

# The impact on farm household welfare of large irrigation dams and their distribution across hydrological basins: Insights from northern Nigeria

*Hiroyuki Takeshima, Adetola Adeoti, and Oluwafemi Adebola Popoola*

## TABLE OF CONTENTS

Abstract .....	1
1. Background .....	1
2. Possible mechanisms for the distributional effects of large irrigation dams across hydrological basins .....	2
3. Empirical framework .....	3
4. Data and descriptive statistics .....	4
4.1 Analytical Variables .....	5
4.2 Analytical Samples .....	6
4.3 Descriptive statistics .....	8
5. Results .....	9
5.1 Main findings .....	9
5.2 Robustness checks .....	11
6. Conclusions .....	12
Appendices .....	14
Appendix A: Estimation of average treatment effects on the treated (ATT) through the multiple-treatment inverse probability weighting method .....	14
Appendix B: Multinomial logit model results .....	15
References .....	16

## LIST OF TABLES

Table 1—Descriptive statistics for mild or no-drought sub-sample of households .....	8
Table 2—Descriptive statistics for high drought sub-sample of households .....	9
Table 3—Effects of switching household residence to a different type of hydrological basins, estimated as average effect of treatment on the treated .....	10
Table 4—Effects of drought on growth rate of real per capita income and of food consumption among comparable drought-defined sub-samples within downstream basins .....	10
Table 5—Share of sample households planting crops in each month (waves 1 and 2 combined), January to August, percent .....	11
Table 6—Estimated coefficients of multinomial logit model (base outcome = downstream basin) .....	15

## LIST OF FIGURES

Figure 1—Classification of hydrological basins in Nigeria based on large irrigation dams .....	5
Figure 2—Number of different types of dams completed in Nigeria, 1960 to 2010, cumulative .....	7

## ABSTRACT

Despite substantial past investment and continued interest in irrigation dam construction in Nigeria, evidence on the impact of such dams on household welfare is generally scarce. In particular, relatively few studies have been done on the geographical scope that their benefits may reach, despite growing evidence from elsewhere that the benefits of large irrigation dams can extend beyond the districts or hydrological basins that contain them, reaching particularly to hydrological basins located downstream. This study assesses the short-term effects of large irrigation dams on household consumption in the northern part of Nigeria. Using two rounds of the Nigeria LSMS survey, we apply multinomial logit inverse probability weighting (MIPW) methods to construct matching samples across three different types of hydrological basins – dam basins, which are basins that contain large irrigation dams and the area upstream of such dams; downstream basins, which are located downstream of large irrigation dams; and non-dam basins, which are not associated with large irrigation dams. Our analyses particularly focus on the benefits provided by such dams for mitigating the drought risks faced by farm households. Drought is an important factor that affects adversely the welfare of farm household in Nigeria. Supplemental irrigation is often used during drought to provide water to crops.

We find that in 2010 and 2012 farm households in downstream basins were relatively less affected by drought and enjoyed relatively stable between-season growth rates of real per capita income and food consumption compared to comparable households in dam basins or in non-dam basins. These findings are consistent with the hypothesis that such households are able to limit the damage drought might cause on their own food production. We link these findings to the following policy messages: (1) downstream hydrological basins are important geographical units to include in assessing the benefits from construction of large irrigation dams in Nigeria; and (2) decisions on where to construct new large irrigation dams or which existing dams to rehabilitate should be partly guided by information on the agricultural productivity and income levels of households in downstream hydrological basins relative to the basins within which construction or rehabilitation of such dams is planned.

## I. BACKGROUND

Agriculture provides an important source of income and food security in low-income countries worldwide. Irrigation is an important agricultural technology that enables both agricultural intensification and rainfall risk mitigation. Globally, the share of cultivated land that has been equipped for irrigation increased from 13 to 21 percent between 1970 and 2012 (FAO 2014). In many countries in sub-Saharan Africa, like Nigeria, however, the share has remained much lower. Despite an irrigation potential of 3 million ha (You et al. 2011) or 10 percent of Nigeria's cultivated area, only about one percent of cultivated land in Nigeria is irrigated.

Dams are important storages of irrigation water globally. In 2000, approximately 30 percent of the irrigated area in the world was irrigated by dams (Duflo and Pande 2007). There has been continued interest by donors in investments to improve the utilization of irrigation dams in sub-Saharan Africa countries like Nigeria – for example, in 2014 the World Bank started in Nigeria a \$500 million project entitled “Transforming Irrigation Management in Nigeria”. Consequently, deeper knowledge about the impact of dams on household welfare remains critical in better informing such efforts.<sup>1</sup>

An important element in assessing the household level impact of large irrigation dams is the spatial distribution of their benefits. Such effects operate both at micro-level within, for example, a particular irrigation scheme and at macro-level across districts or hydrological basins. At the macro-level, downstream regions often benefit more than the areas that contain dams. The literature shows that such distributional effects are clearly observed through district level indicators of agricultural production and incomes (e.g., Duflo and Pande 2007) or land-pixel level indicators of agricultural production (e.g., Strobi and Strobi 2011). These studies suggest that the distributional effects of large irrigation dams can extend beyond districts or hydrological basins that contain the dams and well beyond the area of the irrigation schemes that are typically located next to the dams. Specifically, these studies reveal that a substantial share of the benefits of large irrigation dams accrue to the hydrological basins that are located downstream of the basins containing the dams.

---

<sup>1</sup>Investigating the effects of irrigation dams on household welfare is also important because dams have not always been a major source of irrigation growth. For example, in Japan, rivers account for 70 to 80 percent of the total irrigated area (Gomi 1978), despite the fact that Japan has more than 1,500 large agricultural dams with heights of 15 m or more, and more than 200,000 small reservoirs. Seventy percent of irrigation water is provided by other sources than dams worldwide. Expansion of irrigation has sometimes been led by private sector investments into river diversions or through the use of motorized pumps to extract underground water (Goldman and Smith 1995; Villholth 2013; Takeshima 2016b). In such circumstances, the impact of dams on household welfare across basins is likely to be weaker.

However, in Nigeria, historically, the construction and rehabilitation of large irrigation dams had involved little discussion on the geographical areas that may benefit from the dams. Oftentimes, irrigation dams were built with an insufficient understanding of engineering and local hydrology (Andræ and Beckman 1985). The characteristics of Nigeria's hydrological basins, as in many other countries in Africa, are not well known or have only recently been investigated. Most large dams on the continent were built before 1990. Clearly, there was a limited understanding of where and to what sort of households the potential benefits of such dams would extend. In a recent donor report summarizing the Nigerian government's water management strategies (FAO et al. 2014), it was observed that the locations of the intended beneficiaries of these strategies are not clearly identified, other than those resident in the irrigation schemes typically located immediately next to the dams. Little to no attention is paid to the potential benefits such large irrigation dams may provide households located in basins downstream of the dam.

It is therefore important to better understand how the benefits of large irrigation dams may be distributed across geographical units like hydrological basins. One of the primary benefits from such dams for downstream areas may be the role the dams play in mitigating the effects of drought. Irrigation plays an important role in mitigating the effects of drought. Although much of the irrigation out of such dams occurs in the dry season, the ability to use the water stored in the dams to supplement the limited rainfall obtained by crops during drought periods is an important secondary benefit these dams provide.

Evidence is relatively scarce on the macro-level distributional effects of large irrigation dams on household level welfare and how these dams mitigate weather-related risks, such as droughts, that remain a major threat to agricultural productions in sub-Saharan Africa. We address some of this knowledge gap by using data from the Nigeria Living Standard Measurement Study (LSMS). From both policy and research perspectives, Nigeria is an appropriate case for examining the household welfare impact of irrigation dams in sub-Saharan Africa. Nigeria has more than 200 dams, of which 57 are large irrigation dams, one of the largest concentration of such dams in Africa (FAO 2015; FMWR 2007). Substantial shares of the population in Nigeria live in hydrological basins associated with these large irrigation dams. In addition, the total reservoir capacities of dams in Nigeria is greater than that found in relatively more advanced countries in Africa, like South Africa and Morocco. Nigeria invested 3 billion US dollars between 1970s and 1980s for the constructions of dams (Pradhan 1993, p.21). While many of these dams still continue to operate, little rigorous assessment of their impact on surrounding populations has been conducted. Lastly, the LSMS data from Nigeria are appropriate for such a study as they allow comparisons of households in dam-related basins across wide geographical regions. This improves the external validity of the results, enabling the result to be more broadly applied across Africa.

The fact that most large irrigation dams in Nigeria were built before 1990 poses some challenges in identifying their impacts on household welfare using recently collected LSMS data. We address this issue by constructing suitable counterfactual samples with varying spatial relationships to dams, so that these comparisons can mimic comparisons of households before and after dam construction. Specifically, following Strobi and Strobi (2011), we identify three types of hydrological basins in Nigeria – dam basins, which are basins that contain large irrigation dams and the area upstream of such dams; downstream basins, which are located downstream of dams; and non-dam basins that are not associated with dams. We apply a multiple-treatment inverse probability weighting method (Cattaneo 2010; Cattaneo, Drukker and Holland 2013) to jointly estimate the treatment effects for a household of switching from one type of basin to the other. We further improve the quality of our sample matching by focusing on the first differences of outcome indicators by exploiting the panel structure of waves 1 and 2 of the LSMS survey.

Our analyses suggest that in 2010 and 2012, households in downstream basins were relatively less affected by drought and enjoyed relatively stable between-season growth rates of real per capita income and food consumption, compared to what households in other types of basins experienced. This is likely to be partly due to irrigation limiting the damage of drought on food production for households in downstream basins.

Our results contribute to the literature in several ways. First, they support earlier findings that dams in developing countries can exert considerable distributional effects on household welfare across wide geographical areas, such as hydrological basins, transferring greater benefits to downstream basins than to other types of hydrological basins. Second, they show that these effects may materialize both at the regional level and at the household level in ways that partly shield households from drought risks, an aspect of household welfare that has not been studied in the literature. In addition, our results offer important policy guidance, as is discussed in detail in the conclusion section.

## **2. POSSIBLE MECHANISMS FOR THE DISTRIBUTIONAL EFFECTS OF LARGE IRRIGATION DAMS ACROSS HYDROLOGICAL BASINS**

While our analyses focus on the effects of large irrigation dams on household incomes across hydrological basins and, consequently, cannot directly assess the mechanisms of such effects, here we briefly discuss potential mechanisms to aid in the causal interpretations of our results.

Large irrigation dams can stabilize water flow to downstream basins by artificially controlling discharge. Indeed, dam systems are often best suited to stabilize downstream water flow, rather than to improve water availability in areas surrounding the dams. While it is possible to improve water supply to areas surrounding dams, doing so is costly because the beneficiaries are generally not located in areas to which water flows naturally. Consequently, greater costs must be incurred for conveying water to them.

During droughts, dams stabilize water flows to downstream regions and help maintain local water table levels. These functions of such dams enable beneficiary farmers downstream to obtain irrigation water at lower costs than otherwise would be the case. The higher water table and greater water volume in the river allows irrigation through groundwater pumping or through river diversion of a greater area than would be possible without the dam. Moreover, the higher water table also generally reduces irrigation water extraction costs, saving fuel or making feasible the use of cheaper, lower capacity pumps.

Not all farmers downstream of such dams will adopt irrigation because many farmers may still find the costs of irrigation to be too high. However, even under rain-fed conditions, the higher water table underground downstream of such dams can facilitate water extraction by plants from the soil, allowing the plants to better survive the drought period. This can lead to greater and more stable production even under farmer producing crops exclusively under rain-fed conditions. While some plants, such as certain varieties of cowpeas that have deeper roots (Thomas and Adams 1999) may be generally resilient against declining water table during drought, those varieties are not always available to farmers everywhere.

Stabilizing the water supply downstream of dams through changes in water discharge regimes will also affect hydrological conditions surrounding the dams or upstream of them, but often in an opposite direction. For example, during drought, more water may be discharged from dams, benefiting those needing additional water downstream. However, this action can result in a lowering of the water table in areas surrounding or upstream of the dam. Providing more water to areas downstream of the dam can exacerbate the negative effects of drought in areas surrounding or upstream of the dam.

However, these mechanisms may not always significantly affect household incomes. Consequently, whether large irrigation dams have significant effects on household incomes across hydrological basins is ultimately an empirical question. Farmers, for example, may have alternative ways to insure themselves against drought, such as through livestock rearing. Many farmers also may be growing traditional crops or varieties that may be low yielding but often resistant to drought, so do not realize significant benefits from the large irrigation dams. Alternatively, hydrological conditions in certain areas of Nigeria may be such that dams simply do not stabilize the water supply in intended locations. Our study is motivated to provide some insights into this empirical question.

### 3. EMPIRICAL FRAMEWORK

The general structure of our empirical model is,

$$Y = f(B, S|X) + \varepsilon \tag{1}$$

where a set of outcome variables of interest ( $Y$ ) are a function of the type of hydrological basin in which a study household is resident ( $B$ ) and drought shock ( $S$ ), given key exogenous household characteristics ( $X$ ) and idiosyncratic error ( $\varepsilon$ ).

Attributing variations in key outcomes,  $Y$ , for a household to the type of hydrological basin,  $B$ , in which a household is located raises several methodological challenges.<sup>2</sup> First, households compared across different types of basins may have inherently different characteristics, so that differences observed in the outcomes of interest cannot be readily attributed to the basin type. This can be partly mitigated by focusing on first-differenced outcome variables (such as changes in an outcome of interest between particular time periods), rather than on levels. First differencing minimizes bias due to unobserved time-invariant household fixed effects, which may be correlated with  $S$  or  $X$ . We exploit the within-wave panel structure of the Nigeria LSMS data to construct first-differenced outcome variables. Comparing the difference in these first-differenced outcome variables across  $B$  allows us to employ difference-in-differences (DID) impact evaluation methods.

Second, even after unobserved household-specific characteristics are eliminated through first differencing, variations in the first difference across basin types may be still caused by other factors that may be correlated with the basin types. This can happen if, for example, the first-differenced idiosyncratic error ( $\Delta\varepsilon_{it}$ ) is still correlated with  $B$ . In such

<sup>2</sup>Since our study focuses on short-term effects, we do not worry about endogeneity associated with the constructions of dams per se, a concern highlighted by Duflo and Pande (2007), Sarsons (2015), and Strobi and Strobi (2011). However, different types of endogeneity remain, as described in this section.

case, first differencing is insufficient for effectively addressing the endogeneity of  $B$ . We still need to compare households in different hydrological basins with similar characteristics.

In order to address the second challenge, we both select and stratify our analytical sample of study households and apply a multiple-treatment inverse probability weighting method (MIPW) to generate comparable samples across basins. In the impact evaluation literature, inverse probability weighting is often used to improve the balancing properties of samples across treatment levels (Wooldridge 2007).<sup>3</sup> In the context of our study on the household-level impact of irrigation dams in Nigeria, MIPW attaches greater analytical weights to households in each type of basin whose characteristics are such that they are less likely to be found in the particular basin. The logic behind the use of inverse probability weighting is that the outcomes or behaviors of households in downstream basins, for example, whose particular characteristics are shown to not be closely associated with being a resident in downstream basins (and thus more likely to be in other basins), provide more useful information about how households in other basins benefit from switching to downstream basins than do other more typical household residents in downstream basins. This is because there are more households with such characteristics in other basins, for whom this information is relevant. Conversely, typical households in downstream basins, who characteristically are more likely to be in downstream basins and less likely to be in other basin types, offer relatively less useful information about the household-level benefits of switching their residence from other basin types to downstream basins. This is because there are few households with such characteristics in the other basin types in the first place.

We call our specification DID-MIPW, as it combines both DID and MIPW. Appendix A describes in detail the estimation procedure of MIPW.

## 4. DATA AND DESCRIPTIVE STATISTICS

Household-level outcome variables and household characteristics for this analysis were largely obtained from the Nigeria LSMS data. These data have been collected jointly by the National Bureau of Statistics of Nigeria (NBS) and the World Bank and are nationally representative. The 5,000 household sample was selected through multiple-stage stratified random sampling based on NBS-defined enumeration areas (EA). Approximately 3,000 of the sample households are farm households. Data were collected in two waves, 2010/11 and 2012/13. In each wave, the same households were interviewed twice, once in the post-planting season and once in the post-harvesting season. For the analysis here, the LSMS data were combined with data on hydrological basins, dams, and drought.

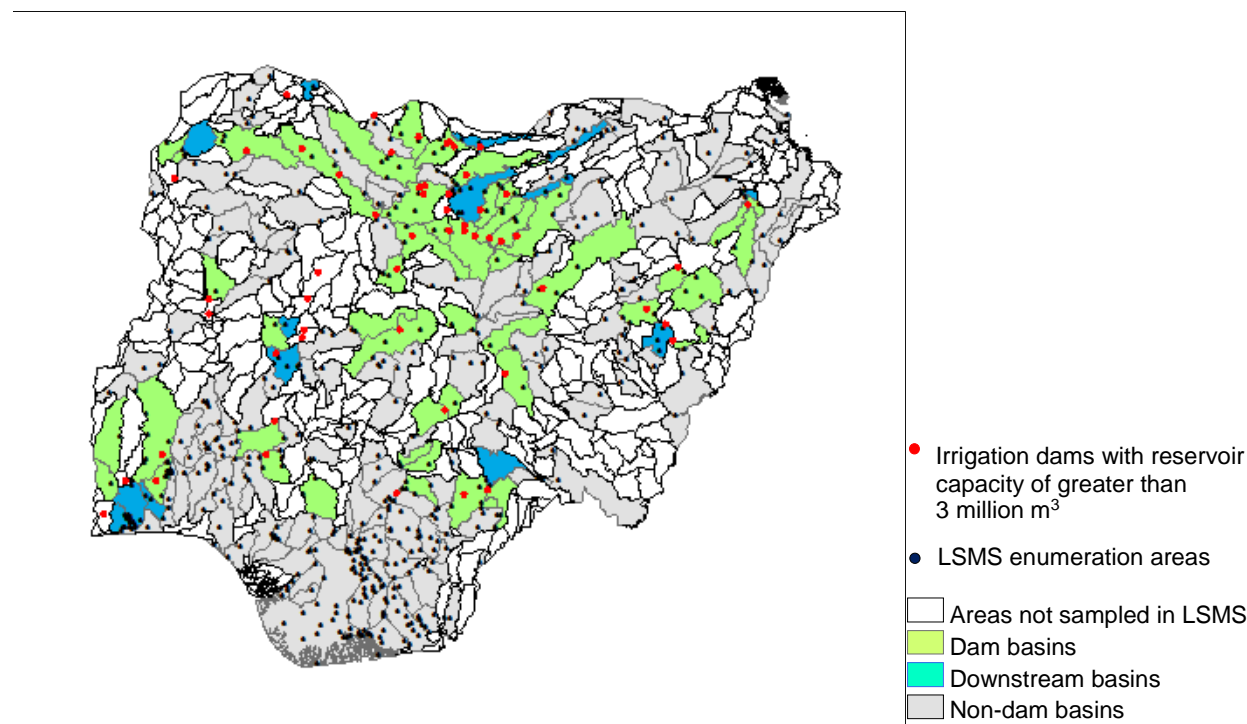
The hydrological basins in Nigeria used in this analysis were extracted from estimations by Masutomi et al. (2009), which improves on the popularly used Hydro1k data set of the United States Geological Survey. Nigeria is estimated to have 565 hydrological basins with an average size of approximately 1,600km<sup>2</sup>, or 40km by 40km.<sup>4</sup> Out of approximately 200 dams in Nigeria, 57 of them are both large with a reservoir capacity greater than 3 million m<sup>3</sup> and were constructed for irrigation purposes, among others (FMWR 2007; FAO 2015). Using the locations of these dams, as well as information provided in Masutomi et al. (2009), we classify each hydrological basin into one of three categories – (1) dam basins, which are areas supplying water into the reservoir of a large irrigation dam; (2) downstream basins, which are downstream from at least one large irrigation dam; and (3) non-dam basins which are areas not associated with any irrigation dam.

---

<sup>3</sup>IPW relies on the ignorability assumption, in which the probability of being found in particular types of basins and the effect of switching the type of basins is independent, given the covariates. As is described in the empirical section, this restriction is somewhat weakened when we focus on average treatment effects on the treated (ATT) rather than average treatment effects (ATE).

<sup>4</sup>While the average size of hydrological basins in Nigeria seem large, these are smaller than those used in an India study by Duflo and Pande (2007). Moreover, some studies suggest that dams can affect hydrological conditions across basins. For example, Thompson and Polet (2000) suggest that, in part due to upstream water development, the maximum extent of flooding in the Hadejia-Jama'are area in Nigeria has declined by about 60 percent since the 1960s.

**Figure 1—Classification of hydrological basins in Nigeria based on large irrigation dams**



Based on this spatial division of the country into hydrological basins, we then determine in what sort of basin each of the EAs use for selecting the LSMS sample households was found. LSMS sub-samples were created based on whether the sample household was resident in a dam basin, a downstream basin, or a non-dam basin. Figure 1 illustrates the locations of the different types of basins, large irrigation dams, and the LSMS EAs across Nigeria.

A drought index is based on soil moisture observations and ranges between 0 and 100, with 0 indicating the most severe drought and 100 indicating no drought. Drought indices for Nigeria were obtained from the African Flood and Drought Monitor (Sheffield et al., 2014; AghaKouchak 2015). From the daily drought index values extracted for each LSMS EA, we calculated average drought index between June and October, the period which typically covers the major cropping season in the North Central, North West and North East geo-political zones of Nigeria.

Based on the drought index, we further stratify the LSMS sample into two groups, a mild or no-drought group and a severe drought group. We use the median value of the index (52.4 in our combined samples from 2010 and 2012 LSMS) as a threshold.<sup>5</sup> We then conduct analyses separately for these groups. Past studies used a similar stratification approach based on weather and climatic variables. For example, Takeshima and Yamauchi (2012) stratified Nigeria into several analytical categories based on historical variation in annual rainfall.

## 4.1 Analytical Variables

**Outcome variables.** Our outcome variables of interests are growth rates between the post-planting and the post-harvest season in per capita household income, the value of per capita household food consumption, and the value of per capita household food consumption that comes from own production. All variables are measured in monetary terms, with nominal values being converted to real values using a price index based on average local prices reported in the post-planting season for rice, maize and sorghum, the three main staple food crops in Nigeria. This index is used to partly control for spatial variations in prices.

Household income is proxied by total household expenditure, which in developing countries often provides a more accurate assessment of household income than do direct measures of income (Deaton 1997). This variable is calculated by aggregating the value of food consumption (home consumption converted into expenditure values using market prices, as well as actual eating-out expenditures), which is also our second outcome variable, with expenditure on non-durable consumption goods, education for household members, health expenditures, net purchase of livestock, net purchase of household assets, housing expenses including utilities (water, electricity, fuels, land and mobile phones, refuse disposal, and rent payments), net cash lending, net purchase of agricultural equipment, net of other unearned income, and remittances received. Consumption and expenditure figures in each category are reported in the LSMS over

<sup>5</sup>We use the median among downstream basins, in order to make sure there are enough samples from downstream basins in each group (because downstream basin sample is the smallest in the data).

different time-periods. We therefore converted each component of household consumption and expenditure to its 12-month equivalent.

The third variable, the value of per capita household food consumption from own production, is used to proxy for food crop production. We measure food crop production through consumption figures rather than direct production, because in sub-Saharan Africa countries like Nigeria home-consumed food crops tend to be harvested over a long period. Consequently, reported production quantities may not accurately measure the total quantities of food crops produced (Deininger et al. 2012).

**Explanatory variables used in the first stage multinomial logit of DID-MIPW.** In quantitative impact evaluation research approaches that use methods that involve computing the probability of belonging to particular treatment regime (this includes propensity score matching approaches), covariates  $X$  are key socio-economic or agro-ecological characteristics of households that are associated with the probability of the household being resident in a certain type of hydrological basin  $B$ .

Nevertheless, as far as estimating the effect of a household switching residence from one type of hydrological basin to another is concerned, our main interest, our approach works exactly in the same way as in typical impact evaluation studies by constructing suitable counterfactuals for observed outcomes. The set of variables  $X$  are therefore selected in the same way as in typical impact evaluation case. The number of variables, however, is limited to avoid the curse-of-dimensionality problem that can often lead to loss of efficiency (Zhao 2008), particularly with a small sample size, as is the case in our study.

The set of explanatory variables  $X$  used in our study consist of (1) agroecological conditions; (2) household demographics and human capital; (3) household wealth; (4) access to infrastructure and relevant public services; and (5) LSMS data wave dummy variables. The agro-ecological variables used include the Euclidean distance to the nearest major river, expressed as the natural log of geographical arc-minutes (Lehner, Verdin, and Jarvis 2006); the aforementioned drought index; the size of the hydrological basin in which a household is resident expressed as the natural log of square kilometers (Masutomi et al. 2009); the historical average annual rainfall in millimeters; and the elevation above sea level of the residence of the household expressed as the natural log of meters of elevation (USGS 1996).

We also included among our agro-ecological variables the share of areas in the hydrological basin with gradients between 1.5 and 3.0 percent and 6.0 percent or above. Areas with gradients between 1.5 and 3.0 percent are most suited for building irrigation dams, as the land is not entirely flat so that water replenishes the dams at a sufficiently fast pace, but also the land is not too steep (not more than 3.0 percent) so that the reservoir area is sufficiently large relative to the dam height, which is ideal from the perspectives of economic storage capacity (Duflo and Pande 2007). In contrast, areas with sufficiently steep gradients (6.0 percent or above) enjoy the lower costs of producing hydroelectricity, so are suited for constructing dams primarily for hydroelectricity. These slope conditions are often suitable for using dams for both irrigation and hydroelectricity purposes (Duflo and Pande 2007; Sarsons 2015). While our focus is on irrigation, suitability for hydroelectricity in the basins may also affect the decisions to construct irrigation dams within the same basins (as dam constructions often involve significant local resources, and often force evictions of residents). A slope variable for our analyses was constructed using the GTOPO30 global digital elevation model (USGS 1996). This variable is used as an instrumental variable for dam construction in other studies. We include this variable in  $X$  so that the estimates can be interpreted more clearly as the effect of dams in areas that are similarly suitable for dam construction.

The household demographic and human capital explanatory variables used include the age of the household head in years, household size expressed in natural log form, and the percentage share of working-age members with formal education. The wealth-related variables include farm size expressed as the natural log of square meters, a binary variable indicating whether the household has off-farm income sources (1 = yes), and the value of per capita non-agricultural household assets expressed as the natural log of their values. Included in the explanatory variables related to access to infrastructure and relevant public services are distance to the nearest major road expressed as the natural log of kilometers; distance to the nearest population center expressed as the natural log of kilometers; and a binary variable indicating whether the household has access to electricity (1 = yes). Access to electricity is included because hydroelectricity provided by large irrigation dams is often a primary source of electricity (Barrios, Bertinelli and Strobl 2010).

## 4.2 Analytical Samples

We focus on certain sub-sets of our sample households that are homogeneous in terms of some key analytical variables in order to improve the balancing properties of the samples across basins. Specifically, we reduce the size of our primary analytical sample of study households from the LSMS based on Nigeria's six geopolitical zones and specific states within those zones, some exogenous household characteristics, and local soil characteristics.

We focus our analysis on the North Central, North West, and North East geopolitical zones of Nigeria, excluding the South East, South South, and South West zones. This is because the majority of large irrigation dams in Nigeria are

located in the Northern zones, and, in consequence, it is difficult to find sufficient dam basins and downstream basins in the southern zones. Moreover, there are sharp differences in household characteristics between the southern and northern zones. In addition, our analytical samples are confined to the following nine states all of which have at least some observations in downstream basins; Adamawa, Bauchi, Benue, Borno, Jigawa, Kano, Kebbi, Niger, and Sokoto. Limiting the analysis to these states was found to substantially improve the balancing properties of the sub-samples in our analysis. This potentially is because the presence of downstream basins within state boundaries may have resulted in a more similar set of agricultural policies and programs in these states than is the case in states without downstream basins.

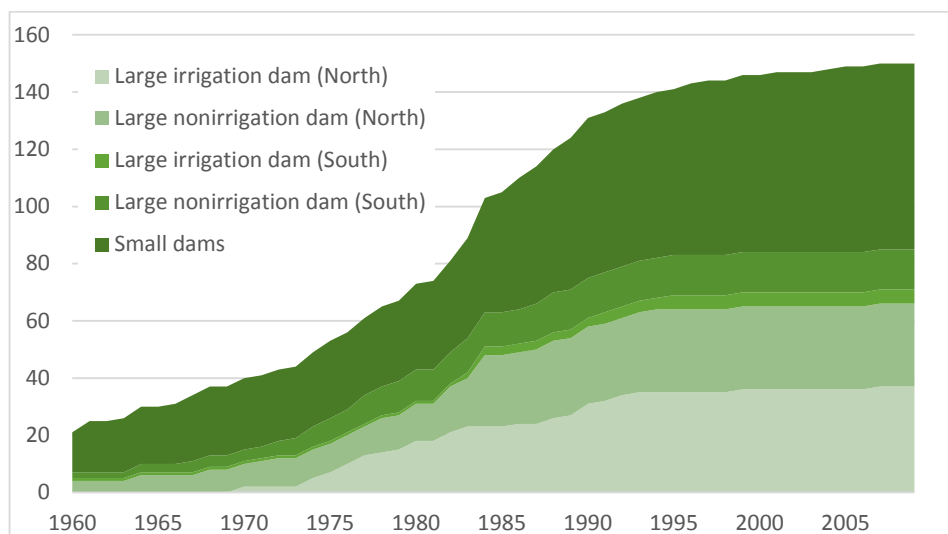
We further limited our samples based on the following criteria for inclusion:

- (a) farm households cultivating at least one farm plot;
- (b) households with at least one working-age member with at least primary education completed;
- (c) households without members employed in the government sector;
- (d) EAs in which at least one sample household does not have access to electricity;
- (e) for the mild or no-drought sub-sample, EAs with historical average annual rainfall of greater than 550 mm;
- (f) areas with elevation no greater than 500 m above sea level; and
- (g) hydrological basins with areas no greater than 4915 km<sup>2</sup> (8.5 in natural log).

Criterion (a) ensures that we are not comparing farm households with households specializing in non-farm activities, the characteristics are which are likely to be very different. Criterion (b) ensures that we are not comparing households with very different educational backgrounds. Criterion (c) excludes households who may be different from other households in our sample in terms of access to various government programs. Criterion (d) excludes areas with high accessibility to electricity, as this is likely to differ across the types of basins given the role of dams in hydroelectricity supply. Criterion (e) excludes arid areas where household characteristics are found to be systematically different from the rest. Criterion (f) ensures that any effects due to differences in elevations across basins are minimized, which is particularly important given that downstream basins are often located at lower elevations than dam basins. Finally, criterion (g) tries to minimize effects on households that may be due to differences in basin sizes.

Soil type, which often significantly affect agricultural productivity and local economies based on farming, can vary considerably across different types of hydrological basins. Therefore the analyses focus on areas with similar soil types so that differences in outcomes are not driven by differences in major soil types. We limit our analysis to the areas in which six types of soils predominate, namely Arenosols, Fluvisols, Gleysols, Leptosols, Lixisols, and Nitisols, and where at least some observations are found in downstream basins. The analysis can be complicated if soil types themselves are affected by dam constructions, which can happen if dams affect inter-regional soil movement, particularly the alluvial movement of soils. We therefore rely on the Harmonized World Soil Map (FAO et al. 2012), which is based largely on soil surveys conducted in the 1960s and 1970s (FAO/UNESCO 1977). As is shown in Figure 2, a majority of the large irrigation dams in northern Nigeria were built after the mid-1970s, while earlier dams were mostly built for hydroelectricity (Akanmu, Remi-John, and Ekpo 2011). Therefore, one can use the FAO et al. (2012) soil map assuming that local soil conditions were mostly exogenous to the decisions to construct a dam.

**Figure 2—Number of different types of dams completed in Nigeria, 1960 to 2010, cumulative**



Source: Authors' compilation based on FAO (2015) and FMWR (2007).

Note: The figure is based only on those dams for which the year of completion was known.

This process of defining quite specific sub-samples to improve the matching quality of the final analytical samples considerably reduces the final sample size. The total sample size of 733 is approximately 20 percent of the original sample of farm households in the North Central, North West, and North East geopolitical zones in the two waves of the LSMS combined. However, this small sample size is of less concern because our analyses essentially focus on averages across hydrological basins, so that the statistical degrees of freedom for each hydrological basin is the sample size in each minus only one. In addition, our analyses are still meaningful. Because the number of large irrigation dams is relatively large in Nigeria compared to other African countries, and LSMS samples of households are located across regions with diverse agroecological and socioeconomic characteristics, we can create counterfactual samples of workable size. Similar analyses are likely to be more difficult in other sub-Saharan Africa countries, where the number of large irrigation dams is often too small.

### 4.3 Descriptive statistics

Descriptive statistics (sample averages) of the explanatory variables used in the analysis for each type of basin are shown in Table 1 for the mild or no-drought sub-sample and Table 2 for the high drought sub-sample. We show both unweighted raw statistics, as well as inverse-probability weighting statistics in which samples in dam basin and non-dam basin are weighted accordingly so that sample distributions in these basins are similar to the sample distribution in downstream basins. The inverse-probability weighting statistics are based on the probability estimated through a multinomial logit model which is discussed in the next section.

**Table 1—Descriptive statistics for mild or no-drought sub-sample of households**

Categories	Unweighted				Inverse probability weighted			
	Dam	Down-stream	Non-dam	F-test	Dam	Down-stream	Non-dam	F-test
<b>Sample size</b>	51	89	83		51	89	83	
Euclidean distance to nearest major river, ln(arc-minutes)	-4.190	-4.343	-4.345		-4.215	-4.343	-4.432	
Drought index, 0 = severe; 100 = no drought	37.271	41.502	41.146	***	36.240	41.502	42.087	**
Size of hydrological basin, ln(km <sup>2</sup> )	7.699	7.663	7.473	**	7.617	7.663	7.453	
Historical average annual rainfall, mm	1310.6	1255.2	1095.9	***	1230.5	1255.2	1134.0	
Elevation, ln(m above sea level)	5.448	5.274	5.270	***	5.380	5.274	5.228	
Natural log of share of areas with suitable gradient, ln(percent)	-0.0283	.407	-.494	*	-.563	.407	.628	
Age of household head, years	46.8	51.2	44.6	***	51.9	51.2	53.8	
Household size, ln(number of members)	1.745	1.829	1.939		1.771	1.829	1.929	
Share of working-age household members with formal education	.657	.768	.629	***	.759	.768	.726	
Farm size, ln(m <sup>2</sup> )	9.094	9.270	9.506		9.239	9.270	9.091	
Have off-farm income sources, 0/1	.569	.382	.566	**	.360	.382	.426	
Natural log of non-agricultural asset value, ln(equivalent of kg of cereal, per capita)	4.060	3.855	4.270		3.793	3.855	3.834	
Distance to the nearest major road, ln(km)	1.707	1.169	1.912	**	1.075	1.169	.952	
Distance to the nearest population center, ln(km)	3.017	3.422	3.378	***	3.324	3.422	3.256	
Access to electricity, 0/1	.078	.090	.325	***	.115	.090	.097	
2012 LSMS wave dummy variable, 0/1	.569	.494	.458		.554	.494	.499	

Source: Authors calculation based on multinomial logit analysis of Nigeria LSMS data. Asterisks indicate the joint difference in means across basin types at given statistical significance: \*\*\* 1%, \*\* 5%, and\* 10%.

**Table 2—Descriptive statistics for high drought sub-sample of households**

Categories	Unweighted				Weighted			
	Dam	Down-stream	Non-dam	F-test	Down-stream	dam	Non-dam	F-test
<b>Sample size</b>	<b>236</b>	<b>81</b>	<b>193</b>		<b>236</b>	<b>81</b>	<b>193</b>	
Euclidean distance to nearest major river, ln(arc-minutes)	-4.410	-4.173	-4.235	***	-4.272	-4.173	-4.459	
Drought index, 0 = severe; 100 = no drought	70.043	68.457	67.949	*	68.498	68.457	66.424	
Size of hydrological basin, ln(km <sup>2</sup> )	7.988	7.462	7.892	***	7.466	7.462	7.488	
Historical average annual rainfall, mm	814.2	728.9	915.0	***	739.3	728.9	727.5	
Elevation, ln(m above sea level)	6.024	5.922	6.014	**	5.953	5.922	5.823	
Natural log of share of areas with suitable gradient, ln(percent)	4.745	4.768	4.754		4.761	4.768	4.771	
Age of household head, years	49.2	43.1	47.7	***	45.4	43.1	44.4	
Household size, ln(number of members)	1.903	1.833	2.028	***	1.823	1.833	1.861	
Share of working-age household members with formal education	.776	.848	.735	***	.882	.848	.817	
Farm size, ln(m <sup>2</sup> )	9.192	9.106	9.374	*	9.112	9.106	9.102	
Have off-farm income sources, 0/1	.682	.716	.819	***	.632	.716	.769	
Natural log of non-agricultural asset value, ln(equivalent of kg of cereal, per capita)	3.859	4.262	4.492	***	4.166	4.262	4.024	
Distance to the nearest major road, ln(km)	2.238	2.508	1.792	***	2.588	2.508	2.139	*
Distance to the nearest population center, ln(km)	2.996	3.003	3.102		2.835	3.003	3.002	
Access to electricity, 0/1	.144	.111	.212	*	.121	.111	.177	
2012 LSMS wave dummy variable, 0/1	.538	.593	.606		.569	.593	.700	

Source: Authors calculation based on multinomial logit analysis of Nigeria LSMS data. Asterisks indicate the joint difference in means across basin types at statistical significance: \*\*\* 1%, \*\* 5%, and \* 10%.

One way of checking the balancing properties of the samples is by examining whether the means of covariates are significantly different statistically across groups with different treatment status (Cavattasi et al. 2011; Takeshima 2016a). Since our analysis deals with three treatment groups, we conduct an *F*-test. The unweighted columns in Table 1 and Table 2 suggest that there are significance differences in the raw sample household characteristics across different types of basins, since the sample means are statistically significantly and jointly different across basins types for a majority of variables. However, in the inverse-probability weighting adjusted samples, the differences for most variables become statistically insignificant. In each of the tables, only one out of 16 variables in each sub-sample still exhibits statistically significant differences when weighted, which is expected under the null hypotheses that these samples are balanced across basin types. These conditions ensure that differences identified through inverse-probability weighting in outcome variables are in fact due to the change in the types of basins, rather than any other confounding factors.

The lack of evidence of the violation of balancing properties may be partly due to the small sizes of the sub-samples used in our analysis. However, small sample size also lowers the power of the test to detect the significant effect. Nevertheless, as is shown in the next section, we find statistically significant effects on outcomes. This indicates that the estimated effects are significantly different from idiosyncratic errors arising due to potential imbalances between the sub-samples.

## 5. RESULTS

### 5.1 Main findings

Table 3 presents the effects on the household welfare outcomes of switching residence across the different types of hydrological basins estimated using the Difference-in-Differences - Multiple-treatment Inverse Probability Weighting (DID-MIPW) analytical method.

**Table 3—Effects of switching household residence to a different type of hydrological basins, estimated as average effect of treatment on the treated**

Outcomes	Drought-defined sub-sample	Downstream vs. Dam basins	Downstream vs. Non-dam basins	Dam vs. Non-dam basins
Growth rate in real per capita income	mild or no drought	-.181** (.079)	-.066 (.069)	.088 (.089)
	severe drought	.213* (.126)	.113 (.141)	-.101 (.169)
Growth rate in real per capita food consumption	mild or no drought	-.247*** (.084)	-.132 (.081)	.093 (.072)
	severe drought	.311* (.147)	.191 (.177)	-.120 (.204)
Change in real per capita consumption of food from own production over 6 months, in kg of staple foods	mild or no drought	-2.331 (16.047)	28.636 (25.815)	30.996 (24.838)
	severe drought	65.977** (28.985)	128.562* (77.026)	62.584 (79.499)

Source: Authors' estimations.

Note: Treated = Downstream basin; Control = Dam basin. Numbers in parentheses are standard errors based on Huber/White/Robust sandwich estimators.

In terms of the growth rates of both real per capita income and real per capita food consumption, households in downstream basins have benefited from greater growth (or less negative growth) where drought was severe, compared to what they would have experienced had they been resident in a dam basin, by 21.3 and 31.1 percentage points higher growth rates, respectively. However, these effects are reversed where the level of drought was mild or did not occur. In areas with severe drought, the greater effect of downstream basin residence may be partly due to greater food crop production. In terms of real per capita food consumption, downstream basin households in areas with severe drought produced 66 kg more food per capita over a six months period from their own agricultural production than what households in dam basins produced.

The negative effects of being in downstream basins when droughts were mild or did not occur are somewhat surprising. However, these results make better sense if we also assess the effect of drought on downstream-basin households. Table 4 presents the results of another set of Inverse Probability Weighting estimates conducted among samples within downstream basins, where treatment and control are now mild or no-drought and severe drought, respectively.<sup>6</sup> These results suggest that in the downstream basin households the effects of drought on the growth rates of real per capita income and food consumption were insignificant. These results, when combined with results in Table 3, weakly suggest that the downstream basin households are less affected by drought than households in other types of hydrological basins.

**Table 4—Effects of drought on growth rate of real per capita income and of food consumption among comparable drought-defined sub-samples within downstream basins**

Outcomes	Growth rate of real per capita income	Growth rate of real per capita food consumption
Estimated effect and standard error	.247 (.167)	.322 (.211)

Note: Treated = Mild or no drought; Control = Severe drought. The sample is limited to the households resident at an elevation below 500m above sea level. Numbers in parentheses are Huber/White/Robust sandwich standard errors.

A caution is needed when interpreting these results. The estimates are average treatment effects on the treated (ATT) rather than average treatment effects (ATE). Here households in downstream basins are considered “treated” households, so that estimates are relevant only to the treated households. ATT is often preferred to ATE because it relies on less restrictive assumptions than ATE (see Appendix A for a detailed discussion). This advantage of ATT, however, limits the scope of interpretation. For example, while we can interpret that households in the downstream basins experienced faster growth rates in these outcome measures compared to what they would have experienced if they were in dam basin, we cannot assess whether or not a dam basin household experienced slower growth rates in the outcomes of interest compared to what they would have experienced if they were resident in a downstream basin.

As was described above, the multinomial logit analysis in our case is not a structural equation, and we cannot make any causal interpretations. Therefore, we only show the multinomial logit results in Appendix B (Table 6). The significance and magnitudes of estimated coefficients from this analysis weakly indicate how each variable affects the

<sup>6</sup>Sample characteristics differ slightly across drought levels, and some treatment samples had to be dropped because their probability of treatment was too close to 1, violating the required conditions for an Inverse Probability Weighting methodological approach.

weights attached to each observation. For example, households located closer to a major river have higher probabilities of belonging to dam basins or non-dam basins, compared to downstream basins. If these households are actually in dam basins or non-dam basins, their outcomes receive lower weights than those located further away from the nearest major rivers in the estimation of ATT for switching their residence to downstream basins from dam basins or non-dam basins. As several variables are statistically significant, our MIPW approach significantly improves the balancing properties across many dimensions of household characteristics.

## 5.2 Robustness checks

A series of modifications were made to the analysis to assess the robustness of the results obtained through the preferred analytical approach described above.

**Different patterns of drought in the hydrological basins.** Strobi and Strobi (2011) suggests that, in addition to drought conditions in downstream basins, drought conditions in dam basins can also affect downstream basins, particularly if drought conditions in dam basins and downstream basins are quite different. In our case, however, the drought conditions are generally highly correlated between dam basins and downstream basins with correlation coefficients of around 0.8. Therefore, this factor is not seen to significantly affect our results.

**Role of political factors.** Water supply to the downstream basin can be affected by political factors, such as when a river crosses national borders (Strobi and Strobi 2011). In Nigeria, we consider similar situations in which water supply to the downstream basin is affected if the river crosses state borders. Specifically, following Strobi and Strobi (2011), we calculate for each basin the share of upstream dams that are beyond state borders, and add them as additional variables. We find the results obtained from doing so are qualitatively similar to the results presented in Table 3 and Table 4.

**Attrition in post-harvesting season.** If attrition rates in the panel survey (non-response rates in post-harvesting season) are high and related to the outcome variables, not incorporating such attrition can lead to biased estimates. However, the extent of attrition in our sample was found to be very small. Therefore, this factor is unlikely to affect our results.

**Assumptions about the effect of the timing of drought.** Our analysis assumes that the effect of drought during the production season is more pronounced in the post-harvesting season, rather than in the post-planting season. Some households can better predict the drought and cope with it by cutting consumption in the post-planting season, in which case the actual drought realization would be correlated with post-planting season consumption. However, the extent of this occurring may be limited due to several reasons. The ability to smooth consumption by reducing consumption during the post-planting season is generally limited because it is often the lean season of the year when consumption levels already are low compared to the post-harvesting season. This is particularly so if households are credit-constrained. Similarly, consumption smoothing through local social network will also be limited under such conditions because weather-related events like drought affect all households in the local community. The direct effect of drought on consumption during the post-planting season is also limited because consumption during this period is mostly from the harvest of the previous year.

In addition, the major staple crops – rice, maize, and sorghum – are harvested starting in October in Northern Nigeria, when most of the post-planting LSMS survey rounds were completed. As seen in Table 5, a majority of rice, maize and sorghum are planted in May or June in Northern Nigeria, which will be harvested after 5 months in October or November. Traditional, long-duration varieties with growing periods of 5 months or more still account for considerable shares of crops planted in northern Nigeria.

**Table 5—Share of sample households planting crops in each month (waves 1 and 2 combined), January to August, percent**

Crop	Month							
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Rice	0	0	1	6	30	41	17	4
Maize	0	0	2	15	31	41	8	2
Sorghum	0	0	1	8	42	39	8	1

Source: Authors.

Nevertheless, we checked the robustness of the results of our analysis by rerunning the model using different reference periods for the drought index variable. Specifically, instead of the average drought index for the period June through October, we tried using the average drought index for September and October. We found that our results were generally robust, not varying significantly with changes in the average drought index reference period.

## 6. CONCLUSIONS

Despite considerable past investment and continued interest in large irrigation dams in Africa, empirical evidence is still scarce regarding their benefits. In particular, relatively little is known about how these dams might affect the effects of weather-risks, like drought, on household consumption in the short-term, and how such effects may differ between households resident in dam basins, downstream basins, or non-dam basins. In sub-Saharan Africa countries like Nigeria, one of the challenges in assessing the short-term household-level effects of irrigation dams across basins is that most of the currently existing large irrigation dams were constructed before extensive household surveys like LSMS were commonly conducted. Consequently, little baseline information exists today that would allow us to compare the welfare levels of households before and after the construction of a dam near their place of residence.

Applying the inverse-probability weighting model to Nigeria LSMS household survey data, we partly overcome this problem. We create analytical samples of household in each of three types of hydrological basins which are similar in terms of key exogenous household characteristics. Such an approach allows us to identify differences in household outcomes that are solely attributable to the differences in the corresponding types of hydrological basin in which a household resides. As these basins are defined according to how they are associated with large irrigation dams, these results provide insights on the household-level effects resulting from the construction of such dams.

We find that, households resident in downstream basins, for whom suitable matches can be found in dam basins and non-dam basins, the other types of basins, are relatively less affected by drought and enjoy relatively stable between-season growth rates in real per capita income and food consumption compared to what they would have experienced had they been resident in either dam basins, which are hydrological basins that contain large irrigation dams and the area upstream of such dams, or non-dam basins, which are not associated with dams. This is likely to be partly enabled by the ability of households resident in downstream hydrological basins to better limit the damage of severe drought on their own food production relative to household resident in the two other types of basins considered.

The findings have important implications on the role of irrigation dams in mitigating the short-term effects of droughts. Our results suggest that irrigation dams in Nigeria do mitigate such effects, but more in downstream basins rather than in the dam basins that feeds the reservoir of the dam. Both in India and in sub-Saharan Africa, large dams have been found to exhibit long-term distributional effects across districts or hydrological basins, with downstream districts or basins being the primary beneficiaries. Our findings in Nigeria show that large irrigation dams in sub-Saharan Africa may also distribute household-level benefits across hydrological basins. These primarily take the form of short-term insurance benefits against drought, with households resident in downstream basins as the primary beneficiaries.

The findings also have important policy implications. Hydrological basins across Nigeria have only been recently characterized relatively accurately, while many of the irrigation dams in place were built before 1990. There had been a limited understanding of which regions would benefit from large irrigation dams. Our analyses show that hydrological basins are one of the important geographical units that can be used to identify areas that are likely to benefit from large irrigation dams. In order to assess the likely welfare impact of dams, even before conducting benefit-cost analyses, first the relevant geographical areas that are likely to be affected by the large irrigation dams need to be identified. This has rarely been done in the past in countries like Nigeria.

Furthermore, our analyses suggest that it is typically the hydrological basins that are located downstream of large irrigation dams that benefit most, rather than the basins that contain the dams. Even though accurate ex-ante benefit-cost assessment of construction or rehabilitation of irrigation dams is challenging, decisions on the feasibility of such investments should be partly guided by information on agricultural productivity and income levels of households in downstream hydrological basins relative to the basins within which construction or rehabilitation of a large irrigation dam is planned.

Of course, there is potentially considerable heterogeneity within hydrological basins, which also would need to be investigated in future studies. However, our analyses at least show that the type of hydrological basins in relation to the location of the dams in which households reside significantly affects their welfare when all other factors are controlled for. Given the methodology with which such effects are estimated, our results suggest that construction of new large irrigation dams in hydrological basins are likely to improve the welfare of households in downstream basins to a greater extent than that of households residing in the basins where dams are built.

While the welfare effects on downstream households that we identified in our analyses are limited to drought risk mitigation, this is still important as drought is likely to have affected household welfare in Nigeria in 2010 and 2012. Our results in Tables 3 and 4 suggest that, where drought was severe (average drought index of less than 50 during the production season), households residing in downstream basins experienced about a 20 percentage point greater expenditure growth between post-planting and post-harvesting season compared to households with similar characteristics in dam-basins. Where drought was mild or absent, households in downstream basins experienced about 20 percentage point lower expenditure growths than households in dam-basins. Households in downstream basins were also unaffected

by drought (Table 4). This implies that households in dam-basins with similar characteristics to households in downstream basins, might have experienced almost 40 percentage point lower expenditure growths between post-planting and post-harvesting if the drought was severe than when drought was not severe. Drought is likely to have significantly and adversely affected the welfare of household resident in either dam-basins or non-dam basins relative to what was experienced by households in downstream basins.

Based on the definition of drought severity used in our analyses, close to 40 percent of farm households in Nigeria experienced severe drought in 2010 and 2012. Drought risk is substantial as correlation in household drought indices between the two years was generally low (correlation coefficient of around 0.20). About 30 percent of regions with mild or no drought in 2010 experienced severe drought in 2012, while about half of the regions with severe drought in 2010 experienced mild or no drought in 2012. With such yearly fluctuations in severity, drought is likely to be a significant risk factor for the often risk-averse farmers of northern Nigeria.

Lastly, our analyses cannot directly inform whether dams should be built in particular locations or if the benefits justify the costs of rehabilitating an existing dam. However, several studies suggest that large irrigation dams in Nigeria may be still fewer than is required, at least relative to what is observed in many other countries, given the relatively large size of Nigeria's arable land. For example, based on Tharme (2003, Table 1), the number of large dams for both irrigation and other purposes is substantially higher in several other countries – compared to the approximately 80 such dams in Nigeria, there are more than 20,000 in China, more than 6,500 in the USA, more than 4,200 in India, 625 in Turkey, and 530 in both South Africa and Mexico. In several of these countries, the arable land is generally smaller than what is found in Nigeria. Therefore, returns are likely to be high for investing in research to identify areas of Nigeria where new irrigation dams should be built, so that the greatest benefit for household welfare can be extracted, also taking into account the distributions of such benefits across hydrological basins. Growing uncertainty in the frequency and severity of drought due to climate change also suggests that returns from large irrigation dams may be rising in countries like Nigeria.

## APPENDICES

### Appendix A: Estimation of average treatment effects on the treated (ATT) through the multiple-treatment inverse probability weighting method

The multiple-treatment inverse probability weighting method (MIPW) is estimated through an exactly identified generalized method of moments (GMM) approach based on a set of equations of moment conditions. (The exposition here draws largely on Cattaneo, Drukker, and Holland (2013) and STATA manual for command `teffects`).  $i$  denotes the observations, and subscript  $i$  denote values for each  $i$  of corresponding variables. The first set of equations are score functions  $\psi_{ML,i}(\cdot)$  associated with the multinomial logit model,

$$\psi_{ML,i} = \psi_{ML,i}(B_i, S_i, X_i, \hat{\gamma}) \quad (2)$$

in which  $\hat{\gamma}$  are a set of estimated parameters.

The second set of equations are Weighted Least Squares,

$$\psi_{WLS,i}(\hat{y}_{B,i}, w_i, I_{B,i}, \hat{\beta}) = w_i(\hat{y}_{B,i} - I_i \hat{\beta}) I_i' \quad (3)$$

in which  $\hat{y}_{B,i}$  is the observed outcome,  $I_{B,i}$  is the indicator variable which takes the value of 1 if the household  $i$  belongs to basin  $B$ , and 0 otherwise.  $I_i$  is  $i$ -th row of a matrix consisting of  $I_{B,i}$ , with control group indicator variable replaced by 1 for all observation.  $I_i'$  is the transpose of  $I_i$ .  $\hat{\beta}$  are the vector of parameters estimating the ATE or ATT of each  $B$  relative to the control regime, and  $w_i$  is the weight that is a function of the probability of belonging to particular basins ( $p_{B,i}$ ) estimated through the multinomial logit model.

The set of equations associated with the multinomial logit model and those of the Weighted Least Squares set of equations are solved by GMM using the moment conditions  $E[\psi_{ML,i}(\cdot)] = 0$  (a property of score functions), and  $E[\psi_{WLS,i}(\cdot)] = 0$ , which are expectations over  $i$ . As is seen,  $\hat{\gamma}$  and  $\hat{\beta}$  affect each other through  $w_i$ .

We consider downstream basins as the treatment. Then, our  $w_i$  is the treatment-adjusted inverse-probability

$$w_i = \frac{p_{\text{downstream},i}}{\hat{p}_{B,i}} \quad (4)$$

in which  $\hat{p}_{B,i}$  is the probability that household  $i$  belongs to the observed  $D_i$ . Our  $w_i$  (ATT weights) differs from weights for the ATE (ATE weights) which use  $w_i = 1/\hat{p}_{B,i}$ . Multiplying ATE weights by  $p_{\text{downstream},i}$  creates a weighted sample of untreated units that has the same covariate distribution as that in the treated group (Pan 2015). Modifying the weights in this way affects the solution of weighted GMM as above, often leading to different ATT than ATE.

We focus on ATT instead of ATE as ATT generally relies on fewer assumptions (Caliendo and Hujer 2006). For example, under the standard one-treatment case, the unconfoundedness assumption for ATE requires that the hypothetical outcomes in each  $B$  ( $y_B^*$ ) satisfy  $y_B^* \perp B \mid X, S$  for all  $B$ , but the unconfoundedness assumption for ATT does not require  $y_B^* \perp B \mid X, S$  for treatment  $B$  (downstream basin). Similarly, for overlap properties, while ATE requires  $0 < \hat{p}_{B,i} < 1$  for all  $B$ , ATT does not require  $p_{\text{downstream},i} > 0$  for downstream-basin samples. Overlap properties for ATT are less restrictive than for ATE because, as is indicated by ATT weights, the value of  $p_{\text{downstream}}$  close to zero does not inflate the weights as it does for ATE weights, avoiding the “identified-at-infinity” problem raised by Khan and Tamer (2010). While the literature does not explicitly discuss the advantage of ATT over ATE in multi-valued treatments case, ATT is likely to be no more restrictive and can be less restrictive than ATE in general.

Our empirical analysis uses a multinomial logit approach to jointly estimate the propensity of households of particular characteristics being found in each of three types of basins. This is more efficient than estimating the pairwise propensity scores from only two types of basins, as multinomial logit approaches use information from the samples in all three types of basins. It also ensures consistency across the three pairs – difference between downstream basins and dam basins, between downstream-basins and non-dam basins, and between dam basins and non-dam basins.

The efficiency of the estimates can be described by the following example. If  $p_{\text{nondam}}$  is high,  $p_{\text{dam}}$  and  $p_{\text{downstream}}$  are lower, their outcomes in dam basins and downstream basins are weighted more if weighted by the inverse of  $p_{\text{dam}}$  and  $p_{\text{downstream}}$ . Conversely, if  $p_{\text{nondam}}$  is low and  $p_{\text{dam}}$  and  $p_{\text{downstream}}$  are higher, their outcomes in dam basins and downstream basins are weighted less. Since in MIPW with three regimes, both the information of difference between a particular pair of basins as well as information on their respective differences with the third type of basin are utilized to obtain the estimates. In the case of high  $p_{\text{nondam}}$ , that information considerably influences the estimation of differences between dam basins and downstream basins.

## Appendix B: Multinomial logit model results

Table 6—Estimated coefficients of multinomial logit model (base outcome = downstream basin)

Variables	Dependent variable = 1 if the household is in particular type of hydrological basin			
	High drought index		Low drought index	
	Dam basin	Non-dam basin	Dam basin	Non-dam basin
Euclidean distance to nearest major river, ln(arc-minutes)	-.839***	-.565**	-.432	-.245
Drought index, 0 = severe; 100 = no drought	.036**	.014	-.118***	-.003
Size of hydrological basin, ln(km <sup>2</sup> )	.624***	.075	-.260	-.688*
Historical average annual rainfall, mm	.004***	.007***	.005***	.002*
Elevation, ln(m above sea level)	1.731***	1.743***	2.258**	-.627
Share of areas with suitable gradient, percent	-.438	-.192	.107	-.122
<b>Demography and human capital</b>				
Age of household head, years	.027**	.020*	-.028	-.047***
Household size, ln(number of members)	.074	.549	-.831*	.093
Share of working-age household members with formal education	-2.267***	2.626***	-3.461***	-2.174***
<b>Wealth</b>				
Farm size, ln(m <sup>2</sup> )	.071	.179	-.145	.168
Have off-farm income sources, 0/1	-.237	.253	1.503***	.343
Natural log of non-agricultural asset value, ln(equivalent of kg of cereal, per capita)	-.401***	.024	.222	.191
<b>Access to infrastructure, public services</b>				
Distance to the nearest major road, ln(km)	-.344*	-.585***	.540***	.223*
Distance to the nearest population center, ln(km)	.111	.295	-1.001***	-.275
Access to electricity, 0/1	-.051	.462	-.847	1.725***
2012 LSMS wave dummy variable, 0/1	-.781**	-.607	2.525***	.444
Constant	-25.198***	-22.630***	-9.569	5.024
<b>Sample size</b>				
Dam basins	510		223	
Downstream basins	236		51	
Non-dam basins	81		89	
	193		83	
p-value (H <sub>0</sub> : overall insignificance)	.000		.000	

Source: Authors' estimation. Asterisks indicate the statistical significance of the joint difference in means across basin types at given statistical significance: \*\*\* 1%; \*\* 5%; and \* 10%.

## REFERENCES

- AghaKouchak, A. 2015. "A multivariate approach for persistence-based drought prediction: Application to the 2010–2011 East Africa drought." *Journal of Hydrology* 526: 127–135.
- Akanmu, J. O., O. Remi-John, and I. E., Ekpo. 2011. "Overview of Dam Development in Nigeria." *Geotechnical and Geological Engineering* 29 (6): 953–960.
- Barrios, S., L. Bertinelli, and E. Strobl. 2010. "Trends in rainfall and economic growth in Africa: A neglected cause of the African growth tragedy." *Review of Economics and Statistics* 92 (2): 350–366.
- Caliendo, M., and R. Hujer. 2006. "The microeconomic estimation of treatment effects—an overview." *Advances in Statistical Analysis* 90 (1): 199–215.
- Cattaneo, M. D. 2010. "Efficient semiparametric estimation of multi-valued treatment effects under ignorability." *Journal of Econometrics* 155: 138–154.
- Cattaneo, M. D., D. M. Drukker, and A. D. Holland. 2013. "Estimation of multivalued treatment effects under conditional independence." *Stata Journal* 13: 407–450.
- Deaton, A. 1997. *The Analysis of Household Surveys*. Washington DC: World Bank.
- Deininger, K., G. Carletto, S. Savastano, and J. Muwonge. 2012. "Can diaries help in improving agricultural production statistics? Evidence from Uganda." *Journal of Development Economics* 98 (1): 42–50.
- Duflo, E., and R. Pande. 2007. "Dams." *Quarterly Journal of Economics* 122 (2): 601–646.
- FAO (Food and Agriculture Organization)/UNESCO (United Nations Educational, Scientific and Cultural Organization). 1977. *Soil map of the world: Volume IV Africa*. Rome and Paris: FAO and UNESCO.
- FAO (Food and Agriculture Organization), IIASA (International Institute for Applied Systems Analysis), ISRIC (International Soil Reference and Information Centre), ISSCAS (Institute of Soil Science – Chinese Academy of Sciences), and JRC (Joint Research Centre of the European Commission). 2012. *Harmonized World Soil Database (version 1.2)*. Rome and Laxenburg, Austria: FAO and IIASA.
- FAO. 2014. AQUASTAT. *Area equipped for irrigation and part irrigated by groundwater*. Rome: FAO.
- FAO. 2015. AQUASTAT: *Geo-referenced Database on Dams*. Computer Disk. Washington DC: FAO.
- Federal Ministry of Agriculture and Water Resources (FMAWR). 2007. *Compendium of Nigerian Dams: Their Briefs, Statistics & Purposes. Department of Dams & Reservoir Operations*. Abuja: FMAWR, Federal Government of Nigeria.
- Goldman, A., and J. Smith. 1995. "Agricultural transformations in India and Northern Nigeria: Exploring the nature of green revolutions." *World Development* 23 (2): 243–263.
- Gomi, T. 1978. "Nihon ni okeru nogyo risui no tokushitsu: kangaikeitai to sono bunpu." *Bulletin of the Faculty of Education, Kanazawa University: Social science and the Humanities* 25: 91–101.
- Khan, S., and E. Tamer. 2010. "Irregular identification, support conditions, and inverse weight estimation." *Econometrica* 78 (6): 2021–2042.
- Lehner, B., K. Verdin, and A. Jarvis. 2006. *HydroSHEDS Technical Documentation*. Washington, DC: World Wildlife Fund US. Available at <http://hydrosheds.cr.usgs.gov>.
- Masutomi, Y., Y. Inui, K. Takahashi, and Y. Matsuoka. 2009. "Development of highly accurate global polygonal drainage basin data." *Hydrological Processes* 23 (4): 572–584.
- Pan, W. 2015. *Propensity Score Analysis: Fundamentals and Developments*. New York: Guilford Publication.
- Pradhan, P. 1993. "Research Priority of Irrigation Management in Nigeria." In *Irrigation Research Priorities for Nigeria: Proceedings of a National Seminar Held at the University of Ilorin, Nigeria from 20 to 23 April, 1993*, edited by E. U. Nwa, and P. Pradhan.. Kano, Nigeria: International Irrigation Management Institute. 19–32.
- Sarsons, H. 2015. "Rainfall and conflict: A cautionary tale." *Journal of Development Economics* 115: 62–72.
- Sheffield, J., F. F. Wood, N. Chaney, K. Guan, S. Sadri, X. Yuan, L. Olang, A. Amani, A. Ali, S. Demuth, and L. Ogallo. 2014. "A drought monitoring and forecasting system for sub-Sahara African water resources and food security." *Bulletin of the American Meteorological Society* 95 (6): 861–882.
- Strobl, E., and R. O. Strobl. 2011. "The distributional impact of large dams: Evidence from cropland productivity in Africa." *Journal of Development Economics* 96 (2): 432–450.
- Takeshima, H., and F. Yamauchi. 2012. "Risks and farmers' investment in productive assets in Nigeria." *Agricultural Economics* 43 (2): 143–153.
- Takeshima, H. 2016a. "Custom-hired tractor services and returns to scale in smallholder agriculture: A production function approach." *Agricultural Economics*, forthcoming.
- Takeshima, H. 2016b. "Chapter 4: Policy options for inducing a supply response", in *The Nigerian Rice Economy: Policy Options for Transforming Production, Marketing, and Trade*, edited by K. Gyimah-Brempong, M. Johnson, and H. Takeshima. Philadelphia: University of Pennsylvania Press.
- Tharme, R. E. 2003. "A global perspective on environmental flow assessment: emerging trends in the development and application of environmental flow methodologies for rivers." *River Research and Applications* 19(5-6): 397–441.

- Thomas, D., and W. Adams. 1999. "Adapting to dams: agrarian change downstream of the Tiga Dam, northern Nigeria." *World Development* 27 (6): 919–935.
- Thompson, J., and G. Polet. 2000. "Hydrology and land use in a Sahelian floodplain wetland." *Wetlands* 20: 639–659.
- USGS (U.S. Geological Survey). 1996. *GTOPO30*. Sioux Falls, SD: USGS Center for Earth Resources Observation and Science.
- Villholth, K. 2013. "Groundwater irrigation for smallholders in Sub-Saharan Africa – a synthesis of current knowledge to guide sustainable outcomes." *Water International* 38 (4): 369–391.
- Wooldridge, J. 2007. "Inverse probability weighted estimation for general missing data problems." *Journal of Econometrics* 141 (2): 1281–1301.
- You, L., C. Ringler, U. Wood-Sichra, R. Robertson, S. Wood, T. Zhu, G. Nelson, Z. Guo, and Y. Sun. 2011. "What is the irrigation potential for Africa? A combined biophysical and socioeconomic approach." *Food Policy* 36 (6): 770–782.
- Zhao, Z. 2008. "Sensitivity of propensity score methods to the specifications." *Economics Letters* 98 (3): 309–319.

## About the Authors

**Hiroyuki Takeshima** is a Research Fellow in the Development Strategy and Governance Division of the International Food Policy Research Institute (IFPRI). **Adetola Adeoti** is a Senior Lecturer in the Department of Agricultural Economics of the University of Ibadan. **Oluwafemi Adebola Popoola** is a PhD Student in the Department of Agricultural Economics of the University of Ibadan.

---

## Acknowledgements

This paper has been prepared as part of the Feed the Future Nigeria Agricultural Policy Project. We are grateful to the United States Agency for International Development (USAID) for providing financial assistance to conduct this study. We also thank the Federal Ministry of Water Resources for sharing with us information about the types and locations of dams across Nigeria and Margaret Adesugba of the IFPRI Abuja Office for acquiring the data. We also appreciate Hua Xie for constructive comments on the study. All remaining errors are our own.

## INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

2033 K Street NW | Washington, DC 20006-1002 USA | T+1.202.862.5600 | F+1.202.457.4439 | Skype: ifprihomeoffice | ifpri@cgiar.org | www.ifpri.org

## IFPRI-ABUJA

c/o International Fertilizer Development Center | No.6 Ogbagi Street | Off Oro-Ago Crescent | Garki II, Abuja | Nigeria  
ifpri-nigeria@cgiar.org | nssp.ifpri.org

This publication of the Nigeria Strategy Support Program is made possible by the generous support of the American people through the United States Agency for International Development (USAID) in connection with the Feed the Future Nigeria Agricultural Policy Project. It has not been independently peer reviewed. The contents are the responsibility of the author(s) and do not necessarily reflect the views of USAID, the United States Government, or the International Food Policy Research Institute.

Copyright © 2016, International Food Policy Research Institute. All rights reserved.  
To obtain permission to republish, contact ifpri-copyright@cgiar.org.