Water Economics and Policy, Vol. 4, No. 3 (2018) 1850007 (28 pages)

DOI: 10.1142/S2382624X18500078

Impacts Assessments without True Baselines: Assessing the Relative Effects of Training on the Performance of Water User Associations in Southern Tajikistan

Soumya Balasubramanya*, Joseph P. G. Price and Theodore M. Horbulyk

International Water Management Institute
127 Sunil Mawatha, Pelawatte, Sri Lanka
**s.balasubramanya@cgiar.org*

Received 23 June 2017
Revised 23 September 2017
Accepted 15 December 2017
Published 19 January 2018

Conducting rigorous evaluations of whether the process of creating new institutions affects their performance of mandated duties presents several challenges. Not only is assignment to process often not random, but when the process of creating new institutions starts, outcomes and other performance-influencing covariates are not measurable because the yet-to-be created institutions are not functioning at baseline. This paper compares the performance of 74 'treated' water user associations (WUAs) in Tajikistan that were created using a longer training process with 67 'control' WUAs that were created using shorter training, to assess the impact of training on WUA performance of mandated duties. First, propensity scores were constructed to estimate the probabilities of being 'treated' by treatment status. These results guided the application of the difference-in-difference technique with right-hand side covariates in a context where field measures of outcomes and other performance-influencing covariates were made after the new institutions were created and functioning. The first measures were taken within 12–18 months of the new institutions being functional and the second measures were taken 24 months after the first. This choice of methods introduces a bias due to measurement error causing an underestimate of the treatment effects, while controlling for biases due to time-invariant and time-varying unobservables. An alternative method that only compared the differences in outcomes at a single point in time after the new institutions were created would have provided an inaccurate estimate of the effects of the intervention. This is a context in which methods such as synthetic controls are

*Corresponding author.

impossible to employ due to the nature of the intervention, other macroeconomic structural changes, and severe data restrictions. The methodology employed here generates evidence that, while biased toward generating an underestimate of effect, can still be useful and informative for policy and management purposes, and for evaluating the impact of process on the functioning of new institutions in transition settings.

Keywords: Water user associations; performance; difference-in-difference; propensity scores; Tajikistan.

1. Introduction

Institutions for water management are rapidly transforming in Central Asia, as larger political and economic changes take place. Water user associations (WUAs) are now the legally mandated form of organization that provides irrigation services to agricultural land in Tajikistan (Republic of Tajikistan 2006). Some of these WUAs were created using a longer and more gradual process than others. The recent literature on institutional reform, which has focussed on *process*, rather than *design* (Denyer Willis and Mota Prado 2014) suggests that WUAs that have been created through a longer process may be more successful in performing their legally mandated functions than those created with a shorter process (Yap-Salinas 1994; Mukhtarov *et al.* 2015).

Quantitative evidence on whether longer training affects performance is important for shaping government programs that are creating new and strengthening existing WUAs (e.g., the ongoing Tajikistan Second Public Employment for Sustainable Agriculture and Water Resource Management Project).[1] However, testing the effects of training on WUA performance in Tajikistan is complicated by the fact that the length of training was not randomly assigned. Observable characteristics of the geographies where WUAs were created and unobservable characteristics of the same geographies that either are constant or are varying over time may have determined selection into longer and shorter training, consequently confounding the effects of training.

Although well-established approaches for constructing an appropriate counterfactual in the absence of randomized assignment to an intervention are available, they require treatment and control units to exist in the pre-intervention period. Two more popular methods consist of synthetic controls (Abadie and Gardeazabal 2003) and propensity scores (Rosenbaum and Rubin 1985). Both methods rely on using pre-intervention data (either covariates and outcomes or covariates alone) to select control units that look similar to the treatment units on observed

[1]Project documents may be found at http://projects.worldbank.org/P133327/tajikistan-second-public-employment-sustainable-agriculture-water-resources-management-project?lang=en.

characteristics. For example, Abadie *et al.* (2010) studied the effects of a tobacco control program in California by constructing a control group comprised of a weighted average of other US states that did not have any tobacco control programs using 19 years of pre-program data on covariates and outcomes. Rubin (2001) used propensity scores to design studies to estimate the causal effects of smoking and tobacco-company misconduct by matching smokers and never-smokers on pre-intervention covariates that were not affected by the intervention.

When treatment and control units do not exist during the pre-intervention period, experimental economics (especially laboratory experiments) could be used to construct counterfactuals; however, these methods are generally used to study strategic behavior in response to exogenous or endogenous policy changes. For example, Tellez Foster *et al.* (2017) conducted laboratory experiments to examine groundwater pumping decisions under alternative energy subsidy scenarios. In their study, all participants play a status-quo round first, after which 'treatment' participants were randomly assigned to play one of the alternative situations, and 'control' participants were randomly assigned to play another status-quo round.[2] Fischer *et al.* (2004) constructed 'treatments' to examine altruistic behavior under differential resource growth rates in the presence/absence of intergenerational links in laboratory settings.

This paper examines whether WUAs created via longer training processes perform their legally mandated functions as well as, or better than, those created via training processes with a shorter duration. This occurs in a non-strategic setting and context where WUAs did not exist before training began. The paper considers the population of WUAs in 20 districts of Southern Tajikistan and employs a difference-in-difference (DID) technique with right-hand side covariates. Since WUAs did not exist in the pre-training period, the DID technique is used with *both* sets of data collected in two time periods *after* the WUAs were created. This modified DID approach departs from the standard approach where one set of observations would have been collected at the time when WUAs were created and the other later. Though the modified approach introduces a bias in the estimated effects of training (explained further in methods section), this approach is able to control for time-invariant selection effects. This would not have been possible if the standard approach were used in this context, since all measures of WUA performance at the time of WUA creation would have taken values of zero.[3] Including covariates that affect performance on the right-hand side may help control for unobservable effects that vary over time.

[2]Tellez Foster *et al.* (2016) also conducted field experiments and compared results between the laboratory and field settings.

[3]Since WUAs did not exist at the start of the intervention, they were not performing any mandated functions.

This choice of methods is guided by the context of the intervention in question. Synthetic controls and experimental methods could not be used here due to the non-existence of pre-training data on covariates and outcomes and due to the non-strategic setting. WUAs created with longer and shorter training components were established in geographies with gravity irrigation schemes of similar characteristics that were not influenced by the treatment[4]; however, they were created in different river basins (see results section). This assignment of WUA program groups across different river basins likely lowered the costs of the WUA creation program for the government (which created WUAs using a shorter training period) and the United States Agency for International Development (which created WUAs using a longer training period); but needs to be controlled for when assessing impacts, in order to address selection effects.

This paper contributes to the literature in three key ways. First, there are many cases in natural resource management where units of treatment are created by a new program or intervention; consequently, units did not exist in the pre-intervention era. Such cases are often not considered as plausible candidates for quantitative assessments due to difficulty in controlling for selection effects and constructing a baseline. Consequently, many assessments of institutional transformations are qualitative (Theesfeld 2004; Zhou 2013; Mustafa *et al.* 2016). Quantitatively assessing performance of new institutions in their early years can yield important evidence that contributes to adaptive management, but also supports construction and testing of theories of change for the future. The modified DID technique with right-hand side covariates used in this paper can be used to estimate impacts (albeit with an underestimate of effects) when treatment and control units do not exist before the intervention begins.

A second contribution of the paper is to explore the size and direction of the resulting bias in performance estimates that is associated with the (requisite) delay in collecting the first set of data. As both treatment and control units did not exist in the pre-training era, most outcomes of interest take values of zero at the start of the intervention. With the modified DID technique, the resulting estimate is biased. However, it is less likely to be confounded by time-invariant selection effects than a simpler estimate that only examines the difference in the performance of the treatment and control groups at one point in time after the intervention is completed.

[4]These gravity irrigation schemes were constructed before 1991 in the Soviet era. Since the collapse of the Soviet Union, no expansion of gravity schemes has been undertaken. Tax revenues are not sufficient to cover the costs of such an expansion, and any expansions would require loans. Development banks such as the World Bank and the Asian Development Bank have been keen to increase the efficiency of existing surface water schemes before new schemes are constructed or existing schemes are expanded.

A third contribution of this paper is the choice of performance indicators it implements. Rather than using standards of performance such as efficiency and equity that may be more suitable while assessing established institutions, the paper uses indicators to reflect the functions that the WUA law (Republic of Tajikistan 2006) requires WUAs to perform. This approach draws from Nagrah *et al.* (2016) who examine the correlation between farmers performing mandated tasks and scheme characteristics and management perceptions. An examination of mandated functions in the case of new institutions can yield important information for adaptive management with longer term effects.

The results suggest that WUAs created using longer duration training processes perform mandated functions better than WUAs that were created using shorter training processes. Perhaps, the most significant empirical result is that WUAs created using longer training recovered membership fees from 19% more of their members; and they were also 10% more likely to hold board meetings for planning activities before the start of the irrigation season. These results contribute to the literature on institutional design by providing new evidence on the relationship between length of training and performance of mandated functions. From a policy perspective, longer training can increase the functionality of new WUAs that are being created in Tajikistan.

A limitation of this paper is that it is not possible to test formally whether the outcomes for WUAs created with longer and shorter training would have followed the same trend in the pre-training period. WUAs did not exist before the training began and consequently were not functional. The modified DID technique is followed in this paper to control for time-invariant unobservables; and the use of covariates that affect performance on the right-hand side to control for time-varying unobservables is employed in a context where many other standard methods of controlling for selection effects are unsuitable. The empirical methods used here are easy to apply during the early years following creation of new institutions. They provide timely data-driven information on the way institutions are functioning, especially in contexts in which it is not possible to separate the intervention from the units of intervention.

The remainder of the paper is organized as follows. Section 2 provides a context for examining WUAs in Tajikistan. Section 3 presents a conceptual framework and reviews literature concerning related aspects of WUA performance globally and in Central Asia. Section 4 describes the methodology, while Section 5 presents the summary statistics and results of the quantitative assessment. Section 6 aggregates the results and discusses them in relation to the existing literature and the broader context of evaluation of water institutions.

2. Context

Land-locked Tajikistan lies in the semi-arid region in Central Asia. Over 70% of its population lives in rural areas (United Nations 2015), with almost one-third living below the national poverty line (World Bank 2015). Agriculture — especially cotton cultivation — remains the main source of gross domestic product. Around 95% of crop production takes place on irrigated land (FAO 2012). Cotton and wheat cultivation is predominant in Southern Tajikistan; in the Soviet era, this was Central Asia's main cotton-producing hub (FAO 2012). Tributaries of the Amu Darya River flow into the fertile Khatlon Province in the southwest, home to the majority of the country's population.

The introduction of WUAs in Tajikistan is best considered in the context of distinct periods that characterize the country's history. During the Soviet era, large collective institutions dominated the political and economic scene. An extensive system of irrigation canals was constructed as part of the Soviet economic plan to provide water to collective farms (cooperative and state enterprises; O'Hara 2000). When the Soviet Union collapsed in 1991, 'Tajikistan did not have chance to become a state before it descended into political violence and civil war' (Matveeva 2009). The Civil War (1992–1997) undermined public services and infrastructure, especially irrigation services.

After the civil war, collective farms were decollectivized into *dehkan* (private) farms. The district irrigation departments — called the *vodkhozes* — which were responsible for providing irrigation services to the Soviet collectives were not able to cope with the challenge of providing water to the thousands of private farms. A lack of Soviet subsidies and the departure of Russian irrigation specialists compromised clarity on jurisdiction and roles, budgets for maintenance and operations, and the availability of technical skills (Gunchinmaa and Yakubov 2009; Shahriari 2009). As one response, the government enacted the *WUA Law* (Republic of Tajikistan 2006) and named WUAs as the institution that would henceforth be responsible for delivering irrigation services to the *dehkan* farms. International assistance to create these new institutions was also requested. Consequently, WUAs were piloted in the late 2000s by several international organizations, and a countrywide program to create new WUAs gained momentum in 2011–2012.

In Southern Tajikistan, WUAs were created in 2012 and 2013.[5] The WUA boards and their members were provided training from international agencies and the national government in water and financial management and in conflict

[5]All WUAs were officially registered by the agency creating them, about one to two years before the community-level process of creating them began. Consequently, the registered age of the WUA may be greater than the number of years of its physical existence.

resolution. However, one key difference among these new WUAs is the duration of time over which they were trained when they were created (ADB 2012; USAID 2014). Specifically, WUAs created through a longer training process were sponsored by USAID, whereas those created through a shorter training process were sponsored by other international agencies. On average, USAID WUAs were created with 1.89 years of training, whereas non-USAID WUAs were created with 0.73 years of training (Balasubramanya *et al.* 2016, p. 44).[6] The content of the training provided to both groups of WUAs was similar, in part because USAID-designed WUA training materials were used by the government as well.

3. Conceptual Framework

3.1. *Review of performance assessment literature*

Irrigation systems provide a prime example of common pool resources (CPRs). The constituent users of irrigation systems — water and infrastructure — have the defining characteristics of being difficult to exclude and are subject to competition among themselves when seeking to access the subtractable resources (Hamidov *et al.* 2015). Rivalry among water users may be due to inadequate water supply, issues of water distribution, or timing of water delivery. Given the distinct features of CPRs, scholars have theorized about which forms of institutional design are most likely to enable community-level institutions to perform their legally mandated functions successfully (Kazbekov *et al.* 2009).

In her seminal work, Ostrom (1990) offers a blueprint of eight core design principles that would enable CPR institutions to exist independently among local users, without being privatized or managed by central government. While acknowledging this body of literature (including Meinzen-Dick *et al.* 2002), assessing the extent to which different groups of WUAs fit Ostrom's CPR institutional design principles would be beyond the scope of this paper.[7] Such analysis would likely require techniques such as ethnographic observation, which involve considerable time. New WUAs may have to be closely observed over several years to gage if the core design principles are correlated with successful performance.

New institutions for resource management in quasi-democratic and liberalizing political economies often face a number of challenges (Skaperdas 2001).

[6]Before WUAs were created, irrigation water for farms was supplied by the state irrigation departments; the latter cannot serve as an appropriate control group to examine WUA performance of mandated duties because the role of supplying irrigation water to the new *dehkan* farms has been changed and legally handed over to WUAs. There are more than 300 WUAs across Tajikistan.

[7]In Tajikistan, all WUAs may align with a number of Ostrom's design principles. For example, user participation in decision-making, monitoring by elected members, and accessible conflict resolution mechanisms are key design features of WUAs, especially those created by USAID (USAID 2014).

New bodies may not be perceived as legitimate, and low social capital may hamper cooperation and trust (Sehring 2007; Hamidov *et al.* 2015). For example, in Tajikistan, government officials have underlined the difficulty in introducing monetary mechanisms, such as water-related fees, because of a so-called 'Soviet mentality' that expects and may encourage non-payment for services including water (Sehring 2007; Matveeva 2009). New bodies such as WUAs may find it difficult to perform their legally mandated functions.

Barcellini *et al.* (2015) stress that the duration of the process for creating participatory institutions warrants investigation because it can influence performance of mandated functions by young institutions. This focus on process partially stems from the idea that rapidly applying a universal 'blueprint' design may not fully appreciate the important contextual nuances of community water management (Smith 2008; Mukhtarov *et al.* 2015; Thiel *et al.* 2015; Ricks 2016). Longer training periods may foster productive relationships and improved functioning, leading to stronger collective action and more serviceable institutions (Nagrah *et al.* 2016). Empirical studies have also demonstrated the importance of training on WUA performance. Batt and Merkley (2010) attributed the shortcomings of WUA performance in Egypt to inadequate training. This echoes Mukhtarov *et al.*'s 2015 case studies in Turkey and Azerbaijan, where WUAs were deemed unsuccessful in performing their functions due to inadequate preparation for taking over irrigation management responsibilities. In Uzbekistan, Wegerich (2000) and Hornidge *et al.* (2013) reviewed the limited or unclear duration of training provided, which they correlated to unreliable water delivery and poor WUA performance. In contrast, Yap-Salinas (1994, p. 126) concluded that 'the longer the project, the greater the chance of success,' after an eight-year WUA project in the Dominican Republic. Also, Johnson and Stoutjesdijk (2008) drew attention to WUAs in Kyrgyzstan, which after receiving training over four years were able to improve performance of mandated functions, improving water delivery and fee collection.

3.2. *Hypotheses*

Short-term analysis of the performance of community-based irrigation management organizations is usually assessed by considering mandated functions (Yakubov 2011). For example, collection of irrigation and/or membership fees and dispute arbitration functions of WUAs have been examined in Jordan, China, and Turkey (Çakmak *et al.* 2004; Huang *et al.* 2010; Mustafa *et al.* 2016). Nagrah *et al.* (2016) consider performance indicators such as maintenance of irrigation watercourses, dispute resolution, collection of water charges, and success of monthly meetings when explaining functioning of water user groups in Pakistan.

Table 1. WUA Performance Indicators

Water delivery
WUAs had a delivery schedule
WUAs had at least two board meetings per year
Routine cleaning and maintenance of irrigation infrastructure
WUAs conducted two pre-irrigation maintenance sessions for secondary canals per year
WUAs conducted two pre-irrigation maintenance sessions for tertiary canals per year
Collect irrigation fees and transfer to government
Share of members paying irrigation service fees
Share of collected irrigation fees transferred to the government
Collect WUA membership fees
Share of members paying membership fees
WUA arbitrates disputes
Likelihood WUA arbitrates dispute between farmers

Source: Authors.

Building upon this literature, this paper considers the following indicators and introduces a number of hypotheses regarding the effects of longer training. A list of indicators is provided in Table 1.

(1) *Water delivery planning*: WUAs in Southern Tajikistan are supposed to create water delivery schedules, preferably by irrigation season. The WUAs are also supposed to hold at least two board meetings every calendar year, in which delivery schedules are finalized and other important decisions such as setting of membership fees and other water deliver-related challenges are discussed. WUAs with longer training are hypothesized to have a greater probability of having seasonal water delivery schedules and of holding two annual board meetings than those with shorter training.
(2) *Repair and maintenance of irrigation infrastructure*: WUAs are supposed to conduct routine cleaning and maintenance of secondary canals (distributaries) and tertiary canals (watercourses) at the start of each irrigation season (there are two cultivation seasons every calendar year). It is expected that WUAs with longer training are more likely to conduct two pre-irrigation season rounds of maintenance annually.
(3) *Collecting irrigation service fees and transferring them to the government*: The WUA is mandated to collect irrigation service fees from all members seasonally and transfer all irrigation fees to the government. Longer training is likely to increase the share of members from whom irrigation fees are recovered and the share of collected irrigation fees that is transferred.
(4) *Collecting WUA membership fees*: WUAs are authorized to collect annual membership fees from all members to finance day-to-day operations. Longer

training is likely to increase the share of members from whom the WUA recovers membership fees.

(5) *Arbitrating disputes*: WUAs are authorized to arbitrate disputes between members, and it is expected that WUAs with longer training have a higher probability of being the body that settles disputes.

4. Methodology

The DID technique calculates the effect of an intervention by comparing the average change over time in the outcome variable for the treatment group to the average change over time for the control group, in effect controlling for differences in the starting points and common time-trend effects in one or more measures of outcome between the two groups.

4.1. *The DID estimator treatment and control units exist in the pre-intervention period*

Consider the (standard) case in which outcomes are measured at the start of the intervention. To describe the estimator mathematically, standard notation that uses two time-periods (where time is regarded as a discrete variable) is used.

Let the policy maker be interested in understanding the effects of the intervention two years after it was implemented (short-term effects).[8] Let t refer to the number of years since the intervention, where $t \in \{0, 2\}$. Let T_t be a categorical variable that denotes the two years, with $T_2 = 1$ if $t = 2$ and $T_0 = 0$ if $t = 0$.

Let d refer to treatment state, where $d \in \{0, 1\}$, where "0" indicates the non-treated/control state and "1" indicates the treated state. In this paper, the treatment is the longer duration of training for some WUAs. Let D_i be a categorical variable that denotes the treatment status of WUA$_i$, with $D_i = 1$ if $d = 1$ and $D_i = 0$ if $d = 0$.

Let i denote a specific WUA, where: $i \in \{1, 2, 3, \ldots, k$ when $D_i = 1$; and $i \in \{k+1, k+2, \ldots, k+m\}$ when $D_i = 0$; that is, there are k treatment WUAs and m control WUAs.

Consider the following equation:[9]

$$Y_{it} = \mu + \gamma D_i + \theta T_t + \beta(D_i {*} T_t) + \vartheta_{it}, \tag{1}$$

where Y_{it} refers to a specific outcome variable for WUA i at time t; and β reports the causal effect of the treatment, which controls for pre-existing differences in

[8]In development practice, short-term effects are generally studied two years after the intervention was completed. This is a commonly observed norm and is not necessarily supported by any theoretical model.

[9]The error structure is assumed to follow: $E(\vartheta_{it}|D = 1, T = 1) = E(\vartheta_{it}|D = 0, T = 1) = E(\vartheta_{it}|D = 1, T = 0) = E(\vartheta_{it}|D = 0, T = 0) = 0$. This is because the identifying assumption is that by explicitly accounting for, the errors are uncorrelated with D and .

outcomes at the baseline.[10] Following the standard approach (Woolridge 2002) to describe the DID estimator $\hat{\beta}$:

$$\hat{\beta} = \{E(Y|D=1, T=1) - E(Y|D=0, T=1)\} - \{E(Y|D=1, T=0) - E(Y|D=0, T=0)\}. \tag{2}$$

Therefore

$$\hat{\beta} = \left(\frac{1}{k}\sum_{i=1}^{k} Y_{i2} - \frac{1}{m}\sum_{i=k+1}^{k+m} Y_{i2}\right) - \left(\frac{1}{k}\sum_{i=1}^{k} Y_{i0} - \frac{1}{m}\sum_{i=k+1}^{k+m} Y_{i0}\right). \tag{3}$$

Alternatively:

$$\hat{\beta} = \frac{1}{k}\left(\sum_{i=1}^{k} Y_{i2} - \sum_{i=1}^{k} Y_{i0}\right) - \frac{1}{m}\left(\sum_{i=k+1}^{k+m} Y_{i2} - \sum_{i=k+1}^{k+m} Y_{i0}\right). \tag{4}$$

The DID technique controls for time-invariant unobservables (μ), such as any WUA-specific, area-specific, or agency-specific fixed effects that are constant over time but may drive differences in level of performance.

4.2. *The DID estimator when treatment and control units did not exist in the pre-intervention period*

To demonstrate the DID technique when outcomes for treatment and control groups take values of zero at the start of the intervention, time is now defined as a continuous variable t_r, where $t_r \in (0, \infty)$.

Let $A(t)$ represent outcomes for the treatment unit, which vary with time t. Let $C(t)$ represent outcomes for the control unit, which also vary with time. Due to the nature of the outcome, both $A(t) \geq 0$ and $C(t) \geq 0$.

Let the difference in magnitude of outcomes between the treatment and control units at any time t_r be defined as $\delta_r = A(t_r) - C(t_r)\, \forall t_r \in (0, \infty)$.

Let the intervention be implemented at t_0 for the treatment unit. Since treatment and control units did not exist at the start of the intervention, $A(t_0) = C(t_0) = 0$. Therefore

$$A(t_0) - C(t_0) = \delta_0 = 0. \tag{5}$$

Let the policy maker be interested in measuring subsequent performance at time t_e, where $t_0 < t_e$. For the purposes of a clearer demonstration, one unit of observation

[10]The DID technique eliminates μ and γ because it involves using the difference $(Y_{i2} - Y_{i0})\,\forall i$. The coefficient θ controls for any year-related effects.

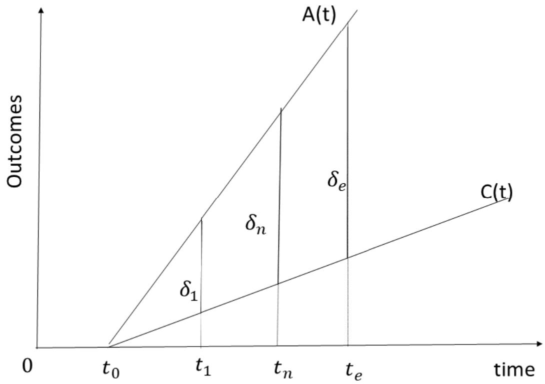

Figure 1. The Standard and Modified DID Technique When the Treatment Unit Performs Better than the Control Unit

from the treatment group and one unit of observation from the control group are considered.

(a) First, consider the case in which the treatment unit performs better than the control unit after t_0; presented in Fig. 1.

The standard DID estimator would measure changes in outcomes for the treatment and control groups between t_e and t_0. That is:

$$\text{DID} = (A(t_e) - C(t_e)) - (A(t_0) - C(t_0)) = \delta_e - 0 = \delta_e. \tag{6}$$

As treatment and control units did not exist before t_0 and $A(t) \geq 0$, $C(t) \geq 0$, and $A(t_0) = C(t_0) = 0$, it is not clear whether the standard DID technique would eliminate the time-invariant unobservables represented by μ in Eq. (1).

Suppose now that the first measures of outcomes are made at $t_1 > t_0$. In this case, the 'modified' DID estimator is

$$\text{DID}' = (A(t_e) - C(t_e)) - (A(t_1) - C(t_1)) = \delta_e - \delta_1. \tag{7}$$

Then

$$\text{DID} - \text{DID}' = \delta_e - (\delta_e - \delta_1) = \delta_1 > 0. \tag{8}$$

This implies that $\text{DID} > \text{DID}'$; that is, the modified DID underestimates a positive effect of the intervention.

Generalizing, let the first measurement of outcomes be made at t_n such that $t_0 < t_n < t_e$. Then, $\delta_n = A(t_n) - C(t_n)$. Then, the 'modified' estimator is

$$\text{DID}^n = (A(t_e) - C(t_e)) - (A(t_n) - C(t_n)) = \delta_e - \delta_n. \tag{9}$$

Let the error

$$\varepsilon(t) = \text{DID} - \text{DID}^n = \delta_e - (\delta_e - \delta_n) = \delta_n > 0. \tag{10}$$

As $A(t)$ and $C(t)$ are diverging, this implies that

$$\mathrm{DID} > \mathrm{DID}^n. \tag{11}$$

As $t_n \to t_0, \delta_n \to 0$. Therefore, $\varepsilon(t) \to 0$ as $t_n \to t_0$.

In this case, if one wishes to assess the effects of the intervention at t_e with baseline outcomes measured after t_0, then measuring farther from t_0 would further underestimate the actual effect of the intervention.

(b) Now, consider the case in which the control unit performs better than the treatment unit after t_0; presented in Fig. 2.

The standard DID estimator is

$$\mathrm{DID} = (A(t_e) - C(t_e)) - (A(t_0) - C(t_0)) = -\delta_e - 0 = -\delta_e. \tag{12}$$

As in case (a), as treatment and control units did not exist before t_0; $A(t) \geq 0$, $C(t) \geq 0$, and $A(t_0) = C(t_0) = 0$. Therefore, it is not clear whether the standard DID technique would control for time-invariant unobservables that may be driving performance of the two groups.

Consider the 'modified' DID estimator:

$$\mathrm{DID}' = (A(t_e) - C(t_e)) - (A(t_1) - C(t_1)) = -\delta_e - (-\delta_1) = -\delta_e + \delta_1. \tag{13}$$

Similar to case (a), and demonstrated in Eq. (8):

$$\mathrm{DID} - \mathrm{DID}' = -\delta_e - (-\delta_e + \delta_1) = -\delta_1 < 0.$$

In this case, $\mathrm{DID} < \mathrm{DID}'$; that is, the modified DID underestimates the magnitude of a negative effect of the intervention.

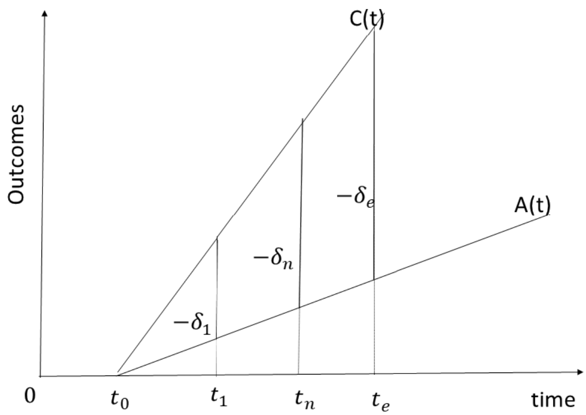

Figure 2. The Standard and Modified DID Technique When the Control Unit Performs Better than the Treatment Unit

Generalizing, let the first measure of outcomes be made at t_n such that $t_0 < t_n < t_e$. Then, $\delta_n = A(t_n) - C(t_n)$.

$$\text{DID}^n = (A(t_e) - C(t_e)) - (A(t_n) - C(t_n)) = -\delta_e - (-\delta_n) = -\delta_e + \delta_n. \quad (14)$$

Similar to case (a) and as demonstrated in Eq. (10)

$$\varepsilon(t) = \text{DID} - \text{DID}^n = -\delta_n < 0.$$

Since $A(t)$ and $C(t)$ are diverging, this implies that

$$\text{DID} < \text{DID}^n. \quad (15)$$

As $t_n \to t_0, \delta_n \to 0$. Therefore $\varepsilon(t) \to 0$ as $t_n \to t_0$, similar to the result found in case (a).

In this case, if one wishes to assess the effects of the intervention at $t_e$with baseline outcomes measured after t_0, measuring farther from t_0 would also underestimate more greatly the magnitude of the actual effect of the intervention.

(c) Adding across cases (a) and (b)

The net effect of the intervention will be biased under the modified DID technique, with the direction of bias unknown, since the modified DID approach underestimates both positive and negative effects. In this context, if some WUAs experience a positive effect of longer training duration and others experience a negative effect, then summing across WUAs provides biased aggregate measure of program effect, where the direction of bias is not known. However, as $t_n \to t_0$, $\varepsilon(t) \to 0$. Therefore, measuring as close as possible to t_0 would minimize the bias and would address time-invariant selection effects.

5. Data

There are four river systems within Tajikistan — Vakhsh, Kafarnihon, Pyandj (which are tributaries of the Amu Darya Basin), and the Syr Darya. Irrigated agriculture is practiced on only 4% of Tajikistan's land.[11] To design a study that was representative of WUA populations in gravity irrigation schemes, 10 such schemes that supply water to 164 *jamoats*[12] in 20 districts were identified.[13]

[11]About 96% of the land area is covered by mountains, where irrigated agriculture is not practiced.

[12]A jamoat is an administrative area within a district (in other countries, the equivalent of a jamoat would be a sub-district). A collection of jamoats makes a district, and a collection of districts make a province. Tajikistan is divided into four provinces and 400 jamoats.

[13]Considerable shares of area in other agricultural jamoats were irrigated by lift irrigation schemes, the operational costs of which are much higher than gravity schemes. This study covers almost every WUA in a gravity scheme that was legally registered and functioning in 2014.

A census of WUAs was conducted in late 2014 in these 164 jamoats, and 150 legally registered and functioning WUAs were identified. The chairs or managers[14] from 141 of the 150 WUAs agreed to participate in the study. Of these, 74 WUAs were created with a training period of 22–24 months and 67 WUAs that were created with a shorter training period of three months.

The chair/manager was requested to donate around 1.5 h of their time twice to provide the survey team with data. The first round of data described the 2014 calendar year and was collected in early 2015. The second round of data described the 2016 calendar year and was collected in early 2017. Respondents were not compensated monetarily for their participation. Instead, manuals on better agronomic practices, such as the use of improved seeds and application of fertilizers, were provided to be shared with the larger community.

Questions in the 2014 survey were repeated in the 2016 survey; the difference being the year about which the respondent was asked to report. The performance indicators for which data were collected reflect the various functions that WUAs in Tajikistan are mandated to perform by the *Water User Association Law* (Republic of Tajikistan 2006; Table 1).

In addition to these performance indicators, data on covariates that might affect performance of mandated functions were also collected for the 2014 and 2016 calendar years. Data collected included: the physical attributes of the WUA such as the type and magnitude of irrigation infrastructure; the area covered by gravity and lift irrigation infrastructure;[15] the service area under irrigation; the area of cotton, wheat, and fodder cultivated; the number of members in the WUA; the size of the WUA board and its gender composition; the number of disputes; and perceptions about the quality of the main, secondary (distributary), and tertiary (watercourse) canals. Measures of covariates pertaining to type and magnitude of irrigation infrastructure and the area covered by gravity and lift irrigation infrastructure remained constant between the two surveys.[16] Measures of other covariates generally varied between the two surveys.

6. Results

As assignment of WUAs to shorter versus longer training periods was not random, propensity scores were constructed by regressing each WUA's training status

[14]If these individuals were not available, then the WUA engineer or treasurer was asked to respond.
[15]Very small areas in only a couple of WUAs were irrigated by lift irrigation. This was as expected because the study was designed to focus on gravity irrigation schemes.
[16]No new irrigation infrastructure was developed nor was existing infrastructure expanded between the two surveys.

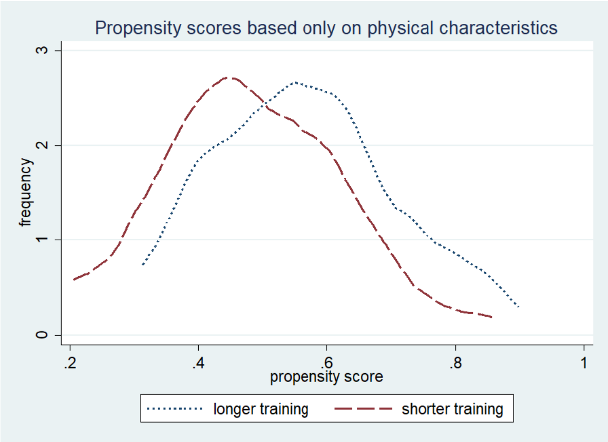

Figure 3. Propensity Scores Using Time-invariant Physical Characteristics (Excluding District and River)

(1, longer training; 0, shorter training) on a set of covariates. These included: location of the WUA (head/middle/tail) on the canal; the number of main, secondary, and tertiary canals within the WUA's command area; the share of the command area covered by gravity and lift irrigation infrastructure; the number of pumping stations within the command area; the number of drainage collectors within the command area; the number of irrigation wells within the command area; the number of years the WUA has been registered (but may not have existed on the ground); the district the WUA was located in; and the river that was the source of water for the irrigation system. These covariates were chosen because they are unlikely to have been influenced by the WUA program. Irrigation infrastructure was developed in the Soviet era and has not expanded since 1990. This is reflected in the fact that in 1994, 720,000 ha of land were irrigated, which marginally increased to 742,000 ha in 2009 (FAO 2012).[17]

Figure 3 presents results for the frequency distribution of the propensity scores associated with longer and shorter training, which were constructed using all covariates, except the district and the river. WUAs with longer and shorter training had overlapping distributions (indicating common statistical support). This implies that both groups of WUAs looked similar in terms of: location (head/middle/tail) on the canal; the number of main, secondary, and tertiary canals within the command area; the share of the command area covered by gravity and lift irrigation infrastructure[18]; the number of pumping stations, drainage collectors, and irrigation

[17]The latest year for which official data are available is 2009.

[18]There is almost no use of groundwater, because aquifers are located in deep, hard rock layers, which increase pumping costs.

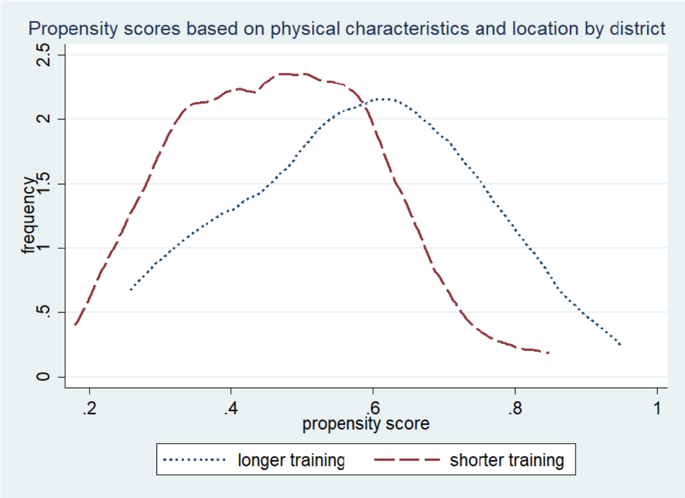

Figure 4. Propensity Scores Using Time-Invariant Physical Characteristics (Including District)

wells within the command area; and the number of years the WUA had been registered (but may not have existed on the ground). Propensity scores were then constructed by including all of those attributes plus dummies for the districts in which the WUA was located (Fig. 4). The extent of overlap between the frequency distributions is still considerable, but less than in Fig. 3. Finally, when propensity scores were constructed by also including dummies for the river that was the source of water (Fig. 5), the two frequency distributions had limited common support. Figures 3–5 suggest that while WUAs with longer and shorter training are not different in terms of their physical characteristics, they differ in terms of the river basins in which they were created. As WUAs with longer and shorter training were created by USAID and the government, respectively, program costs would have been lower with both agencies concentrating their efforts on specific river basins rather than each agency covering every river basin.

The following sections report the results from the modified DID technique with right-side covariates, listed in Table 2. There were minor changes between years in the share of irrigated area; the share of irrigated area under wheat, cotton, and fodder production; and perceptions on the quality of the primary, secondary, and tertiary canal. Covariates such as the number of registered members in the WUA, the size of the WUA board, and the number of disputes varied considerably between 2014 and 2016. Covariates such as the physical infrastructure, the share of command area served by gravity and lift irrigation, the district the WUA is located

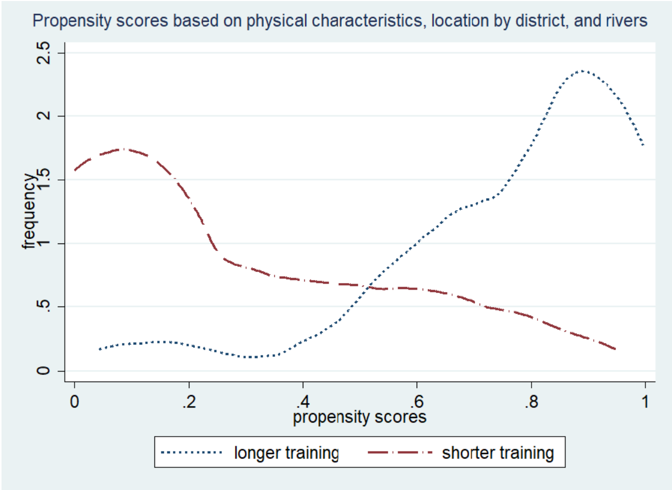

Figure 5. Propensity Scores Using Time-invariant Physical Characteristics (Including District and River)

in, and the river that is the source of water did not vary between the two years. Only covariates whose measures varied between 2014 and 2016 are reported, because those whose measures did not vary between 2014 and 2016 (such as the physical infrastructure, river, and district) are eliminated by the DID technique.

Tables 3–7 report the effects of having longer training duration on WUAs' performance of mandated duties, while including covariates (Table 2) that may also affect performance. In each table, the reported values express the modified DID estimate as a percentage difference, followed in parentheses by the standard error.

6.1. *Water delivery*

WUAs with longer training were 4% more likely to have a seasonal water delivery schedule in 2016 and were 9% more likely to have at least two board meetings in 2016, but these results are not statistically significant at the usual levels (Table 3). A 1% increase in area under cotton cultivation between 2014 and 2016 was associated with a 28% increase in the probability of having a seasonal water delivery schedule in 2016 ($p < 0.05$). This is consistent with the observation that cotton is a water-intensive crop that is cultivated in the dry season (summer). Also, a 1% increase in the area under wheat production between 2014 and 2016 was associated with a 27% decrease in the probability of having a seasonal irrigation schedule in 2016 ($p < 0.1$). This is consistent with the observation that wheat is a winter crop, which is mostly rainfed, with minimal supplemental irrigation. One

Table 2. Summary Statistics for Covariates that Could Affect Performance (Included as Right-hand Side Variables in the DID)

	2014: Survey 1				2016: Survey 2			
	Treatment		Control		Treatment		Control	
Indicator (Y)	*N*	Mean (Std Dev)	*N*	Mean (Std Dev)	*N*	Mean (Std Dev)	*N*	Mean (Std Dev)
% WUA service area irrigated	73	0.96 (0.10)	67	0.92 (0.17)	73	0.94 (0.12)	67	0.94 (0.12)
% Irrigated area producing wheat	73	0.23 (0.19)	58	0.28 (0.29)	73	0.23 (0.21)	58	0.25 (0.20)
% Irrigated area producing fodder	73	0.07 (0.16)	58	0.07 (0.13)	73	0.08 (0.06)	58	0.10 (0.13)
% Irrigated area producing cotton	73	0.47 (0.24)	58	0.42 (0.42)	73	0.49 (0.20)	58	0.34 (0.25)
No. of members in the WUA	74	240.62 (230.08)	67	285 (295.21)	74	687.41 (3424.94)	67	584.22 (1,468.40)
No. of disputes	73	24.22 (31.53)	67	32.61 (44.67)	73	14.66 (22.06)	67	33.91 (68.76)
No. of board members	73	5.28 (1.18)	67	4.10 (2.08)	73	9.32 (7.22)	67	7.17 (6.65)
% of female board members	73	0.07 (0.11)	67	0.08 (0.15)	73	0.09 (0.13)	67	0.09 (0.14)
Quality of main canal (1–5)	74	2.91 (0.58)	67	2.98 (0.67)	74	2.54 (0.92)	67	2.57 (0.96)
Quality of secondary canal (1–5)	74	3.08 (0.52)	67	2.87 (0.72)	74	3.03 (0.92)	67	2.69 (0.85)
Quality of tertiary canal (1–5)	74	3.09 (0.48)	67	2.79 (0.66)	74	3.12 (0.72)	67	2.81 (0.78)

Source: Data collected by the authors.

Table 3. Effects of Longer Training on Water Delivery

	WUA had a delivery schedule in 2014 and 2016	WUA had at least two board meetings in 2014 and 2016
Longer training	0.04 (0.07)	0.09 (0.06)
% WUA service area irrigated	0.33 (0.21)	−0.22 (0.20)
% Irrigated area producing wheat	−0.27 (0.13)*	−0.02 (0.12)
% Irrigated area producing cotton	0.28 (0.13)**	0.15 (0.12)
% Irrigated area producing fodder	−0.00 (0.22)	−0.16 (0.02)
No. of members in WUA	0.00 (0.00)	−0.00 (0.00)
No. of disputes	−0.00 (0.00)	−0.00 (0.00)
No. of board members	0.01 (0.00)**	0.00 (0.00)
% of female board members	0.23 (0.20)	0.28 (0.189)
Quality of main canal (1–5)	−0.04 (0.03)	−0.02 (0.03)
Quality of secondary canal (1–5)	−0.03 (0.03)	0.00 (0.03)
Quality of tertiary canal (1–5)	0.01 (0.04)	0.04 (0.04)
Constant	0.78 (0.06)***	0.77 (0.05)***
N	131	131
R^2	0.14	0.11
F-statistic	$F(12,118) = 1.59$	$F(12,118) = 1.29$
Prob $> F$	0.10*	0.29

Notes: The reported values express the modified DID estimate as a percentage difference, followed in parentheses by the standard error.
***Difference is significant at 1%.
**Difference is significant at 5%.
*Difference is significant at 10%.

additional member on the board between 2014 and 2016 increased the probability of having a seasonal water delivery schedule by 1% in 2016 ($p < 0.05$).

6.2. *Routine cleaning and maintenance of irrigation infrastructure*

WUAs with longer training were 6% more likely to conduct pre-irrigation seasonal maintenance of secondary canals in 2016 and 3% more likely to conduct the same for tertiary canals, but these results are not significantly different from zero at the usual levels (Table 4). A 1% increase in the area under fodder cultivation between the two years reduced the probability of conducting pre-irrigation seasonal maintenance of tertiary canals by 21% in 2016 ($p < 0.1$), consistent with the observation that fodder cultivation in Tajikistan is mostly rainfed.

6.3. *Irrigation fee collection and transfers*

WUAs with longer training had 9% more members paying their irrigation fees in 2016 than in 2014 when compared with WUAs with shorter training, although

Table 4. Effects of Longer Training on Routine Maintenance of Irrigation Infrastructure

	WUA conducted pre-season maintenance of secondary canals	WUA conducted pre-season maintenance of tertiary canals
Longer training	0.06 (0.04)	0.03 (0.04)
% WUA service area irrigated	−0.11 (0.15)	0.07 (0.12)
% Irrigated area producing wheat	0.05 (0.09)	0.05 (0.08)
% Irrigated area producing cotton	0.06 (0.09)	0.11 (0.08)
% Irrigated area producing fodder	−0.20 (0.15)	−0.21 (0.12)*
No. of WUA members	−0.00 (0.00)	0.00 (0.00)
No. of disputes	0.00 (0.00)	0.00 (0.00)
No. of board members	0.00 (0.00)	0.00 (0.00)
% of female board members	0.01 (0.14)	0.04 (0.12)
Quality of main canal (1–5)	−0.01 (0.02)	−0.01 (0.02)
Quality of secondary canal (1–5)	−0.00 (0.02)	0.00 (0.02)
Quality of tertiary canal (1–5)	0.03 (0.03)	−0.02 (0.03)
Constant	0.05 (0.03)	0.03 (0.03)
N	131	131
R^2	0.06	0.09
F-statistic	$F(12,118)=0.60$	$F(12,118)=1.01$
Prob $> F$	0.84	0.44

Notes: The reported values express the modified DID estimate as a percentage difference, followed in parentheses by the standard error.
*Difference is significant at 10%.

these results are not significant at the standard levels (Table 5). However, longer training has no effect on the change in the share of irrigation fees transferred to the government. WUAs with longer training also experienced an increase in indebtedness between 2014 and 2016 of TJS 27,691 (∼$ 3,146),[19] effect of which was not statistically significant from zero at the usual levels. A 1% increase in the cultivated sarea of wheat was associated with a reduction of debt of the magnitude of TJS 25,952 (∼$ 3,377), reflecting the fact that wheat cultivation requires considerably less water. A 1% increase in area under fodder cultivation resulted in 53% increase in the share of members paying their irrigation fees ($p < 0.05$), consistent with the fact that since fodder cultivation is mostly rainfed, low (zero) water use results in low (zero) irrigation charges, which more farmers could afford to pay.

6.4. *WUA membership fees collection*

WUAs with longer training were able to increase the share of members from whom they recovered membership fees by 19% between 2014 and 2016 ($p < 0.01$;

[19]All numbers adjusted for inflation and converted at the exchange rate of 1 Tajikistan Somoni (TJS) ≈ 0.13 US$ in 2016–2017.

Table 5. Effects of Longer Training on Irrigation Service Fee Recovery and Transfer

	Share of members from whom irrigation service fee is recovered	Share of collected irrigation service fees transferred
Longer training	0.09 (0.07)	0.00 (0.11)
% WUA service area irrigated	0.21 (0.22)	−0.08 (0.33)
% Irrigated area producing wheat	−0.18 (0.17)	0.33 (0.26)
% Irrigated area producing cotton	0.08 (0.13)	−0.02 (0.18)
% Irrigated area producing fodder	0.53 (0.22)**	−0.24 (0.34)
No. of WUA members	−0.00 (0.00)	0.00 (0.00)
No. of disputes	−0.00 (0.00)	−0.00 (0.00)
No. of board members	−0.00 (0.00)	−0.01 (0.01)
% of female board members	−0.07 (0.19)	0.24 (0.31)
Quality of main canal (1–5)	0.01 (0.03)	0.01 (0.05)
Quality of secondary canal (1–5)	−0.02 (0.04)	0.02 (0.06)
Quality of tertiary canal (1–5)	0.04 (0.04)	−0.04 (0.07)
Constant	0.08 (0.06)	0.94 (0.09)***
N	101	102
R^2	0.13	0.05
F-statistic	$F(12,88)=1.09$	$F(12,89)=0.38$
Prob $> F$	0.38	0.97

Notes: The reported values express the modified DID estimate as a percentage difference, followed in parentheses by the standard error.
***Difference is significant at 1%.
**Difference is significant at 5%.

Table 6). In addition, a 1% increase in irrigated area between 2014 and 2016 was associated with a 37% increase in the share of members from whom the WUA was able to recover fees during the same period ($p < 0.1$). Adding an additional WUA member reduced the share of members from whom the WUA would be able to recover fees ($p < 0.01$), reflecting that having more members makes monitoring more difficult; however, the magnitude of this coefficient is very small (0.01%).

6.5. *Dispute arbitration*

WUAs with longer training were 6% less likely to be the arbitrator of disputes between 2014 and 2016 (Table 7); however, this result is not significant ($p = 0.46$). A 1% increase in the area under cotton cultivation increased the probability of the WUA arbitrating disputes from 2014 to 2016 by 28% ($p < 0.1$), consistent with the higher demands on irrigation associated with cotton cultivation.

Table 6. Effects of Longer Training on WUA Membership Fee Recovery

	Share of *dehkan* farms from whom WUA membership fee is recovered
Longer training	0.19 (0.06)***
% WUA service area irrigated	0.37 (0.21)*
% Irrigated area producing wheat	−0.14 (0.13)
% Irrigated area producing cotton	0.07 (0.12)
% Irrigated area producing fodder	0.21 (0.21)
No. of WUA members	−0.00 (0.00)***
No. of disputes	0.00 (0.00)
No. of board members	0.00 (0.00)
% of female board members	0.12 (0.18)
Quality of main canal (1–5)	−0.01 (0.03)
Quality of secondary canal (1–5)	−0.06 (0.03)
Quality of tertiary canal (1–5)	0.02 (0.04)
Constant	−0.27 (0.05)***
N	115
R^2	0.26
F-statistic	$F(12,102)=2.77$
Prob $> F$	0.00***

Notes: The reported values express the modified DID estimate as a percentage difference, followed in parentheses by the standard error.
***Difference is significant at 1%.
*Difference is significant at 10%.

Table 7. Effects of Longer Training on the Likelihood of a Dispute Being Arbitrated by the WUA

	WUAs settled disputes in 2014 and 2016
Longer training	−0.06 (0.08)
% WUA service area irrigated	0.24 (0.27)
% Irrigated area producing wheat	−0.07 (0.17)
% Irrigated area producing cotton	0.28 (0.16)*
% Irrigated area producing fodder	−0.27 (0.27)
No. of WUA members	0.00 (0.00)
No. of disputes	0.00 (0.00)
No. of board members	−0.00 (0.01)
% of female board members	0.06 (0.25)
Quality of main canal (1–5)	0.00 (0.04)
Quality of secondary canal (1–5)	0.01 (0.05)
Quality of tertiary canal (1–5)	−0.00 (0.05)
Constant	0.35 (0.07)***

Table 7. (*Continued*)

	WUAs settled disputes in 2014 and 2016
N	131
R^2	0.05
F-statistic	$F(12,118)=0.48$
Prob $> F$	0.92

Notes: The reported values express the modified DID estimate as a percentage difference, followed in parentheses by the standard error.
***Difference is significant at 1%.
*Difference is significant at 10%.

7. Discussion

Experimental methods such as randomized controlled trials in the field and laboratory experiments have the ability to control for selection effects and to provide unbiased estimates of effects of interventions, but have limitations related to the generalizability of results under different circumstances. When an appropriate control group is difficult to identify due to the scale of the intervention, synthetic controls can be used to construct a counterfactual, but they require the treatment unit and the potential control units to exist and be functioning for a considerable duration before the intervention commenced.

Assessment of the impact of large-scale development programs, such as those that create new institutions for community-based resource management, is challenging for two reasons. First, the intervention often creates the units of interventions themselves. This means the units of intervention did not exist — and consequently their covariates or outcomes were not measurable — before the intervention, thus ruling out the application of methods such as synthetic controls. Secondly, assignment to treatment is not random, and because all covariates and outcomes took values of zero at the baseline, a standard DID approach may not eliminate time-invariant unobservables.

In this paper, the DID technique is employed in a context where measurements of outcomes and covariates were made in two time periods after the intervention was implemented, because, at baseline, both took values of zero. This modified DID technique with right-hand size variables eliminates bias due to time-invariant unobservables and controls the effect of time-varying unobservables, but introduces a measurement error. Despite this bias, the estimates provide empirical evidence on how the intervention influences the performance of new institutions in their early days, which could be highly valued for adaptive management. An alternative available estimate — the measured difference between outcomes in the

later period — is likely to be less informative because it cannot control for selection effects that influence outcomes in each group.

A note of caution is warranted when using the modified DID technique when the treatment and control units do not exist before the intervention. In this paper, the outcomes being examined — service delivery and financial functions — are typically slow to change, especially when institutions are new. In cases in which the outcomes respond more rapidly to an intervention, the measurement bias may outweigh the reduction in bias achieved through the elimination of time-invariant unobservables. Practitioners using the modified DID technique would be advised to consider the nature of the outcomes under consideration.

In this paper, the results demonstrate that longer training enabled WUAs in Southern Tajikistan to perform their mandated functions better between 2014 and 2016. This was particularly the case for recovering membership fees from *dehkan* farm members, where WUAs with longer training increased their membership fee recovery by 19% in the two-year period. The existing literature on CPRs and institutional reform suggests that the length of the training process may be an important factor in determining institutional performance. The results in this paper add support to that claim, providing some evidence that additional training may allow WUAs to perform functions better. Follow-up measurement in subsequent years may present a more complete picture of the differences in this effect.

In countries such as Tajikistan where political and economic transitions are giving rise to new institutions, understanding performance in the short term can provide timely evidence for adaptive management. In this case, the results suggest that the government should focus on longer training if it expands its WUA program. Equally usefully, data-based comparisons such as those reported here can indicate conditions in which there are no radical differences in response to the program intervention that might necessitate immediate action. Where some analysts may be reluctant to conduct empirical analysis in the absence of well-established treatment and control groups, such as with the formation of WUAs, quantitative methods of assessment can be applied in less-than-ideal conditions to generate evidence that can prove useful for policy purposes.

Acknowledgments

This research is funded through a project titled ‘Impact of Water Users Associations on Water and Land Productivity, Equity and Food Security in Tajikistan’ funded by the United States Agency for International Development under USAID

Grant Number AID-BFS-G-11-00002. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Agency for International Development. The authors would like to thank Sahana Balasubramanya and Marie-Charlotte Buisson for suggestions; Sohrob Aslamy and Katherine MacDonald for supervising data collection; and two anonymous referees for extensive comments that helped us improve the paper. Any errors are ours.

References

Abadie, A, A Diamond and J Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105, 493–505.

Abadie, A and J Gardeazabal (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*, 93, 113–132.

Asian Development Bank (ADB) (2012). Tajikistan: Irrigation Rehabilitation Project completion report. Validation report, Independent Evaluation Department. Manila: ADB.

Balasubramanya, S, M-C Buisson, P Saikia, K MacDonald, S Aslamy, T Horbulyk, C Hannah, M Yakubov and A Platonov (2016). Impact of water user associations on water and land productivity, equity and food security in Tajikistan. Baseline report, prepared for the United States Agency for International Development (USAID), Colombo: International Water Management Institute, pp. 1–127. https://agrilinks.org/sites/default/files/resource/files/Feed%20the%20Future%20Tajikistan%20Water%20User%20Associations%20Impact%20Evaluation%20-%20B.pdf (last accessed 15 June 2017).

Barcellini, F, L Prost and C Marianne (2015). Designers' and users' roles in participatory design: What is actually co-designed by participants? *Applied Ergonomics*, 50, 31–40.

Batt, HA and GP Merkley (2010). Water management and user association analysis for irrigation improvement in Egypt. *Irrigation and Drainage*, 59, 150–160.

Çakmak, B, M Beyyribey and S Kodal (2004). Irrigation water pricing in water user associations, Turkey. *Water Resource Development*, 20, 113–124.

Denyer Willis, G and M Mota Prado (2014). Process and pattern in institutional reforms: A case study of the Police Pacifying Units (UPPs) in Brazil. *World Development*, 64, 232–242.

Fischer, ME, B Irlenbusch and A Sadrieh (2004). An intergenerational common pool resource experiment. *Journal of Environmental Science and Management*, 48, 811–836.

Food and Agriculture Organization of the United Nations (FAO) (2012). *AQUASTAT: Tajikistan*. http://www.fao.org/nr/water/aquastat/countries_regions/TJK/ (last accessed 9 January 2017).

Gunchinmaa, T and M Yakubov (2009). Institutions and transition: Does a better institutional environment make water users associations more effective in Central Asia? *Water Policy*, 12(2), 165–185.

Hamidov, A, A Thiel and D Zikos (2015). Institutional design in transformation: A comparative study of local irrigation governance in Uzbekistan. *Environmental Science and Policy*, 53, 175–191.

Hornidge, AK, L Oberkircher and A Kudryavtseva (2013). Boundary management and the discursive sphere — Negotiating realities in Khorezm, Uzbekistan. *Geoforum*, 45, 266–274.

Huang, Q, J Wang, KW Easter and S Rozelle (2010). Empirical assessment of water management institutions in Northern China. *Agricultural Water Management*, 98, 361–369.

Johnson, SH and J Stoutjesdijk (2008). WUA training and support in the Kyrgyz Republic. *Irrigation and Drainage*, 57, 311–321.

Kazbekov, J, I Abdullaev, H Manthrithilake, A Qureshi and K Jumaboev (2009). Evaluating planning and delivery performance of water user associations (WUAs) in Osh Province, Kyrgyzstan. *Agricultural Water Management*, 96, 1259–1267.

Matveeva, A (2009). The perils of emerging statehood: Civil war and state reconstruction in Tajikistan. Working paper, Number 46, Series 2. London: Crisis States Research Centre, pp. 1–59.

Meinzen-Dick, R, KV Raju and A Gulati (2002). What affects organization and collective action for managing resources? Evidence from canal irrigation systems in India. *World Development*, 30(4), 649–666.

Mukhtarov, F, S Fox, N Mukhamedova and K Wegerich (2015). Interactive institutional design and contextual relevance: Water user groups in Turkey, Azerbaijan and Uzbekistan. *Environmental Science and Policy*, 53, 206–214.

Mustafa, D, A Altz-Stamm and L Mapstone Scott (2016). Water user associations and the politics of water in Jordan. *World Development*, 79, 164–176.

Nagrah, A, AM Chaudhry and M Giordano (2016). Collective action in decentralized irrigation systems: Evidence from Pakistan. *World Development*, 84, 282–298.

O'Hara, SL (2000). Lessons from the past: Water management in Central Asia. *Water Policy*, 2, 365–384.

Ostrom, E (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*, pp. 1–294. Cambridge, UK: Cambridge University Press.

Republic of Tajikistan (2006). Water user association law. *Law of the Republic of Tajikistan*, pp. 1–12.

Ricks, JI (2016). Building participatory organizations for common pool resource management: Water user group promotion in Indonesia. *World Development*, 77, 34–47.

Rosenbaum, PR and DJ Rubin (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39, 33–38.

Rubin, DJ (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2, 169–188.

Sehring, J (2007). Irrigation reform in Kyrgyzstan and Tajikistan. *Irrigation and Drainage Systems*, 21, 277–290.

Shahriari, H (2009). Activities, water and gender in Tajikistan's rural sector: A social assessment of Konibodom, Bobojon Ghafurov, and Yovon. Social Development

Department, Europe and Central Asia series. Washington, DC: World Bank. http://documents.worldbank.org/curated/en/364511468117242593/pdf/514450WP0TJ0So10Box342028B01PUBLIC1.pdf (last accessed 15 June 2017).

Skaperdas, S (2001). The political economy of organized crime: Providing protection when the state does not. *Economics of Governance*, 2, 173–202.

Smith, JL (2008). A critical appreciation of the ‘bottom-up’ approach to sustainable water management: Embracing complexity rather than desirability. *Local Environment*, 13(4), 353–366.

Tellez Foster, E, A Rapoport and A Dinar (2016). Alternative policies to manage electricity subsidies for groundwater extraction: A field study in Mexico. UCR SPP Working Paper Series WP#16-05. Riverside, California: School of Public Policy, University of California-Riverside.

Tellez Foster, E, A Rapoport and A Dinar (2017). Groundwater and electricity consumption under alternative subsidies: Evidence from laboratory experiments. *Journal of Behavioral and Experimental Economics*, 68, 41–52.

Theesfeld, I (2004). Constraints on collective action in a transitional economy: The case of Bulgaria’s irrigation sector. *World Development*, 32(2), 251–271.

Thiel, A, F Mukhtarov and D Zikos (2015). Crafting or designing? Science and politics for purposeful institutional change in social-ecological systems. *Environmental Science and Policy*, 53, 81–86.

United Nations (2015). Tajikistan. http://data.un.org/CountryProfile.aspx?crName=TAJIKISTAN (last accessed 22 June 2017).

United States Agency for International Development (USAID) (2014). *USAID Family Farming Program — Quarterly Report: Year Four — First Quarter*. Unpublished document, pp. 1–46. Washington, DC: USAID.

Wegerich, K (2000). Water user associations in Uzbekistan and Kyrgyzstan: Study on conditions for sustainable development. Occasional paper — Water Issues Study Group, London: School of Oriental and African Studies, University of London, Number 32, pp. 1–23.

Woolridge, J (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

World Bank (2015). Tajikistan. http://data.worldbank.org/country/tajikistan (last accessed 22 June 2017).

Yakubov, M (2011). Assessing irrigation performance from the farmers’ perspective: A qualitative study. *Irrigation and Drainage*, 61(3), 316–329.

Yap-Salinas, LH (1994). Converging factors in the successful transfer of irrigation management responsibilities to water users’ associations in the Dominican Republic. Paper presented at the Int. Conf. Irrigation Management Transfer, 20–24 September, Wuhan, China, pp. 117–127.

Zhou, Q (2013). Decentralized irrigation in China: An institutional analysis. *Policy and Society*, 32, 77–88.