wide range of stakeholders (guideline interviews with representatives of the DDA and district administration, field visits by local teams and external consultants, debriefing of engineers and social mobilisers) there are two main explanations to the positive assessments: firstly, most infrastructure projects have been and still are used according to design; there are only few examples of defunct infrastructure (mostly smaller roads) or wrongful use of some buildings (see also "*Immediate Outcome 1: operational key infrastructure*"). Secondly, owing to the participatory procedure of project identification and implementation, there appears to be a feeling of responsibility and local ownership for the infrastructure. We have little evidence of infrastructure that proved divisive or conflicted.

SURVEY-INDEPENDENT INDICATORS FOR ROBUSTNESS CHECKS

Note that we use the perception-based indicators for SPNA immediate outcomes as principle explaining independent variables in the analytical section below. Because of the limitations of

perception-based observational data explained above we also used non-survey based explaining variables for the quality of the DDAs (see above *DDA capacity*) and the rollout of infrastructure projects as robustness checks to our models.

PROCESS INDICATORS AND MODERATORS

More participatory modes of aid delivery and better maintenance after the handover of projects are expected to strengthen the impact of SPNA on stabilisation. We expect a similar effect for better function development shuras at village level. For different sectors of infrastructure implementation we do not have explicit impact expectations.

The immediate results of SPNA are measured by the previously discussed **visibility** and **assessment of DDAs** and of **infrastructure projects**. These are our principle predictors of programme-induced stabilisation in the districts. However, SPNA places a principle emphasis on the delivery process, on procedures, on learning by doing, on capacity-building, participation, communal ownership and responsibility. Hence, we need to account for variation in the process of SPNA delivery, for the kind of SPNA infrastructure delivered, and for the village-level organisational capacities to absorb SPNA-induced development in the districts.

We expect that (a) more or less participatory modes of delivery, (b) arrangements for maintenance after project completion, (c) different project sectors and (d) the functionality of the communal (village and cluster levels) shura structure moderate the impact of SPNA outcomes on district-level stability. We introduce these aspects of infrastructure projects and their social context as **moderators** or, more formally, **as interaction variables** in our regression analysis.

MODES OF DELIVERY

Working through capacitated DDAs, SPNA is, in principle, a highly participatory development programme. Before DDAs had been trained on the procedures, however, initial project identification and delivery was often less participative. Additionally, in some districts the implementing partner decided to bypass the DDAs altogether. This was mostly the case in two districts of Baghlan. In other cases, yet again, the DDA only existed “on paper” and trainings were all but a formality, without any results for proper project implementation (this relates mostly to districts in Takhar with Warsaj being an obvious negative example already described in previous reports).

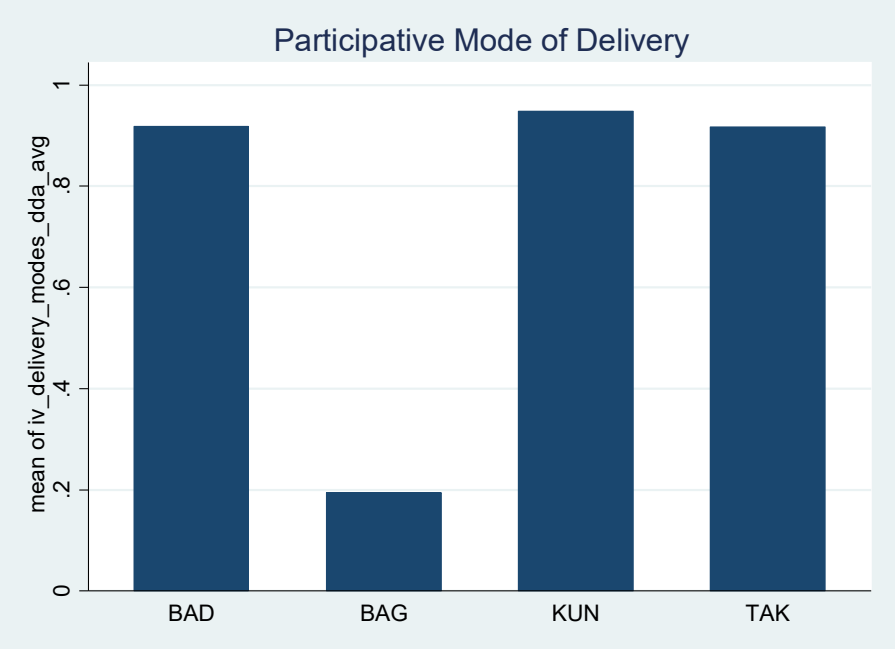


Figure 71: All waves, 25 districts - participative modes of delivery

MAINTENANCE

Arranging the maintenance of infrastructure projects as part of the handover process has been an integral part of SPNA implementation since 2012. It is usually taken care of by forming maintenance committees on the side of the beneficiaries and making Operation and Maintenance (O&M) kits as well as small O&M funds available. Larger repair and constructional issues need to be addressed to the respective line ministries, a step some communities have taken either via the DDA or less frequently directly without the involvement of the DDA.

Despite the common procedure introduced in 2012, we do find variation in the state of maintenance recorded during project-specific interviews with DDA and community representatives by our field teams, with Baghlan scoring highest and Takhar lowest. Despite the relative differences in ranking, arrangements for maintenance are fairly widespread. The main problem is, however, as our interview partners pointed out, that even with the SPNA Phase V O&M funds – which not all districts had received – more substantial repairs are difficult to carry out because of costs or lack of expertise needed (e.g. with Jurm DDA and maintenance committee on 4 Jul 2018). Districts without an O&M component have virtually no means to implement more demanding maintenance activities. The responsible line ministries are usually not able or willing to allocate budgetary resources for repair work of SPNA-financed infrastructure (or in fact for any kind of infrastructure constructed in their districts). According to DDA representatives interviewed in June 2018 in Badakhshan, they prefer to spend this money on infrastructure directly financed from the state budget.

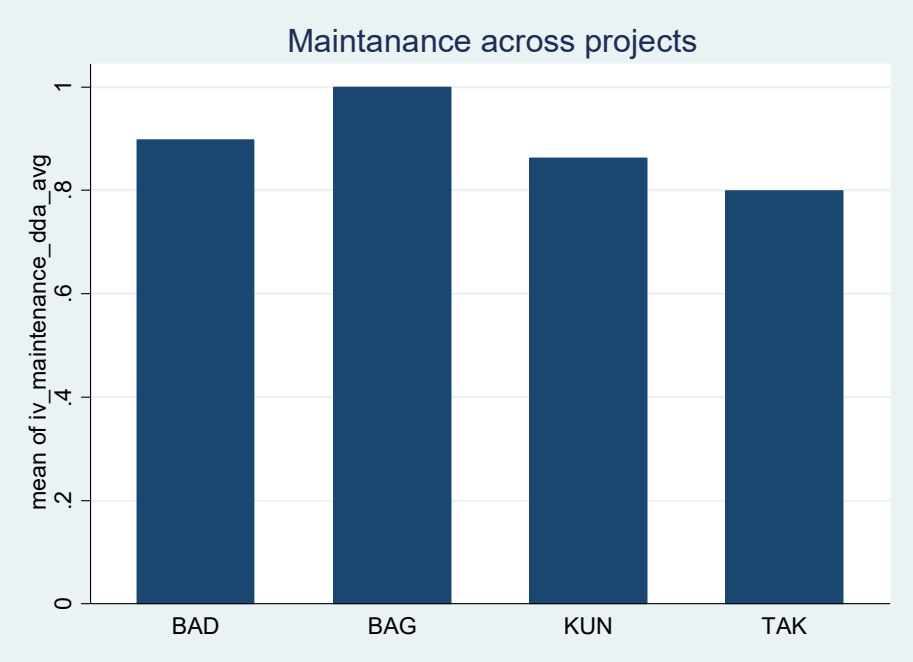


Figure 72: All waves, 25 districts - project maintenance

INFRASTRUCTURE SECTOR

As next, we are interested to find out if the respective sectors of infrastructure projects influence the general effect of SPNA on stability in different ways. Most projects implemented fall into the sectors education, agricultural infrastructure (e.g. irrigation canals), road infrastructure and health. The mix of projects prioritised and implemented differs between districts. Hence, we have sufficient variation in the mix of sectors across individual districts to test if different sectors interact in different ways with the effect our main explaining variables (immediate SPNA outcomes) have on stabilisation. We will use this in the analytical models below.

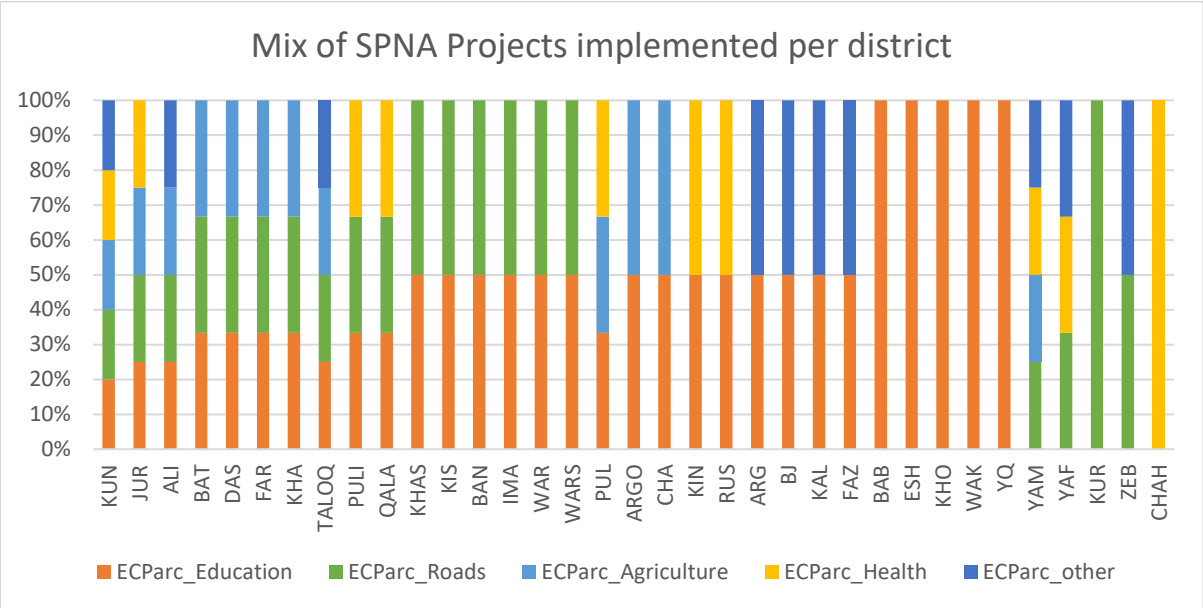


Figure 73: All waves, 25 districts – sectors of projects implemented per district (ECP stands for expert coded projects)

SHURA STRUCTURE FUNCTIONALITY

A further moderator relates to the shura structure. As was discussed, the functioning of the shura structure is an essential precondition for an efficient and functioning DDA. The more functional and inclusive in particular the intermediate CLDC level is, the more local communities represented by their CDCs have a chance to communicate their problems to the district level and partake in the benefits of SPNA.

We developed an index of shura structure functionality based on three questions in the village profile: V29 asks whether the CDC is gaining or losing influence; V30_3 asks if CDC representatives participate in CLDC meetings; and V30_5 asks if the CDC representatives are aware of the work of the DDA. We coded the responses and arrived at a functionality index where 10 is the highest rank and 0 the lowest.

Kunduz, Baghlan and Badakhshan all reach average values between seven and eight on the functionality scale. Takhar, however, only scores 4.8. This fits well with the findings presented in the qualitative section, which found that, most likely as a result of a provincial level DRRD interpretation of the DDA structure, CLDCs in Takhar only count the three top office holders (head, deputy and treasurer) as actual members and only invite them to CLDC meetings. In contrast, CLDCs in Baghlan and Kunduz, and to a lesser extent in Badakhshan, are more inclusive, giving more CDCs the chance to participate in the shura structure.

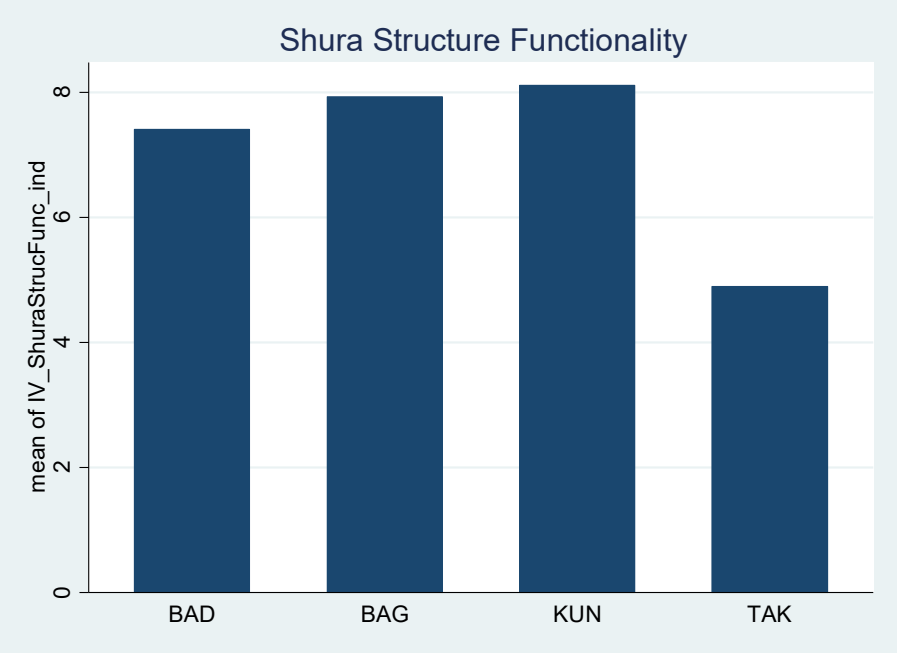


Figure 74: All waves, 25 districts - shura functionality index

DEMOGRAPHIC CONTROL VARIABLES

ETHNIC SELF-DESCRIPTION

As expected, the distribution of ethnic self-description does not change much between waves. Even adding 10 new districts in Wave 4 does not really change the balance. Based on past experience, we cannot take this for granted, because ethnic belonging is to some extent a fluid and situational

category in Afghanistan (as in many other places). In a survey before 2010 we experience significant changes not only of villages (e.g. communities that had identified as Hazara under more secure conditions changing to Tajik in the subsequent survey) but even of whole districts (e.g. respondents initially identifying as Tajiks changing to Aimaq in a subsequent survey). However, on average and with the larger sample of 25-35 districts, such changes do not seem to change the general picture of ethnic distribution in the SPNA region.

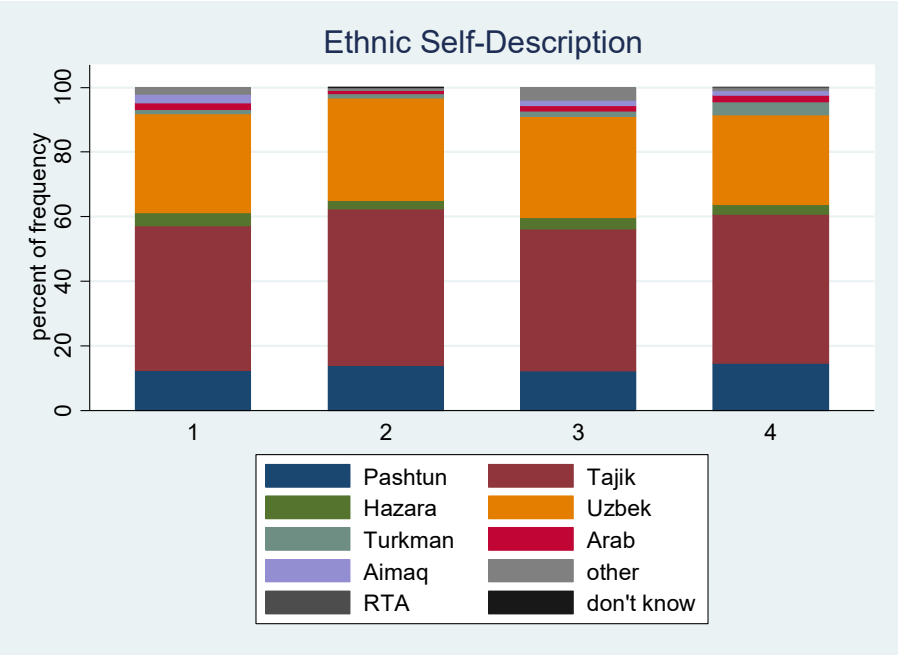


Figure 75: All waves, 25 districts – ethnic self-identification by wave

In terms of differences between provinces, Kunduz clearly is the most ethnically diverse province and the only province where Pashtuns form a relative majority of around 37%. Baghlan has an even higher share of Pashtuns (over 40%), but they are only the second largest ethnic group in our survey after Tajiks. Takhar divides into an Uzbek majority and a large Tajik minority. The dominant group in Badakhshan are Tajiks with a sizable Uzbek minority.

We have used ethnic belonging and ethnic diversity as control variables in past research. In this report we reduced the controls to the regionally relevant question of whether the respondent identifies as Pashtun or not.

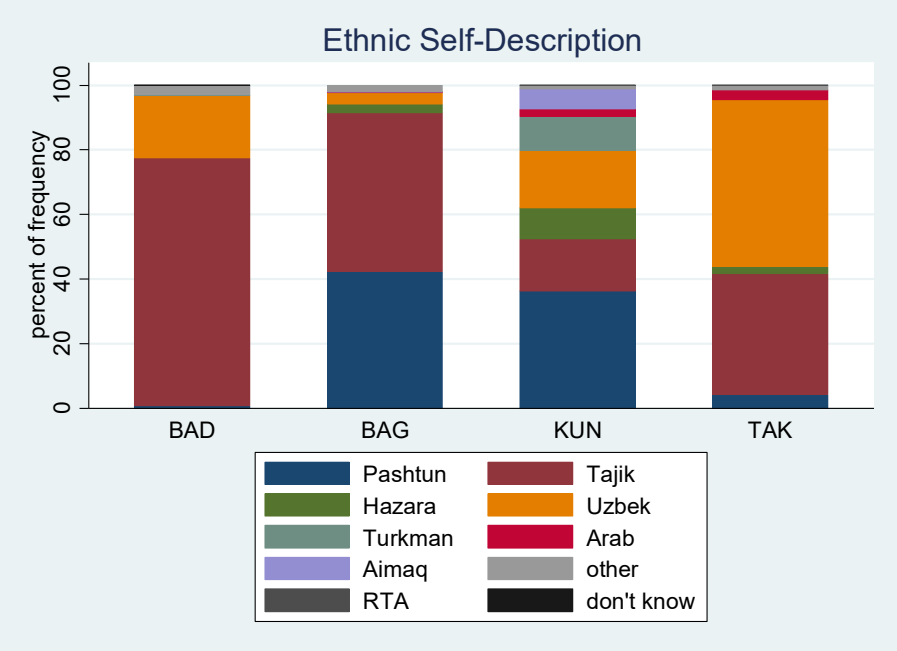


Figure 76: All waves, 25 districts – ethnic self-identification by province

RESPONDENT AGE

Because we are doing heads of household interviews, respondents are with a few exceptions senior adult men. They are on average between 40 and 50 years old, ranging from 15 to 100 years (we recoded three biologically impossible outliers down to 100 years). There are slight differences between waves, but we need to note that age, even more than ethnicity, is not a fixed category in rural Afghanistan. As a rule, it is a self-estimate and situative as many individuals of the older generations are simply not aware of the precise year of their birth. Few people thus know (or care) when exactly they were born.

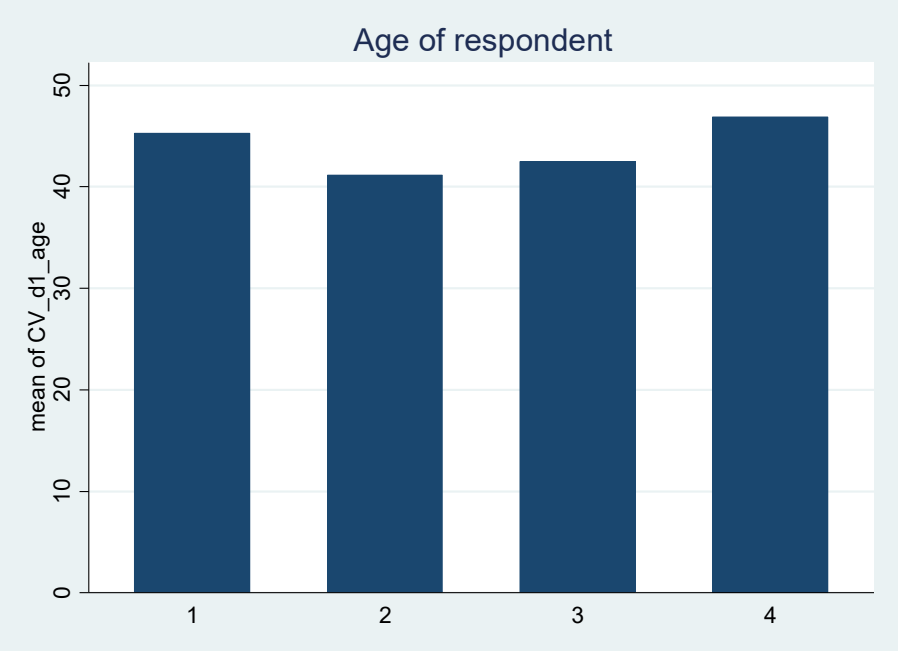


Figure 77: Age of respondent by wave

LITERACY

Education, however, is not fluid and people know and do care about the education they have received. Education is often also a matter of considerable pride and prestige. Literacy rates are increasing across waves, from 25% in Wave 1 to 39% in Wave 4.

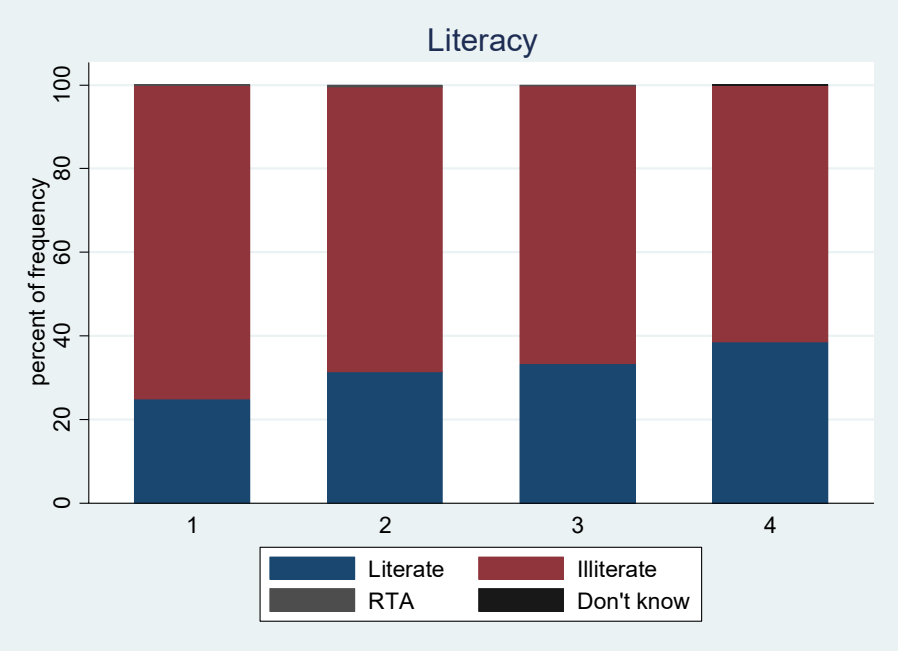


Figure 78: Literacy of respondent by wave

YEARS AT NON-RELIGIOUS SCHOOL

Improvements in education are even more tangible when we look at how many years respondents report having spent at school. On average, this index grows from just above two years in Wave 1 (driven by 72% reporting to have spent 0 years at school) to 4 years in Wave 4 (with only 62% reporting not to have spent any time at school).

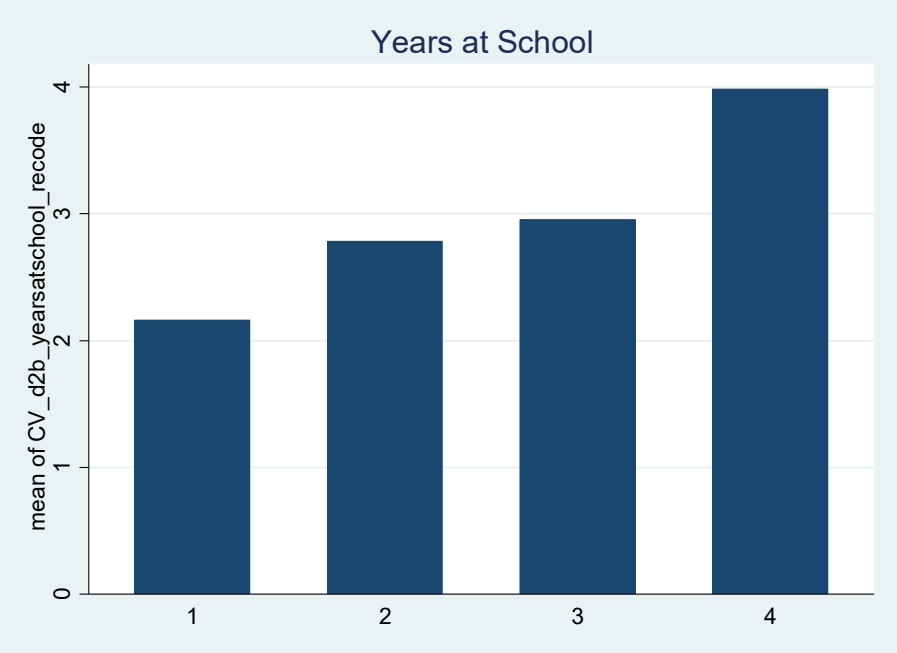


Figure 79: Years respondent spent in school by wave

SUMMARY OF RESULTS OF THE DESCRIPTIVE STATISTICS

The descriptive statistics presented above provide evidence of how the different dimensions of stabilisation trended since SPNA took off, how immediate SPNA outcomes developed over time and how the modes of delivery of SPNA evolved. The most important observations with implications for SPNA's impact on stability are summarised below.

Firstly, "objective" (incident-based) **security** improves towards the endline, but the quality of security changes. In contrast to the earlier improvement observed in 2012, when state and international forces managed to push pack the armed insurgency, the number of incidents in 2017 dropped as a result of prior gains of the insurgency and government inactivity. The result is a stalemate. Under these conditions, subjective security perceptions and trust in the state to provide security in the future on average across the sample deteriorated further (for differences between provinces see detailed discussion above).

Secondly, we note a deterioration across most stabilisation indicators in the field of **governance**. In particular, the indicators for state presence, state legitimacy and official state power stagnate or worsen. The decline towards the endline assessment reflects a trend towards state retreat and government fragmentation.

Thirdly, subjective indicators of general **development** activities decline towards the endline. Our "objective" (project count at village level) indicator trends vary widely across the provinces. In comparison to the baseline, they have declined in Kunduz, stagnated in Takhar, but increased in

Baghlan and particularly in Badakhshan. Compared to the baseline, the district, village and household **economy** indicators have either increased (cars, self-assessed material well-being, size of district bazaars), or, after an initial deterioration, recovered to baseline level (access to land). Compared to the baseline values mechanisation of agriculture (tractors) stagnate on average.

Fourthly, in the field of **adaptation/modernisation** we find a decrease in the ratio of public schools as compared to religious schools. Perceptions of international development aid as a value threat improved towards the endline (less threat).

Fifthly, for **immediate SPNA outcomes** we note an increase in DDA visibility towards the endline, while the assessment of the DDAs' contribution to district development declined. Regarding the infrastructure component, visibility of projects declined slightly over the last years of the SPNA lifespan, while the evaluation of the utility for the district remained very high (close to 99% positive responses) but declined slightly for the household utility question.

Lastly, at the time of the endline, most infrastructure projects financed by SPNA were being **maintained** to some degree (reportedly between 80% in Takhar and 100% in Baghlan). Most projects had been identified in a **participatory** manner, i.e. with the involvement of the DDA (above 80% in all provinces but Baghlan, which is the exception here with just under 20% participatory project identification). In terms of the functionality of the **CDC-shura structure**, upon which DDAs are founded, Kunduz, Baghlan and Badakhshan score between 70% and 80% on the functionality index; here Takhar is the exception with a score just below 50%.

Thus far we looked at these trends for stability, SPNA outcomes and modes of delivery separately. It is the task of the following analytical section to enquire as to how the different dimensions of stability, SPNA outcomes and modes of delivery relate to each other. In more formal terms, we will use statistical tests to verify the hypothesis that immediate SPNA outcomes and modes of delivery have a significant effect on the different dimensions of stability identified in the models.

ANALYSIS: STABILISING EFFECTS OF SPNA

INTRODUCTION TO THE ANALYTICAL STRATEGY APPLIED

Ideally, impact assessments of any programmes would adhere to what has become known as the *Gold Standard* of impact evaluations in development (Gaarder/Annan 2014; Grävingholt et al. 2012; OECD/DAC 2012). They would start before programme implementation, work on the basis of an explicit theory of change and deal with the fundamental problem of any impact assessment – attribution – via randomization of activities as far as operationally and ethically possible.

Attribution means that change observed on the dependent variables (the intended stabilisation results described above, *Indicators for SPNA's immediate outcomes*) can be traced back to programme measures. There are two principle problems that prevent us from simply treating a correlation between programme outputs and changes on the dependent variables (in the case of SPNA: stability measurements) as programme effects. The first one is that factors not related to the programme may have caused the observed change. In the case of SPNA, it is obvious that changes observed in security, governance, economic development and ability to adapt are influenced by a number of such factors, which may at same time also have affected SPNA implementation. And the second is the possibility of reverse causation and selection effects, i.e. that differences in the programme outputs may have

followed rather than caused differences in stability. This could be the case if programme measures clustered, for example, only in more secure, better governed, better resourced or culturally more inviting communities.

The best way to deal with the attribution problem would be a random assignment of development measures to beneficiaries. In the presence of truly effective randomization and a sufficiently large sample size, randomization of access to development measures would allow us to simply compare if those who received the development measures are on average different on the dependent (stabilisation) variables than those who did not receive those measures. The difference between the two would represent the effect of the development intervention. Statisticians call the group exposed to the measures the "treatment group" and those who did not receive the measures the "comparison" or "control group". The idea is that because of the randomized treatment, both groups are on average similar on all potential characteristics – except for their access to project activities. Thus, the control group represents an effective approximation of a counterfactual, i.e. the situation of the "treated" group had they not received the treatment (here: the development measures). If both groups differ in terms of the dependent variable of interest, we can be confident that this is a consequence of the project activities rather than of any other, unobserved difference between both groups.

However, randomising the development measures is often not possible; it may be ethically problematic to identify beneficiaries by randomization rather than on a basis related to the programme goals only (e.g. need or capacity assessments). Randomization may also contradict other important aspects of the roll-out strategy and increase costs because of logistic issues (e.g. less-efficient clustering of beneficiary communities).

In the specific case of the SPNA, it was not feasible to design it as an experiment. The reason is twofold. First, capacity-building measures benefit whole districts. Here, a comparison group could have been formed only from other districts. Funding limitations, however, prevented us from expanding in-depth quantitative and qualitative data collection efforts beyond the treatment districts. Second, and conceptually more important, for the infrastructure component the identification of projects and beneficiaries was largely left to the DDAs – and this was an intentional and defining part of the stabilisation approach. Hence, randomization of the infrastructure component on the sub-district level was simply impossible: SPNA capacitated DDAs for the very reason that they take informed decisions via transparent, competent and fair procedures on who receives what and when. Randomization instead would have been counter-conceptual to SPNA and to the very idea of stabilisation via institutionalised procedures.

Hence, we had to apply what was, from a methodological viewpoint, a second best approach relying on observational data. Specifically, we implemented a combination of the following strategies:

Cross-sectional design with controls for alternative explanations

To overcome the lack of randomization of capacity-building and infrastructure measures, we decided to proxy household and community exposure to SPNA measures using the intended immediate programme outcomes. In each district, one immediate consequence of SPNA capacity-building and infrastructure provision should be increased visibility of the capacitated institution (the DDA) and of the infrastructure prioritised and implemented in transparent and participatory procedures. A second intended consequence of SPNA should be increased appreciation of the DDA as development provider and of the positive effects attributed to the SPNA infrastructure projects. In other words, we assume that more and better SPNA measures are directly linked to higher visibility and appreciation of the DDA and the implemented infrastructure projects. The SPNA impact hypothesis then states that

variation on visibility and appreciation of the DDA and the provided infrastructure has an effect on a number of stabilisation indicators specified in SPNA's theory of change. The relationship between SPNA measures/inputs and the immediate outputs is both plausible and has been tested for the endline: SPNA was the major programme training and capacitating DDAs in the assessment districts and higher capacitated DDAs are more visible and predict higher levels of acceptance (see above *Immediate Outcome 2* and below *Alternative channels*).

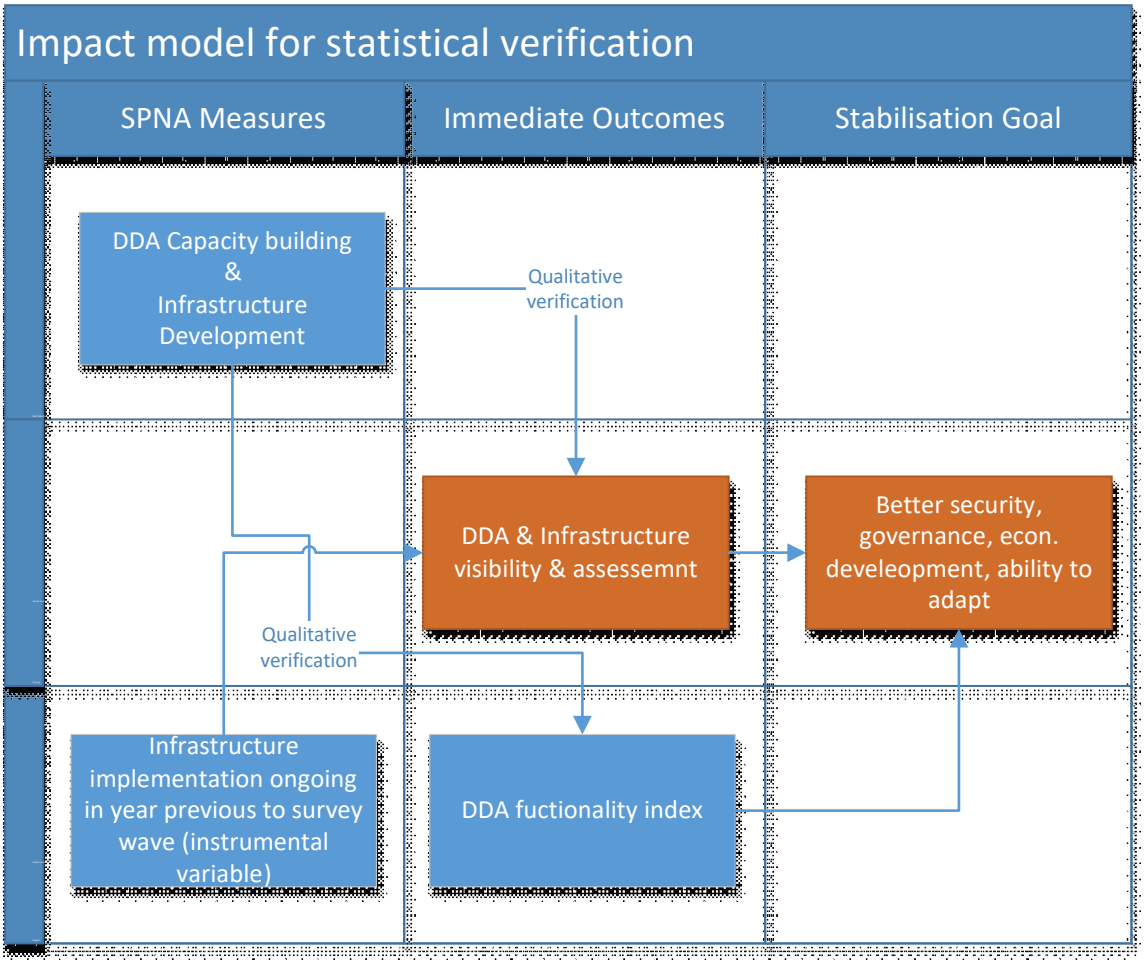


Figure 80 Schematic representation of impact models - we focus on the organge model of causation and use the other relationships depicted as robustnes checks

We have been dealing with the challenges of attribution in the following ways.

- **Alternative explanations for variation on the dependent stabilisation variables:** We know that a number of factors other than SPNA have a strong influence on the stability that the programme intends to foster. We are statistically controlling for the most obvious factors that may potentially correlate with both the SPNA implementation and stabilisation variables: the level of stability at the time of the baseline, the objective security situation, the question of territorial control and the presence of state and non-state armed forces, other development

activities and a general ideological worldview. Additionally, we are controlling for districts; this means that we limit our comparisons to villages/households that located within the same district. This is a precaution that helps us reduce the risk that differences across districts drive observable correlations between our SPNA predictors and the stabilisation variables. Introducing "district dummy" variables does not exclude this risk, but controls for all those otherwise unobserved factors that do not vary within one and the same district.

- **Reverse causation:** In order to exclude reverse causation, we need to be sufficiently confident that our measurements for SPNA are not driven by the stabilisation variables defined as outcomes in our theory of change. We know that SPNA has, for the most part, been implemented equally across all districts, irrespective of their level of security, governance-situation, remoteness and openness to modernisation. The 35 districts covered in this research received in principle a very similar package of SPNA measures. Hence, we have no reason to expect that more or less SPNA is driven by variation of stability within the treated districts covered.

  It is, however, not clear, whether we can exclude this for the within-district variation of the SPNA proxy indicators defined as predictors in our models (visibility and assessment of DDA and infrastructure). The short answer is that without randomization, we cannot formally exclude reverse causation. However, reverse causation is less plausible for those stabilisation fields that are not directly linked to the functions of the DDA and the infrastructure processed through the DDA. Security and the performance of government (e.g. the provision of public services) are not responsibilities of the DDA. Hence, feeling more secure or having greater political trust in the government does not obviously cause higher visibility and greater appreciation of the DDA and the infrastructure it delivers. For economic development and openness to modernisation the separation between more general positive attitudes and the specific attitudes to the DDA (but not for specific infrastructure or visibility) may be more difficult.

In Wave 3, we included a survey experiment in the questionnaire that added information on development activities by the government, international development agencies and local NGOs as well as no additional information. We then asked questions relating to political trust, perceptions of corruption and attitudes in regard to civil obligations like taxes for services and voting. We intended to check if information on development by various development providers influences political trust. The experiment did not show any statistically significant differences between the three different "treatment" groups and the comparison group that did not receive any additional information. However, given the nature of the experiment, the null finding applies only to the role of the provision of information – rather than potential effects of actual experiences of DDA activities. In other words, it does not mean that SPNA failed to influence these outcomes, but, rather, that the specific treatment did substantially influence individual political attitudes.

GROUPING VARIABLES – DISTRICT EFFECT CLUSTERS

The search for specific effect clusters via districts grouped according to similar conditions with regard to their long-term exposure (a) to violence, (b) to the state and (c) to modern road infrastructure produced some results that may inform more tailor-made stabilisation strategies in the future. The approach helps to identify more consistent framework conditions that may have an influence on the significance, strength and even direction of programme-specific stabilisation effects.

PANEL/TIME SERIES DESIGN WITH CONTROLS FOR ALTERNATIVE EXPLANATIONS

For the endline we extended the inferential statistical methods further. The four waves gave us sufficient observations at village level to conduct a time series analysis where the villages surveyed representatively in each wave formed a panel. This panel analysis adds value to the impact assessment because it compares the variation on the SPNA predictor/explaining variables and stabilisation outcome variables for one and the same village over time, while controlling for alternative time sensitive variables. The advantage is that village specific, but time invariant factors are automatically controlled for. What does not change in the village over time, cannot explain observed changes on the stabilisation variables, and is hence excluded as an alternative explanation for those changes.

The panel analyses allow for a more effective consideration of potential alternative explanations. It does not, however, deal any better with the second issue, i.e. the possibility of reverse causality. In other words, if increased security perceptions predicted higher visibility and appreciation of the DDA and the SPNA infrastructure projects rather than the other way around, we would still run the risk of misinterpreting the results. To further reduce this risk, we used time shifts in SPNA roll-out as an *instrumental variable*. Instrumental variables correlate directly with the explanatory variable but only indirectly with the stabilization outcome (through its correlation with the explanatory variable). The temporal sequence of the SPNA roll-out was mainly determined by administrative decisions, rather than by socio-economic and political conditions in the target districts. The roll-out, in turn, certainly determined SPNA activities but it had presumably no direct effect on the stabilisation variables defined in the model (other than through its effects on SPNA activities). We use the exogenous variation of the explanatory variable induced by the SPNA roll-out with aim to estimate more unbiased effects on the stabilization outcomes.

As we will see, the results of these additional analyses are in line with the other estimation approaches presented, making us somewhat more confident that correlations do in fact represent effects of project measures on stabilization outcomes – rather than the other way around. However, a simple coding of ongoing project implementation in the previous year is not an ideal instrumental variable. It focusses on infrastructure only, and does not differentiate between projects not yet started and already finalised more than a year ago. Neither can it exclude that in some (few) cases, project implementation may have, indeed, been delayed by reasons relating to the security component of stabilisation. Thus, overall, the instrumental variable strategy can only provide another piece of evidence on the SPNA effects, rather than a representing a solution to the attribution problems highlighted above.

ROBUSTNESS CHECKS: NON-SURVEY BASED PREDICTORS

We took a further important step to minimise the risk of a misleading attribution of observed stabilisation effects to SPNA. Above we described how we used survey-based responses on visibility and appreciation of the DDAs and SPNA infrastructure as proxy indicators for SPNA's immediate outcomes. These are, however, the opinions of respondents.

We also wanted to know if more objective measurements of the quality and functionality of SPNA's principle counter-part, the DDA and the CDCs on which it is based, confirm the stabilisation effects we observed for our subjective predictors.

SPNA does not only consist of hard (infrastructure projects) and soft (participatory project selection, monitoring, conflict resolution) activities that should be visible to a large section of a district's population (and thus to our respondents). As we increasingly realised in the course of our qualitative research, DDAs also conducted many activities behind the scenes, which likely remained invisible to all, but the most interested and engaged respondents. Such activities relate to various forms of advocacy DDAs conducted on behalf of local communities, e.g. conveying community-level concerns about teacher absenteeism to the district administration and the district education department, protesting about arbitrary detentions by the police, lack of medicine in local healthcare units, etc. At endline, we attempted to construct variables that could – to some extent – measure such *invisible effects*, i.e. effects that respondents did not consciously associate with the DDA.

Hence, we ran the same regression analysis with expert-coded predictor variables at CDC (functionality index) and DDA (maturity index) levels. As we will see, both tests confirmed our earlier results. As such, we found consistent positive associations between an intact chain linking CDCs via CLDCs to DDAs with peoples' assessment of power-holders, positive assessments of development, self-reported material well-being and – to a lesser extent – on people's assessment that the woliswali cares for the need of the people.

Lastly, we used our qualitative research to trace the realisation of the theory of change from outputs (delivery of infrastructure measures, trainings and capacity-building) via intermediate outcomes (the actual operationality of delivered infrastructure and its use; the application of capacity-building trainings by DDAs and the actual services DDAs provided to the populations of their districts) and the statistically measured stabilisation impacts. These results lend further support to the argument that correlations do in fact represent causal effects of SPNA measures.

How to read the following analysis and where to look for further statistical evidence

The actual analysis consists of two distinct sections: (a) the cross-sectional analysis, and (b) the panel analysis. We introduce them separately.

The cross-sectional analysis

In the following section we **first** present the findings from the two cross-sectional analyses – one of the endline and the second of the pooled analysis across all waves. These cross-sectional analyses compare households with one another, rather than comparing the same households at different points of time (see notes on panel analysis below). The results of the main effects of immediate SPNA outcomes on stability are grouped according to the four fields of intervention – security, governance, economy/development and adaptation/modernisation.

Subsequent to the discussion of main effects, we look at how maintenance (see section "Maintenance"), participation (see section "*Modes of delivery*"), the different developmental sectors (see section "*Infrastructure sector*") and the shura structure (see section "*The shura structure*") interact with the main effects of SPNA: we investigate the question whether they have a moderating effect, and whether they strengthen or weaken the way SPNA affects stabilisation.

Finally, we investigate if specific district clusters react differently to SPNA outcomes in the four different fields of stabilisation. The chapter **concludes** with a summary of findings. Note that in this

chapter we simplified the technical presentation of results and tried to keep the narration of results as accessible as possible. Detailed tables and more technical descriptions of statistical methods are in the annex.

The panel analysis

The following analytical chapter – the panel analysis – is an important and comprehensive robustness check for the cross-sectional models discussed before. Contrary to the cross-sectional analyses, the panel analyses focusses on the village level rather than the household level. Rather than comparing different households, it compares individual villages with themselves at different points in time. This chapter is more technical and formal. Since it mostly confirms the results of the cross-sectional analysis, it should be interpreted as a robustness check that lends additional support to the main findings of the cross-sectional analyses.

CROSS-SECTIONAL ANALYSIS

EXPLANATION OF THE MODELS

We estimate sets of models with three different specifications. Each of the sets looks from a different angle at the question of likely effects the SPNA programme had on district-level stability across the four fields of intervention.

Cross-sectional regression: The **First Set of regression models** only looks at the situation at the Wave 4 endline assessment. We compare respondent views on twelve stabilisation variables (discussed above**,** *Indicators / dependent variables and main dynamics in the four fields of stabilisation*) between those with more (treatment group) or less (comparison group) exposure to one of the two components of SPNA (DDA visibility; infrastructure visibility) and with higher opinions about the immediate outcomes of SPNA components (DDA and infrastructure assessment).

We estimated three sets of models: as **first step**, we estimate models considering households in all 35 districts included in Wave 4. As **second step**, we then add the village-average values of the dependent stabilisation variables at the baseline in 2010/11 to control for what the respective dimension of stability was like before SPNA started. Since baseline values from Wave 1 are not available for the 10 districts that were added in Wave 4, those observations are not included in these models. Finally, as **third step**, we add four moderating variables to the regression; they are related to the SPNA process and we expect them to interact as "amplifiers" for observed main effects of SPNA predictors (visibility, assessment) on stabilisation dependent variables. These are participatory modes of delivery, maintenance after project handover, specific sectors that are possibly more effective than other sectors, and, finally, the village-level functionality of the shura structure, which is an important lowest-level receptor of changes SPNA induced at district level. We have discussed these variables in section Process indicators and moderators.

Pooled cross-sectional regression: The **Second Set of regression models** takes a very similar approach as the first one but considers households across all four waves across all 35 districts resulting in a total sample of 24.000 observations. While the first set of models informs us about likely effects of SPNA on stability at the time of the endline, this second model takes the whole implementation period of SPNA into consideration. We compare respondent opinions on stability within individual waves (see annex for further explanations). T

Simplified cross-sectional regression for subset analysis: The **Third Set of regression models** analyses if SPNA stabilisation effects vary between **specific district clusters**. We initially defined these clusters in 2015 (Wave 3) based on impact theory paired with the empirical results of our initial and inconclusive impact assessment of Wave 2. For the endline we refined and simplified the clusters to (a) districts with continuously higher and lower levels of negative security incidents over the full assessment period; (b) districts with continuously more and less state presence over the full assessment period; (c) districts with more or less access to modern infrastructure, i.e. main *lines of communication* (LOCs) passing through or passing in the vicinity of the district centre. Note that all three criteria take the full assessment period into consideration, and are based on expert-coded data rather than respondents' opinion (incidents, governance zones, paved roads). In order to simplify the analyses, we use statistical procedures to reduce the number of dependent variables (our measurements for SPNA-affected stabilisation) from 12 to 4, i.e. one for each field of SPNA intervention.

In all models of the cross-sectional analysis, we cluster standard errors at the village level and compare only respondents within one and the same district (see above; this implicitly controls for any unobserved variation across districts)[13]. In successive steps we control for relevant demographic factors like age, ethnic self-description and the level of education of the respondent. We also control for context variables that may represent alternative explanations for variation in our dependent stabilisation variables (most importantly for the number of projects received at village level, i.e. not SPNA-related development inputs, the objective security proxied by incident counts in a 3 km radius around the villages, and the governance context proxied by our governance zones).

In the main text we provide a narrative discussion of results across the models. Contradictory findings between models are highlighted. Tables provided use simple colour coding to illustrate statistically significant correlations explained in the narrative.

More detailed summary tables of statistical results are available in the annex.

RESULTS

Main effects

For the discussion we focus on the results of the most robust model, i.e. the one including the control for baseline levels of the dependent variables.[14]

Table 7 summarises the main results for SPNA's immediate outcomes across the four fields of stabilisation. We use a simple colour coding indicating results that are statistically significant at the conventional 95 percent level or higher (for details see annex). The colour coding shows the direction of the significant relationships – green for positive, red for negative stabilisation effects. We will discuss the results for each field. We refer to the results of the additional models of the pooled cross-sectional analysis as appropriate.

[13] The exception are models scrutinising district effect clusters, where we interact with district clusters and can, hence, not control for individual districts.

[14] Note that we treat the results for infrastructure assessment at district level with caution because, as mentioned earlier, we have too little variation in the responses to generate reliable results (almost all respondents indicated very favourable assessments).

	DDA vis	DDA ass gen	DDA ass sect	Inf vis	Inf ass dis	Inf ass household
Sec1 (security perception)						
Sec2 (security provision)						
Gov1 (district admin. cares)						
Gov2 (social cohesion)						
Gov3 (most powerful actor)						
Gov4 (corrupt conflict resol.)						
Dev1 (general development)						
Dev2 (state dev. Contribution)						
Dev3 (material well-being)						
Dev4 (access to agricult. land)						
Mod1 (int. dev. as value threat)						
Mod2 (school curriculum dissat.)						

Table 7: Wave 4 (35 districts) main effects with all controls (35 districts) and with all controls plus baseline control (25 districts).

Security

SPNA's governance component – DDA visibility and assessment – correlates positively with respondents' security perceptions of household and district-level security. This is consistent with the findings in Wave 3. At endline, only infrastructure visibility relates positively to both security perception indicators in a statistically significant way. We note more negative effects on the confidence that the state will be able to provide security in the future. People with a more positive opinion of the DDA's contribution to development and about SPNA infrastructure projects tend to be more sceptical about the state providing security in the future. This is an important change to the Wave 3 assessment and to the cross-sectional pooled assessment covering all four waves: at the time of the Endline the effect of DDA and infrastructure assessment on confidence in the ability of the state to provide security turns from positive to negative.

This is an important and problematic change that does, however, fit into the general security and governance trends described above. By 2017/18 development activities and the societal institutions providing them (the DDA) still improve general security perceptions, but they fail to convince people that the state will be able to deal with security threats in the future. It seems that people aware of and appreciative of development are increasingly looking elsewhere for their future security.

Governance

Consistent with findings reported in Wave 3 we see a positive and statistically significant relationship between most immediate SPNA outcomes and the perception of a district administration that cares for the needs of the local community. The visibility of a DDA that is perceived as active and useful for

district-level development and that facilitates the provision of visible and useful infrastructure are good news for the state administration at district level. Even under conditions where trust levels in the government are low and further deteriorating, these SPNA-specific immediate outcomes seem to increase political trust in the district-level state government.

We find negative effects on the perception of who is the most powerful person in the district (official representatives of the executive coded as best, informal gunmen worst, modern and traditional unofficial powerholders in between). This means that more favourable assessments of the DDA as well as of infrastructure and its visibility correlate with less formal traditional powerholders and with members of the legislative as opposed to the government-appointed executive authority. An Analysis of Variance (ANOVA) shows that in Wave 4, the result is driven by respondents considering elected representatives of the legislative (members of parliament, the provincial council or the DDA), and informal traditional powerholders (like maliks, arbobs, the ulemma) most powerful; these respondents have, on average, higher opinions about DDA contributions to development.

Importantly from the stabilisation perspective, it is not driven by the lowest option for powerholders – commanders, gunmen and Taliban – which correlates with lower average values for DDA appreciation than the highest option, i.e. official powerholders. This finding, if only rather weak, is relevant for SPNA, since one of its aims is to strengthen the relationship between formal societal organisations (the DDA) and the official district-level state administration. The finding also indicates that at the time of the endline, the visibility and appreciation of infrastructure projects was higher in areas with a higher degree of informal, but nonetheless state-associated local power relations.

In the pooled cross-sectional regression (across all waves) perceptions of the district administration as caring were also positively influenced by SPNA, while the perceptions of formal authority, similar to the Endline, were not. In terms of power perceptions (but not legitimacy) some SPNA outcomes seem to increase the perceived influence of non-executive and informal authorities more than of the official state representatives of the executive branch. There are two possible explanations for this. It may relate to the fact that DDAs increasingly specialised as mediators or transmission belts not only between communities and the state, but also between communities and different kinds of non-state authorities. Hence, a more visible and active DDA would also strengthen the perception of authority of such informal strongmen.

However, at the time of the endline the relationship between immediate SPNA outcomes, and perception of who is in charge in a district, are mostly driven by the infrastructure component. This observation may point to a qualitative finding reported above, according to which in some districts elected representatives seek to please their electorate by providing infrastructure projects to loyal communities, in the sense of typical pork-barrel clientelistic politics. If this interpretation is correct, it would be a case of reverse causality in terms of the impact hypothesis: here, informal rule would cause a higher incidence and appreciation of SPNA infrastructure projects. This interpretation would only be plausible under the following assumption: Since SPNA identifies infrastructure projects via the DDA, it would imply that this kind of informal rule influences the decisions of the DDA while making the DDA more visible as service provider in some sectors (which would explain the negative correlation between sector-specific DDA assessment and formal rule in Table 7).

In other words: at the time of the endline, with formal state rule retreating across many districts, patronage via elected officials is growing in importance as development provider or facilitator. Qualitative evidence, however, suggests that this influence is driven by political priorities and ignores best development practice, hence running against SPNA governance and stabilisation goals.

At this point, we are unable to be certain with regard to which interpretation of causation is more plausible – some SPNA outcomes may foster perceptions of informal authority; informal authority may, however, also in some cases (political patronage) foster visibility and appreciation of some SPNA outcomes. However, in terms of authority and in stark difference to the consistently positive link between SPNA and political trust in formal government institutions – there is an unintended relationship between SPNA outcomes and informal rule.

As in past analysis, the visibility of infrastructure projects has a negative effect on corruption perceptions, while better DDA and infrastructure assessments correlate positively with lower corruption perceptions. This finding is in line with expectations, since people often associate infrastructure projects with corrupt practices. Hence, we would expect that the assessment of the quality of infrastructure and the facilitating organisation – the DDA – would be more effective in reducing perceptions of corruption.

Compared to the pooled cross-sectional analysis, the consistent negative SPNA-associated effect on local social cohesion of prior waves all but vanishes in Wave 4. Only DDA visibility and general assessment show a negative correlation which, however, turns insignificant once we introduce controls for baseline values (see comprehensive table in annex). Before, DDA visibility, DDA assessment and infrastructure assessment produced significant negative correlations, indicating a likely negative effect of immediate SPNA outcomes on local social cohesion, measured by the willingness to participate in community level hashar. We interpreted this as a possible modernisation effect, resulting from an increasing takeover of more local obligations by the district level. With the general decline of both state-provided and DDA-provided district-level governance functions outlined above (*Indicators/dependent variables*) this change fits the picture. This association contrasts, however, with our qualitative observation of large-scale and frequent hashar activities organised by DDAs and CLDCs. We are still seeking to better understand these two seemingly contradictory but nonetheless credible observations.

Development & Economy

DDA visibility and sector-specific assessments correlate positively with all four dependent variables defined in the stabilisation field of development and (household) economy. Respondents knowing the DDA and appreciating its contribution to one or more development sectors are more likely to have a generally positive view of development and of government contribution to development; they are also more likely to evaluate the material wellbeing of their household more positively and are less likely to be landless. At the time of the Endline in Wave 4, the effects clearly cluster with the DDA component of SPNA; in addition, only the visibility of infrastructure shows a statistically significant and positive correlation with self-assessed material well-being.

The pooled assessment across all waves mostly coincides with the cross-sectional analysis at the endline. DDA visibility and sector-specific assessment show the most consistent positive effects, with two interesting qualifications: in prior waves DDA visibility did not correlate significantly with general development perceptions and both sector-specific as well as general DDA appreciation correlates negatively with access to land. This may imply that, under conditions of a general downturn in development activity by the time of the endline, the DDA is increasingly associated with the activities that remain. Access to land / landlessness seems to be an inconsistent dependent variable in the cross-sectional regression models (where respondents are compared across districts). The underlying

hypothesis that SPNA-induced DDA competence and activity may have a (marginal) effect on fairness of access to land, i.e. may reduce the trend towards land concentration, is more convincingly checked when we compare changes within villages over time (see *panel analysis* below). Increased access to a capacitated and competent DDA, while access to land increases at the same time in the same villages, would be an indication supporting the hypothesis.

A further difference in the pooled assessment worth mentioning is that infrastructure visibility and assessment positively correlate with a wider range of dependent variables relating to development and the economy in the models (see annex). It seems that the effects of SPNA infrastructure on these stabilisation variables fade towards the endline.

Adaptation/modernisation

At the endline, the former positive (reducing) effect on perceptions of foreign development as a value threat turns negative for DDA sector-specific assessment and for infrastructure visibility (the latter only before adding the baseline values as a control variable). Respondents with a more positive evaluation of the DDA's contribution to development in different sectors, now tend to hold more critical opinions about the negative effects of foreign development aid on local norms and Islamic values. No other relationship is statistically significant. The mitigating factor of the DDAs – pronounced in Wave 3 and visible in the pooled cross-sectional analysis – seems to have gone by the time of the Endline. In earlier waves, more positive assessments of the DDA and higher visibility of the DDA predicted lower value-threat perceptions with regard to international development aid. However, before the endline, higher appreciation of SPNA infrastructure was associated with a stronger sense of value threat, a correlation that subsequently turned insignificant. These observations are difficult to explain. Increasing Taliban influence may foster more conservative interpretations of development. It is also possible, as we have discussed it in the qualitative section on immediate outcomes (see sections on "*DDA meetings and female participation*" and "*Conflict resolution*"), that as a result of trainings, DDA members adopt more moderate positions on certain issues such as women's participation in public life or the authority of household heads. This moderate shift in values might lead to certain tensions between DDAs and "ordinary" respondents (male heads of households) which is expressed in doubts regarding the compatibility of development with traditions where DDAs are more active and visible.

In sum we, nonetheless find that **a clear majority of observed SPNA effects** at the time of the endline **is positive**. This is particularly consistent for subjective local security, for the perception of the local administration as caring, and for all development variables, both for development aid and for the household economy. These are vital dimensions of stabilisation that SPNA intended to influence in positive ways. We record likely positive net effects of SPNA even for those dimensions that are in decline in absolute terms, that is which show a negative trend over the waves (e.g. subjective district and household security assessment started rather positively in Wave 1 and deteriorated since then in most provinces continuously. In fact, compared to the baseline all indicators with the exception of self-assessed material wellbeing, have declined).

However, we see fewer statistically significant positive effects for endline than for the whole period taken together. This is mostly driven by the more consistently positive results in Wave 3 2015 and reflects the different trends we have reported in earlier reports for the prior waves. In the first follow-up (Wave 2) we did not identify any consistent effects across all districts. Instead, effects were

fragmented, in part contradictory and district-specific. Our main impression was that consistent effects under rather different framework conditions take time to unfold. Our recommendation was to fine-tune SPNA to the specific potentials and needs of different districts and district clusters. Hence, if in some districts, infrastructure visibility added to perceptions of corruption or value threat, but not in other districts, we felt that a stabilisation programme in its early stages needed to be flexible enough to adapt to such findings and tailor the approach to the local context.

The second follow-up (Wave 3) confirmed this early impression. By 2014/15 we identified a range of effects associated with SPNA that showed statistically significant and mostly positive SPNA-induced effects across the whole sample, confirming that consistent stabilisation effects needed time to emerge. This was the time when DDA organisation was at its peak, the Afghan state was still rather actively taking on the security challenges posed to it after the wind-down of foreign troop presence, and many Afghans in the assessment region were hoping that a political peace process with the Taliban might be possible (Wave 3 was implemented before the first fall of Kunduz to the Taliban). It was, however, also a time in which the importance of informal powerholders increased while concurrently district administrations grew weaker.

Moderators

Alongside the main effects of SPNA's immediate outcomes on stabilisation reported above (Table 7) **we want to know** (a) **if differences in the process of delivery influence these stabilisation effects** (maintenance and participatory procedures), (b) **if different sectors of infrastructure projects strengthen or weaken main effects** and (c) **if the degree to which the village-level shuras are functional and upwardly connected affect stabilisation** of SPNA.

Expressed in statistical terms, we are looking at the interaction of those moderating variables with the main predictors (the variables measuring SPNA's immediate outcomes). Technically, we are interested in whether the interaction effect of the moderator is statistically significant, and if it strengthens, weakens or even flips the main effect. We are also interested in how the relationship between the main predictor variables and the dependent stabilisation variables change at different values of the moderating variable (see annex for an explanation of the statistical procedure and the detailed tables summarising the results of the regression models with interaction variables).

Note that for maintenance, participative modes of delivery and the sectors, we have variation only at district level. The variables *maintenance and participation* represent averages of dummy-coded (yes and no) SPNA projects per district; and the variable *sectors* represent a dummy code per district that tells us if a district implemented an SPNA project belonging to a specific sector. These variables can only provide a rough approximation of the concepts we seek to measure. A further (analytical) problem is that in the majority of districts all projects are being maintained (22 of 35) and have been implemented by a participatory procedure (24 of 35). Hence, we have limited variation and results may be driven by only a few districts. The functionality of the village-level development shuras (CDC) is coded based on three questions of the village profiles (influence of CDC; connection to cluster level shuras; knowledge about work of DDA). This moderating variable is, hence, measured at CDC level and is thus more robust in its variation.

The picture emerging for our moderating variables is mixed. If we only take the strongest and statistically significant results into account, we can identify the following patterns:

Maintenance increases the already mostly negative relationship between SPNA outcomes (a positive DDA assessment) and the perception of official representatives of the state being the *most powerful person in the district* (Gov3). (As mentioned, this does not lead to the appreciation of the most negative types of actors such as commanders or the Taliban, but of the intermediate category comprising members of the legislative or traditional informal authorities). For low maintenance districts this relationship is still positive, for high maintenance it flips to negative. We find a similar dynamic for the effect of the general assessment of the DDA on *perceptions of corrupt conflict regulation* (Gov4): in districts with low maintenance scores this effect is positive, in high maintenance districts more positive general perceptions of the DDA predict worse perceptions on fair conflict processing. This pattern is repeated for our *poverty indicator "access to land"* (Dev4). In districts with low maintenance scores, positive DDA perceptions go together with more self-reported access to land, and vice versa for districts with high maintenance rates. Only for the effects of DDA visibility and infrastructure assessment on value-threat perceptions is the interaction effect for maintenance positive and strong, turning a negative in to a positive main effect.

The results should, however, be interpreted with caution because of the aforementioned limitation of variation in the data at time of the endline (i.e. maintenance was very widespread). Low maintenance rates (score equal or below 0.5) are driven by only five districts (Chah Ab, Taloqan, Qala-i Zal, Wardooj and Yamgan, four of which are fully or at least mostly not under government control). Only 11 districts are below the average of 0.87 on the maintenance scale.

Participation mostly increases stabilising effects. Of the statistically significant effects 13 are positive, and six are negative. However, as explained above (*Modes of delivery*), the effects are also driven by a few districts (mostly the four Baghlan districts as well as Fayzabad and Taloqan) with low scores for participative project selection. As expected, for most DDA-related SPNA predictors, participation increases positive effects on stabilisation. The picture is more mixed for the infrastructure-related predictors (infrastructure visibility and household level assessment).

The **shura functionality index**, capturing variation at village rather than district level, shows clear and expected results: in 14 statistically significant cases, better upward integration of village shuras boosts the stabilising effects of SPNA, while only in four cases do we record negative interactions. The negative interactions relate to the DDA assessment's effect on village-level social cohesion (hashar obligations), to perceptions of who is the most powerful person in the district, and to general development perceptions. In these cases, a better integration of a village organisation seems to weaken effects achieved via district-level DDA capacity-building.

This may imply that the villages are more active and autonomous to deal with hashar and implement projects, and do not rely that much on the DDA. In all other cases, a better integrated village shura increases the stabilisation effects of SPNA programme outcomes, both for the infrastructure and capacity-building component. The effects are most consistent for security perceptions, for corruption perceptions and for self-assessed material well-being. For these dimensions of stabilisation, the better the village shura functions and is linked to higher levels of the shura structure, the higher are the stabilisation effects of SPNA's immediate outcomes.

Finally, we want to know if we can identify patterns for how different **sectors** of infrastructure projects implemented by SPNA interact with the main effects of DDA and infrastructure visibility and assessment. We look at the four most common sectors – schooling, roads, health and agriculture. Statistically this part of the analysis is weaker than the other three interactions reported, for two reasons: first, most districts implemented projects form different sectors; this means we only code if

a specific sector was present in a district while ignoring what other sectors have or have not been covered. Second, sectors themselves may not be independent/neutral in terms of our predictor, specifically infrastructure visibility (different sectors may have different implications with regard to the visibility of the infrastructure created). Hence, with a note of caution, we only summarise the strongest and most consistent findings for the sectors here and ignore the infrastructure visibility predictor.

Schools seem to positively influence the effect of DDA assessment on social cohesion (hashar), general and state-associated development perceptions, and self-assessed material well-being. For the latter, the interaction effect of DDA visibility is, however, negative. In other words, for districts where schools were implemented, a more visible DDA is associated with a lower self-assessment of the material well-being of households in comparison to districts where no schools were implemented. For the infrastructure component, the effects of positive assessments on security perceptions and poverty (landlessness) are stronger in districts that implemented SPNA schools. At the same time, the association between infrastructure assessment and the quality of official power in the district is more negative in these school districts. Note that most districts implemented SPNA schools and, hence, the interaction results are driven by a relatively small number of districts that did not implement schools (see above, Figure 73).

Road infrastructure increases the effect of the capacity-building component (DDA visibility and assessment taken together here) on security, perceptions of the district administration as caring, the state contribution to development as well as material well-being and less dissatisfaction with the public school curriculum. On the other hand, road infrastructure interacts negatively with effects on hashar compliance and general perceptions of development. It seems that road projects are good for boosting the relationship between DDA-related and state-related stabilisation outcomes. For infrastructure assessment, road projects increase the effect on the view of the district administration as caring and on lower perceptions of corruption in conflict regulation.

For the capacity-development component of SPNA (the DDA) *health* infrastructure interacts positively only on perceptions of more formal actors being the most powerful people in the district. On the other hand, districts with SPNA health projects tend to have weaker effects of DDA assessment on security perceptions and state contribution to development (in the latter case the interaction even switches the main effect from positive in non-health districts to negative in districts with health projects). The interactions of health projects with the effects of infrastructure assessment on perceptions of a caring woliswali and the school curriculum are negative.

The only noteworthy interaction for infrastructure projects in the *agricultural* sector is a positive effect on the relationship between the general assessment of the DDA and a perception that power in the district is exercised by official authorities (in districts with agricultural projects this relationship is positive, in those without it is negative).

District effect clusters

In the subset analysis of SPNA effects on stability, we use a statistically simplified model with only four dependent variables – one for each field of intervention – measuring stabilisation. Like in the previous model, we look at the relationship between SPNA and stability measurements across **all** waves, using the full sample.

We then look at how the main effects of our six SPNA predictors on the four dependent stabilisation variables change between different district clusters. These clusters are defined (a) by the degree of violent fighting they experienced over the whole assessment period (1 is low, 3 is high), (b) by the degree to which state presence was felt throughout the whole period (1 is low, 3 is high) and (c) by their access to modern road infrastructure (0 is no, 1 is yes). In other words, we want to know if effects are stronger, weaker or just different between districts consistently exposed to more or less violence, where the state is consistently more or less present, and between districts that have consistently more or less access to modern road infrastructure.

Most **main effects** of SPNA in this simplified model are positive and statistically highly significant. The reduction of dependent variables does not weaken, but rather focusses the earlier results with 12 stabilisation outcome variables. Table 8 shows that DDA and infrastructure assessment relates to positive subjective security perceptions, DDA and infrastructure visibility as well as DDA assessment have positive net effects on our new composite governance and development perceptions, as well as on a reduction of value-threat perceptions. Only the effects of infrastructure assessments do not correspond to this consistent positive trend and are even negative for value threat perceptions.

	DDA vis	DDA ass gen	DDA ass sect	Inf vis	Inf ass dis	Inf ass hh
Security						
Governance						
Development / Economy						
Modernisation / Adaptation						

Table 8: Summary table for main effects of immediate SPNA outcomes on integrated/simplified stabilisation variables

In order to better understand under what conditions SPNA outcomes are more likely to achieve maximum stabilisation effects, and under what conditions these effects may be less pronounced or even negative, we group the districts into clusters according to three criteria that we believe can influence stabilisation effects (see above, *security*, *state presence* and *connectedness*). We then compare how SPNA effects on the different dimensions of stability are different between those groups. As before, we only report results that are statistically significant at the 95 percent level or above (see annex for a detailed explanation with example on how to interpret the marginal effects).

We will now report the most relevant results in accordance to the expectations we had when we defined the district effect clusters in the first place.

Tables Table 9 to Table 11 summarise the reported results. The symbols mean the following:

ECD means Expert Coded District cluster

↘* = means that the interaction with the main effect is statistically significant and negative (more violence, more state presence or access to modern roads decreases the main effect on stability).

↗* means that the interaction with the main effect is statistically significant and positive (more violence, more state presence or access to modern roads increases the main effect on stability).

(+) = means that at different or all levels of the interaction variable (e.g. low, medium, or high state presence) the correlation between the SPNA predictor variable (e.g. DDA visibility) and the stabilisation dependent variable (e.g. governance/political trust) is statistically significant and positive.

(-) = means the same for a statistically significant but negative correlation.

Exposure to violence

Security: We tentatively expect that positive SPNA effects on security perceptions are higher in areas affected by high levels of violence, for more consistent state presence and better-connected districts. The reasoning is that meaningful development work in highly insecure areas is rare and a psychological effect on security perceptions should be greatest here. Where we do observe meaningful interaction effects for **violence** they support our expectations – the effects are stronger in more violent districts (see the example explained above in detail).

Governance: Based on earlier studies we expect higher levels of continuous **violence** to influence the effect of SPNA on governance perceptions negatively. The analysis of marginal effects generally supports this assumption. The only exception is a positive effect of infrastructure assessment at household level on governance perceptions in more violent districts.

Economy / development: We would expect SPNA effects on this core dimension of stability that development work can hope to influence (rural development and poverty reduction) to be strongest under peaceful conditions (**security subset**) with strong state presence and in well-connected districts. The results here are mixed, possibly since the dependent stabilisation variable contains two dimensions: development satisfaction and the material well-being of the household. DDA visibility has the highest effect in peaceful and medium violence affected districts, while the DDA assessment shows best results in violent districts or across all three district clusters. However, none of the interaction effects themselves are statistically significant.

Value threat/modernisation: SPNA, when implemented as planned, has a high potential to reduce religiously informed, sometimes ideologically driven fears that foreign development may be a threat to local norms and Islamic values (main effect). It works via procedures owned by representative local bodies, is participatory and interacts in a meaningful and transparent way with societal as well as governmental institutions. Hence, we expect SPNA's effect on openness to development to be strongest in less violent (and, hence, less ideologically contested) areas, in areas of high state presence and in well-connected rather than remote districts. However, the DDA effect on a reduction of value threats appears to be strongest in violent districts. The infrastructure component works differently altogether. Infrastructure visibility in peaceful and medium peaceful as well as remote districts has a positive effect. Infrastructure assessment, however, has a strong negative effect across all levels of violence, in districts with medium and high state presence and in remote as well as connected districts.

	DDA vis	DDA ass gen	DDA ass sect	Inf vis	Inf ass dis	Inf ass household
Security						
ECD Violence	(-) for medium	(+) all	(+) all	(-) all	(+) for high ↗*	(-) peaceful (+) violent ↗*
Governance						
ECD Violence	(+) all	(+) peaceful & medium ↘*	(+) peaceful & medium ↘*	(+) all		(+) violent ↗*
Development / Economy						
ECD Violence	(+) peaceful & medium	(+) violent	(+) all			
Modernisation / Adaptation						
ECD Violence	(+) medium & violent ↗*	(+) medium & violent ↗*	(+) violent	(+) peaceful & medium ↘*	(-) peaceful & medium	(-) all

Table 9: Interaction effects for District Effect Cluster Violence

State exposure

Security: It is SPNA's intention to improve the interaction between societal institutions and state bodies for development; hence, continuous state presence is a precondition to achieve this goal also for the case of subjective security. For state presence the result is even stronger. More consistent state presence over time increases positive effects of SPNA on subjective security.

Governance: State presence should in theory have a positive effect, unless the state is only trusted while it is far away, as a promise or idea. The analysis of marginal effects generally supports this assumption. Partial effects are generally stronger in more state-exposed areas.

Economy/development: Following a similar logic as for the violence exposure subset, we would expect SPNA effects on economy/development to be strongest under conditions of strong state presence. **State presence** works for the most part as expected and shows only positive significant interaction effects (increase in state presence increases positive effects of SPNA on development/economy). What is interesting, however, is that in two cases the direction of statistically significant relationships switches direction between district clusters: for medium state presence, DDA assessment has a negative effect on development perceptions, for high state presence the relationship is positive (for low it is not significant); for infrastructure assessment[15] the effect is negative for low state presence and positive for high state presence.

[15] Note again that we only report district-level results if household level shows similar results, since the variation for district level responses is too low to produce robust results on its own.

Adaptive change/modernisation: For the reasons outlined in the previous subsection, we expect SPNA's effect on openness to development to be strongest in areas of high state presence. For state presence, the results are as expected for the predictors relating to the capacity-building component of SPNA, i.e. to the DDAs.

	DDA vis	DDA ass gen	DDA ass sect	Inf vis	Inf ass dis	Inf ass hh
Security						
ECD StatePres	(-) for low		(+) for high ↗*	(-) for low & medium ↗*	(+) for medium ↗*	(+) medium & high ↗*
Governance						
ECD StatePres	(+) low & high ↗*	(+) all	(+) low & high ↗*	(+) medium & high ↗*	(-) low & medium (+) high ↗*	(+) high ↗*
EconDev						
ECD StatePres	(+) low & medium	(-) medium (+) high ↗*	(+) all		(-) medium (+) high ↗*	(-) low (+) high ↗*
Modern						
ECD StatePres	(+) medium & high ↗*	(+) all ↗*	(+) high ↗*			(-) medium & high ↘*

Table 10: Interaction effects for District Effect Cluster State Exposure

Access to lines of communication

Security: We think that the effect of SPNA-induced development on security perceptions could be higher in well-connected rather than remote areas because remote areas may feel that more international development activity may endanger their relative local peace. The result for access to **modern road infrastructure** is more mixed – DDA assessment and infrastructure assessment at district level are stronger and positive in their security effect in remote areas, infrastructure assessment at household level in well-connected districts.

Governance: Access to **modern road infrastructure** should increase governance perceptions. The analysis of marginal effects does not fully support this assumption. Most partial effects are similar and positive both for remote and for connected districts.

Economy/development: As outlined for the previous subsets, we expect strongest results regarding economy/development in well-connected districts. The picture is, however, mixed. For the three positive main effects (DDA visibility, DDA assessment and infrastructure visibility) the effect seems to

be stronger in remote areas. However, the interaction effects themselves, where significant, show that more connected districts may increase main effects where either both clusters show significant results, or where remote clusters show negative results.

Adaptive change/modernisation: Lastly, we expect stronger SPNA results in well-connected rather than in remote districts. For modern road infrastructure, the results are as expected for the predictors relating to the capacity-building component of SPNA, i.e. to the DDAs.

	DDA vis	DDA ass gen	DDA ass sect	Inf vis	Inf ass dis	Inf ass hh
Security						
ECD ModInf	(-) for remote	(+) for remote ↘*		(-) both	(+) remote (-) connect ↘*	(+) connect ↗*
Governance						
ECD ModInf	(+) both	(+) both	(+) both	(+) both	(-) remote	(+) connected
EconDev						
ECD ModInf	(+) remote	(-) remote (+) connected ↗*	(+) both ↗*	(+) remote	(-) remote (+) connected' ↗*	(-) remote (+) connected ↗*
Modern						
ECD ModInf	(+) both ↗*	(+) both	(+) connected ↗*	(+) remote	(-) remote	(-) both ↘*

Table 11: Interaction effects for District Effect Cluster Modern Road Infrastructure

IMPLICATIONS OF THE CROSS-SECTION ANALYSES FOR SPNA

The implications of the three different analytical perspectives presented above can be summarised as follows. **First**, **the two-pronged stabilisation approach of SPNA** - capacitating DDAs and providing infrastructure through a transparent participatory procedure - **worked**. With high likelihood the immediate outcomes of SPNA have had mostly ***positive net effects*** **in** the stabilisation **fields of security, governance, development** and **openness to modernisation**. ***Negative effects*** **cluster around** the normative ranking of **who is most powerful** in the district and, **local social cohesion** proxied by how mobilised the community is to do collective unpaid work (hashar obligations). In the fields of security and modernisation, the results are more mixed: confidence in the state as future security provider turned negative in Wave 4, and there were mixed effects for reduction of perceptions that international development work is a threat to local values and Islamic norms, as well as for rejection of the public school curriculum.

Second, most of the **positive effects increased and became more consistent over time**. In the first follow-up in 2012 we observed no consistent effects across all districts. In 2015 effect patterns had already emerged that consolidated by the time of the endline. Considering effects that are consistent across all three perspectives (Wave 4, all waves and the simplified models), the **capacity-building component** for DDAs seems to **produce most lasting (longer-term) positive results**. The effects of infrastructure projects seem to fade more quickly with time passing (weaker in Wave 4 when most projects had already been completed for some time). However, positive effects on state provided future security and value-threat perceptions, manifest in 2015, had vanished by 2017. The same is true for initially negative effects on local social cohesion and a separate, though rather strong, negative correlation between the visibility of SPNA infrastructure and perceptions of corruption in dealing with conflicts.

Third, and highly important, **we record these mostly positive results and trends for SPNA stabilisation effects against overwhelmingly negative general trends in all fields after 2012**. In other words, despite a general increase of instability, a further fragmentation of local governance, an institutional retreat of the state from the hinterland and even some administrative centres, and an increase in the ideological challenge to the international engagement, the net effects of SPNA remain positive and have even consolidated over time. This is remarkable. However, it is also obvious that even consolidated positive net effects of SPNA are far too weak to turn the tide of general disintegration in North Afghanistan. What can be claimed with some confidence is that **without SPNA the situation would have deteriorated even further**; to what extent exactly, we are, however, unable to judge.

Fourth, observable effects of **SPNA outcomes** at the time of the endline are **mostly positive, despite years of institutional limbo of** the main local counterpart, **the DDAs**. This has been a different – and arguably even more consequential – challenge to SPNA effectiveness than the deteriorating framework conditions on the ground; it has also been more avoidable. To put it bluntly, despite the explicit commitments made in the Tokyo declaration of 2012, the Afghan government has been unwilling or unable to develop, let alone implement, a coherent local governance policy in accordance with the provisions of the constitution. As a consequence, governance below the provincial level has been in limbo. Non-survey data clearly shows that this uncertainty very much affected and weakened the DDAs. They remain dependent in their core functions (development) on SPNA or similar external inputs (Takhar, Badakhshan, few exceptions in well-connected districts).

Fifth, we find **indications of functional adaptation of the DDAs**, possibly explaining the positive correlations between and across SPNA outcomes and more informal and traditional power structures of districts. While we find consistent evidence that the SPNA approach fosters the output legitimacy of the state (the state administration cares) we do not find similar evidence for SPNA strengthening perceptions that the state is effective and powerful in the districts. The DDA, equipped with a procedure to deliver infrastructure, works **as a transmission belt between the people it represents and any kind of local authority**; i.e. not only **state authority but also commanders or the Taliban**.

Sixth, while the capacity-building effects of SPNA to the DDAs seem to be more lasting than those of the infrastructure component, **DDAs** nonetheless **need the incentive of development funds to remain active.** DDAs of districts in which SPNA has concluded before Phase V, tended to be more often inactive or even fully defunct. In contrast, DDAs with still ongoing projects – even if only O&M funds – met more regularly, were more active and, while regularly meeting to discuss the implementation of ongoing projects, also fulfilled governance functions not directly linked to the project itself.

PANEL ANALYSIS

EXPLANATION OF THE MODELS

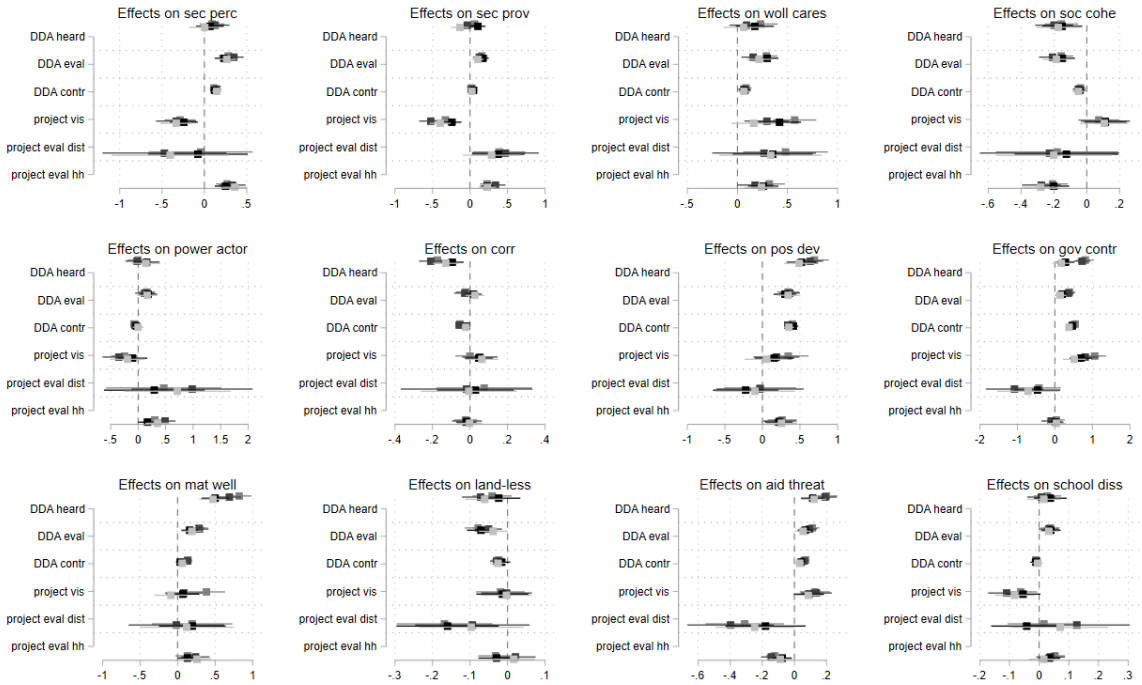
In addition to cross-sectional and pooled cross-sectional analyses presented above, we have also estimated village panel analyses to generate an additional layer of quantitative evidence on SPNA effects. These panel analyses rely on the fact that our village sample has remained constant across survey waves (with the exception of a systematic expansion of districts to be included into the evaluation). While we have not been able to interview the same household sample in every survey wave, we have been able to survey the same villages at multiple points in time. Thus, rather than comparing villages with one another, we can compare villages with themselves across time.

This approach brings about an important methodological advantage: if we find correlations between SPNA interventions and outcomes of interest, we can be more confident that unobserved village-differences do not confound these associations. The main disadvantage results from the fact that we have collected most relevant information on the household rather than the village level. In order to consider this information in the panel analysis, we have to aggregate data (usually, by estimating the mean of household-characteristics per village). This aggregate information is accurate only to the extent that surveyed households are representative of the total village population, which we tried to achieve by proper proportional random sampling of the villages. A lack of representativeness, in turn, may bias the results of our analysis. Thus, simply stated, **the village panel analysis is stronger than previous analyses in terms of identifying correlations** but weaker in terms of ensuring accurate measurement of key variables of interest. Thus**, the panel analyses** should not necessarily be seen as the strongest and most reliable element of the analysis but as **an additional contribution to methodological triangulation.**

RESULTS OF LINEAR REGRESSIONS

Figure 5 summarizes the main results of the village panel analyses. Individual figures plot the coefficients of linear regressions (Ordinary Least Squares) of the twelve outcomes of interest on the six main explanatory variables: (1) whether households have heard of the DDA, (2) how they evaluate the DDA generally, (3) how they evaluate the DDA per seven sectors (maximum score in any sector) (4) whether they are aware of SPNA projects in their districts, (5) how they evaluate the projects' impact for the district, (6) and how they evaluate the projects impact for their household. We have estimated models comparing villages in the same district and wave (district fixed-effects) and models only comparing villages across time (village fixed-effects). We have estimated both types of models in two different versions (see annex for details on the variables used): with and without a set of variables that represent potential alternative determinants of the outcomes of interest.[16]

[16] Models also control for the state of the respective outcome variable in the preceding wave (t-1) to account for time dependencies: current conditions may be a consequence of previous conditions rather than of changes in SPNA-related factors. The time-lagged outcome variable controls for such dependencies.



Results of OLS estimations; from top to bottom coefficients represent models (1) without controls and district FE, (2) without controls and village FE, (3) with controls and district FE, (4) with controls and village FE.
All models control for outcomes in the previous wave. Horizontal lines represent 95% confidence intervals.

Figure 81: Overview of results from panel regression models

We are confident that the explanatory variables correlate significantly with the outcomes of interest, when the horizontal lines (95% confidence intervals) do not cross the dashed vertical zero line. We assume that effects are consistent if this is the case for all horizontal lines per explanatory variable – representing the results of the four different models described above.

The results indicate inconsistent SPNA effects on security perceptions. A positive DDA appraisal is associated with positive security perceptions and perceptions of state-based security provision. DDA awareness does not seem to have any similar effects. Project visibility, on the other hand, correlates negatively with both security-related outcomes, while a positive project assessment seems to go hand in hand with a more positive assessment of security perception and future provision.

Effects on political trust (*woliswali* cares) are much more consistent: our estimates indicate positive effects of all four explanatory variables – they are most robust for DDA appraisal, project visibility and household-level appraisal. Five out of the six explanatory variables correlate negatively with our measure of social cohesion – three of them in a significant and robust way: DDA awareness and appraisal. Only one of the four explanatory variables (positive for project assessment on household level) correlates with peoples' assessment of powerful actors at conventional levels of statistical significance. For DDA evaluation, only models with additional controls show significant positive relationship. For interviewees' perceptions of corrupt conflict resolutions, only awareness of the DDA is associated with more negative assessments of levels of corruption.

People who are aware of the DDA, and hold a positive general as well as sector-specific view of the DDA, tend to have a more positive perception of development interventions. For infrastructure projects, only the evaluation at household level relates to more positive views of development in general. The pattern for positive development attributed to the government is slightly different: for

people's view of the government's contribution to development DDA and infrastructure visibility as well as DDA general and specific appraisal predict more positive results.

As for our second dimension of economic development, self-reported material well-being, the effects are positive for all three DDA-related predictors and household-level infrastructure appraisal. We do not find any strong evidence of SPNA effects on landlessness, only that people who hold a positive general view of the DDA seem to have a higher likelihood of being landless.

Estimations produce relatively consistent evidence that DDA awareness and appraisal as well as project visibility are all associated with a lower likelihood of people thinking that aid represents a threat. However, the actual appraisal of projects seems to correlate negatively with this outcome (robust only at household level). Finally, we do not find any consistent correlations between project indicators and people's approval of the public school curriculum.

Taken together, these analyses produce **two key findings**: (1) **positive effects of the capacity-building** component of SPNA across the implementation period **are broadly spread and outweigh by far any negative effects**. They are strongest for general satisfaction with development, but also consistent for material well-being, reduced value threat, better perceptions of the district administration and more positive subjective security perceptions. **Negative effects** are consistent for village-level **hashar compliance** (social cohesion) and, driven by DDA visibility, **for perceptions of corrupt conflict resolution**. (2) **Positive effects of the infrastructure dimension** of SPNA are **less conclusive**. This relates in part to the fact that project evaluation at district level is close to invariant positive and, hence, results are most often statistically insignificant and are driven by the relatively few observations. More importantly, however, is the observation that project visibility and assessments at district and household levels often pull in opposite directions. When project visibility predicts more negative stabilization results (subjective security, power actor, agreement with public school) the project assessment tends to predict positive results; we also find opposing prediction when project visibility relates to positive stabilization outcomes (social cohesion and value threat). Only for woliswali cares do we find that higher visibility and positive assessments both predict higher scores for a caring district administration.

Results of Instrumental Variable Regressions

In order to account for methodological challenges discussed above, we also implemented a so-called "instrumental variable" approach. This estimation strategy requires variables that are "exogenous," in the sense of not being influenced by other variables in our statistical models. These variables need to have an effect on the explanatory variables (DDA/project visibility and appraisal). Moreover, they may influence the outcome variables only through their effect on the explanatory variables. Instrumental variable estimations aim at using the exogenous variation induced into the explanatory variables to estimate unbiased effects on the outcome variables. Identifying truly exogenous instruments that meet all methodological requirements is extremely challenging. We use an instrumental variable that can only approximate these requirements: the presence of an ongoing SPNA project in a given district and year.

The presence of a project should influence project visibility and assessment. Moreover, we expect projects to stimulate DDA activity and thereby increase people's awareness and positive assessment of the DDA. Our estimations rest on the assumption that project presence influences our outcomes of interest only through its effects on DDA awareness and appraisal as well as project visibility and

evaluation. As most districts in the sample have received SPNA projects across all waves, the remaining variation of this variable is mainly driven by specific patterns of the SPNA rollout and termination that we assume to be exogenous to other variables in our model - this means that the decision on the temporal sequence of the roll-out was primarily driven by internal organisational criteria rather than by factors that relate to the SPNA outcome variables.

We first test whether this instrument does in fact correlate with our six main explanatory variables in a meaningful way. The results of these "first stage regressions" show that this is the case for DDA visibility and sector-specific contribution for the capacity-building component of SPNA, as well as project appraisal on district and household levels for the infrastructure component. People residing in a district with an active SPNA project in the preceding survey wave are on average more aware of the DDA, and hold a better view of the DDA's contribution to development; they also hold a better view of project effects on their districts and households than interviewees in other districts. Additional investigations indicate that "ongoing SPNA projects" in the preceding wave constitute a sufficiently strong instrument only for DDA visibility as well as people's view of project effects on their districts and households.

We exploit the instrument to re-estimate effects of these three explanatory variables on our stabilization outcomes. By and large, the findings support the results summarized above. The only substantive differences indicate more positive SPNA effects on respondents' security perceptions (awareness of DDA and project contribution to district) and assessments of power holders as well as more negative effects on perceptions of state contributions to security (awareness of DDA, project contribution to district and household), respondent's corruptions perception (project contribution to district and household) and people perceiving aid as a threat (awareness of DDA). The fact that the results mostly mirror those of the preceding analysis, makes us more confident that findings reported above were not driven by reverse causality and omitted variable bias. As such, the instrumental variable analyses generate additional evidence on the robustness of the main results of the panel analyses.

Alternative Channels of SPNA Effects

Previous analyses presented above measure SPNA output based on respondents' awareness and assessment of the DDA and district-level projects. It is possible, however, that SPNA produces additional positive stabilization effects through channels that remain unnoticed by individual households. In particular, support to DDAs may increase their capacity and performance in undertaking problem-solving activities that are not visible to many ordinary citizens and thereby do not influence respondents' awareness of the DDA and their evaluation of DDA performance. Neglecting such effects would lead us to underestimate potential SPNA stabilization effects.

In order to consider such contributions, we assess the effects of three variables that aim at capturing individual villages' access to the DDA and its services/activities: whether villages are in direct contact with the cluster, whether they participate in cluster activities, and whether they are aware of the DDA. Importantly, we do not derive this information from household survey data, but from qualitative interviews with village-level officials. This allows us to measure villages' potential access to DDA services independently of individual households' awareness of these activities. In many instances, villages have been deprived of any impactful DDA interventions as a consequence of a lack of a direct channel of communication and interaction with the DDA. In many instances, such channels were interrupted by low functionality and inclusiveness of village clusters that represent the primary

intermediary institutions linking villages to the DDA. The breakdowns of cluster functionality and inclusiveness, in turn, can be the consequence of a variety of factors – some of which are highly idiosyncratic such as cluster heads being sick, moving or passing away. This random dimension provides an important methodological advantage for estimations of direct effects of villages' access to the DDA: it helps alleviating so-called "endogeneity" concerns described above. When we know that villages' access to the DDA has often been determined randomly, then we can be more confident that any observable correlations of this variable with our stabilization outcomes are (1) not the result of effects of stabilization measures on villages' access to the DDA and (2) not primarily driven by any unobserved factor that determines both stabilization measures and villages' access to the DDA.

We create a simple additive index that combines the three measures of villages' access to the DDA mentioned above, and estimate correlations with all twelve stabilization outcomes, using the models introduced above. The results indicate consistent positive associations with peoples' assessment of power-holders, positive assessments of development, self-reported material well-being and – to a lesser extent –with people's assessment that the woliswali cares for the need of the people. The explanatory variables correlate negatively, on the other hand, with people's assessment of the state's contribution to security provision.

As a robustness-check we followed a similar approach for the expert-coded functionality of DDAs (DDA maturity index consisting of 11 criteria, see above *DDA capacity*). We find a significant positive correlation with security perception, most powerful person in the district, general development perceptions (but not development attributed to the government only), and material well-being, and a statistically significant and negative correlation for corrupt conflict-processing and compliance with community-level hashar obligations (cohesion).

Implications of the panel analysis for SPNA

The panel analysis supports the results of the three cross-sectional analysis presented earlier; hence, the implications for SPA are identical as summarised at the end of that section (see above, *Implications of the cross-sectional analysis*). To recap, we find that the majority of statistically significant effects are positive, that this is especially the case for development-related stabilisation outcomes, for self-assessed material well-being, for trust in a caring district administration and for subjective security perceptions. The results are less conclusive for trust in the state as future security provider, for views on who holds power in the district, for corruption in dealing with conflicts and for openness to modernisation. Positive stabilisation effects are more consistent for the capacity-building component of SPNA (i.e. the DDA-related variables) than for the infrastructure component.

Beyond the confirmation of earlier results, the integration of an instrumental variable and the investigation of alternative channels of SPNA impact add additional analytical value to this impact assessment.

Firstly, we get confirmation that our survey-independent "instrumental" variable of SPNA project exposure, indeed, influences our survey-based SPNA output proxies in a way we would expect. More exposure in the year before any survey wave leads to higher scores on four of our six main predictors for SPNA. From the qualitative research, we know that SPNA has been the principle aid programme influencing DDA performance and visibility in the districts. It is, however, not the only programme influencing the survey-based DDA predictors and this needs to be taken into account.

Secondly, the application of this external variable as a more or less random "switch" for more and for less SPNA exposure (=projects ongoing vs projects not ongoing in the previous year) between districts confirms most of the previous findings, increasing our confidence in the results. However, effects on security perceptions and district-level powerholders seem to increase in positive terms: effects on viewing the state as capable to provide future security, on corruption and on foreign aid as value threat increase in negative terms.

Third, the inspection of alternative, indirect channels of SPNA impact show that more or less random disruption of interaction between village communities and district-level governance institutions via the shura structure negatively affects three important aspects of stability SPNA seeks to strengthen: a positive view of development, self-assessed material wellbeing, the view that state officials hold power in the district, and, finally, trust in a caring district administration.

UNDERSTANDING THE IMPACT OF SPNA

SUMMARY OFANALYTICAL MODELS

The analysis of SPNA-induced variation of stabilisation outcomes between the same villages over time shows very similar results to the pooled cross-sectional analysis, where we compared variation between respondents residing in the same district. The directions for all pairs in each stabilisation field are almost identical with some differences in the degree to which they are statistically significant. We will briefly highlight the main trends and identify relevant differences.

With one exception in the panel analyses, statistically significant effects of SPNA in **the field of security** are positive. The only difference is that project visibility is statistically not significant in the cross-sectional analysis, but negative for both security outcome variables (subjective security and future security provision by the state) in the panel analysis.

In **the stabilisation field of governance**, patterns are very similar, with all positive effects on the perception of a caring district administration, and mostly negative effects on village-level social cohesion (with the exception of infrastructure visibility). The panel analysis shows less negative effects of SPNA on the perception of who is the most powerful actor in the district – but most correlations are not significant at the 95% confidence interval. For corruption we have the only real contradiction between panel and cross-section – in the panel analysis, DDA visibility has a clear negative effect on the perception of corrupt conflict regulation, while the cross-section shows a positive relationship. Also DDA assessment and contribution pull in opposite directions between the two types of regression analyses.

The panel and cross-section produce identical and positive results for all significant relationships in **the field of development** as well as self-assessed material well-being. Only for self-reported access to land is the effect of DDA assessment and DDA contribution negative in both analytical approaches.

For **modernisation,** the results are identical: positive effects of all DDA-related predictors as well as infrastructure visibility on a reduction of the perception that international development threatens local values and Islamic norms; but a negative relationship for infrastructure assessment. The relationship between SPNA predictors and disagreement with the public school curriculum point in opposite directions: negative for DDA sector-specific assessment and infrastructure visibility, but positive for infrastructure assessment and DDA-general appraisal.

How did the impact unfold

The SPNA strategy has two components – capacitating the DDAs and providing smarter infrastructure via a DDA-centred procedure. For analytical purposes, we divided these components into the dimensions of visibility and assessment. We also looked into differences across the process of SPNA implementation that would affect the impact of SPNA outputs on stability. Taken together the different components and dimensions of SPNA had consistently positive effects on stability in the target districts, manifest in the regression results since the second follow-up assessment in 2014/15. However, the causal mechanisms for the two components and the two dimensions are different for the different contextual conditions under which SPNA worked since 2010. We will now elaborate on those mechanisms that we find most relevant, based on a triangulation of the statistical results with the qualitative research conducted.

Security

SPNA is certainly not the main factor influencing subjective security and trust in the state as future security provider. However, our systematic guideline interviews clearly show that since 2012, people are increasingly worried about a drawdown of civil foreign engagement along with the wind-down of the military presence. While the opinions about the security impact of the military presence turned increasingly negative and divided, the opinions about the need for continued foreign development aid presence and input remained high. Processing this aid via Afghan institutions with a high degree of community representation at district level – and here the DDA is the only such institution since proper district councils have as yet not been formed according to constitutional requirements– is popular, and makes the ongoing locally adapted foreign engagement visible. Hence, we think that the net effect of SPNA on subjective security is mostly psychological, and grew in importance with the objective decline of other development efforts in villages and districts.

Regarding trust in the capacity of the state to provide security in the future, we find a much more sceptical assessment after the fall of Kunduz, with the fragmentation and partial withdrawal of the state from many districts, and with the increasing importance of militias and the Taliban. In this situation, it is mostly societal institutions such as the DDA that mediate between the different conflict parties and the people. DDAs and the shura structure have on numerous occasions negotiated limited ceasefires (e.g. around the Jurm bridge) and other concessions from the armed factions that make life a lot easier – and more secure – for the civilian population. Moreover, a number of interview partners also mentioned conflict resolution by the DDA (and the broader shura structure) as important for defusing conflicts before they become violent. Therefore, DDAs do not just have a psychological effect, but also a tangible impact on security. This effect is, however, unrelated to trust in the state as security provider.

This highlights a specific risk for stabilisation programmes like SPNA: the programme cannot control to whom positive effects are attributed. In terms of security we have persuasive qualitative evidence that improvements in the perception of household and district security are attributed to whoever is in charge – this may be the government but this may also be a powerful local commander or, indeed, the Taliban, who made local security their trademark and selling point. It may also be a strong head of DDA or traditional elder who successfully negotiated with Taliban, militias or state actors – conditions that matter for local security. Examples would be a one-year suspension of collection of ushr by Taliban and pro-government militias in Qala-e Zal (as far as we know only the militia under the command of Nabi agreed to this). Another example relates to compromises that were necessary for conducting Citizen Charter elections in Taliban-controlled areas of Kunduz (but not Baghlan). Often,

but not always, DDAs also succeeded in keeping access open for children to go to schools and for citizens to go to hospitals even in Taliban or commander-controlled areas.

Put bluntly, some stabilisation effects may also legitimize commander or Taliban rule. The difference between local security perceptions, and confidence in the state to provide security in the future by the time of the endline assessment, highlights this risk.

Governance

SPNA's strategy to foster the output legitimacy of the state through service delivery and what political scientists call "throughput legitimacy" through participatory and locally owned procedures of infrastructure delivery, works consistently after initial contradictory results of the first follow-up (Wave 2). SPNA has positive net effects even in times of a downturn in absolute scores of the trust in a **caring district administration,** and a general retreat and fragmentation of state presence by the time of the endline assessment. In other words – without SPNA this trend would have been even worse.

The assessment also shows that in terms of legitimacy there is no zero-sum game (one loses what the other gains) between local societal organisation (shura complex) and the state administration. Rather, higher visibility and more functional DDAs have a positive effect on perceptions of a caring district administration. This is an important and consistent finding that counters – at least at the level of subnational government – the concern of the central government that any devolution or even decentralisation of some responsibility to lower levels of citizen representation would reduce the legitimacy of the central state.

A key question regards the causal mechanisms through which DDAs manage to make the woliswali more responsive. As our qualitative interviews show, in a growing number of districts, DDAs began to represent local issues and grievances towards the district administration. As such they intervened in cases of unlawful arrests, absent teachers or when medicines were found to be expired in the local health centre. It was usually along the upwardly integrated shura structure (CDC-CLDC-DDA) that local communities communicated their grievances to the DDA, who in turn contacted the woliswal and / or relevant line departments. Communication the other way round is, however, also important. Through the DDA, the district-level government is also able to effectively communicate with the population of a district, explaining its behaviour and reaching out to the population for the distribution of for example food or agricultural inputs. This "transmission belt function" of the DDA may explain the consistently positive effect of SPNA capacity-building on a caring district administration.

The causal mechanisms at work are, however, not so clear when it comes to the question of how SPNA influences the **perception of state authority** in the districts. Similar to the argument put forward for security perceptions, performing and visible DDAs seem to work with different forms of local authority; they work with state administrations, but if those are fading or captured they also work with unofficial authorities, informal authorities, commanders, or the Taliban. However, we find that the statistically weak, but significant and consistently negative effect of SPNA immediate outcomes on perceptions of district authority is mostly driven by a correlation of higher DDA and infrastructure scores, in districts where elected representatives were most influential (parliamentarians, members of the Provincial Councils or, indeed, DDA representatives) or where informal authorities were powerful (elders, the Ulama, maliks or arbabs). Respondents considering Taliban or commanders as most powerful had, on average, lower opinions about DDA and SPNA infrastructure.

Regarding local **social cohesion** and collective mobilisation, district-level organisational capacities and village-level capacities do seem to contradict each other to some extent. Strong village shura structures and stronger district-level organisational capacities (DDA) in terms of organising Hashar seem to weaken each other. This could be a modernisation effect in the sense that some responsibilities move upwards and weaken certain aspects of community self-reliance. At the same time, this finding contradicts our observation that DDAs and their constituent CLDCs organise frequent, large-scale and, from a perspective of the common good, important hashar activities. With these large-scale district-level or cluster-level hashars communities clear major canals, maintain roads, clean snow to make remote communities accessible in the winter or react to natural disasters such as landslides, floods or earthquakes. In other words, DDAs and the associated shura structure continue to frequently organise larger and smaller scale hashars, and so far our interviews have not picked up complaints about low or insufficient participation. The association of higher DDA and SPNA infrastructure activity is thus without doubt associated with lower *perceived* participation in hashar, while hashar activities continue to be frequent where the shura structure is still active and capable. As of now this remains a riddle that we cannot solve.

In terms of **corruption** in conflict resolution, SPNA would expect that capacitated DDAs and better procedures in development delivery would lower perceptions of corruption. This does not always seem to be the case. At the endline effects for project visibility point to increased perceptions of corruption and the panel analysis records a similar result for the visibility of the DDA. DDA and infrastructure assessments are more consistently positively related to a lower perception of corruption. This finding does make sense and implies that only the quality of SPNA outputs, and not the visibility, predicts positive effects on corruption. Some DDAs are more partisan and may increase divisions and corruption, and infrastructure implementation may be prone to corruption. An additional mechanism at work is that higher involvement of communities at the politically relevant level of districts is raising awareness of corruption-related grievances.

Moreover, where the shura structure functions, it is not just the project selection procedure by 20-30 DDA representatives that is transparent and rational, but news of the decisions is quickly disseminated throughout the district. Following important district-level decisions, DDA representatives, who are by definition also the leading elected officials of their clusters, inform their constituent CDCs, who in their turn inform their CDCs' population – usually after the Friday prayer when all men of a community gather. This way, news, for example about the selection of an SPNA project including the procedure that lead to the decision, is rapidly disseminated throughout a district, leaving less room for suspicion and rumours.

Development & local economy

SPNAs effects on the core business of development works: Our measurement of ***general*** **development perceptions**, as well as of **government-associated development**, improve through SPNA. This is not as trivial as it may look at first sight. It would be straightforward and rather endogenous to find that SPNA improves perceptions of the aid it delivers via the DDAs. Similarly, improvements attributed to NGOs or international development actors would be directly affected by SPNA. However, we are looking at general development satisfaction across all actors (excluding the DDA to avoid direct endogeneity), statistically controlling for the specific effects attributed only to international development organisations and controlling for simple exposure to aid (project counts). Thus, what remains should be indirect effects of SPNA on general perceptions of development and on the specific,

state-associated development induced change. In other words, SPNA produces knock-on effects on development perceptions not directly related to the development aid it provides.

The causal mechanism behind SPNA and increased **self-reported material well-being** is even more indirect and may relate to objective improvements associated with improved development. The cause for objective improvements of the household economy is most likely driven by other factors, but a subjective association of improvements with visibility and quality of SPNA-induced change could explain this relationship.

District-level governance institutions have the potential to increase fairer **access to land** by resolving land disputes fairly and preventing excessive land grabbing by local strongmen and their associates and clients. The decrease of reported access to land between the baseline in 2010/11 and Wave 3 in 2014/15 and a subsequent increase at the endline reflects an initial increase of land concentration in the hands of fewer actors, and a reversal of this trend in the past three years. Our conflict case studies and qualitative interviews do indicate that land tenure conflicts and land grabs by more powerful groups, often with their own gunmen and commanders, has been a key grievance for many local communities in most affected districts. However, the reversal of such trends is rarely associated with DDAs or the state administration, but rather with other local patrons and increasingly – at least for Pashtun communities in Kunduz – with the Taliban (for example Khanabad district). This would explain why the SPNA association with self-reported access to land is partly negative. We cannot, however, exclude the possibility that there is a degree of reverse causality in the relationship and more vulnerable and poorer households tend to be more dependent on development efforts and hence have a higher appreciation of the DDA and successful infrastructure implementation.

Modernisation

With regard to the openness to development-driven modernisation, causal mechanisms for SPNA seem to have changed over the years. At the time of the endline, they appear to be even more influenced by ideology. Up to and including Wave 3 in 2014/15 the DDA seems to have mitigated the perception that international development aid may be a threat to local values and Islamic norms (while it may be of material benefit). This would have been the expectation from the SPNA programme as supporting inclusive, locally rooted structures. However, by the time of the endline, this positive effect seems to have ended. This may be explained by two dynamics: first, a diversification of tasks of DDAs that increasingly diverts the DDAs from the development focus they originally had. This is in part in line with SPNA's strategy to train DDAs on governance tasks that go beyond narrow development facilitation. At the same time the DDAs evolve in a changing social context that is increasingly informed by islamist hardliners, including the increasingly influential Taliban; but also local commanders under the influence of former jihadi warlords like Abu Sayyaf (an example would be Commander Qari Abdul Wardood in Shohada) or Hekmatyar (Commander Basir in Chahab). As a result of trainings and exposure visits to more liberal Muslim societies, many DDAs moved towards a more liberal interpretation of Islam. This does not mean that they moved towards a liberal "NGO-understanding" of Islam, but they opened up ideologically and this is visible in a number of conflicts, in which DDAs intervened for example to prevent the marriage of underage girls. Many accept female participation in the DDA meetings. That is – within the local context – they moved away from strict old traditions which the Taliban understand as the true understanding of the Prophet. The statistical results, however, provide some indication that this increasing openness did not moderate or change increasing conservative interpretations of foreign modernising influence.

Maintenance & participation

So far, we have simply coded whether SPNA infrastructure is being **maintained** or not (yes/no coding). Using this coding we calculated district averages calculating the ratio of maintained to not-maintained SPNA infrastructure projects. This approach was shown to be methodologically problematic as by Wave 4, almost all SPNA projects had some form of maintenance resulting in very little variance between the districts. In the end only five – in their own way rather atypical districts – had any projects with no maintenance at all. Likely as a result of very limited variance, results for maintenance/lacking maintenance in a district led to rather counter-intuitive results.

However, our further qualitative investigation of the maintenance arrangements has shown significant variance – not regarding whether maintenance is carried out or not – but *how* maintenance is being done. In some cases, local CDCs and CLDCs maintain projects on their own, in other cases, maintenance is the task of responsible line departments only, in yet other cases CDC/CLDC representatives jointly carry out maintenance tasks with representatives of line departments. We also found significant variance with regard to the involvement of DDA in maintenance activities. We are now working on identifying a new, more nuanced coding and subsequent statistical modelling to better understand the impact of different modalities of maintenance on stabilisation outcomes.

Understanding the impact of **participatory project selection procedures** – or the lack of it in cases where government officials decided on which projects should be implemented with SPNA funds – faces similar challenges, as mentioned for maintenance. Non-participatory project selection is mostly limited to the districts of Baghlan Province as well as to Taloqan and Fayzabad Districts. The DDAs interviewed by us in Baghlan Province strongly resented their exclusion from project selection and this was surely communicated down the shura structure. In Taloqan and more so Fayzabad, the DDA is largely dysfunctional and strongly captured by local power-brokers. Most other projects in most other districts were selected in a participatory manner.

The results of the statistical analysis show mostly positive effects in particular related to the DDA (visibility and assessment) and less so to SPNA infrastructure measures (visibility and assessment). Negative effects relate in particular to participation in hashar. Remarkably, transparent participatory project selection even turned the negative affect of project visibility on perceived corruption in conflict resolution into a positive one.

Our qualitative analysis shows in particular the following benefits of participatory selection by via the DDA: (a) it helps target the most urgent infrastructure gaps, (b) it does so in an accepted and generally legitimate fashion, (c) it empowers communities to take development and the future of their district in their own hand, and, (c) by disseminating the outcomes of the project selection procedure through the shura structure, it informs the population of a district about it. As a result, for SPNA we noted hardly any complaints and suspicion about the preferential treatment some communities vis-à-vis others, as is so common with other infrastructure projects. Taken together, the participatory procedure raises the profile of the DDA and shows it to be a just and transparent forum for development related decision-making.

Shura structure functionality

We have significant variation regarding shura structure functionality within our sample, allowing us to feel confident about the results. The statistical analysis shows overwhelmingly positive results

affecting almost all stabilisation variables. However, contrary to participation, which produced affects mostly related to DDA visibility and assessment, the partial effects of the shura structure cluster around SPNA infrastructure.

The probable explanation is that where the shura structure functions better, i.e. it is better integrated and feeds better into the DDA, communities are more likely to be able to communicate their development needs, shape implementation more according to their needs and influence the operationality of the completed in more positive ways. Thus while *participation* is more about the fairness of procedure, the empowerment of the DDA and the dissemination of results, the shura structure's moderating influence is likely more about the actual satisfaction of developmental and service delivery (governance) needs.

Sectors

Regarding sectors, once again we find comprehensive variance. In the statistical analysis, positive partial effects dominate (23 positive vs. 15 negative), though not as forcefully as for example regarding the shura structure. Positive outcomes dominate relating to the sector-specific assessment of DDA contribution, while negative assessments regarding infrastructure visibility. In terms of sectors, most positive results relate to roads (11 correlations) followed by schooling (8). Roads are likely successful because they potentially benefit very large population numbers, while schools were the most popular project type supported by SPNA (114 out of 203 SPNA projects were in the field of education and schooling) thus once again affecting large population numbers (not per project as roads but in total). Most negative effects were associated with healthcare, possibly linked to the difficulties associated with finding qualified staff.

In a nutshell: we gathered compelling evidence that the SPNA investment in the DDA shura structure via capacity-building and the introduction of procedures for smarter and more inclusive infrastructure projects has been a well-placed investment in relative, or net stabilisation. The programme worked, but did not change the course of history.

PRACTICAL LESSONS LEARNED

Over the past eight years, we had the privilege to assess an ambitious, but at the same time still rather explorative development programme that aimed at achieving stabilising effects at district level in a volatile and changing security, political, economic and social environment. We were free to apply qualitative and quantitative methods as we saw fit to observe and measure changing stability and attribute these changes – where plausible and evident – in part to SPNA activities, outputs and outcomes. From the beginning, AKF (Afg) advised us on the questions they, as the lead implementing organisation, were most interested in, but it did not interfere in any way with the methodology applied or with the reporting of results. This high degree of scholarly freedom also means that all technical flaws or methodological shortcomings are the sole responsibility of the evaluators or of circumstances.

The task was demanding and not in all cases were we able to answer all questions posed to us in a fully satisfactory way. Here we want to give a short overview of the most important practical lessons we learned in the course of the assessment.

Lesson 1: Start with the programme

Any impact assessment should start at the latest before practical programme implementation takes off – and ideally already during the programme creation phase. At this early stage a meaningful and accessible theory of change can still be developed, key indicators can be defined and sufficient resources in the budget can be set aside for impact monitoring as well as evaluation purposes. It is more difficult to accommodate sensible changes at a later date.

Lesson 2: Find a balance between being as relevant as possible to the needs of the client (implementing organisation) but keep the assessment technically as independent as possible

A close coordination between the organisation conducting the impact assessment and the development organisations implementing the stabilisation programme is important. In the present case, briefings and exchanges of ideas with the financing agencies in Germany, authorised by the implementing organisation in Afghanistan AKF (Afg) were also helpful. What is key is that those financing and implementing the stabilisation effort explain their information needs and expectations, and that the organisation conducting the impact assessment is transparent about what can be realistically achieved by the assessment. Also, some information on specific indicators can only be obtained by or through the facilitation of the implementing partner. A sensitive aspect are information needs from within the implementing organisation that may cause extra work for the field offices. In our experience this kind of information is not always easily obtained and thus not readily available even to the monitoring and evaluation units of the implementing organisations. In some cases we found that such internally produced information did not correspond to the reality or was otherwise unreliable (e.g. information on the DDA and implementing challenges of infrastructure projects).

While a close cooperation with the client is paramount, the evaluation should be technically as independent as possible. The credibility of the results depend on this kind of independence. Hence, evaluators and evaluating agencies should have a reputation in conducting this kind of research in a professional and academically independent manner. To this end, for-profit organisations may only be the second best choice – whenever possible and feasible a long-term cooperation with an academic institution, think-tank or institute may be considered. To some extent, ARC represents a somewhat equivalent option through the rootedness of its members in academic institutions and their background in social scientific research.

Lesson 3: Deal with fluctuation of personal and organisational changes on the side of the implementing organisation

In the case of longitudinal impact assessment over many years, the fluctuation of staff and changes in the organisational setup of the implementing and/or financing organisations may cause problems. In our case, the priorities of the assessment and the interest in integrating results of the assessment into further programming changed over time.

Lesson 4: Never underestimate the unreliability of maps and fluctuation of names and places in Afghanistan (or similar fragile state contexts with low administrative capacities)

A serious challenge to any systematic longitudinal assessment in Afghanistan is the fluctuation of lower-level units of analysis identified for repeated waves/assessments. District borders change or are often still contested and do not coincide with official maps; village names and organisational forms change frequently; households migrate. Since consistency in the units of analysis is an important condition to a number of statistical methods used, this challenge needs to be addressed and extra care needs to be taken that the field and survey teams identify and assess the same units in each wave.

Lesson 5: The pros and cons of large and changing vs. small and consistent teams

Large teams produce more individual uncertainty and mistakes, but individual mistakes have less profound consequences on average. Working with smaller and more consistent teams increases the reliability and competence of those members (see also Lesson 4) but individual mistakes have a more profound impact on the overall results. This is why we lost one question in wave one, cohort one (mistake in the way the question on self-assessed material well-being was asked by two team members) and why we had to re-do two districts in the endline assessment because of mistakes committed by three interviewers with regard to some of the more sensitive questions in Taliban controlled areas.

Lesson 6: Increasing the utility and use of insights from the impact assessment during programme implementation

This assessment shows that there are different implications regarding the use of results at different stages of programme implementation. Initially, the baseline and first follow-up assessment provide in-depth qualitative knowledge about the socio-political structure and challenges of districts and sub-districts. These initial assessments also provide statistical information on specific challenges and needs of individual districts in terms of intended stabilisation effects. This information could be used to tailor and focus the stabilisation programme in a more flexible way to the specifics of districts and district clusters. For example, some districts may show a strong link between infrastructure measures and perceptions of corruption while others show a strong correlation between bad perceptions of DDA performance and low trust in the district administration. Specifically, in the early stages, more flexible and adaptive programme implementation may make a lot of sense, not least to set credible incentives for the performance of the partner structure (here: the DDA). Another example relates to the CLDCs as a key organisational level in the transmission of SPNA activities from the DDAs to the communities (CDC-level) and vice-versa, from the CDCs to the DDAs. Increased focus on the CLDC level might have enhanced SPNA results.

Lesson 7: Coordinating monitoring & evaluation and the impact assessment

Linked to the above “Lesson 6”, the M&E of the implementing organisations should also be coordinated with and linked to the impact assessment. The benefit is that the monitoring could

provide indicators that help the impact assessment to trace the performance of outputs and intermediate outcomes as necessary steps to fulfil the theory of change of the programme.

Moreover, in order to make the most from the information obtained in the early impact assessment stages, it would make sense to set up in-house monitoring capacities that accompany the impact assessment, and are able to react to results produced from the start. This should be an additional part of the impact assessment design, aiming at lasting capacity/organisational development on the side of the implementing organisation.

Lesson 8: the importance of qualitative research (a mixed method approach)

Implementing a complex qualitative research is particularly important in cases where a fully robust experimental survey design is not possible. Qualitative research can help trace intermediate outcomes on the ground and is also essential for identifying and coding for factors that were not originally foreseen, or were not originally thought important. A case in point is the importance of the cluster structure in linking communities to the DDA and thus ultimately to the district administration.

Lesson 9: Methodological improvements when experimental designs are not feasible

For future impact assessments of similar stabilisation approaches that do not allow for proper randomization of measures, data gathering on potential instrumental variables should be included as far as possible (the identification of suitable instrumental variables is extremely challenging) right from the outset of the assessment.

RECOMMENDATIONS ON WHERE TO EXPECT IMPACT UNDER WHAT FRAMEWORK CONDITIONS

GENERAL

Stabilisation effects via capacity-building of DDAs and the provision of infrastructure projects through them are real and – with a time lag of two to four years – consistent. Based on the findings, the implications for SPNA and the lessons learned presented above we are suggesting the following general recommendation on achieving, increasing and sustaining impact of programmes like SPNA or building up on SPNA.

IMPACT REQUIRES TIME

For SPNA, it took four years for consistent stabilisation effects to materialise. This has a number of important implications for similar future programmes.

Manage and focus expectations. The political agency behind such stabilisation programmes need to understand that this is not a “quick impact approach,” but a development approach with a focus on mid-term stabilising effects. During the first two years, effects may be messy, local and contradictory.

This brings us to the next point: **be flexible and adaptive** in the first phases. During the first phase, we saw that different districts reacted very differently to the SPNA measures they were exposed to. It makes sense to tailor the programme more to the specific needs and challenges of individual districts. A competent pre-assessment makes sense. In districts with low levels of trust and high levels of corruption particularly in the development sphere, this should be taken into account (in the first follow-up we identified a number of districts in which SPNA infrastructure related negatively to corruption perceptions). Districts with high levels of grievances from pressing demands for *ushr* and infighting militias are a completely different environment. The degree of physical presence of the state varies greatly. All this would need to be taken into account. A one-fits-all approach for this initial phase may not be the best approach. This flexibility should be highlighted already in the initial offer to the donor. Initial flexibility, tailoring programme measures to different local contexts, and factoring in of a time lag in terms of consistent impact are operational challenges many development programmes face (and are often unable to address); for stabilisation programmes focussing on district-level governance and infrastructure development under conditions of violently contested state formation, it seems to be a more pressing priority.

The initial disparity of SPNA effects at the time of the first follow-up was the reason that we introduced the District Effect Clusters in the second follow-up. The idea was to systematically asses if a greater or lesser level of violence, a greater or lesser level of state penetration and a greater or lesser level of connectedness between district clusters explain in a more general way how SPNA has different effects under different conditions. We did find that **there are** indeed **differences** in how certain SPNA measures work in certain subsets. We also showed that **all three** subsets are relevant as categories for understanding differences in impact that could inform the choices and priorities a stabilisation programme can take. Nevertheless, we also found more consistent effects across all districts from the second follow-up onwards. Hence, a further investigation into what clusters of districts are more receptive for different stabilisation instruments and strategies makes sense for any continuation of the approach after the first dust has settled. The recommendation would be to further investigate the implications of such effect clusters – in close cooperation with AKF (Afg) staff and programme designers – to be as practical and concrete as possible.

For the initial implementation period of two to three years, however, a generalisation of likely effects under only a few different framework conditions is not promising – here a district-by-district analysis of specific needs, risks and potentials may be more constructive.

A third recommendation is to **clearly communicate the rules of engagement** for the programme to the principle local partner. If those conditions are not met during the first phase, it makes sense to withdraw from that district or to change to a mode of limited engagement. What should not happen is what happened with Warsaj. Here SPNA was implemented without adaptation from 2010 to 2013 despite the fact that neither DDA nor shura structure was operational at all. Warsaj may be the district with the worst examples of (non-violent) dysfunctional Kabul-level patronage. Positions in the government administration in this district have been and still are held by a former Head of the Independent Administrative Reform Commission who originates from the district. He has an astonishing track record of appointing and retaining completely incompetent people in positions of consequence. The head of DDA is a case in point. It is non-existent but at least on paper it received all trainings and participated in project identification.

Lobby government and its international backers to push for institutionalisation of district-level community representation

As of the finalisation of the report, there is a massive gap in the sub-national governance system on the district level and below. The new sub-national governance policy confirms CDCs as the main village-level governance bodies. However, and in contrast to the 2010 sub-national government policy, by not mentioning DDAs, it basically relegates this – as our analysis shows – crucially important governance body to informality and thus condemns it to a slow death. In its place it foresees the establishment of elected district councils, which should have been elected in October 2018.

However, with the cancellation of the district council elections and the de facto dissolution of DDAs, there is now a huge gap in the sub-national government system in terms of connecting and linking local communities to the district-level state. With their demise, several key aspects of governance will suffer: the delivery of district-level development infrastructure, maintenance of already existing infrastructure, a structured and effective hierarchy of conflict resolution, an institutionalised channel for communities to communicate their issues to the state – and vice versa, for the state to contact its population.

Given the current grave situation, we strongly suggest (1) to use the results of this robust finding to advocate with the Afghan government to reinstate – if necessary by decree, and if wished only as a temporary mid-term measure – the DDA and the associated cluster structure to act as a de facto district council. (2) to lobby for the design of a citizen charter type of grand programme for DDAs and clusters, and (3) in the meantime to continue supporting existing, now informal, DDAs from own and donor funds.

The alternative, i.e. to connect villages directly to the provincial level of government, would risk ignoring one historic and one contemporary challenge to state-formation in Afghanistan. Historically, political power in Afghanistan has been highly centralised only on paper. In reality, central state power tended to stop at the provincial administrative centres and hardly ever reached the communities in meaningful ways. Subnational governance around the NSP-shura structure had started to change this. The second challenge is that the state is more fragmented and in retreat than at any other time in the past decade. Clipping away the district-level representation of communities and hoping that villages and clusters will link with the state in the provincial centres is not a very likely scenario.

Be conscious of the limitations to controlling all aspects of impact

A programme like SPNA can stabilise not only official patterns of authority, but also informal or even Taliban authority. When the state retreats and concedes territorial control, security and other governance functions to local commanders, the communities themselves or even the Taliban, the DDAs and shura structure will work with whoever is in charge. SPNA has no control over the way communities and their representatives are using the capacities it built in order to survive under changing framework conditions. This is not bad per se, but the agencies involved should be conscious of this risk.

DO NOT COMPROMISE THE GOVERNANCE FOCUS FOR QUICKER AND TECHNICALLY BETTER RESULTS ON THE INFRASTRUCTURE COMPONENT

Do not bypass the principle partner organisation – the DDAs – as has been done in Baghlan. As a stabilisation programme, this link has been vital and cannot be replaced by working directly through the administration. Participatory approaches were shown to likely deliver a broad range of positive stabilisation affects.

RECOMMENDATIONS FOR THE FOUR FIELDS OF STABILITY

SECURITY

The security effect of SPNA is complex, indirect and divides into three main dimensions with different implications for recommendations.

Firstly, there is the practical dimension of increasing the **DDAs' capacity to mediate in conflicts and across frontlines**. Here, improved security is tangible and the result of improved local security governance. Mediation via councils of elders (in form of shuras or jirgas) is based on traditional local knowledge and is institutionalised in Afghanistan. This is not something Afghans need to learn, though the institutionalised pyramidal setup of the shura structure for dealing with conflicts that cannot be solved on one level, as well as some techniques, such as documenting agreements as signed contracts and gathering them for example in the DDAs book, likely make DDA-facilitated conflict resolution more efficient. However, initially the NSP-shura structure has been more about successful development and less about political issues like conflict regulation. With the drawdown of foreign development organisations' presence in the region, community representation has again become more traditional and political. SPNA capacity-building serves to combine a range of qualities and resources in an officially accepted institutional framework. DDAs combine the pre-existing mediating and political qualities of elders or jirgamars (among Pashtuns) with modern knowledge, procedures and administrations. They combine this with incentives to provide meaningful infrastructure. **Hence, capacity-building and working with the DDAs improved the impact on local security**. *This link between improving and gently modernising and institutionalising pre-existing local governance capacities to mediate and incentivise this process with new ideas of fair procedures and infrastructure provision has been a successful and promising strategy. This should inform future programmes*.

Secondly, there is the context-specific **psychological effect of continued international awareness and support**. We know from our qualitative interviews that a key concern of Afghan communities in 2012 was that they would be forgotten again once the international military leaves. While many interview partners agreed that the withdrawal of foreign forces was in principle the right decision, a majority qualified this statement by adding that it will have a positive security effect only if international support for the government and communities continued. Clearly, SPNA infrastructure projects have been a visible sign of continuing presence and support. *In the context of conditions in Northern Afghanistan, where foreign aid has been widely appreciated and where foreign support of the state has been an ideological issue only for Taliban and IS, this psychological effect is important and can be expected to be achieved by programmes like SPNA.*

Thirdly, and more ambivalently, there is the **connection that the DDA establishes between communities and the local political power structures**. During times of relative calm and while the state was extending its reach and territorial control, this was intended by SPNA and unproblematic. However, now that the state is fragmenting and retreating from many districts, improved security

provision via the DDA may be to a lesser extent attributed to the state, and more to whoever is in charge or whoever is the local counterpart for negotiations with the DDA. Under such conditions, from the perspective of local communities, SPNA may still improve stability in the field of security, but at the same time contribute to the acceptance of non-state or even Taliban local authority. In the absence of an officially accepted political peace process that would define the relationship to non-state armed groups, this is a politically sensitive issue. De facto, even the state has already established fields of mutual functional cooperation and at least temporary separate peace arrangements, where schools and hospitals in Taliban-controlled areas are receiving funding, and some development work is ongoing.

The recommendation here is to raise awareness of this risk and make clear that societal and representative organisations like the shura structure, even when backed by smart and well-resourced stabilisation programmes, cannot replace the core functions of the state in terms of presence, authoritative rule and physical security provision.

GOVERNANCE

The main issue for SPNA here is that the full potential of **stabilisation effects** on the legitimacy and authority of official institutions is **reduced by political failures of the central state** itself. These failures relate, **first**, to the fact that the government has been unable to put forward an authoritative and consistent **subnational governance policy**. The subnational governance policy it presented is now, with regard to the district level, after the cancellation of the district council elections, obsolete without a fall-back option. This vacuum of elected representation contradicts the constitution, it contradicts the interests of the state that is challenged to its core by a grassroot Islamist insurgency; and it contradicts international obligations the Afghan Government signed up to in the Tokyo Declaration in 2011. This situation leaves district-level governance in limbo. **Second**, under the National Unity Government the state **fragmented** and **withdrew** from many districts in the assessment area of North-East Afghanistan. The dynamics and reasons for this have been analysed above. This, however, means for SPNA that any **lasting sustainable governance effect is questionable**, despite the fact that we have been recording the strongest and most consistent SPNA-stabilising effects with regard to the output legitimacy of district-level state administration.

The recommendation here is the same as above: lobby and convince the central government, most importantly the IDLG, MRRD and the presidential administration, to take the district level more seriously and fix the legal limbo. Dropping the DDAs without having a functional replacement – as it looks now – is a serious mistake.

DEVELOPMENT & LOCAL ECONOMY

For the development dimension of this field of stabilisation the message is clear: providing development infrastructure via a participatory and transparent procedure locally "owned" by the DDAs results in **visible and highly appreciated infrastructure projects**. We simply have no evidence of any "development ruins" (in a handful of cases projects had been – temporarily – taken over either by the Taliban or by the military – those cases were located in Dasht-i Archi, Yangi Qala, Yamgan, Qala-i Zal and Wardooj). What is important here is that appreciation of SPNA-induced development work has a clear and consistent **effect on generally more positive development perceptions** and on **development associated with the government**. *The recommendation here is simple: advertise this*

approach, specifically to the Afghan government, emphasising that SPNA and the DDA are not competing with, but amplifying the effect of government-provided development.

In terms of **poverty reduction**, the results are consistently positive for self-assessed material well-being of the households and more negative for self-reported access to land. We are not sure if SPNA, indeed, has a direct effect of objective poverty reduction. The expected effect would be long term an indirect one (better education, better access to medical services etc.). However, improvements in the subjective material situation of the household seem to be associated with more visible and better functioning DDAs and successful infrastructure projects. While this effect could be one of reverse causality (better-off households are more likely to know and appreciate the DDA and know about and appreciate the infrastructure delivered) this would then contradict the results of our second measurement for poverty, i.e. no access to land. *Hence, we find it difficult to deduce a clear recommendation based on these findings.*

MODERNISATION & ADAPTATION

The drivers of openness towards development-induced modernisation seem to have changed over the years of SPNA implantation. The **discourse got increasingly ideological** it seems. We already noticed an increasing influence of ideology for this measurement in a longitudinal study (2007-2017) around the year 2010. In guideline interviews of 2012 and 2015 we did not see a pattern of concrete threats associated with development (very few examples) but an increase in scepticism with regard to the respect and fairness of foreign non-Muslims towards the Umma worldwide. Hence, value-threat perceptions seem to have more of a detached ideological interpretation today. This is much more difficult to influence via a programme like SPNA compared to concrete experiences with poorly implemented development aid. As long as perception of development-induced change is concrete and specific, SPNA had a high potential for mitigating fears via DDAs and participatory transparent project implementation (this worked until 2015). Ideological and abstract threat perceptions cannot be targeted at this level. We would need to investigate the sources of ideological influence more closely to see if they connect more to (international) media, i.e. social media via smartphones, radio and TV, or, as it used to be, to local ideological influencers, like mullahs, elders or community intellectuals.

CONCLUDING REMARKS

SPNA invested 105 million Euros over a time span of 8 years, with the goal of achieving meaningful stabilisation effects across a total of 52 districts of Northern Afghanistan. The programme set out to achieve this goal by supporting district-level institutions and providing high-impact infrastructure via those capacitated institutions.

HAS THIS GOAL BEEN ACHIEVED?

The task of this assessment was to facilitate the discussion of the above question by supplying evidence in a methodologically sound and transparent way. We did so by assessing what SPNA did over time in both qualitative and quantitative terms, how stability – divided into four distinct subfields of stabilisation – changed, and how these two observations and measurements of change likely relate to each other.

OUTPUTS AND IMMEDIATE OUTCOMES

When we look at **what SPNA did**, we see a programme that invested systematically in training its key partner at the district level, the DDA, and in providing infrastructure via procedures that for the most part involved the DDA as a key decision-making body.

For the *capacity-building component of SPNA,* our coded qualitative indicators (DDA functionality index in the 25 survey districts observed by us since the baseline) shows moderate improvements during the time of the most intensive interaction with SPNA (Wave 2 to Wave 3) relating to the preparation and implementation of block grants for infrastructure. The DDA functionality index increased by 18 percentage points from 48% in Wave 2 to 66% in Wave 3. However, when the main activities of SPNA in most survey districts were winding down between Wave 3 and Wave 4, this index dropped by 11 percentage points to 55%. The survey-based indicators for the intended immediate SPNA outcomes initially show little change in the visibility of the DDAs (Wave 2 and Wave 3 around 60%) but a strong increase at the time of the baseline (Wave 4 with 79%). The two DDA-assessment indicators we used, based on both general and sector-specific survey questions, show a moderate decline between the waves. The general assessment of the DDAs' contribution to district development remains, however, high (2.49 where 3 is the maximum score), the sector-specific assessment of DDA contribution, to the contrary, is consistently low (between 2.13 in Wave 2 and 1.74 in Wave 4 where 4 is the maximum score).

The development and sustainability of DDAs was clearly influenced by forces outside the control of SPNA, most importantly by the legal and organisational limbo the DDAs faced particularly towards the end of SPNA's lifespan, and their continuing dependence on development programmes' agreeing to work with them. However, across the waves there was a consistently low level of attribution to the DDAs of positive developments within individual development sectors (e.g. schooling, drinking water supply, healthcare, etc.), which shows that, despite capacitation and trainings, specific positive changes were rarely credited to the DDAs - in contrast to the much higher general positive assessment of DDAs' development impact on districts.

For the *infrastructure component of SPNA,* our coded qualitative indicator (operationality of infrastructure projects according to CDC and CLDC guideline interviews) indicate that most are fully operational (68%) or mostly operational (9%). Only 2% of coded infrastructure measures were non-operational for reasons other than that they were cancelled or were still under construction.

The qualitative assessment also provides a strong indication that projects processed in a participatory manner, via the DDA, are more likely to be coded as fully operational, than those that were not selected via the DDA. Our survey-based indicators show that the visibility of SPNA infrastructure first moderately increased between Waves 2 and 3 (from 53% to 57%) and then declined, as most projects in the survey districts had been concluded between Wave 3 and the endline Wave 4 (a drop to 43%). The general assessments of the utility of the infrastructure projects for the district are almost invariably positive across waves (between 2.98 and 2.99 when 3 is the maximum value). The assessment for the utility of the project for the household of the respondent is also highly positive (between 2.64 in Wave 2 and 2.43 in Wave 4 on a scale where 3 is the best value).

Under the given framework conditions, the infrastructure component is a substantive achievement. It implemented infrastructure that is visible, non-divisive, appreciated and (mostly) works according to design.

CHANGES IN THE CONTEXT

When we look at how the four fields of **stabilisation** developed, we see a deterioration of the security context between 2009 and 2011, which, after brief improvement, worsened again between 2013 and 2015. The decline in security incidents in 2012 is related to active counter-insurgency measures, while the decline after 2015 coincides with a decline in state presence and reflects a stalemate after the Taliban regained and expanded territory lost to the COIN operations of 2010-12. The governance context reflects these changes – formal state rule has been under pressure since the baseline in 2010/11 – first the COIN strategy replaced Taliban-controlled areas and contested areas not with stable state rule, but rather with a reappearance of the informal commander and militia system in 2012. Since then, the Taliban extended their territorial control substantially (Kunduz, Baghlan and Badakhshan). In Takhar, commanders and their militias remained dominant. However, towards the endline, state rule fragmented and retreated from many areas, including some remoter district centres. Our context indicators for development indicate a drop in village-level project activities compared to the baseline, but an increase in the number of cars per village (an indicator of economic activity). For the modernisation context, we note a decrease in the ratio between public and religious schools.

Turning to those stabilisation indicators that SPNA may influence (which are different from the general state-related and security forces-related context indicators above) we observe the following: Subjective security perceptions, as well as confidence in the state's ability to provide security to the population in the future, decline. Indicators related to the legitimacy and effectiveness of the state decline, or return to baseline values at the endline (after an initial improvement). Perceptions of corrupt conflict resolution fall below baseline values after initial improvements. Social cohesion at village level remains – on average – stable (the trends are, however, different across the provinces – Badakhshan improves while Kunduz declines).

In terms of development perceptions, overall satisfaction with development across the different sectors declines rather steeply after Wave 2, while positive developments attributed to the state suffer only a minor drop between baseline and endline. In terms of openness to development-induced modernisation, we record a decrease in perceptions of international development aid as a value threat.

Taken together, with the exception of the local economic situation (household and village) and value-threat perceptions, most of our stabilisation indicators either stall or worsen between baseline and endline assessment.

DOES THIS MEAN SPNA FAILED TO HAVE A POSITIVE EFFECT ON STABILITY?

The descriptive statistics only show how SPNA, and the stabilisation indicators it intended to positively influence, trended separately under the given context conditions. However, our statistical models show consistently, after controlling for context factors, that more SPNA exposure predicts positive net effects on most of the stabilisation indicators defined in the four fields of intervention. In other words, without SPNA, the decline in stability would have been worse. We also find that the defining aspects of the SPNA approach to delivering infrastructure – community participation via the shura structure – interacts positively with the stabilisation effects observed. Put differently, participation and the functionality of the development shura structure boost the effects of SPNA on stabilisation. These are key components of AKF (Afg)'s philosophy of doing development and engaging with beneficiary populations.

Does the impact accounted for justify the investment or meet the expectations in terms of stabilisation results?

This is likely a political question, and the answer depends on the expectations different stakeholders have about what can and should be achieved in terms of development-induced stabilisation. With the exception of one high-ranking Afghan government representative, no one we spoke to expected the SPNA to have a decisive effect on stability trends in absolute terms. Donors, implementing organisations and beneficiaries are aware that there are factors beyond the influence of SPNA that have a stronger impact on many of the dimensions of stability than a programme like SPNA. At the same time, it is an open question whether mostly positive net effects that only mitigate overall negative stabilisation trends justify such a substantial investment as was made for SPNA (note that we are only looking at stabilisation effects here, not at the more general developmental effects the programme had – as a district-level infrastructure programme the outcome is quite obviously successful).

We believe that there are a number of inhibiting factors with more general implications that decreased the potential impact of SPNA on stability – factors that were beyond the control of SPNA and not predictable at the start of SPNA.

Firstly, the continuing legal limbo of district-level subnational governance impaired the ability of SPNA to foster a process of institutionalisation of representative community-based institutions at this level. The Afghan Government has had a number of different plans, but down to the present day has failed repeatedly to formulate and implement a comprehensive sub-national governance policy.

Secondly, and in part relating to the first point, for their functioning and operating the DDAs depend not only on know-how and capacities, but first and foremost on development programmes working with or through them. However, in recent years fewer programmes have been implemented at district level and, generally, few programmes work via the DDA; furthermore there is no legal requirement for them to do so. Since the DDAs have no budget, there is little incentive, or, indeed, ability for them to continue regular operations once projects dry up.

Thirdly, the outcome of the 2014 presidential elections, resulting in a year of political stalemate and an uncomfortable US-negotiated political compromise, resulted in inconsistencies and contradictions at sub-national level which – in combination with an escalating insurgency – increased the fragmentation and eventual withdrawal of the state from many areas within districts, and even led to the full withdrawal of the state from some districts.

These factors are exogenous to district-level stabilisation dynamics and would have to be addressed at a higher level by different stakeholders.

QUOTED SOURCES

AKF-A 2010: Technical Proposal. Stabilisation Programme Northern Afghanistan: Kunduz, Takhar, Badakhshan, and Northern Baghlan (31.03.2010), Kabul, Aga Khan Foundation (Afghanistan).

AKF-A 2016: Stabilisation Programme in Northern Afghanistan : SPNA Phase V (2016-2018) Proposal, Kabul, Aga Khan Foundation Afghanistan.

BMZ 2010: Bundesentwicklungsminister Dirk Niebel: Entwicklungsoffensive in Afghanistan, in: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, 05.05.2010, at: http://www.bmz.de/de/presse/aktuelleMeldungen/2010/mai/20100505_afghanistan/; 30.03.2013.

BMZ 2014: Neue entwicklungspolitische Strategie für die Zusammenarbeit mit Afghanistan im Zeitraum 2014 - 2017. Verlässliche Partnerschaft in Zeiten des Umbruchs (BMZ-Strategiepapier), Bundesministerium für Wirtschaftliche Zusammenarbeit und Entwicklung, at: http://www.bmz.de/de/mediathek/publikationen/reihen/strategiepapiere/Strategiepapier342_03_2014.pdf; 11.09.2014.

Chiari, Bernhard 2014: Searching for the right audiance: The military's ISAF experience, in: Chiari, Bernhard (Ed.): From Venus to Mars? Provincial Reconstruction Teams and the European Military Experience in Afghanistan, 2001-2014, Freiburg i.Br./Berlin/Wien, Rombach Verlag, 11-28.

Collier, Paul/Sambanis, Nicholas 2005: Understanding civil war: evidence and analysis, Washington, D.C., World Bank.

Foundation, The Asia 2007: Afghanistan in 2007. A Survey of the Afghan People, Washington, DC.

Gaarder, Marie/Annan, Jeannie 2014: Impact evaluation for peacebuilding: challenging preconceptions, in: Andersen, Ole Winckler/Bull, Beate/Kennedy-Chouane, Megan (Eds.): Evaluating Methodologies for Aid in Conflict, London, New York, Routledge, 130-153.

Giustozzi, Antonio/Mangal, Silab 2015: JTIC Brief: Assessing the Taliban's capture of Kunduz and its ongoing implications (15.10.2015), Jane's Terrorism & Insurgency Monitor.

Grävingholt, Jörn/Leininger, Julia/Haldenwang, Christian von 2012: Effective statebuilding? A review of evaluations of international statebuilding support in fragile contexts (Evaluation Study, June 2012), German Development Institute; Ministry of Foreign Affairs of Denmark.

Islamic Republic of Afghanistan/IDLG 2010: Sub-national Governance Policy, (Islamic Republic of Afghanistan, IDLG), at: http://jawzjan.gov.af/Content/Media/Documents/SNGP-English-Afghanistan307201192625245553325325.pdf; 19.02.2013.

Juan, Alexander De/Gosztonyi, Kristóf/Koehler, Jan under review: Conditional Effects of Aid on Political Perceptions – Mixed Methods Evidence from North East Afghanistan, in: World Development.

Koehler, Jan 2014: The Afghan perspective on ISAF – changes and trends in North-East Afghanistan, in: Chiari, Bernhard (Ed.): From Venus to Mars? Provincial Reconstruction Teams and the European Military Experience in Afghanistan, 2001-2014, Freiburg i.Br./Berlin/Wien, Rombach Verlag, 65-86.

Koehler, Jan 2015: Institution-Centred Conflict Research – a Methodological Approach and its Application in East Afghanistan, in: Sökefeld, Martin (Ed.): Spaces of conflict, Bielefeld, transcript, 86-113.

Koehler, Jan/Gosztonyi, Kristof/Feda, Basir/Jawad, Hayatullah Tavakoli 2015: Stability Impact SPNA 2010-2015 : Assessing Stability in 25 SPNA Target Districts of Badakhshan, Baghlan, Kunduz, and Takhar, Berlin, Mazar, ARC-Berlin for AKF-A, unpublished report.

McChrystal, Stanley A. 2009: Commander's Initial Assessment (unclassified with removal of enclosure) (30.08.2009), Kabul, Headquarters International Security Assistance Force.

Ministry of Education Islamic Republic of Afghanistan/UNICEF/Hall, Samuel 2018: All in School and Learning: Global Initiative on Out-Of-School Children – Afghanistan Country Study, Kabul,

USAID, UNICEF, at: https://www.unicef.org/afghanistan/sites/unicef.org.afghanistan/files/2018-05/afg-report-oocs2018.pdf; 06.01.2019.

Murtazashvili, Jennifer Brick 2016: Informal Order and the State in Afghanistan, Cambridge, Cambridge University Press.

NATO 2013: Inteqal: Transition to Afghan lead, in: http://www.nato.int/cps/en/natolive/topics_87183.htm; 19.09.2013.

Nixon, Hamish 2008: The Changing Face of Local Governance. Community Development Councils in Afghanistan, Kabul, Afghan Research and Evaluation Unit (AREU), at: http://www.areu.org.af/Uploads/EditionPdfs/802E-Changing%20Face%20of%20local%20Governance-WP-print.pdf.pdf; 09.02.2013.

OECD/DAC 2012: Evaluating Peacebuilding Activities in Settings of Conflict and Fragility. Improving Learning for Results (DAC Guidelines and Reference Series), Paris, OECD, at: http://dx.doi.org/10.1787/9789264106802-en; 28.01.2013, OECD Publishing.

Saltmarshe, Douglas/Medhi, Abhilash 2011: Local Governance in Afghanistan: A View from the Ground, Kabul, Afghanistan Research and Evaluation Unit (AREU).