WPS6567

Policy Research Working Paper 6567

Is Small Better?

A Comparison of the Effect of Large and Small Dams on Cropland Productivity in South Africa

Elodie Blanc
Eric Strobl

The World Bank
Development Economics Vice Presidency
Partnerships, Capacity Building Unit
August 2013

Policy Research Working Paper 6567

Abstract

This study estimates and compares the effects of small and large irrigation dams on cropland productivity in South Africa. To this end, a panel data set of South African river basins is constructed. The econometric analysis reveals that although large dams increase cropland productivity downstream, they have a negative effect on cropland within the vicinity. However, their existence can enhance the relatively small positive impact of local small dams. Although a cost-benefit analysis of irrigation benefits shows that small dams may be more viable than large ones, large dams can play a potentially important role within a system of both types of dams.

This paper is a product of the Partnerships, Capacity Building Unit, Development Economics Vice Presidency. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at is eric.strobl@polytechnique.edu.

Is Small Better? A Comparison of the Effect of Large and Small dams on cropland productivity in South Africa

Elodie Blanc and Eric Strobl[1]

JEL Classification codes: O20, Q19

Sector Boards: WAT, ENV, EPOL, ARD

Keywords: large dams, small dams, agricultural productivity, South Africa

[1] Elodie Blanc is a research scientist at the Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology. Eric Strobl (corresponding author) is a professor at Ecole Polytechnique and IPAG Research Lab; his email address is eric.strobl@polytechnique.edu. Eric Strobl is grateful for financing from La Chaire Développement Durable of the Ecole Polytechnique.

The use of dams to facilitate irrigation through water storage in times of scarcity in developing countries has had a long and varied history. Traditionally, much of the funding by international donors was for large dam construction (Hathaway and Pottinger 2008). However, in the 1990s, considerable concern emerged over the fact that the distributional environmental and economic impacts of large dams had been largely ignored. This situation eventually culminated in the creation of the World Commission on Dams to assess these infrastructures globally. The conclusions of the World Commission on Dams were rather damning, stating that "a lack of equity of distribution of benefits has called into question the value of many dams in meeting water needs…when compared to alternatives" (WCD 2000, p. xxviii). Importantly, this report and the consequent media attention not only resulted in the considerable reduction of donors' funding for large-scale infrastructure projects but also sparked the birth of a substantial (and ongoing) contingency of lobbyists arguing that small dams, which cause little environmental degradation and population displacement, are an obviously superior solution to the water scarcity problem of agriculture in developing countries.[1]

A closer examination of the large versus small dam controversy, however, reveals that the underlying issues are much more complex than often portrayed. Although it is undeniable that small dams result in relatively negligible negative environmental and economic spillovers, it must also be recognized that large dams provide substantially greater storage capacity, operate at a lower per-storage cost, and lose less water owing to evapotranspiration than their small counterparts (Keller, Sakthivadivel, and Seckler 2000). Moreover, in addition to irrigation, large dams can be multipurpose, providing flood control as well as generating substantial amounts of hydropower and, hence, offsetting some or all of the irrigation investment costs. Importantly, the relative weight of the advantages and disadvantages of large versus small dams depends upon quantitative assessments of the various factors involved. In a study of potential large- and small-scale irrigation investment in sub-Saharan Africa, Young (2008) find that the relative cost-benefit ratio crucially depends on assumptions regarding per unit capacity investment, and these assumptions vary widely. Moreover, it appears that the question may not be a matter of the choice of large versus small but rather the choice of a "continuum of options of large and small dams…(as) part of a more efficient irrigation system" (Foster and Briceño-Garmendia 2010). In such a system, small dams can serve small, immediate water needs, whereas large infrastructures ensure long-term availability and can support nearby smaller dams (ICID 2000).

Despite the continuing debate over whether small dams are preferable to large ones, a perusal of the academic literature unearths essentially no statistically sound quantitative estimates of the relative benefits or costs. For example, in terms of agricultural production, Duflo and Pande (2007) examine the case of India and find that dams benefit agricultural production only in downstream, not in nearby, communities. This result is echoed for the case of the African continent by Strobl and Strobl (2010). These studies, however, only examine the case of large irrigation dams. The only exception is the study by Ersado (2005), which looks at small irrigation dams for Ethiopia and finds that households located closer to a small dam enjoy greater agricultural yields. However, the author does not provide comparative results for being closer to a large dam, and the econometric estimation does not control for the likely endogenous location of dams, which, as Duflo and Pande (2007) and Strobl and Strobl (2010) demonstrate, can substantially bias the derived benefits, at least for large dams.

In this article, we address the paucity of research on the small versus large dam controversy by estimating and comparing the cropland productivity effects of small and large irrigation dams in South Africa. We believe that our article provides the first comparative quantitative study of the relative impact of large and small dams. In this regard, South Africa is arguably a particularly suitable case study. South Africa has a fairly varied climatology over time and space, allowing for considerable interbasin water transfers via irrigation dams. Additionally, after decades of investment, the country stands out for its operation of the largest number of dams on the African continent, although it is still argued to be operating below its potential, with only approximately 10 percent of the cultivated area equipped for irrigation (Foster and Briceño-Garmendia 2010).[2]

To complete the task at hand in this article, we assemble a 20-year annual panel data set of large and small irrigation dams and cropland productivity for river basins in South Africa. Additionally, we employ the instrumental variable (IV) approach suggested by Duflo and Pande (2007), which allows us to address the inherently endogenous nature of dam location. Our analysis reveals a number of interesting results. In line with the literature cited above, our estimates suggest that large irrigation dams improve the productivity of downstream cropland. However, the impact of local large dams crucially depends on controlling for small dams and their interaction with these dams. More specifically, large dams reduce cropland productivity within the vicinity on their own, and they can further augment the relatively smaller positive impact of local small dams.

The remainder of the article is organized as follows. In the next section, we outline the empirical framework and econometric specification employed in the analysis. Section III provides a description of our data and some summary statistics. Our econometric results are contained in section IV. We use our estimates to infer the economic significance of large and small dams in section V. Concluding remarks are provided in the final section.

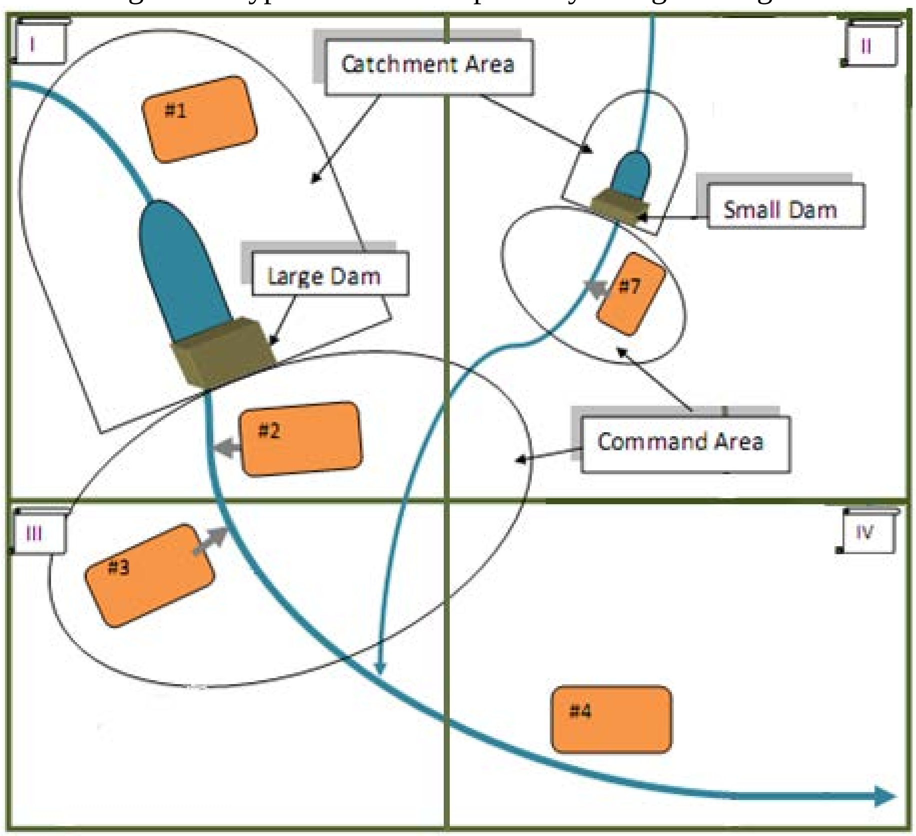
I. ESTIMATING THE DISTRIBUTIONAL IMPACT OF LARGE AND SMALL DAMS

The distributional impact of large and small dams is estimated econometrically. The empirical framework and econometric specification employed in the analysis are described below.

Empirical Framework

An intrinsic part of estimating how dams may affect the productivity of an agricultural area is the geographical location of the area of interest relative to the dam(s). More specifically, consider, as in figure 1, a hypothetical region consisting of one large and one small dam and seven different agricultural plots. For each plot, the nature of dam engineering and hydrological features imply a priori expectations regarding the impact of these dams. Plot #1 is assumed to be located entirely within the large dam's catchment area and is expected to lose productivity as a result of water seepage from the reservoir, increased water logging, and soil salinity as well as possible water restrictions to ensure maximum water storage. In the command area of the large dam (that is, Plots #2 and #3), agricultural land is likely to benefit from irrigation because the fixed cost of accessing irrigation within the canal network of the command area is generally lower than for other forms of water harvesting. Although one may assume that downstream Plots #4 and #5 are too far away to benefit from the irrigation canal network, they may still benefit from the dam if it is used to prevent floods and droughts by regulating the flow of water downstream (Strobl and Strobl 2010).[3] In contrast, the catchment area of the small dam is presumed to be small enough to not contain any agricultural plots, although we also assume that a small dam can provide some irrigation benefits to nearby cropland (that is, Plot #7). Finally, although one might expect no direct effect on the non-downstream Plot #6 from either the large or the small dam, there may be general equilibrium effects because farmers may migrate closer to a dam to take advantage of the new irrigation system or lower their production in the face of higher input and lower output prices.

Figure 1. Hypothetical Example – Hydrological Regions



Our hypothetical example suggests that, to completely disentangle the distributional effect of large and small dams on cropland productivity within this framework, it would be ideal to know a number of geographical attributes of the dams, particularly their location, the extent of their catchment and command areas, and the location of cropland. Unfortunately, detailed, comprehensive spatial data on all dams' catchment and command areas do not exist for South Africa. Nevertheless, as Strobl and Strobl (2010) note, one can still investigate distributional effects by identifying river basins and determining their upstream/downstream relationships and locating agricultural plots and dams within these basins. In this regard, it is helpful to consider the four basins in figure 1, noting that Basin I lies immediately upstream from Basin III (and, hence, Basin III is immediately downstream from Basin I), and Basin IV, although it does not neighbor Basin I, lies (further) downstream from it. In contrast, although Basin II neighbors Basin I, it lies neither downstream nor upstream from it. We assume that the catchment area always lies completely within the basin in which the dam (small or large) is located, and a large dam's own as well as its downstream regions may contain some of its command area.

In considering the effect of a dam within this upstream/downstream breakdown of our hypothetical region, the net a priori effect in a large dam's own basin would be ambiguous because we assume that it contains all of the catchment as well as part of the command area. In contrast, for a small dam, as long as its catchment area is small enough, it is unlikely to have much effect on the productivity of nearby cropland unless it can directly benefit from the irrigation network. We also assume that there is no effect further downstream from the small dam. In downstream Basin III, one can expect a positive effect of the upstream large dam as long as it benefits from an irrigation network. As argued earlier, further downstream plots in Basin IV may indirectly benefit through flood- and drought-induced operational rules. Finally, Basin II, although neighboring Basin I, is neither upstream nor downstream from it, but it could be subject to general equilibrium effects, as mentioned above.

<<**B**>>*Econometric Specification*

Our main goal is to determine the effects of small and large dams on cropland within the same basin or immediately downstream from it. Our base specification focuses on the direct effects of dams:

$$CP_{it} = \alpha + \beta_1 D(L)_{it} + \beta_2 D(S)_{it} + \beta_3 UD(L)_{it} + \beta_4 X_{it} + \varepsilon_{it} \,, \qquad (1)$$

where *CP* is basin *i*'s cropland productivity, *D(L)* refers to the number of dams within the basin *i*, *D(S)* is the number of small dams in *i*, *UD(L)* is the number of large dams located upstream from *i*, *X* is a vector of other explanatory variables, *t* is a year indicator, and ε is the unexplained error term. As noted above, a priori, the expected sign of the coefficient β_1 on large dams in an agricultural area's own region is ambiguous, the expected sign of β_2 on small dams is likely to be positive, and β_3, representing the effect of upstream large dams, is also anticipated to be positive.

It should be noted that the estimated coefficients on *D(L)*, *D(S)*, and *UD(L)* would only be unbiased if dams were randomly allocated across South Africa. Instead, one could reasonably expect that dams have been strategically placed according to cost-benefit considerations and influenced by political factors. Thus, unless the set of controls *X* captures all factors that determine dam allocation and cropland productivity, a simple ordinary least squares regression of (1) is likely to produce biased and inconsistent estimates of β_1, β_2, and β_3. Moreover, these biases may be upward or downward depending on the political or economic motivation behind dam placement.

<<**B**>>*Instrumental Variable Strategy*

In their seminal article on the impact of large dams in India, Duflo and Pande (2007) suggest an innovative IV strategy to address the endogeneity problems intrinsic in (1). More specifically, they calculate the share of dams located in a state (which is at a more aggregate level than their unit of analysis of districts) prior to their sample period and multiply this share by the total number of dams in India. They then regress the district level of dams on this state-level variable interacted with district-level river gradient indicators to predict arguably exogenous district-level dam construction proxies, which they use as instruments. They use the river gradient proxies on the grounds that certain river gradients are more appropriate for dam construction than others. In contrast, the use of the state share of dams implicitly assumes a state-level variation in the policy of dam allocation and assumes that this can be proxied by the exante state share interacted with time dummies.

Here, we adopt the Duflo and Pande (2007) IV strategy of jointly using river gradients and dam allocation policy changes to isolate exogenous variation in dam location, which we alter to fit the South African context in a number of ways. First, as will be discussed in greater detail below, our unit of analysis is based on a hydrological delineation of river basins that allows us to determine their upstream-downstream relationship rather than an administrative breakdown. Second, we use the fall of the apartheid regime and subsequent water policy changes to construct the arguably exogenous policy change variable used in the IV strategy. More specifically, as

Lustenberger (2010) notes, prior to the change of regime, water transfers were politically motivated by apartheid principles in which water was provided for the vast regions held by white farmers rather being used than to develop activities in the homelands, the nominally independent regions set aside for black South Africans. The groundwork that enabled such discriminatory practices was laid as early as the beginning of the 20th century with the introduction of the Water Act of 1912, which determined that customary rights were dedicated to riparian rights. Water and its uses were assumed to belong to the owners of the land crossed by the rivers, mainly the white minority. Thus, water use was intrinsically linked to land ownership, with no constraint on use. As Conca (2005) notes, the link to the control of irrigation necessarily allowed better control of rural areas. More precisely, given that a sequence of legislative pieces before and during apartheid essentially placed 87 percent of the land into the hands of the white minority, water rights and uses were mainly in the control of this group. This legal entity established the stepping stone for the apartheid regime, which took power in 1946, to introduce a policy of "hydrohegemony," under which water transfer and irrigation systems were constructed not only on the basis of geographical and economic considerations but also in terms of political motivations that mainly benefited the white population. This aspect was further solidified in the 1956 Water Act, which explicitly distinguished between "public" and "private" water and prioritized irrigation water use above all other uses.

When the African National Congress came into power at the end of apartheid in 1994, it intended to radically break with the former policy of water transfers. More specifically, it sought to redress all discriminatory aspects, particularly the waste of agricultural water distributed at low prices to white farmers and the relative lack of water infrastructure available to disadvantaged black populations. This approach culminated in the promulgation of the 1998 Water Act, which has two important aspects with regard to the current context. First, it provided for the nationalization of water resources, thereby superseding the principle of riparian rights as established under the Water Act of 1921. Second, the emphasis was switched from water supply management to water demand management, which involved the reallocation of water rights, particularly with regard to agriculture. As Stein (2000) notes, the 1998 Water Act "obliges repositories of decision making power to take account and give effect to the fundamental principles and objectives" (p. 290). Thus, arguably, the change in water policy since 1994 put an end to aspects of water infrastructure construction that discriminated against nonwhite South Africans, a large part of which resided in the former homelands.

To capture this change in policy, we construct a step variable for river basins located in former homelands that takes the value of zero until 1994, one until 1998, two thereafter, and zero for all basins in nonhomeland areas.[4] This policy proxy is intended to capture the likelihood that former homelands were less disfavored for dam construction after the fall of apartheid, particularly after the implementation of the Water Act 1998.

Third, unlike Duflo and Pande (2007), we want to separately identify the effects of large and small dams. Thus, we need some difference in the instruments used for these. In this regard, one should note that small dams in South Africa are typically earth-fill dams, which are constructed as a simple embankment of well-compacted earth. In fact, in our database, more than 92.1 percent of dams classified as small are earth-fill dams. Importantly, small earth-fill embankment dams are often built on ephemeral (seasonal) rivers (see RAIN (2007) and Lasage et al. (2008) In contrast, such rivers are generally unsuitable for large dams, which, in South Africa, are almost exclusively buttress, arch, or gravity dams.[5] We thus use the interaction term of our policy variable with perennial (year-round) river gradients as instruments for both large and small dams, but we also use ephemeral length within a basin (interacted with the policy variable) to separately identify the potentially different effect of small dams on agricultural productivity.[6]

Our specification for exogenously predicting dam location relies on the use of differences in dam construction because of the change in water infrastructure policy after the end of apartheid as well as differences in the geographic suitability of river basins to predict the number of dams:

$$D(sz)_{it} = \partial_1 + \sum_{k=2}^{4} \partial_2 \left(RGr(k)_i * P_{it} \right) + \partial_3 \left(M_i * P_{it} \right) + \sum_{k=2}^{4} \partial_4 \left(RGr(k)_i * l_t \right) \\ + \partial_5 ERLENGTH_i (sz = small) * P_{it} + \omega_{it} + v_i + \mu_{it} \qquad , (2)$$

where $D^s{}_{it}$ is the number of dams in river basin i at time t and of size $sz = \{SMALL, LARGE\}$, v_i are river basin fixed effects, ω_{it} is a vector former homeland and former nonhomeland year interaction terms, and μ_{it} is the error term. P_{it} is the policy proxy described above. The river gradient variable, $RGr(k)_i$'s, $k = \{2, \ldots, 4\}$, represents the fraction of perennial rivers in basin i with a 1.5–3 percent (moderate), 3–6 percent (steep), and over 6 percent (very steep) river gradient. $ERLEGNTH_i$ is the length of ephemeral rivers within a basin and is not used when $sz=LARGE$. Our source for all geographical information on rivers is the Food and Agriculture Organization's Rivers of Africa geospatial database (FAO 2000). To calculate all other geographical features, we use the GTOPO20 (USGS 2011a) data set. The interaction of year dummies, l_t, with the river gradient variables is intended to capture country-wide, time-varying effects of river gradients that may affect dam construction, whereas M_i is a vector consisting of river-basin-specific, time-invariant measures of elevation, gradient, river length, and river basin area.[7]

To derive the effect of own basin and upstream dams on cropland productivity, we estimate the following specification:

$$\begin{aligned} CP_{it} &= \alpha + \beta_1 \hat{D}(LARGE)_{it} + \beta_2 \hat{UD}(LARGE)_{it} + \beta_3 \hat{D}(SMALL)_{it} \\ &\quad + Z_{it}\beta_4 + Z_{it}^{U}\beta_5 + X_{it}\beta_6 + \nu_i + \mu_{mt} + \omega_{it} \end{aligned} \quad , \qquad (3)$$

where Z_{it} represents the right-hand-side explanatory variables in (2), except for the $RGr(k)_i * P_{it}$ interaction terms. Z_{it}^{U} are the same control variables as Z_{it}, but they are calculated in terms of a river basin's upstream basins. X_{it} is a vector of other available control variables. To estimate (3), we follow Duflo and Pande (2007) and use the estimated number of large dams from (2), $\hat{D}(LARGE)_{it}$, its upstream equivalent, $\hat{UD}(LARGE)_{it}$, and the estimated number of small dams, $\hat{D}(SMALL)_{it}$, as instruments. We thus have three instruments for three endogenous variables.

One should note that, as Duflo and Pande (2007) argue, the approach above, in contrast to standard IV methods, uses all available variation efficiently by using all basins to predict the relationship between basin geographic features and the number of dams (rather than only those upstream) and avoids averaging these features when there are several upstream basins. In fact, as Duflo and Pande (2007) note, if every basin had a single upstream basin, then this approach would be identical to a two-stage least-squares procedure where the interaction of river gradient proxies with predicted dam incidence in homeland/nonhomeland areas are instruments.

It is important to outline the identifying assumption behind estimating (3) using the specified instruments. In particular, unbiased estimates of β_1, β_2, and β_3 implicitly rest on the assumption that agricultural production across former homeland river basins with different river gradients would not have changed systematically from 1994 onward even if no dam construction had taken place. In other words, even if there were other policies favoring or disfavoring agricultural production in former homeland basins and these were implemented during the same 1994 to 1998 period, we assume that these did not differ across our river gradient categorization. One possible culprit might be the Land Reform Program, which was introduced with the new African National Congress of the postapartheid regime to redress the unequal patterns of land redistribution. This program was intended to facilitate land restitution and redistribution in favor of the disadvantaged and poor (that is, the underprivileged black population) on a willing buyer/seller basis but with grant support. Feasibly, if river gradients are directly related to agricultural productivity and more land was redistributed in homelands or nonhomelands according to these criteria, then our identifying assumption may be violated. However, as of the end of 2001, only 2 percent of land had actually changed hands, falling short of the 30 percent goal (Twala 2006).

Finally, it should be kept in mind that our IV strategy only captures the local average treatment effect in the sense of the impact in basins where dams were built because of favorable river characteristics and would not have been built otherwise. Thus, the above strategy provides us only with an estimate of the economic impact of technologically feasible dams. It does not have sufficient power to capture the effect of dams on cropland productivity placed because of political reasons, for example.

II. <<**A**>>*DATA SOURCES*

To proceed to the econometric analysis, we collected data regarding dam location and characteristics, the spatial unit of analysis, the cropland productivity and other factors. The description of these data is provided below.

<<**B**>>*Dams*

To identify and geographically locate dams in South Africa, we use the Department of Water Affairs (DWA) dams database (DWA 2011), which collects detailed information on dams that fall within the dam safety permit system, essentially covering all known dams with at least 5 m height and a reservoir capacity of 50,000 m^3.[8] Importantly, the DWA dams database provides information on the exact latitude and longitude of the location, construction date, reservoir surface area and capacity, purpose, size classification, and, for many dams, height.

Our purpose here is to distinguish between the effects of large and small dams in South Africa. A crucial component in this regard is the definition of dams as large versus small, a classification for which there is no global agreement. In fact, definitions differ widely. For example, the International Commission on Large Dams defines large dams as those with at least 15 m in wall height or a reservoir exceeding 3 million m^3. In contrast, in the United States, dams greater than 90 m are classified as large, whereas in China, the definition is based on a reservoir capacity greater than 100 million m^3. In Switzerland, dams with 40 m wall height or those with less than 40 m but a reservoir capacity greater than 1 million m^3 are considered large.

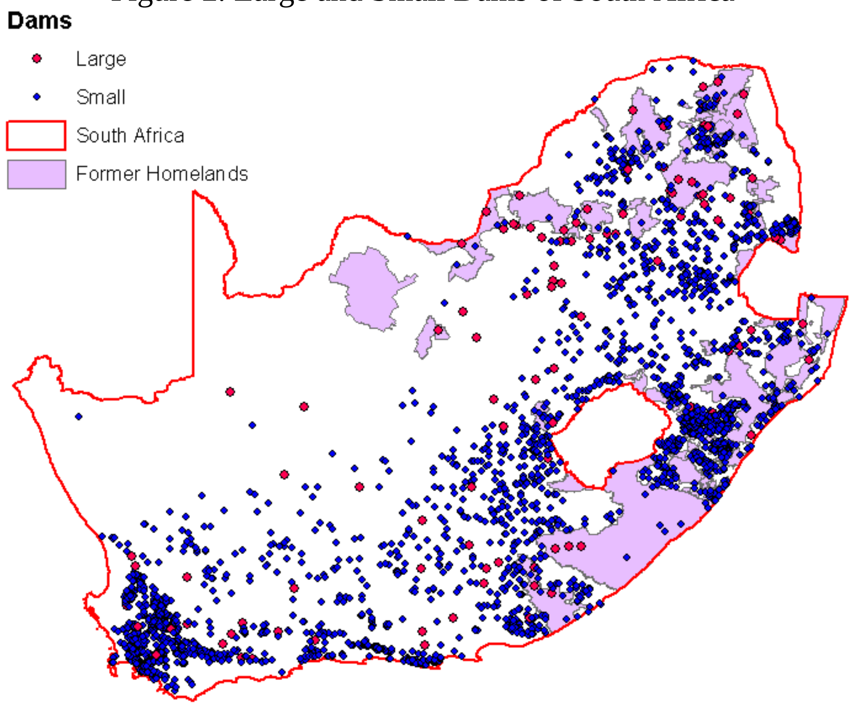
In this paper, we instead adopt a definition of small and large based on a system of dam classification used in South Africa to determine the legislative requirements of dams in terms of design, construction, operation, maintenance, and abandonment. More specifically, one of the main concerns in South Africa has been the potential economic impact of dam failure, and dams have been classified accordingly using their wall height and hazard potential. In this regard, the DWA classifies dam wall height as (a) small, more than 5 m and less than 12 m; (b) medium, more than 12 m but less than 30 m; and (c) large, equal to or more than 30 m. The hazard potential of dams is categorized into (i) low (loss of life, none; potential economic loss, minimal), (ii) significant (loss of life, not more than 10; potential economic loss, significant), and (iii) high (loss of life, more than 10; potential economic loss, great). With these two aspects in mind, dams in South Africa are classified into three possible categories determining the legislative requirements associated with each. Given that small and medium dams (in terms of height) are more likely to be in the same legislative category given their hazard potential and that large and small dams never share the same category regardless of the risk of failure, we group small and medium dams into one group and treat large dams as a separate group. For convenience, we label dams in the former group as "small" and in the latter group as "large" dams. Note that our categorization of dams rests implicitly on the DWA's assessment of potential economic losses in case of failure and thus is based on an evaluation of the local South African context rather than on another relatively arbitrary classification adopted elsewhere. We restrict our analysis to dams whose main purpose is irrigation, constituting 55 percent and 79 percent of all large and small dams in the data, respectively. Thus, we have a total of 114 large and 3,257 small dams (as of the year 2000). Summary statistics provided in table 1 show that large dams are, on average, nearly twice as old as smaller ones. Unsurprisingly, their mean reservoir capacity is substantially larger (over 360 times), constituting 93 percent of the total storage capacity in South Africa. Similarly, the surface area is substantially larger (1,310 Ha for large dams compared to 103 Ha for small dams). Finally, as of the year 2000, most large dams were located in the former nonhomeland territories, whereas most small ones were in the former homelands. We depict the location of large and small irrigation dams in figure 2.

Table 1. Dam Characteristics Means

	Nb	**Age** (in year)	**Capacity** (in '000s of m^3)	**Surface** (in Ha)	**Height** (in meters)	**% Former Homeland**
LARGE	114	42 (25)	128,124 (579,173)	1,310 (4,167)	47 (20)	38.6
SMALL	3,257	23 (19)	353 (727)	103 (3,035)	10 (4)	62.5

Source: DWA (2011)
Note: Statistics for dams existent in the year 2000. Standard deviation is presented in parentheses.

Figure 2. Large and Small Dams of South Africa

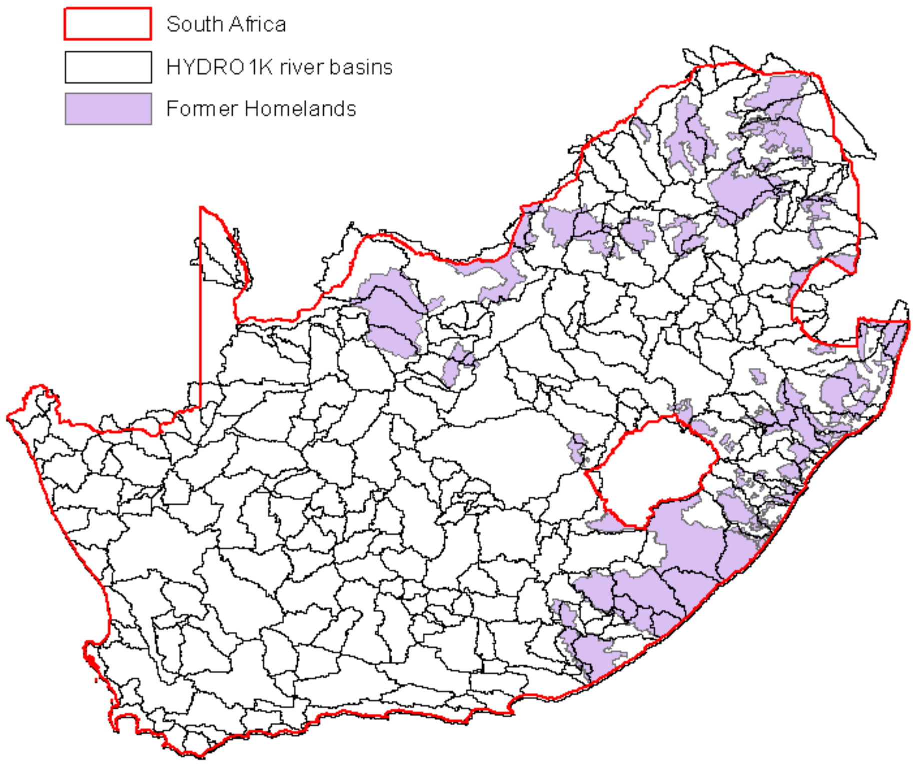


<<**B**>>*Spatial Unit of Analysis*

As outlined earlier, in light of inadequate data on command and catchment areas, we break South Africa down into river basins, allowing us to characterize their hydrological relationship. We follow Strobl and Strobl (2010) and use the HYDRO 1K data set (USGS 2011b), which provides drainage basin boundaries data for the African continent as derived from river network and flow direction data. At its most disaggregated level, this process involves dividing the African continent into 7,131 six-digit hydrological basins with an average area of 4,200 km^2, where each basin can be classified into whether it is upstream, downstream, or not related to another basin in the data set. Of these river basins, 245 lie completely within South African borders, whereas another 98 lie partially within these borders. We consider those in which the majority of the area lies within

South Africa as part of our sample, giving us a total sample of 299 basins. As depicted in figure 3, these basins vary greatly in shape and size, and 31 percent of the basins have a majority of their area located in former homeland territories.

Figure 3. River Basins

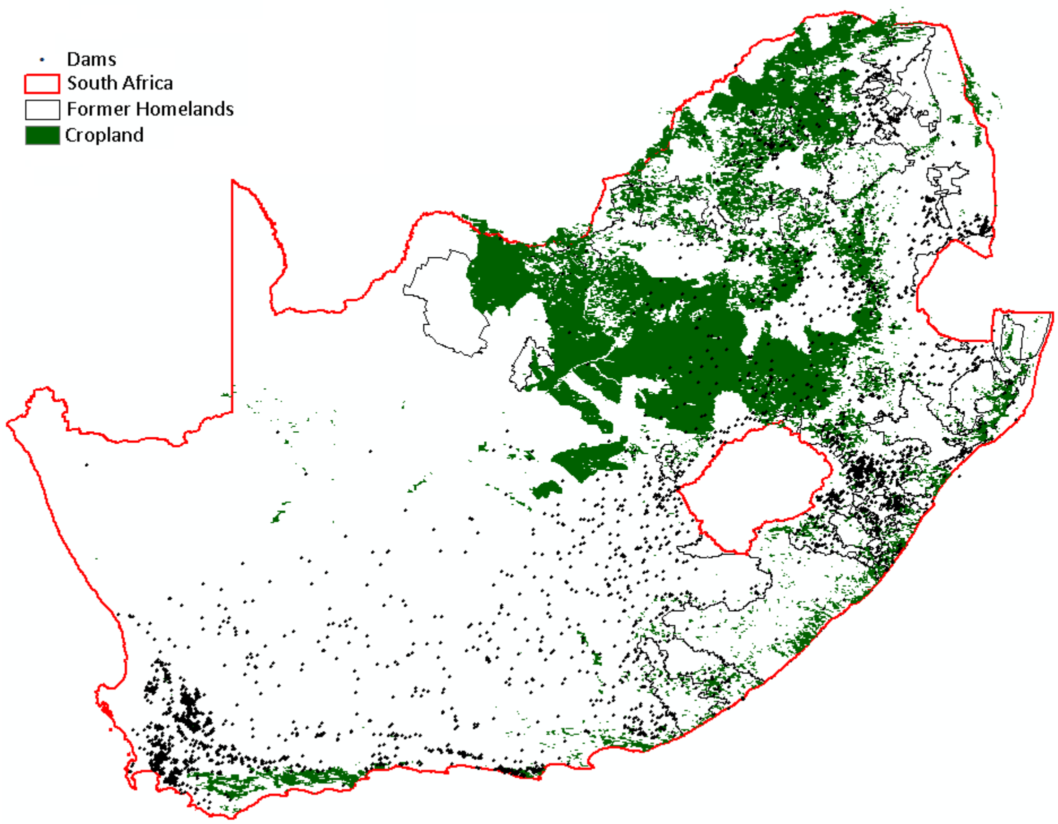


Finally, in terms of a priori expectations regarding the impact of dams within our empirical framework, we assume that the catchment area of dams is likely to be contained within a dam's own vicinity. Although we do not have a geographical delineation for the catchment area of dams, the DWA provides unofficial shapefiles for the area of the reservoir lake for 377 of these. These shapefiles show that 91 percent of dams' catchment areas are completely contained within the river basin where the dam is located, whereas for the remaining dams, 80 percent of the catchment area is contained within the basin. Thus, our assumption about the catchment area being generally contained within a dam's river basin appears reasonable.

<<**B**>>*Cropland Productivity*

Time-varying standard agricultural measures of cropland productivity, such as crop yields, do not exist for South Africa at a spatial resolution as fine as our hydrological regions.[9] We thus follow Blanc and Strobl (2013) and use two satellite data sources. The first is the Global Land Cover 2000 data set (GEM 2011), hereinafter called GLC 2000, which classifies land cover across the globe into 22 distinct land cover categories at a 1 km resolution based on images acquired by the SPOT 4 satellite during 2000. We use the land cover category "Cropland" to spatially capture agricultural crop areas within our basins, although we also experiment with including the categories "Mosaic of Cropland/Shrub or Herbaceous Cover" and "Mosaic of Cropland/Tree Cover/Other Natural Vegetation". The "pure" cropland cells constitute approximately 14 percent (56,000 km^2) of South Africa. According to the GLC 2000, 15 percent of the six-digit code basins did not contain any "pure" cropland cells. Of the 223 basins that did, 77 were located in the former homeland river basins. We depict the identified cropland areas within South Africa in figure 4. The largest areas of "pure" cropland are located in the northern part of South Africa, mainly in nonformer homeland territories.[10] Most dams depicted in the same graph are located near cropland, except in the central part of the country.

Figure 4. Cropland Delineation



To measure the productivity of cropland areas within basins over time, we use as a proxy their net primary production (NPP) as measured by satellite images. NPP quantifies the conversion of atmospheric carbon dioxide into plant biomass, and the resultant values can serve as a proxy of cropland productivity. Examples using NPP to proxy cropland productivity include Heinsch et al. (2005) and (Lobell et al. 2002). As Hicke, Lobell, and Asner (2004) note, one of the advantages of using NPP to proxy cropland productivity over large areas and over time is that, unlike economic data, it provides a common metric among different crop types, thereby facilitating comparisons and aggregation over all types. NPP data for South Africa are the MOD17A2 measures derived from observations of the Moderate-resolution Imaging Spectroradiometer on the

NASA Earth Observing System Terra satellite. The data are available at the 8 km spatial level on an annual basis for the period 1981 to 2000 and are given as grams of carbon per square meter (gC/m^2/year). We use the annual average of NPP to proxy cropland productivity. Although the Moderate-resolution Imaging Spectroradiometer data are at a greater resolution than the GLC 2000 data, this generally did not pose a problem in terms of masking out cropland given that most cropland cells were agglomerated at a much larger scale than 8 km.

The use of the Moderate-resolution Imaging Spectroradiometer data as a proxy for actual cropland productivity has been validated in a number of studies (Zhang et al. 2008; Turner et al. 2006). To roughly verify that this measure can be related to standard agricultural data, we aggregated its value for all of South Africa and correlated it with FAOSTAT (2011) data-derived measures of cropland productivity. The resultant Pearson correlation coefficient was 0.74 and was significant at the 1 percent level, providing evidence that these measures are substantially positively related. We also investigated whether the relationship between NPP and data-derived cropland production might be nonlinear by using a Kernel-weighted local polynomial smoothing estimator, but we found no evidence of this. To provide a “monetary” feel for the rough equivalent of a unit of NPP, we also regressed the total value of agricultural cropland production in South Africa on NPP, additionally controlling for time dummies.

<<**B**>>*Other Explanatory Variables (X)*

Hicke, Lobell, and Asner (2004) note that NPP can change in response to shifts in types of crops, changes in crop management practices (for example, fertilization, irrigation, pest management), and climate-related factors. In this regard, estimating net changes in NPP due to changes in crop management as a result of dam construction is the goal of our exercise, although the results may need to be interpreted in a broader sense if irrigation benefits result in more productive pest management and fertilization. In contrast, failing to take account of crop type shifts may result in overestimating the net benefits of dams in using NPP as a proxy of cropland productivity if, as a result of dam construction, farmers were to switch to crops that benefit more from irrigation. Unfortunately, we know of no time-varying data on crop types at the basin level that would allow us to control for such changes.

In terms of controlling for climatic factors, we use weather data extracted from the CRU TS 2.1 data set (Mitchell and Jones 2005), which provide monthly precipitation and temperature measures at the 0.5 degree level over the entire 20th century. We use these to derive two specific annual river-basin-wide measures of climate, namely, precipitation and reference evapotranspiration. To capture the effect of precipitation, we calculate the local standardized precipitation index, which has been argued to be particularly good at capturing the cumulative effect of reduced rainfall over time in a chosen locality.[11] To capture the effect of temperature on cropland productivity, we calculate reference evapotranspiration, which measures the evaporative demand of the air within a basin, by following Hargreaves and Samani (1985).

Our precipitation and evapotranspiration indices depend on local climate only. However, the hydrological cycle in a region involves the continuous movement of water on, above, and below the surface of the Earth over time, so water available as an input to plant growth may depend not only on current local weather and soil conditions but also on their counterparts considerable distances away. Moreover, particularly within the context of the current article, a dam's irrigation benefit is likely to depend on the available river flow. To capture this river flow within our basins, we employ the Geospatial Stream Flow Model (GeoSFM), which is a semidistributed, physically based hydrological model developed by the United States Geological Survey (USGS (2011b) with particular relevance for Africa's hydrology (Asante, Macuacua, et al. 2007; Asante, Artan, et al. 2007).[12] More specifically, GeoSFM allows one to simulate the dynamics of runoff processes using spatial information on river basin and network coverage, land cover type, soil characteristics, and daily precipitation and evapotranspiration data. To satisfy the model's requirements for soil characteristics (water-holding capacity, hydrologically active soil depth, texture, average saturated hydraulic conductivity), we take data from the Digital Soil Map of the World (FAO 2011) and use the procedure by Schuol and Abbaspour (2007) to derive daily precipitation and evapotranspiration series per river basin using the CRU climate data. Using these inputs in GeoSFM produces the daily stream flow in terms of cubic meters per second (m^3/s), and we sum these to obtain annual equivalents. Importantly, GeoSFM is run solely with the time-invariant geographic characteristics of basins and time-varying climatic information as inputs and does not take into account the influence of flow through dam location and operation rules. Thus, it is the stream flow that would have occurred if water had been allowed to pass through the continent unobstructed. We provide summary statistics of all our variables in table 2.

Table 2. Summary Statistics

Variable	Mean	St. Dev.	Description
NPP(CROP)	1,219	288	NPP in cropland
NPP(MOSAIC)	1,244	286	NPP in cropland and cropland mosaic
NPP(CLOSE)	1,179	282	NPP in cropland and land within 100 km of dams
D(LARGE)	0.39	0.82	Large dams in own vicinity
D(SMALL)	9	19	Small dams in own vicinity
UD(LARGE)	0.20	0.66	Large dams upstream
UD(U2 - LARGE)	0.11	0.80	Large dams further upstream (one level)
UD(NB - LARGE)	2	6	Large dams in neighboring but non-upstream areas
UD(SMALL)	2	6	Small dams upstream
SPI	−0.6	0.5	Standard precipitation index
EVAP	10	1	Evapotranspiration
RF	674	2,040	River flow
RFH	0.14	0.35	High river flow event
RFL	0.05	0.22	Low river flow event
RFHU	0.13	0.34	High river flow event upstream
RFLU	0.07	0.25	Low river flow event upstream
RGr(MOD)	0.16	0.14	Fraction basin with perennial river gradient 1.5-3 percent
RGr(STEEP)	0.11	0.12	Fraction basin with perennial river gradient 3-6 percent
RGr(VSTEEP)	0.05	0.09	Fraction basin with perennial river gradient above 6 percent
ERLENGTH	10.3	54.9	Length of ephemeral rivers in kilometers

Sources: See Section II.
Note: Summary statistics refer to all level-six areas with at least some cropland, except for the river gradient variables, which also include areas upstream from these.

III. <<**A**>>*ECONOMETRIC RESULTS*

We proceed to our econometric analyses by first ensuring that our assumptions hold. We then present the main regression results and, finally, report various robustness checks.

<<**B**>>*Preliminary checks*

To verify that the river gradients matter in terms of predicting the location of dams,[13] we regressed the number of dams as well as the subgroups for our pooled data set over the

1981–2000 period on our three different river gradient dummy variables as well as controls for overall gradient, perennial river length, elevation, total area of a river basin, former homeland based fixed effects, and a set of year dummies.[14] As Duflo and Pande (2007) note, on engineering grounds, low-level river gradients should be preferred for large dam locations relative to both higher-level and very low slopes. In contrast, even fairly low river gradients are often suitable for small dams (see, for instance, Stone 2003; Nissen-Petersen 2006; FAO 2010).

As shown in the first column in table 3, the results for large dams are similar to those of Duflo and Pande (2007) and Strobl and Strobl (2011), where dam construction is more likely in moderately sloped basins and less likely for steeply sloped basins and is more likely for very steeply sloped gradients, all relative to low sloped gradients. As Duflo and Pande (2007) note, the result for the very steeply sloped dummy is most likely due to the often multipurpose nature of large dams.[15] We also include the length of ephemeral rivers to verify that ephemeral rivers are not conducive to large dam construction; as shown, this is indeed the case. We thus exclude ephemeral river length as a control in any subsequent specifications involving large dams to create our set of instruments.

Table 3. First Stage Regression Results, Dependent Variable: D(sz)

	(1)	(2)	(3)	(4)	(5)	(6)
RGr(MOD)	1.005**	−25.42**				
	(0.180)	(4.957)				
RGr(STEEP)	−0.948**	1.831				
	(0.227)	(7.221)				
RGr(VSTEEP)	0.224	−11.56*				
	(0.187)	(5.270)				
P	−0.000343	0.00270*	0.0816**	2.604**		
	(0.000247)	(0.00116)	(0.0113)	(0.371)		
RGr(MOD)×P					−0.342**	6.391*
					(0.0799)	(2.823)
RGr(STEEP)×P					−0.482**	−38.55**
					(0.0898)	(7.205)
RGr(VSTEEP)×P					−0.676**	−7.181*
					(0.199)	(3.320)
ERLENGTH×P						−0.00222**
						(0.000447)
Dams Size	LARGE	SMALL	LARGE	SMALL	LARGE	SMALL
Observations	4,460	4,460	4,460	4,460	4,460	4,460
Basins	223	223	223	223	223	223
F-Test for RGr_{it}	59.82**	44.26**	---	---	13.45**	10.89**

Note: Regressions include homeland, river basin elevation controls (three dummies indicating fraction of area with 250–500 m, 500–1,000, and above 1,000 m elevation), river basin overall gradient controls (three variables indicating the fraction of area with 1.5–3 percent, 3–6 percent, and over 6 percent gradient), area, total perennial and ephemeral river lengths, and time dummies. Standard errors given in parentheses are robust. ** and * are 1 percent and 5 percent significance levels, respectively; RGr(MOD) is the fraction of perennial river gradient at 1.5–3 percent, RGr(STEEP) is the fraction of perennial river gradient at 3–6 percent, RGr(VSTEEP) is the fraction of perennial river gradient at above 6 percent, and ERLENGTH is the length of ephemeral rivers in kilometers. The F-test for RGr_{it} refers to a test on the coefficients of the river gradient variables.

Examining small dams, we find that the presence of irrigation dams is less likely in very steep and moderate river gradient basins, as might be expected a priori. Perhaps somewhat peculiarly, we also find that small dams are more likely to be constructed in basins with larger shares of steeply sloped rivers. Importantly, we find that small dam location is more likely in basins where there are more ephemeral rivers.

We next check whether our policy change proxy serves as a reasonable predictor of dam allocation over our sample period. To this end, we rerun the specification described above but include our policy change proxy, P_{it}, and control for basin-level fixed effects and thus for all geographic and nongeographic time-invariant features. As shown, for both small and large dams, the estimated coefficient suggests that the policy changes from 1994 onward increased dam construction in basins that are predominantly based in former homelands.

Finally, we report the estimated coefficients of $RGr(k)_i * P_{it}$ from the first step regression of (2) for large dams and those of these interaction terms as well as with ephemeral river length for small dams for the pooled sample over all years in Table 3. As shown, the individual interaction terms are significant in all cases. Most important, the joint F-test statistic for the river gradient interaction variables demonstrates that these variables have substantial predictive power in all samples.

Main Results

When estimating equation (3), one concern is that, as Duflo and Pande (2007) note, agricultural data tend to be strongly autocorrelated. Indeed, tests indicated that our measure of agricultural productivity was autocorrelated to the third order. To address this problem, Duflo and Pande (2007) use a feasible optimal IV estimator. In our context, however, NPP data are also highly spatially correlated. A Moran's (1950) test indicates that for each of our 20 years, the null hypothesis of no spatial correlation could be decisively rejected. To address these two issues simultaneously, we implement the nonparametric covariance matrix estimator proposed by Driscoll and Kraay (1998) in an IV framework as in Driscoll (2004), which calculates standard errors corrected for spatial correlation, autocorrelation (to the third order), and heteroskedasticity using an extension of Newey and West's nonparametric variance-covariance estimator.

Estimation results for equation (3), considering all 223 river basins with cropland and including only our climatic controls as well as controlling for basin fixed- and year-specific effects, are given in the first column of table 4 for large dams only. Accordingly, the precipitation proxy, the standardized precipitation index (*SPI*), is significantly positive, confirming that greater rainfall increases cropland productivity. Somewhat surprisingly, the evapotranspiration index is marginally insignificant (at the 5 percent level), although in the latter specifications, it is significant. With regard to the river flow variable, Asante et al. (2008) note that river flow series generated from the GeoSFM are particularly suitable for describing river flow anomalies. In the second column, we experiment with including the dummy variables low river flow event and high river flow event, *RFL* and *RFH*, respectively, indicating when stream flow was one standard deviation below and one standard deviation above its basin-level mean (as calculated over the 1950–2000 period), respectively. As shown, we now find that years in which river flow was below its long-term mean induced a significant fall in cropland productivity, whereas positive anomalies had no impact.

Table 4. Regression Results Considering Large Dams Only, Dependent Variable: CP

	(1)	(2)	(3)	(4)
D(LARGE)			66.58**	321.1**
			(24.74)	(70.34)
UD(LARGE)			36.59	258.6**
			(27.66)	(91.77)
SPI	57.82*	57.88*	57.12*	55.27*
	(28.43)	(28.02)	(28.92)	(26.67)
EVAP	−75.80	−71.42	−72.27*	−74.67*
	(42.60)	(41.15)	(36.47)	(34.66)
RF	0.00856			
	(0.0130)			
RFL		−54.27**	−53.53**	−50.39**
		(19.45)	(19.23)	(17.98)
RFH		8.636	7.659	4.781
		(16.94)	(16.26)	(15.60)
DAMS	LARGE	LARGE	LARGE	LARGE
IV	No	No	No	Yes
Observations	4,460	4,460	4,460	4,460
Basins	223	223	223	223

Note: Driscoll and Kraay (1998) standard errors corrected for spatial and second-order autocorrelation are given in parentheses; ** and * are 1 percent and 5 percent significance levels. CP (dependent variable) denotes cropland productivity. Regressions include river basin fixed effects, former homeland year interactions, interactions of homeland basin predicted dams with river basin gradient, perennial and ephemeral river lengths, river basin area, river basin elevation, river gradients, and river gradient-year interaction terms.

In the third column, we include the number of large dams within a basin as well as the number of large dams located upstream, both in their noninstrumented form (that is, we estimate the ordinary least squares version of equation (3)). Accordingly, we find that large irrigation dams positively increase cropland productivity within their own basin, but they have no effect on cropland located immediately downstream. Importantly, however, instrumenting our dam variables changes these findings. In particular, as shown in the fourth column, the coefficient for upstream dams is now significant, suggesting that these dams do increase productivity. Moreover, the size of the own-vicinity dam coefficient increases multifold and, somewhat contrary to prior expectations, suggests an effect even larger than that of upstream dams. These two results together indicate, as Duflo and Pande (2007) and Strobl and Strobl (2011) show, that not controlling for endogeneity in dam location introduces a downward bias in the estimates. The larger impact of own-vicinity large dams stands in contrast to these previous studies.

The main purpose of this study is, of course, to investigate the possible role of small irrigation dams in aiding cropland productivity relative to large ones. Thus, we reestimate the specification of the last column of table 4 but include the (instrumented) number of small dams within a basin. This approach produces a number of important findings, as shown in the first column of table 5. First, the presence of small dams, *D(S)*, significantly increases cropland productivity within their basin. Second, in terms of the impact of small dam inclusion, one finds that the coefficient on upstream (large) dams only changes slightly in value. In contrast, large dams within a basin have a marginally insignificant (although still positive) effect on cropland productivity (that is, a third of that of upstream dams). This finding importantly suggests that where small dams are present, the identifying assumption of the IV approach developed by Duflo and Pande (2007) may be violated if these are not controlled for in the analysis. This finding is further confirmed by a simple z-test of the equality of coefficients (the z-statistic is 2.96). In terms of the relative impact of small dams, one finds that compared to a large upstream, a small dam has only 4.3 percent of the productivity-boosting effect on cropland.

Table 5. Regression Results Accounting for Large and Small Dams, Dependent Variable: CP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
D(LARGE)	77.40	67.22*	−73.32*	−95.79*	−75.52*	−91.08**	190.6
	(42.48)	(29.26)	(34.52)	(35.29)	(29.70)	(31.81)	(118.6)
UD(LARGE)	246.5*	184.6*	271.0**	234.2*	229.0**	258.9**	−2.805
	(97.42)	(82.28)	(93.79)	(89.14)	(47.19)	(86.80)	(5.198)
D(SMALL)	10.67**	9.533**	6.749**	6.448**	5.593**	6.707**	−2.178
	(2.255)	(2.493)	(2.065)	(1.919)	(1.535)	(2.059)	(9.590)
D(SMALL)×D(LARGE)			17.35**	16.73**	13.05**	16.72**	0.565
			(2.893)	(2.669)	(2.231)	(2.830)	(16.40)
UD(LARGE, NB)		212.4					
		(128.2)					
UD(LARGE, U2)		−84.75					
		(134.2)					
UD(SMALL)		2.256					
		(1.500)					
UDAM(SMALL, NB)		−9.617					
		(16.61)					
UDAM(SMALL, U2)		0.768					
		(0.867)					
SPI	53.95*	53.71*	52.92*	47.91*	32.41	49.64*	51.91*
	(23.64)	(23.77)	(23.41)	(22.57)	(16.66)	(22.29)	(22.86)
EVAP	−76.56*	−77.63*	−79.65*	−90.19**	−48.10*	−76.54*	−81.34*
	(32.81)	(31.02)	(31.45)	(29.76)	(21.52)	(31.33)	(31.36)
RFL	−47.88*	−46.73**	−45.28*	−43.23**	−38.70**	−44.85*	−42.35*
	(17.30)	(16.09)	(16.59)	(13.48)	(11.81)	(16.50)	(16.06)
RFH	0.0810	−0.133	0.598	−0.706	11.04	2.564	0.176
	(15.44)	(12.51)	(15.75)	(15.89)	(10.59)	(15.15)	(15.67)
CL	CROP	CROP	CROP	CLOSE	LC	MOSAIC	CROP
IV	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,460	4,460	4,460	4,460	4,460	4,460	4,460
Basins	223	223	223	223	223	223	223

Note: Driscoll and Kraay (1998) standard errors corrected for spatial and second order autocorrelation are given in parentheses; ** and * are 1 percent and 5 percent significance levels. CP (dependent variable) denotes cropland productivity. Regressions include river basin fixed effects, former homeland year interactions, interactions of homeland basin predicted dams with river basin gradient, perennial and ephemeral river lengths, river basin area, river basin elevation, river gradients, and river gradient-year interaction terms. In the first column, DAM and UDAM refer to the total of own-vicinity and upstream large and small dams together, whereas in the second and third columns, they refer to large dams only. LC excludes cells that had NPP of more than two standard deviations in 2000 compared to 1981. MOSAIC includes cropland mosaic. CLOSE indicates within 5 km of dams. In the last column, reported coefficients are for nonirrigation dams.

One may recall that, conceptually, there are a number of assumptions underlying our base econometric specification. More specifically, in terms of large dams, we abstract from the potential indirect general equilibrium impact of large dams that are neighboring but not hydrologically related as well as the effect of those that are further (than immediately) upstream. For small dams, we assume that they are too small to have any impact outside their own vicinity. To verify that these assumptions are plausible, we introduce the instrumented number of large dams in neighboring nonhydrologically related (*NB*) and in further upstream (*U2*) basins as well as the number of small upstream, neighboring nonhydrologically related, and further upstream. As shown in table 5, none of these additional variables is statistically significant, suggesting that small dams are too small to have any impact outside their own basin and that large dams can have an impact outside their own basin but only on those crops located immediately downstream.[16]

As noted in the introduction, the view of some critics has recently been that perhaps the issue should not be considered a choice between large and small dams; rather, these two types together can provide a more efficiently functioning irrigation system, where large dams can aid the impact of smaller dams. To investigate this possibility, we include their interaction term in the third column. This approach dramatically changes some of the previous findings. In particular, one now finds that the large dam variable is significantly negative, thus implying that a large dam has a negative impact of cropland within its vicinity if there are no small dams present, lending some truth to the assertion that water seepage and soil salinization harms local farmers. Because the average number of small dams is approximately nine, for the average river basin, the net effect of a large dam is reassuringly about the same (57 units of NPP) as when one does not control for the interaction with small dams (as shown in the first two columns). When the interaction term is included, the effect of small dams on their own is approximately 40 percent smaller, implying that some of the small dam impact in the previous results is due to the benefit of being near a large dam. In other words, some of the nonnegative impact for large dams is due to a compensating effect of small dams. Perhaps most important, we find that the positive coefficient on the interaction term suggests that large dams can augment the positive effect of small dams, perhaps by being part of a coordinated system of irrigation.

Robustness Checks

Our interpretation of the negative impact of large dams on NPP is that this impact is likely due to water seepage, soil salinization, and, perhaps, other negative spillovers from the construction of these large infrastructure projects on cropland. If this were indeed the case, one would most likely expect this impact to be particularly prominent for cropland located near the dam. To investigate this possibility, we restrict our dependent variable to capture only cropland within 5 km of any large dam. As depicted in the fourth column of table 5, the negative effect is indeed larger (22 percent) for these areas, although a simple z-test suggests that this difference is insignificant.

A major drawback of our satellite-derived productivity data is that we can only identify cropland that existed in 2000. Thus, we need to assume that these cells were used for cropping throughout our sample period. Apart from general measurement error, this assumption could feasibly introduce systematic misclassification if the presence of dams caused changes in cropland. Thus, for some cells, changes in NPP may be due in part to changes in land use rather than productivity-enhancing or productivity-reducing effects of irrigation dams. To roughly investigate the role of such measurement error, we attempt to disregard grid cells that are more likely to have had a different land use classification prior to 2000. More precisely, we exclude all cells (approximately 8 percent) whose 1981 NPP value was either two standard deviations above or two standard deviations below the mean value of NPP for cropland in 2000. Two standard deviations below corresponds roughly to that what would be the mean value of the category “Sparse Herbaceous or Sparse Shrub Cover,” whereas two standard deviations above is close to the mean value of the category “Tree Cover, Broadleaved, Evergreen.” The results, shown in the fifth column, of using the mean basin values of NPP for this subset of cells demonstrate that the negative effect of own large dams is somewhat smaller, the positive effects of small dams and their interaction with large dams are also somewhat smaller, and the positive impact of upstream large dams is now marginally larger. Again, simple z-tests suggest that these differences are not statistically significant. Thus, there is little evidence that our use of the 2000 land use classification produces noticeable biases in our results.

Thus far, we have restricted our sample to include areas from 1 km^2 grid cells from the GLC 2000 that are characterized as “pure” cropland and to exclude those cells categorized as containing cropland within the mosaic of other vegetation. Feasibly, this approach may ignore smaller cropland and limit the interpretation of our results beyond very large or agglomerated plots. Thus, we reestimate our base specification from specification (13) using the mean NPP value of all cropland, including that classified as mosaic.[17] As shown, there is little change in the coefficients for local small dams or upstream large dams. However, the negative impact of large local dams is again somewhat larger than that for the sample using only pure cropland cells. As before, these differences are not statistically significant.

Finally, one should note that underlying the identification of the IV strategy is the assumption that agricultural production would not have varied in response to the policy changes considered as a function of land suitability in the absence of irrigation dam construction. An insightful placebo test in this regard would be to conduct a similar analysis using all nonirrigation dams in South Africa. More specifically, we employ the same IV strategy as above using nonirrigation dams from the DWA database. There were 90 large and 772 small nonirrigation dams in the data, the main functions of which differed widely, ranging from water and industrial supply to recreational use. The coefficients on the nonirrigation dams are reported in the last column of table 5. As shown, reassuringly, none of these is statistically significant. Hence, these findings provide at least secondary support for the validity of our identifying assumption.

River Flow Dependence

Arguably, the storage of water in dams will be reduced during drought years, resulting in limited release and hence ability of dams to provide irrigation benefits. To this end, Duflo and Pande (2007) and Strobl and Strobl (2011) identify drought events by considering local precipitation. These authors find that the estimated impacts depend on these events. Of course, the river flow local to the dam actually determines the amount of water stored, and, as argued earlier, this depends on the complete hydrological cycle rather than only local precipitation. Therefore, we use our river flow proxy dummies to isolate water-rich and water-poor years in a region. Additionally, we identify high and low river flow events in upstream areas and calculate dummy variables, *RFHU* and *RFLU*, respectively. As shown in the first column in table 6, regressing NPP measure on these as well as all other climatic variables shows that the upstream variables have no direct effect on local cropland productivity, whereas the local low river flow continues to have a negative effect.

To investigate how river flow may determine the impact of dams, we include our instrumented local and upstream dam variables and their interaction terms with the river flow dummies, initially for large dams only. As shown in the second column, although the average positive effect of large local and upstream dams is maintained, their overall impact is dependent, in part, on the available river flow. In particular, we find that low river flow reduces the overall impact of large dams by nearly one-third. In contrast, the upstream large dam effect is enhanced by larger river flow upstream.

One might expect that, given their much lower storage capacity, the benefits from small dams would be less dependent on local river flow. This assumption is confirmed in the third column, where we include the instrumented number of smaller dams, its interaction term with large local dams, and their interaction terms with local river flow anomaly events. Accordingly, the results for large dams as well as their dependence on local and upstream river flow hold as before. In contrast, the impact of small dams on their own does not depend on local annual mean river flow. Moreover, the additional benefits are not dependent on river flow.

Table 6. Regression Results Accounting for River Flow Dependency, Dependent Variable: CP

	(1)	(2)	(3)
D(LARGE)		327.4**	−78.58*
		(67.16)	(37.37)
D(LARGE)×RFH		48.50	157.0*
		(32.51)	(55.79)
D(LARGE)×RFL		−97.29**	−30.12
		(29.81)	(89.03)
UD(LARGE)		262.8**	284.3**
		(87.93)	(88.72)
UD(LARGE)×RFHU		48.73**	50.66**
		(14.83)	(14.72)
UD(LARGE)×RFLU		6.400	−0.502
		(27.15)	(25.03)
D(SMALL)			6.284**
			(1.916)
D(SMALL)×RFH			5.821
			(3.521)
D(SMALL)×RFL			0.529
			(2.502)
D(SMALL)×D(LARGE)			18.46**
			(2.942)
D(SMALL)×D(LARGE)×RFH			−10.86
			(5.599)
D(SMALL)×D(LARGE)×RFL			−5.814
			(9.354)
SPI	57.76	55.74	54.29*
	(31.01)	(27.58)	(23.80)
EVAP	−69.66	−73.31	−74.78*
	(40.16)	(37.73)	(34.81)
RFL	−47.62**	−10.11	−3.762
	(13.51)	(18.27)	(18.08)
RFH	18.34*	−4.976	−47.56
	(7.200)	(16.53)	(22.88)
RFLU	−19.07	−11.13	−10.40
	(17.06)	(14.33)	(13.23)
RFHU	−27.02	−33.56	−33.76
	(25.86)	(24.79)	(25.19)
CL	CROP	CROP	CROP
IV	No	Yes	Yes
Observations	4,460	4,460	4,460
Basins	223	223	223

Note: Driscoll and Kraay (1998) standard errors corrected for spatial and second order autocorrelation are given in parentheses; ** and * are 1 percent and 5 percent significance levels. CP (dependent variable) denotes cropland productivity. Regressions include river basin fixed effects, former homeland year interactions, interactions of homeland basin predicted dams with river basin gradient, perennial and ephemeral river lengths, river basin area, river basin elevation, river gradients, and river gradient-year interaction terms.

IV. COST-BENEFIT ANALYSIS

We now use our estimated coefficients to provide a cost-benefit analysis of large and small dams. Unfortunately, we do not have explicit information on the costs of construction of the dams in our database. However, Alexander and van Wyk (2005) provide estimates of costs for nine dams in South Africa according to their height and capacity, and we use these data to infer the costs of dams in our sample. More specifically, we regress their unit costs in terms of dollar per cubic meter of capacity on dam height and find the following relationship:

$$\textit{Unit cost} = 0.4161921 - 0.0063412 \,^{*}\, \textit{height} \,, \qquad (4)$$

where both coefficients are significant at the 1 percent level. Using the mean values of the capacity and height of small and large dams from our sample as given in table 1, these estimated coefficients suggest that large and small dams have average unit costs of \$0.12/m^3 and \$0.35/m^3 and average total costs of \$15,123,640 and \$124,531, respectively.

To determine the average total benefits of our large and small irrigation dams, we need both an indicator of their average annual benefit to agricultural production and their average life span. The annual benefits can be derived from our estimated coefficients and the fact that one unit of NPP roughly translates into 1 kilocalorie of food. Conveniently, the units of the annual NPP measure (gC/m^2/year) can be roughly converted into kilocalories and, hence, to their nutritional value if the crop involves food products. That is, as Mackenzie et al. (2004) note, organic matter stores approximately 112 kcals of energy per 12 gC; hence, 1 gC is equal to approximately 9.33 kcals. However, some of the energy in NPP is lost postharvest owing to transport and processing, crop residue left in the field, and roots. We thus use the adjustments for each of these factors proposed by Imhoff et al. (2004) to convert our NPP kilocaloric measure into kilocalories of final agricultural production available for human consumption.[18] This approach suggests that 1 gC/m^2/year of NPP can be converted into approximately 1 kilocalorie available for human consumption. Given that global estimates of the cost of the minimum amount of kilocalorie per person per day (2100) is approximately \$1.25,[19] then one unit of NPP would be valued at roughly 0.06 cents. The annual monetary benefits of a dam can then be derived by using our estimated coefficient, which gives the marginal additional annual benefit to each square meter of cropland, and the fact that basins have, on average, 15 km^2 of cropland.

To proxy the life span of a dam, we take advantage of the fact that the South African dam register classifies many dams in terms of their safety risk. More specifically, dams are assessed structurally to determine whether they pose any risk in terms of economic loss and are then categorized accordingly. We take the mean age of those classified as significant risk as a proxy for when they have reached the end of their life span. The mean age of dams within this category in our sample was found to be 45 and 34 years for large and small dams, respectively.[20]

Using our estimated coefficients from the third column of table 5, the estimated life spans, and the assumption that there is cropland downstream from each large dam, we find that the net annual own-vicinity and downstream benefit from a large dam is $1,246,885, whereas the annual (own-vicinity) benefit for a small dam is $58,032. Given their construction costs, this finding suggests that large dams, on their own, have an internal rate of return of 8 percent, whereas the equivalent figure for a small dam is 45 percent. Because the larger dam likely aids the smaller one rather than vice versa, we attribute any additional benefits due to their interaction to the large dam. In this regard, the internal rate of return for a large dam rises by one percentage point for each small dam in the vicinity, suggesting that in the presence of 37 small dams, the internal rates of return would be equivalent. If we consider a scenario in which river flow is high, then the internal rate of return for large dams on their own is 20 percent.

Considering the total benefits accrued over their lifetimes, the benefit-cost ratio over a large dam's lifetime is approximately four and that of a small dam is 16. The latter rises by an additional 0.3 for the former for each additional small dam in the vicinity. The corresponding benefit-cost ratio for large dams under a high river flow scenario is 10.

V. CONCLUDING REMARKS

Our analysis demonstrates that the controversy of small versus large dams is somewhat misdirected when considering irrigation. More specifically, we show that the net impact of local large dams on NPP depends on controlling for small dams and the interaction between these two types. Although large dams can have negative impacts on nearby cropland, in contrast to their positive effect downstream, they triple the relatively small productivity-enhancing effect of small dams. Nevertheless, the overall impact of large dams is somewhat sensitive to available river flow.

A simple cost-benefit analysis demonstrates that smaller dams appear to have a substantially more favorable internal rate of return and benefit-cost ratio than large dams if considered in isolation, but this benefit can be partially offset by considering large dams as part of a system of small dams. However, our calculations should be viewed with some caution because they only consider irrigation benefits and construction costs and, as such, may particularly over- or underestimate the figures for large dams. For example, we do not incorporate population displacement and environmental costs, factors that originally called the feasibility of large dams into question. However, large irrigation dams can also be multipurpose, providing not only irrigation benefits but also energy through hydropower production. Thus, our analysis only provides an incomplete picture of the large versus small dam controversy. Finally, we only estimate the effect on yield, not the effect on profit. In this regard, one might very well expect prices on crops to fall as the yield supplied on the market increases as a result of dams. Moreover, changes in irrigation may change the price that farmers have to pay for irrigation. In principle, then, the increase in net profit could be substantially lower than the increase in yields.

REFERENCES

Alexander, R. A., and D. I. van Wyk. 2005. Dam Development Options and Economic. Permanent Water Commission.

Asante, K., G. Artan, S. Pervez, C. Bandaragoda, and J. Verdin. 2007. Technical Manual for the Geospatial Stream Flow Model (GeoSFM). In *US Geological Survey Open-File Report 2007*. Reston, VA: USGS.

Asante, K., G. Artan, S. Pervez, and J. Rowland. 2008. "A Linear Geospatial Streamflow Modeling System for Data Sparse Environments." *International Journal of River Basin Management* 6: 233–241.

Asante, K., R. Macuacua, G. Artan, R. Lietzow, and J. Verdin. 2007. "Developing a Flood Monitorig System From Remotely Sensed Data for the Limpopo Basin." *IEEE Transactions on Geoscience and Remote Sensing* 45: 1709–1714.

Blanc, E., and E. Strobl. 2013. "The Impact of Climate Change on Cropland Productivity: Evidence from Satellite Based Products at the River Basin Scale in Africa." *Climatic Change* 117 (4): 873–890.

Chen, S., and M. Ravallion 2008. "The developing world is poorer than we thought, but no less successful in the fight against poverty." In *World Bank Policy Research Working Paper* 4703.

Conca, Ken. 2005. "Growth and Fragmentation in Expert Networks: The Elusive Quest for Integrated Water Resources Management." In *Handbook of Global Environmental Politics*. Cheltenham, UK.

Driscoll, J. 2004. "Does Bank Lending affect Output? Evidence from US States." *Journal of Monetary Economics* 51: 451–471.

Driscoll, J. C., and A. C. Kraay. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics* 80: 549–560.

Duflo, E., and R. Pande. 2007. "Dams." *Quarterly Journal of Economics* 122: 601–646.

DWA. 2011. *Department of Water Affairs of South Africa*. http://www.dwa.gov.za/.

Ersado, Lire. 2005. "Small-Scale Irrigation Dams, Agricultural Production, and Health: Theory and Evidence from Ethiopia." World Bank Policy Research Working Paper No. 3494.

FAO. 2000. Rivers of Africa, Geo-spatial database, First edition. http://www.fao.org/geonetwork/srv/en/metadata.show?id=12&currTab=summary.

FAO. 2010. "Manual on Small Earth Dams: A Guide to Siting, Design, and Construction." *FAO Irrigation and Drainage Paper* 64.

FAO. 2011. Digital Soil Map of the World. http://www.fao.org/nr/land/soils/digital-soil-map-of-the-world/en/.

FAOSTAT. 2011. FAO Statistical Databases. http://faostat.fao.org.

Foster, V., and C. Briceño-Garmendia. 2010. Africa’s Infrastructure: A Time for Transformation. Washington, D.C.: The World Bank.

GEM. 2011. Global Land Cover 2000 data set (GLC 2000). Global Environment Monitoring Unit (GEM).

Hargreaves, G.H., and Z.A. Samani. 1985. "Reference Crop Evapotranspiration from Temperature." *Applied engineering in agriculture* (1): 96–99.

Hathaway, T., and L. Pottinger. 2008. "The Great Hydro-Rush: The Privatisation of Africa’s Rivers." In *Electric Capitalism: Recolonising Africa on the Power Grid*, edited by D. McDonald. HSRC Press.

Heinsch, F.A., W.M. Jolly, Q. Mu, and J.S. Kimball. 2005. *Regional Scaling of Cropland Net Primary Production for Nebraska using Satellite Remote Sensing from MODIS and the Ecosystem Process Model BIOME-BGC*: American Geophysical Union.

Hicke, J., D. Lobell, and G. Asner. 2004. "Cropland Area and Net Primary Production Computed from 30 years of USDA Agricultural Harvest Data." *Earth Interactions* 8: 1–20.

ICID. 2000. Role of Dams, Irrigation, Drainage and Flood Control. In *International Commission on Irrigation and Drainage Position Paper*.

Imhoff, M., and L. Bounoua. 2006. "Exploring Global Patterns of Net Primary Production Carbon Supply and Demand Using Satellite Observations and Statistical Data." *Journal of Geophysical Research* 111: 1–8.

Imhoff, M., L. Bounoua, T. Ricketts, C. Loucks, R. Harriss, and W. Lawrence. 2004. "Global Patterns in Human Consumption of Net Primary Production." *Nature* 429: 870–873.

Keller, A., R. Sakthivadivel, and D. Seckler. 2000. "Water Scarcity and the Role of Storage in Development." International Water Management Institute Research Report No. 39.

Lasage, R., J. Aerts, G. Mutiso, and A. de Vries. 2008. "Potential for Community Based Adaption to Droughts: Sand Dams in Kitui, Kenya." *Physics and Chemistry of the Earth* 33: 67–73.

Lobell, D., J. Hicke, G. Anser, C. Field, C. Tucker, and S. Los. 2002. "Satellite Estimates of Productivity and Light Use Efficiency in United States Agriculture, 1982-1998." *Global Change Biology* 8: 1–15.

Lustenberger, F. 2010. "The Problem of Water in Southern Africa and in particular the Management of Water in South Africa." AgroParisTech Working Paper.

Mitchell, T.D., and P.D. Jones. 2005. "An Improved Method of Constructing a Database of Monthly Climate Observations and Associated High Resolution Grids." *International Journal of Climatology* 25: 693–712.

Moran, P.A.P. 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika* 37: 17–23.

Nissen-Petersen, E. 2006. *Water for Small Dams: A Handbook for Technicians, Farmers and Others on Site Investigations, Designs, Cost Estimates, Construction, and Maintenance of Small Dams*: ASAL Consultants Ltd.

RAIN. 2007. A Practical Guide to Sand Dam Implementation. Amsterdam: Rainwater Harvesting Implementation Network.

Schuol, J., and K. Abbaspour. 2007. "Using Monthly Weather Statistics to Generate Daily data in SWAT Model Application to West Africa." *Ecological Modelling* 201: 301–311.

Smil V. 2003. *Energy at a Crossroads*. MIT Press.

Stein, Robyn. 2000. "South Africa's New Democratic Water Legislation: National Goverment's Role as Public Trustee in Dam Building and Management Activities." *Journal of Energy and Natural Resources* 18: 285–295.

Stone, L. 2003. *Earthen Dams for Small Catchments: A Compilation of Design, Analysis, and Construction Techniques Suitable for the Developing World*, Michigan Technological University.

Strobl, Eric, and Robert O. Strobl. 2010. "The Distributional Impact of Large Dams: Evidence from Cropland Productivity in Africa." *Journal of Development Economics* 96 (2): 1–19.

Turner, D., W. Ritts, W. Cohen, S. Gower, S. Running, M. Zhao, M. Costa, A. Kirschbaum, J. Ham, S. Saleska, and D. Ahl. 2006. "Evaluation of MODIS NPP and GPP Products across Multiple Biomes." *Remote Sensing Environment* 102: 282–292.

Twala, W.D. 2006. "Land and Agrarian Reform in South Africa." In *Promised Land: Competing Visions of Agrarian Reform*, edited by P. Rosset. Food First Books.

USGS. 2011a. Global 20 Arc-Second Elevation (GTOPO20). U.S. Geological Survey. http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/Elevation_Products.

USGS. 2011b. HYDRO1k Dataset. U.S. Geological Survey. http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/HYDRO1K.

WCD. 2000. *Dams and Development: A New Framework for Decision-Making. The Report of the World Commission on Dams*: Earthscan Publications Ltd.

You et al. 2010.

Young, L. 2008. Irrigation Investment Needs in Sub-Saharan Africa. In *Background Paper 9*, *Africa Infrastructure Country Diagnostic*. Washington, D.C.: World Bank.

Zhang, Y., Q. Yu, J. Jiang, and Y. Tang. 2008. "Calibration of Terra/MODIS Gross Primary Production over an Irrigated Cropland on the North China Plain and Alpine Meadow on the Tibetan Plateau." *Global Change Biology* 14: 757–767.

Notes

---

1 See, for instance, www.internationalrivers.org/

2 For example, You et al. (2010) estimate that the total irrigated area in South Africa could be increased by a further 566,000 hectares, 33 percent of which would be due to small-scale irrigation dams.

3 Note that there may also be potentially negative effects on further downstream plots if these are of the floodplain type. However, floodplain agriculture is not really a prominent feature of South African agriculture; hence, we do not consider this aspect to be relevant here.

4 We also experimented with using a dummy variable that takes the value of one from 1994 onward, but this approach produced similar results in our econometric analysis.

5 More than 85 percent of large dams in South Africa are of the arch or gravity type, or both.

6 It should be noted that we do not use the river gradient of ephemeral rivers because almost all ephemeral rivers in South Africa were below 1.5 percent, as is usual for ephemeral rivers in Africa.

7 Similar to Duflo and Pande (2007), we define elevation measures as the percentage in a basin falling in the categories of 0–250 meters, 250–500 meters, 500–1,000 meters, and above 1,000 meters. In terms of overall river basin gradients, we use a breakdown analogous to the river gradient categories described above.

8 Dams of these dimensions are required by law to be registered.

9 Even more aggregate regional level agricultural data on any consistent basis only exists for the former nonhomeland areas prior to the fall of apartheid.

10 Unfortunately, the scale of the graph only allows us to observe highly agglomerated cropland areas.

11 The standardized precipitation index is used by the South African Weather Service to track changes in precipitation.

12 GeoSFM is currently used as a component in the United States Geological Survey famine early warning system; see http://earlywarning.usgs.gov/fews/.

13 To predict dam location, we use all relevant South African river basins for our analysis (that is, those that contain some cropland as well as the basins upstream from these).

14 We also attempted to run this specification for each year, excluding the time dummies. The results are qualitatively and, generally, quantitatively very similar across all years.

15 For a few large dams, our dam database lists other purposes in addition to irrigation. However, in many cases, internet searches on the dams that only listed irrigation as their purpose also indicated that these dams were used for other purposes, such as hydropower and water supply. Therefore, it is unlikely that the information on the purpose of dams allows us to truly distinguish between multipurpose and single purpose dams. Nevertheless, we reran our specification for large dams excluding the dams that listed more than one purpose. This approach reduces the size of the coefficient toward zero and, if performed by year, below zero in some cases, although it always remains insignificant.

[16] It is noteworthy that the coefficients, particularly on *UD(LARGE, NB)*, are large, meaning that in case of a type II error, we could not rule out very large positive equilibrium effects.

[17] Of our 223 basins, 960 also contained this type of vegetation.

[18] More specifically, we use multipliers of 0.4, 1.28, and 2, respectively, to systematically subtract likely losses from the given energy value (Imhoff et al. 2004; Imhoff and Bounoua 2006).

[19] See Chen and Ravallion (2008).

[20] The life span of 45 years for large dams coincides well with the average expectancy proposed in Smil (2003).