Dams

Esther Duflo and Rohini Pande*

Draft: April 2005

Abstract

Infrastructure is widely cited as an integral part of the development of any country. However, evidence on the returns to infrastructure, and their distribution in the population, remains limited. This paper studies one of the most important, and controversial, forms of public investment in infrastructure in India – large dam construction. We use the fact that the river gradient in a district affects its suitability for dam construction to construct instrumental variable estimates of the effect of dam construction on economic outcomes. The empirical results suggest that the district in which a dam is constructed sees poverty rise and no improvement in agricultural productivity. In contrast, the districts located downstream witness a fall in poverty and increase in irrigated area and agricultural production. Taken together, these findings support the view that large dam construction in India is a marginally cost-effective investment which has, in aggregate, increased poverty.

*The authors are from MIT and Yale University respectively. Pande thanks NSF for financial support for this project under grant SES-0417634. We thank Abhijit Banerjee, Michael Greenstone, Michael Kremer, T.N. Srinivasan and Chris Udry for comments. We also thank Sara Aronchick, Niki Klonaris and Callie Scott for excellent research assistance, CIESIN at Columbia University for GIS help, and Abhijit Banerjee, Shawn Cole, Seema Jayachandran and, especially, Petia Topalova for sharing their data.

1 Introduction

In 2000, public spending on infrastructure averaged nine percent of government spending in developing countries.[1] Despite the magnitude of such spending, and a widespread belief that infrastructure is integral to development, empirical evidence on how increases in physical infrastructure affect productivity and individual well-being remain limited (Bank 1994). Equally, there is very little evidence on which groups of citizens gains, and which lose, from such investments. This paper examines these questions in the context of large dam construction in India.

Worldwide, over 45,000 large dams have been built and nearly half the world's rivers are obstructed by a large dam. As of the year 2000, dam reservoirs stored roughly 3,600 cubic kms of water, generated 19 percent of the world's electricity supply and provided irrigation for between 30-40 percent of the 271 million hectares irrigated worldwide (World Commission on Dams 2000). India, with over 4,000 large dams, is the world's third most prolific dam builder (after China and the USA) and is the country with the most irrigated area. Large dam construction remains the main form of government investment in irrigation potential in India.[2] Between 1951-1997 public investment on major and medium irrigation projects was approximately 33 billion US dollars (Thakkar 2000).

The question of whether economic and social costs of large dams outweigh its benefits is widely debated. Proponents of large dam construction emphasize the fact that, over the 20th century, large dams have significantly increased irrigated land for agriculture, water and hydropower available for domestic or industrial use and have improved flood control. More controversially, they argue that other forms of water harvesting, such as ground water use and small dykes and dams, while useful, remain relatively cost ineffective, and incapable of meeting the demands of large and growing populations in countries with highly seasonal rainfall (Biswas and Tortajada (2001), B.D.Dhawan (1989)).

[1] This translates into roughly 1.4% of GDP, (IMF Finance Statistics)

[2] Irrigation is the stated objective of over 96 percent of India's dams. Almost all of India's dams are reservoir type storage projects which impound water behind the dam for seasonal, annual and, sometimes, multi-annual storage and regulation of the river.

Opponents point to the fact that improved, and subsidized, water provision associated with large dam construction has caused farmers to change cropping patterns towards water-intensive crops like sugarcane. As a result, they argue, dam construction has enhanced the very water shortage problem it was intended to solve. In addition, they suggest that increases in agricultural productivity are limited by the loss of agricultural and forest land associated with dam construction, and increased salinity and waterlogging in the command area of the dam (McCully (2001) and Singh (2002)).

A separate, and important, concern relates to the regional distribution of the costs and benefits associated with dam construction (J.Dreze, M.Samson, and S.Singh (1997), and Roy (1999)). Irrigation benefits accrue to those living downstream from the dam. In contrast, those living near the dam reservoir bear the displacement costs, and are more exposed to diseases caused by the large-scale impounding of water, such as malaria.[3]

Despite the intensity of this controversy, there are no credible evaluations of how a dam, on average, affects agricultural production and rural welfare.[4] This, in part, reflects the fact that a myriad of factors, such as geographic suitability for dam construction, demand for water storage and the political clout of local governments, affects dam placement. A simple comparison of outcomes in regions with and without dams is unlikely to provide a causal estimate of the effects of dam construction.[5]

In this paper, we address this problem by making use of the fact that districts in which the river flows at an incline are preferred for dam construction. This preference is related to the two main functions of dams, water storage and raising the level of water upstream.[6]

[3]According to World Commission on Dams (2000) global estimates suggest that 40-80 million people have been displaced by reservoirs.

[4]Merrouche (2004) uses Indian state-level panel data to provides an OLS estimate of how dam construction has affected poverty and inequality across Indian states. The study, however, fails to account for endogenous dam placement. Most other existing evaluations are case studies, often limited to the largest dam projects.

[5]For instance, the Indian states of Gujarat and Maharashtra both have the highest dam concentration in India, and are among the richest and fastest growing states. Clearly, their growth experience cannot be entirely attributed to dam construction. Further, it is very likely that their success in dam construction was, at least in part, related to their economic performance.

[6]Raising the water level upstream enables water diversion into a canal, and electricity generation. Hence,

In practice, this translates into a higher incidence of dam construction in river valleys in hilly areas. Hence, controlling for overall gradient and river length in a district, we can use differences in the river gradient to predict how, in a given year, dams constructed in a state are distributed across districts in that state.[7]

Our analysis uses the variation in dam construction induced by differences in river gradient across districts to construct instrumental variable estimates of the economic impact of dams in regressions of district agricultural and poverty outcomes on the number of dams and district characteristics. We control for district fixed effects, state year interactions, and the interactions of most district geography variables with the number of dams in the state in that year. Only the interaction between the fraction district area along the river in different gradient categories and the number of dams in the state in that year is assumed exogenous. Our empirical strategy is robust to a range of omitted variable and possible endogeneity concerns. First, by comparing within state and year cells we can control for differential time trends across states. Second if, within states, districts with greater river presence or relatively more incline have, over time, evolved differently in a way correlated with overall dam construction in the state, this is controlled for by the interaction between the number of dams in the state in that year and these variables.

Our results reconcile the seemingly irreconcilable claims of the proponents and adversaries of dams. Dam construction does not affect agricultural production or wages in the districts where they are located. Instead, rural poverty in these districts rises. In contrast, area irrigated, agricultural production, and wages improve in districts located downstream, and poverty is reduced. Dams also reduce the impact of rainfall shocks in downstream districts. Cost benefit analysis suggests that the increase in overall productivity was large enough to imply that dam construction had the potential of improving overall welfare. The fact that, in reality, overall poverty increased points to inadequate redistribution from the

a narrow gorge with a valley opening upstream (to provide the required storage) minimizes dam construction expenses.

[7] The spirit of our analysis is related to **?**) who use geographic variables to construct instruments for the extent of a country's trade.

winners to losers.

These findings are important in their own right. They also help improve our understanding of the benefits and costs associated with undertaking large public projects which cause substantial displacement (a feature shared by many industrial projects in developing countries). On the methodological front, our estimation strategy can be used to provide convincing estimates of the causal effects of large infrastructure projects, when project location is influenced by geography.

The remainder of the paper proceeds as follows: In Section 2 we provide the context for our study by reviewing the existing literature on the productivity and welfare effects of large dam construction in India. Section 3 describes the data, Section 4 the empirical strategy, Section 5 the empirical results and Section 6 concludes.

2 Background

In this section we describe the main insights offered by the case-study based literature on the economic impact of dam construction in India. As irrigation is the primary purpose of most large Indian dams, we focus on irrigation dams.[8]

Most irrigation dams in India are embankment dams. That is, an artificial wall is built across a river valley, and water is impounded behind this wall in a 'reservoir'. A system of spillways and gates convey normal stream and flood flows over, around, or through this wall, and artificial canals channel water from the reservoir to downstream regions for irrigation. The area upstream from which water and silt flow into the reservoir, and the area submerged by the reservoir, form the dam's catchment area. The area downstream to the reservoir which is covered by the dam's canal network makes up the dam command area.

The main economic benefit of dams has been realized in agriculture. Between 1951 and 2000 India's food grain production quadrupled from 59 to 200 million tonnes. Roughly two-thirds of this increase was in irrigated areas. In 2000, dam irrigation accounted for

[8]The three primary purposes associated with dams are irrigation, hydropower and flood control.

roughly 38% of irrigated area.[9] Hence, it has been argued that roughly a quarter of India's increased food grain production over the last half century comes from dam irrigated areas. The fraction of increased food production in dam irrigated areas that is attributable to dam irrigation rather than, say, the concurrent uptake of mechanized agriculture remains controversial with estimates varying from 10% (R.Rangachari, N.Sengupta, R.Iyer, P.Banerji, and S.Singh 2000) to over 50% (Gopalakrishnan 2000).

Critics of dam irrigation argue that an important reason for limited productivity gains is waterlogging, and increased soil salinity. Seepage from the dam's reservoir and irrigation channels makes neighboring land waterlogged. Further, the rise in the water table increases salt concentration near the soil surface and makes it less productive.[10] According to a 1991 estimate by the Ministry of Water Resources in India, about 2.46 million hectares of command area of major and medium irrigation systems suffered from waterlogging, and 3.30 million hectares were affected by salinity/alkalinity (R.Rangachari, N.Sengupta, R.Iyer, P.Banerji, and S.Singh 2000). This is about a tenth of the area estimated to be irrigated by dams, suggesting that this problem is of serious magnitude.

For a more micro-perspective, Table 1 lists the irrigation benefits associated with four large Indian dams (main source Thukral (1992)). While these dams are relatively large, it is not unusual for a dam's command area to encompass multiple districts and often neighboring states. For instance, almost the entire canal network of the Pong dam, located in Himachal Pradesh, services the neighboring states of Punjab and Haryana.

Dam irrigation benefits mainly accrue to those living in the dam's command area, i.e. downstream to the dam.[11] Singh (2002) provides case study evidence that the advent of dam irrigation increased multi-cropping, and increased cultivation of more profitable water-intensive cash crops such as sugarcane. Case studies also suggest that land prices

[9] Irrigated area make up a third of cultivated area. It is usual to interpret the area irrigated by major and medium irrigation projects as dam irrigated.

[10] When the concentration of soil reaches a range of 0.5 to one percent the land becomes toxic to plant life (Goldsmith and N.Hildyard 1984).

[11] A number of studies suggest that farmers in the upper and middle reaches of the command area typically benefit more than tail-enders (Thakkar 2000).

in downstream areas rise. In contrast, those living in the vicinity of the reservoir and immediately upstream obtain no irrigation benefits.[12] Moreover, to ensure sufficient water flow to the dam there is often a ban on schemes that diverts water upstream of the dam (Tehri 1997).

The area around the reservoir potentially witnesses increased economic activity during dam construction. Reservoirs are often developed as tourism sites, and fishing potential is enhanced by construction of the reservoir. Counterbalancing this, is the submergence of land and forests associated with reservoir construction. Estimates suggest that 4.5 million hectares of forest land was submerged between 1980 and 2000 (R.Rangachari, N.Sengupta, R.Iyer, P.Banerji, and S.Singh 2000). Moreover some claim, that while dam construction takes many years, development activities in the district are often suspended as soon as a dam is planned, and land prices fall.

The largest cost of dam construction arises from the displacement of people living in areas submerged by construction of the reservoir. As can be seen by the example of Hirakud dam in Table 1, the area submerged by the reservoir can reach up to nearly 10% of the district's area. The extent of displacement associated with dam construction varies substantially with population density in the vicinity of the dam – the Nagarjuna Sagar dam, while three times the height of Hirakud, displaced many fewer villages. Dams are typically located in uphill areas in river valleys, where tribal populations are more likely to be resident. Official figures for 34 large dams show that tribals, who make up 8% of India's population, constituted 47% of those displaced (R.Rangachari, N.Sengupta, R.Iyer, P.Banerji, and S.Singh 2000). Tribal populations tend to be economically disadvantaged, and are likely to find adjusting to a new environment particularly hard.

Resettlement is the responsibility of the relevant project authorities, and what a displaced person ultimately gets depends a great deal on his/her political power and organizational abilities.[13] Until the 1990s the main resettlement policy was cash compensation based

[12] Lift irrigation is rarely practiced in India.

[13] The Indian government doesn't have an explicit national policy on resettlement and rehabilitation of the displaced population.

on the Land Acquisition Act of 1894. The compensation level was based on project-specific government resolutions, and was very often below the value of land available for purchase by the government. As a result, the cash often ended up being used as a temporary means of subsistence (J.Dreze, M.Samson, and S.Singh 1997). In addition, rights of the landless, and those without formal land titles, to compensation are typically not recognized.

In practice, resettlement follows an incremental process – people are shifted and resettled according to the construction and demarcation schedule. As a result people are shifted to resettlement sites at the last minute irrespective of whether they have received compensation and whether rehabilitation sites are ready. Once shifted, these individuals lose substantial bargaining power, and are often unable to obtain any outstanding compensation (Thukral 1992).[14]

A number of studies suggest that dam construction can negatively impact the health status of those living in the vicinity of the reservoir. A reservoir provides a natural ground for vector breeding, and hence the area around it becomes a breeding ground for diseases such as malaria, schistosomiasis, filariasis and river blindness (see Sharma (1991)).

To summarize, the role of dams in increasing irrigation potential is largely undisputed. However, whether this increase has translated into substantial productivity gains and the regional distribution of these gains remains controversial. There is also a widespread belief that the government's rehabilitation policies have been largely insufficient to compensate those displaced, and may have left them more vulnerable to future productivity shocks. It is beyond the scope of this paper to estimate all the possible channels, but we will evaluate whether, on net, production increased in the area served by the dams, and whether poverty increased or decreased during the period.

[14]**?**) discuss results from a survey of 602 families displaced by Koyna dam. 20% of those displaced received no land. More than half the families (324) received forest land and roughly 75% of these got land full of stones and tree stumps which was currently not being cultivated. They also found the displaced populations changed their cultivation patterns. Instead of 579 families who used to grow rice only 246 now do. 125 families moved to growing millets (jowar) which was not grown earlier by any family .

3 Data and Descriptive Statistics

Our analysis combines a variety of data drawn from many different sources. Below, we briefly describe these data. Table 2 provides descriptive statistics, and the Data Appendix further details on these data.

3.1 Dams

Our dam data comes from the World Registry of Large Dams, maintained by the International Commission of Large Dams (ICOLD). This registry provides information on dam height, year of completion, purpose (irrigation, electricity or both), river it is built on and nearest city for all large dams in India.[15] Using information on nearest city, we manually identified the district location for the 3,364 dams that had been completed by 1999. We used these data to construct our main regressor of interest – the annual cumulative number of dams completed in each district in each year.

Our analysis spans the period 1970 to 1999 (the precise years covered vary according to the outcome variable being considered). In 1970, the mean number of dams in a district was 2.05, and the median district did not have a dam. Between 1970 and 1999 the number of large dams quadrupled (from 882 to 3,364). Figure 1 shows this increase was accompanied by significant regional variation in dam construction with the increase highly concentrated in the Western region. In 1999 the median district had one dam (46% of the districts had no dams) and the mean number of dams per district was 7.84.[16] Nine of the thirty two states or Union Territories in our sample witnessed no dam construction up to 1999.[17] Seven of these, also saw no dam construction in bordering districts in neighboring states. Our analysis excludes these seven states (which lie in the very mountainous Northeast region),

[15] A dam with a height of 15m or more from the foundation is defined as a large dam. Dams between 5-15m high with a reservoir volume of more than 3 million cubic meters are also classified as large dams.

[16] The median district in the three largest dam constructing states Maharashtra, Gujarat and Madhya Pradesh had 39, 18 and 15 dams, respectively.

[17] These are Arunachal Pradesh, Mizoram, Nagaland, Punjab, Sikkim, Dadra and Nagar Haveli, Daman and Diu, Delhi, Pondicherry. The only big state among these is Punjab.

as well as the two districts with over 100 dams.

3.2 GIS data

We use GIS data to obtain district-wise data on total area, river length, elevation and gradient, both overall, and along river. These data exist polygon-wise with each district comprising of multiple polygons. For each district overall gradient was computed as the percentage of land area (summed across all polygons in a district) in different elevation/gradient categories. Gradient along river was calculated in an identical manner, with the difference that we restricted the calculation to polygons through which the river flowed.

Not all districts with a river have a dam. Among the most striking examples of this are districts in the Gangetic plain, where there are essentially no dams. This may be, in part, related to the region's flatness: most of the Gangetic plain has elevation below zero. The Western regions, where most dams are located, have a relatively large fraction of river length flowing at some gradient.[18]

We used census district maps to identify for each district all neighboring districts, and within these the set of upstream and downstream districts (defined with reference to river flow).

3.3 Agricultural and Rural Welfare Outcomes

Our agricultural data are from the World Bank India Agriculture and Climate data-set which includes 271 Indian districts in thirteen Indian states and covers the years 1957-58 to 1986-87 for production and crop-wise outcomes, and extends up to 1994 for male agricultural wage and area outcomes. We deflate monetary variables using the state-specific Consumer Price Index for Agricultural laborers.

The average share of cultivated area under irrigation in a district increased from 26% to 45% between 1973 and 1995 (net cultivated area remained roughly constant over the

[18] However, southern states (such as Kerala and Karnataka), which also have rivers with a moderate gradient, have fewer dams than the western states, suggesting that geographic potential was not the only determinant of dam construction.

period). The three states with maximum dam construction started from a very low share of irrigated area, and increased rapidly (for example it went from 9% to 31% in Madhya Pradesh). However, we observe parallel increases in irrigation availability in other states as well (the alternative to dams is ground water irrigation).

Our rural welfare measures are constructed from quinquennial household surveys conducted by the Indian National Sample Survey (NSS). The survey covers all Indian states with a rural sample of about 75,000 households. Random selection of households within districts makes the construction of district averages feasible.[19]

We use data from the 1973-74, 1983-84, 1987-88, 1993-94 and 1999-2000 ("thick") NSS surveys. Processed data for the year 1973 was obtained from Jain, K.Sundaram, and S.D.Tendulkar (1988). For 1983 onwards, Topalova (2004) computed district-wise statistics. She follows Deaton (2003a), Deaton (2003b) to compute adjusted poverty estimate. Introduction of a new recall period for household expenditure in the 1999-2000 round implies that these data are not directly comparable to earlier rounds. To achieve comparability with earlier rounds, she follows Deaton and imputes the correct distribution of total per capita expenditure for each district from the households expenditures on a subset of goods for which the new recall period questions were not introduced. The poverty and inequality, and mean per capita expenditure measures were derived from this distribution.

District identifiers are available since 1987.[20] For the years 1973 and 1983, the variables are constructed at the NSS region level (a region is a group of district sharing common characteristics, for which the sample is sufficiently large for the NSS to consider the data "representative" of the region). Murthi, P.V.Srinivasan, and S.V.Subramanian (2001) provide a matching between the 1973 and 1983 NSS regions and the 1981 census district definition, and a matching between the 1981 census definition and the 1991 census definition.[21] We matched the NSS data across rounds at the level of the 1981 district definition,

[19] The NSS Organization does not report district averages, as it considers its district sample size inadequate for obtaining reliable poverty estimates for each district. This does not affect us, since we report regression results for multiple districts and do not make inferences about any particular district.

[20] They had to be recovered from hard copies for the year 1993.

[21] India's districts boundaries changed several time between 1961 and 1991, mostly due to the splitting of

using census maps and other geography indicators.

4 Empirical Strategy

4.1 Sources of Variation

Our unit of analysis is the Indian district. The extent of dam construction in an Indian state and the river gradient in the district jointly determine the number of dams built in a district. A district located in a state which undertakes little dam construction receives fewer dams than its counterpart in a high dam construction state. Within a given state, a district in which the river flows at little to no gradient receives a lower share of the state's total dams than a district in which the river flows at a moderate gradient.

In Table 3 we provide three two-by-two tables to illustrate how our identification strategy exploits these two facts. The three panels in Table 3 show the means of the change over our sample period in the number of dams constructed, the head count ratio and log agricultural production for four state-district combinations. Districts are classified as high or low river gradient districts. A district in which less than 90 percent of river gradient is below 1.5 percent is classified as a high gradient district, and low otherwise. States are classified as high or low dam construction states. A state which built more than a hundred dams by 1999 is classified as a high construction state, and low otherwise. On average, 11.69 dams were built in high construction states and 0.75 in low construction states (the difference being 10.94).

Panel A compares the number of dams constructed between 1973 and 1999 in high and low gradient districts in both types of states. Relative to low dam construction states, the increase in dams is higher in *both* high and low gradient districts in high dam construction states. However, relative to low dam construction states, the difference between high and low gradient districts is higher in high dam construction states. The difference in these differences can be interpreted as the causal effect of a district's river gradient on dam

districts into two parts.

construction, under the assumption that, in the absence of state differences in overall dam construction, districts with the same gradient in different states would receive the same number of dams. Between 1973 and 1999 a high river gradient district, in a high dam construction state, received 7 more dams.

In Panels B and C we examine changes in rural head count ratio and log production. In a high river gradient district located in a high dam construction state the head count ratio increased by an additional 4.5 percent and log agricultural production fell by 0.32. The Wald estimate of the poverty impact of dam construction is the ratio of the difference in difference in Panels A and B, and stands at 0.64.

These tables illustrate our identification strategy, but our results are imprecise in that we only use part of the available information. In addition, these estimates fail to control for other geographical variables which may be correlated with both river gradient and changes in, say, the head count ratio. We, therefore, turn to a more general framework.

4.2 General Framework

The following two equations relate the outcome of interest (e.g. per capita consumption or agricultural production) in district i belonging to state s, in year t, to total completed dams in the district in that year, and to those in neighboring upstream districts.[22] We lack information on the precise catchment and command area associated with each dam. However, given the size of a district, it is reasonable to assume that the catchment area of a dam falls within the district in which it is physically located. In contrast, it's command area is likely to extend to neighboring downstream districts.

$$y_{ist} = \beta_1 + \beta_2 D_{ist} + \nu_i + \mu_{st} + \omega_{ist}, \tag{1}$$

$$y_{ist} = \beta_3 + \beta_4 D_{ist} + \beta_5 D^U_{ist} + \nu_i + \mu_{st} + \omega_{ist} \tag{2}$$

[22] In what follows, we will refer to the "neighboring upstream districts" as "upstream districts", for short.

where ν_i is a district fixed effect, μ_{st} is a state-year interaction effect, and ω_{ist} a district-year specific error term.[23] District fixed effect control for time-invariant characteristics that affect a district's propensity to get dams. The state-year interactions account for annual shocks which are common to districts in a state. Hence, the regressions only exploit annual variation in dam construction across district within a state for identification.

The identification assumption underlying the above regression is that variation in dam construction across districts belonging to the same state within a year is uncorrelated with other shocks which affect these districts. However, there are reasons to expect this assumption to be violated – for example, we may expect relatively greater dam construction in districts which witness rapid agricultural growth and therefore have a higher demand for irrigation. For this reason, we adopt an instrumental variables approach. To construct instruments for dam construction we exploit the fact that preferred dam location is along rivers at points with moderate incline (which increases water flow). At the same time, too steep an incline can make dam construction impossible.

To implement this strategy we first estimate the following regression to predict the average number of dams in district i in year t:

$$D_{ist} = \alpha_1 + \sum_{k=2}^{5} \alpha_{2k}(RSl_{ki} * \overline{D_{st}}) + \sum_{k=2}^{4} \alpha_{3k}(El_{ki} * \overline{D_{st}}) + \sum_{k=2}^{5} \alpha_{4k}(Sl_{ki} * \overline{D_{st}}) + (X_i * \overline{D_{st}})\alpha_5 + \nu_i + \mu_{st} + \omega_{ist} \quad (3)$$

$\overline{D_{st}}$ is the cumulative number of dams in state s in year t. RSl_{ki} is the fraction of river area in gradient category k, and Sl_{ki} is the fraction district area in gradient category k. We consider five gradient categories – gradient less that 1.5%, 1.5-3%, 3-6%, 6-10% and more than 10%. El_{ki} is the fraction of district area in elevation category k. We consider four categories (in meters) - 0-250; 250-500; 500-1000 and over 1000. Finally, X_i is a vector who's elements are district area and river length.

We use this regression to predict the number of dams per district $\widehat{D_{ist}}$ (specifically, it is the predicted value from equation (3)), and for upstream districts $\widehat{D^U_{ist}}$. The latter is the

[23] We control for autocorrelation by clustering the equation at the district-level.

sum of the predicted value from equation (3) for all upstream districts, or 0 if the district has no upstream district.

Denote by Z_{ist} the vector of right hand side variables in equation (3), except for the interactions $RSl_{ki} * \overline{D_{st}}$. Let Z^U_{ist} denote the corresponding variables for upstream districts.[24]

We then augment equations (1) and (2):

$$y_{ist} = \gamma_1 + \gamma_2 D_{ist} + Z_{ist}\gamma_3 + \nu_i + \mu_{st} + \omega_{ist} \tag{4}$$

and:

$$y_{ist} = \delta_1 + \delta_2 D_{ist} + \delta_3 D^U_{ist} + Z_{ist}\delta_4 + Z^U_{ist}\delta_5 + \nu_i + \mu_{st} + \omega_{ist} \tag{5}$$

We estimate equation (4) with 2SLS, using $\widehat{D_{ist}}$ and Z_{ist} as instruments, and equation (5) using $\widehat{D_{ist}}$, $\widehat{D^U_{ist}}$ and Z^U_{ist} as instruments.

The first stage equations are:

$$D_{ist} = \pi_1 + \pi_2 \widehat{D_{ist}} + Z_{ist}\pi_3 + \nu_i + \mu_{st} + \omega_{ist} \tag{6}$$

and:

$$\Delta_{ist} = \phi_1 + \phi_2 \widehat{D_{ist}} + \phi_3 \widehat{D^U_{ist}} + Z_{ist}\phi_4 + Z^U_{ist}\phi_5 + \nu_i + \mu_{st} + \omega_{ist} \tag{7}$$

where Δ_{ist} represent D_{ist} or D^U_{ist}.

For equation 4, this 2-step procedure is identical to running a 2SLS using the interactions $RSl_{ki} * \overline{D_{st}}$ and Z_{ist} as instruments. For equation 5, this procedure uses the entire set of districts to predict the relationship between the district geographical features and the number of dams (rather than the set of districts which are upstream), and avoids averaging the features when there are more than one upstream district.

[24]For multiple upstream districts, we sum river length and district area across all upstream districts, and average the other variables (which are proportions).

5 Results

5.1 Geographic Determinants of Dam Construction

We start by examining the relationship between district geography and dam construction (3). In Table 4 we present results for the 5 year sample for which we have rural welfare outcomes, and the 21 years for which we have agricultural data irrigated.[25]

The estimates of interest are the coefficients on the interaction of the total dams in a state in a given year with district geographic characteristics. These coefficients tell us which geographic variables lead a district in a state to receive more dams, as the number of dams in that state increases. The estimated coefficients are sensible: relative to other districts in the same state, more dams are built in larger districts with greater river presence and a higher fraction of district area of moderate elevation (250 to 500 meters), and with moderate gradient (1.5% to 3%). Important for our purpose, dam construction is also greater in districts where a larger fraction of area around the river is of moderate gradient (1.5% to 3%). More dams are also built in districts where a larger fraction of area along the river is very steep (more than 10%). This, very likely, reflects the construction requirements of hydroelectric dams. Together, the four interactions between the gradient along river and the number of dams present in the district in a particular year are significant (the F-statistics are 2.63 and 3.4, respectively).

Table 5 reports the first stage equations for our 2SLS procedure (equation (6) and (7)). The number of dams in a district is regressed on the predicted number of dams in the district and in upstream districts. Likewise, the number of dams in upstream districts are regressed on the predicted number of upstream and own district dams. The regressions include all control variables except the interaction of river gradient with total dams in the state in that year. Not surprisingly, the coefficient on the predicted dams is close to 1 in columns (1) and (2), with a t-statistic of over 5. The coefficients of the predicted number

[25] At the moment, we do not have production data for the last 7 years in the sample. However, the first stage for this sample is virtually identical to that reported.

of dams in upstream (downstream) districts are also close to 1 and highly significant in the upstream (downstream) regressions.

5.2 Agricultural Outcomes

To examine the impact of dam construction on area irrigated and cultivated we consider gross and net area measures. Gross measures are informative of the prevalence of multi-cropping. Columns (1)-(8) Panel A of Table 6 provides the OLS estimates (equations (1) and (2)), and Panel B the two stage least squares estimates (equations (4) and (5)). Both the OLS and 2SLS estimates suggest that dams do not significantly alter the area irrigated, either gross or net, in the districts where they are built (columns (1) and (3)), but significantly increase both of these in districts located downstream.[26] The 2SLS estimate exceed the OLS estimate, although both are significant and statistically indistinguishable from each other. The point estimate suggests that one more dam increases irrigated area (growth or net) in the downstream district by roughly 1%. The absence of an effect in the district where the dam is built is consistent with the claim of opponents of dam construction that land submergence and degradation around the reservoir limits irrigation gains in the vicinity of the dam. Columns (5)-(8) show that total area cultivated, both gross and net, is unaffected by dam construction. However, the point estimate for the dam's own districts are negative in both panels. Finally in columns (9)-(10) we observe that area under High Yielding Varieties (HYV) crops increases downstream.

Table 7 examines the impact of dam construction on agricultural production and yield for 19 major crops and on fertilizer usage. Here, we have data for the years 1971-1987. Columns (1) and (2) consider total production. Both the OLS and 2SLS estimates suggest that dam construction led to an insignificant decline in overall production in the district where they were built, and a significant increase in the downstream districts; see row 2, both panels. Similarly, agricultural yield shows an insignificant decline in the district where

[26] That is, districts with more dams in neighboring *upstream* districts have more irrigated area: see columns (2) and (4), row 2 in both panels.

dams were constructed, but a significant rise downstream, see columns (3) and (4). The own district results are, again, consistent with the claim that land around dams is degraded. This degradation is, in part, compensated for by increased productivity elsewhere in the district. The downstream districts, that do not bear any of the environmental costs associated with dam construction, and enjoy positive productivity gains. Finally, in columns (5) and (6) we examine fertilizer usage and find and increase in downstream districts.

The evidence on HYV area suggests that dam irrigation has altered cropping patterns. Table 8 provides OLS and 2SLS estimates of the impact of dams on the area cultivated, yield, and production, by major crops or crop groups.[27]

Dam critics have argued that cheap water provision implies that post-dam construction farmers devote larger areas to relatively water intensive cash crops, notably sugar. In line with this, we observe an overall insignificant increase in area devoted to water intensive crops in downstream districts. Among major crop this increase comes at the expense of pulse. The area devoted to wheat, which can also be irrigated, increases significantly as well.

The impact of dams on crop yield is modest, even for highly water intensive crops; see panel D. No crop shows a significant increase in its yield in the downstream district, though the coefficients are positive for the sum of water intensive crops and sugar.

The increase in area devoted to water intensive crop, combined with moderate yield increases, implies a significant increase in the production of water intensive crops in the downstream district. Overall production of water intensive crops increases downstream by 0.7% for each dam built. This is mainly attributable to a large increase in sugar and rice production.

Taken together, these results provide a consistent picture of the impact of dam construction on agricultural outcomes. In the districts where they are built dams have no positive

[27] It should be noted that because all the variables are in log, the observation disappears when a district does not produce the crop at all for a year. These specifications are therefore not suitable to capture the effect on dams on the decision to start to produce a given crop.

impact on agricultural production. In downstream districts, dams enhance agricultural production, both for some cash crop (sugar) and for an important staple (wheat). These results suggest that the dam's impact on welfare are likely to be very different in the dams' own district and in neighboring districts.

Substitution and Cost Benefit analysis

The increase in net irrigated area due to each dam is about 1% downstream and 0.7% in the own district. To what extent does this reflect crowding out of public or private investments in irrigation? These point estimates, combined with the size of the irrigated area in the average district, imply, in 1980 that the net increase in irrigated area due to a dam was 6,000 hectares. The Planning Commission directly estimates the area irrigated by dams. Their estimate was that, 1980, 23.5 millions hectares were irrigated by dams, which corresponds to an average of 7255 ha per dam. These estimates taken together suggest a modest crowd out of other investment in irrigation (of order of 20%). The fact that there is some crowd out is consistent with estimate on public good showing that downstream districts have a little less investment in other water infrastructure, such as tubewells (Table 13).

The cost of dam construction is generally expressed in terms of the cost of irrigating one more hectare through a dam. We therefore base the cost benefit analysis on the combination of the value of additional production per additional hectare irrigated (from our estimates) and the capital and recurrent cost of one additional hectare irrigated by a dam (from the planning commission). Using 1980 means and our estimates, the increase in production due to a dam in a district downstream is about Rs 2.37 annually, or Rs 47 million in present discounted value (assuming a 5% discount rate). This is due to an net increase in 2,530 hectare in irrigated area. The present discounted value of the next increase in production per hectare irrigated therefore turns out to be Rs 17,779. The planning commission estimates that, in 1980, the cost per hectare of dam irrigation was about Rs 16,129 (this include capital cost and annualized maintenance cost of Rs 300 per hectare). The estimates therefore suggests that the net present value of building a dam is just about positive (the rate of returns is about 10%). This analysis is, however, probably an upper bound on the economic value of a dam, since it does not take into account the loss of production in the dam's own district (the point estimate is very imprecise and not significantly different from zero, but it is large). The estimates also suggest that dam construction responsible for between 8% and 10% of the growth in agricultural production over the period.

These estimates are very close to the conclusions reached by the author of the India case study of the World Commission on Dams, which used very different methods: they concluded that 10% of the growth in agricultural production over the period was due to dams, and that the Net Present Value of the average dam was probably slightly negative. Although the World Commission of dams was supposed to be an independent, non-partisan international body put in charge of objectively assessing the costs and benefits of dams, the India estimates were criticized by many proponents of dams as being overly conservative: The international commissions on large dams (ICOLD) issued a response to the WCD India report claiming that dam's contribution to the growth in agricultural production was closer to 80%. Our estimates suggest that the conclusions of the WCD were closer to the truth.

These estimates, however, do not address the main debate between advocates and adversaries of dams: on balance, were they beneficial to the poor, or were the poor hurt by dams?

5.3 Rural welfare

To examine how dams have affected rural welfare, we consider rural consumption, poverty and wage outcomes. The results are in Table 9 – panel A provides OLS estimates (equations 1 and 2) and panel B two stage least squares estimates (equations 4 and 5).

Columns (1) and (2) show the impact of dams on mean capita expenditure. In column (1), both the OLS and the 2SLS estimates suggest that more dams in a district reduce mean capita poverty expenditure (10 more dams lead to a decrease of 3% in the OLS specification, 7% in the IV specification). In column (2), we include dams built in upstream districts as an additional explanatory variable. Dams may have a modest positive impact on per capita expenditure in the downstream districts, but the coefficient is insignificant in both the OLS and 2SLS regressions. The 2SLS point estimate is twice as large, but statistically indistinguishable from the OLS estimate.

We find similar results for poverty outcomes, with the difference that the downstream effects are significant in the 2SLS specification. Columns (3) and (4) consider the head count ratio: dams significantly increase poverty in their own district, and lead to a much smaller decline in poverty in downstream districts. The downstream effect is significant at 10% in the 2SLS estimates. The head count ratio is a relatively crude measure of the extent of poverty. The poverty gap, instead, measures the depth of poverty – specifically, how much income would be needed to bring all the poor to a consumption level equal to the poverty line. Columns (5) and (6) consider poverty gap as the dependent measure, and again find similar effects: dams significantly increase the poverty gap in their own district, and significantly reduce it downstream. The point estimate for the poverty reduction associated with dam construction upstream is a fifth (in the OLS case) to a eighth (in the 2SLS case) of the poverty increase in the dams' own district. In our sample there are, on average, 1.75 district downstream of each dam. This implies that, on balance, the poverty reduction in districts downstream to the district where dams are constructed is too small to compensate for the poverty increase in the dams' own district.

Columns (7) and (8) consider the impact of dams on the Gini coefficient. There is no apparent pattern of an impact of dams on inequality either in their own district or in the neighboring districts (though inequality appears to be reduced in districts upstream to where dams are constructed).

Finally, columns (9) and (10) examine how dam construction affects the male agricultural wages. Annual data are available for 1971-1994, but for fewer states than in the poverty sample.[28] Examining agricultural wages helps us relate our findings for agricultural outcomes to those for poverty and consumption outcomes. One may expect higher land productivity (especially from the production of cash crops) to translate into higher

[28] We get similar results when we restrict the year to the years for which we have NSS data

agricultural wages, and higher wages have, in turn, been shown to be important for rural poverty reduction (Dreze and Mukherjee 1991). We find that wages increase in districts located downstream from a dam. The 2SLS and OLS estimates are very similar and suggest that 10 dams located upstream increase agricultural wages by roughly 0.02 percent. In contrast, wages did not increase in the districts where the dams were located. There appear to have been no economic force at play to compensate for the cost occurred because of the dam construction.

Functional form

Table 10 examines whether the number and the size of the dams affect their impact. Our instruments are too weak to allows estimation of several parameters. However, as the IV and OLS estimates are very similar we focus on OLS estimates.

In Panel A we introduce separately a dummy for whether there is at least one dam in a district, and a linear term in the number of dams present. The dummy for "at least one dam" is insignificant, and we observe no change in the coefficients of the number of dams (in magnitude or size). The effects of the number of dams on economic outcomes appears to be linear in the number of dams.

Panel B breaks up the number of dams variable into the number of small, medium and large dams. We find the effect of dams on poverty is driven by large dams, while the impact of dams on productivity is driven by medium dams. This is in line with the case study literature which suggests the negative impact of dams is the most pronounced for large dams.

Dams and Rainfall Shocks

It is often claimed that the seasonal nature of rainfall in India accentuates the importance of dams as a means of ensuring perennial irrigation and providing water security in the event of floods or droughts.

To examine this thesis, Table 11 examines the role of dams in mediating the effect of rain shocks on poverty and agricultural outcomes. We use annual rainfall data for Indian districts over the period 1971-1999 to construct our measure of rain shock.[29] Our rain shock measure is the fractional deviation of annual rainfall from the district's historical average.

Panel A, Table 11 shows that positive rain shocks improve rural welfare. In addition, we find that, for a given rain shock, dams constructed upstream to a district have a stabilizing effect. That is, having a dam upstream reduces the adverse effect of a bad rain shock on rural welfare. In contrast, dams built in own district are destabilizing.

In Panel B we examine agricultural outcomes. Rain shocks don't have a significant effect on agricultural wages or gross irrigated area. However, a positive rain shock significantly raises agricultural production, and once again we find that dams constructed upstream have a stabilizing effect. This effect is not present for dams built in own district.

Alternative neighborhood measures

In Table 12 we examine whether the impact of dam construction in a district extends to districts other than its immediate downstream neighbor. For brevity, we only report the 2SLS estimates.

In Panel A we examine all neighboring districts. The effect of dams constructed downstream to a district is negative but insignificant, while there is no effect in neighboring districts which are neither upstream or downstream. Panel B examines whether a district benefits from dam construction in districts which are upstream to its upstream neighbors. Here, we find no effect.

Districts characteristics: demographics and public goods

We use census data for the years 1971, 1981 and 1991 to examine whether dam construction affected population distribution, via urbanization or migration. We also examine the impact of dam construction on the fraction rural population working as agricultural laborers and

[29]The rainfall data-set, Terrestial Air Temperature and Precipitation: Monthly and Annual Time Series (1950-1999), Version 1.02 was constructed by Wilmott and Matsuura the Center of Climatic Research, University of Delaware. The rainfall measure for a longitude-latitude node combines data from twenty nearby weather stations using an interpolation algorithm based on the spherical version of Shepard's distance-weighting method.

cultivators. We focus on the rural male population.[30]

The results are in Table 12, panel A. We find no effect of dam construction on fraction rural male population or fraction rural population that are migrants; see columns (1)-(4), both panels. A number of authors have noted that tribal populations are disproportionately affected by dams. The 2SLS estimates suggest a reduction in the rural male tribal population in the district where the dam is constructed. While insignificant at the 10% level, it possibly reflects the greater displacement of the tribal population. Finally, we observe no impact of dam construction on the fraction male agricultural laborers or cultivators.

Panel B shows the relationship between dams and other public goods. Dams are not associated with any decline (or increase) in the availability of other publics goods in the dam's districts Joint test of significance shows that, overall, dams have no significant impact on public goods in their district or in the district downstream. However, dams are associated with a decline in the availability of (non-dam) water infrastructure in the downstream districts, which is consistent with the observation that there appears to have been some substitution of non-dam water related investment in districts where more hectares were irrigated through dams.

6 Conclusion

In this paper we have provided evidence on one particular form of infrastructure, large dam construction in India. We argued that any credible evaluation must address the endogeneity of dam construction. The problem of convincingly estimating the impact of large infrastructure projects extends beyond dams: the placement of any large public capital project, e.g. roads and railroads, reflects regional need and a complicated decision-making process, which makes an evaluation of their impact particularly difficult (Gramlich (1994) discusses these issues in the context of the US empirical literature on infrastructure and

[30]There are two reasons for this. First, female migration tends to for marriage reasons. Second, increases in female participation in agriculture is difficult to interpret. It may be due to higher returns or it may be that worse economic outcomes cause previously not working women to enter the labor force

productivity). While a growing cross-country literature finds that productive government spending enhances growth, these studies are unable to convincingly control for unobserved heterogeneity (see, for instance, Canning, Fay, and Perotti (1994) and Esfahani and Ramirez (ming).

This paper provides a potentially generalizable way of estimating returns to infrastructure investment, the use of geographic suitability for infrastructure. [31] It also points to the importance of identifying who the beneficiaries of investment projects are, and who bears the costs. We find a strikingly unequal sharing of costs and benefits of large dam construction. The results on poverty show that, in India, neither the market (through migration) nor the government (through transfers and compensation) are able to compensate for the inherent inequities in the physical benefits derived from a dams.

[31] Two studies in progress, on railroad in China Banerjee, Duflo, and Qian (2004) and highways in the US (Michaels) use a related approach, where they try to predict railroad or highway construction using the pattern that the grid would have had if it had connected all cities and treaty port (for China) or all big cities on a North-South and East-West axis (for the US).

7 Data Appendix

GIS Data

The data set used was the GTOPO30 (Elevation Data) downloaded from http://edcdaac.usgs.gov/gtopo30/gtopo30.html. The river map (Drainage-network) was downloaded from http://ortelius.maproom.psu.edu/dcw/ under file name 'dnnet'. this data was processed by CIESIN, Columbia University

Agricultural and Rural Welfare Outcomes

The districts in this data-set are defined according to 1961 boundaries. Kerala and Assam are the major agricultural states absent from the data set. Also absent, but less important agriculturally, are the minor states and Union Territories in the Northeastern part of India, as well as the far-northern states of Himachal Pradesh and Jammu-Kashmir. The data-set is downloadable at www-esd.worldbank.org/indian. The agricultural wages series provided is an annual measure of male agricultural wages constructed from monthly wage data collected by the Directorate of Economics and Statistics (Ministry of Agriculture, India). In constructing the annual measure, June and August were weighted more heavily to account for high intensity of field work during these months. The price deflator series was constructed from Indian Labor statistics by **?**) up to 1992, and further updated by **?**).

Rural Welfare Outcomes

In constructing the poverty measures, Topalova (2004) uses poverty lines proposed by Deaton. These are preferable to those used by the Indian Planning Commission, which are based on defective price indices over time, across states and between the urban and rural sector. Poverty lines were unavailable for some of the smaller states and union territories, namely: Arunachal Pradesh, Goa, Daman and Diu, Jammu and Kashmir, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, Tripura, Andaman and Nicobar Islands, Chandigarh, Pondicherry, Lakshwadweep, Dadra Nagar and Haveli. Most of these are excluded from our analysis as they have no dams or we lack other data for them. For those included, we used the neighboring states' poverty line.

References

Banerjee, A., E. Duflo, and N. Qian (2004). The railroad to success: The effect of infrastructure on growth in china. Mimeo, MIT.

Bank, W. (1994). *World Development Report 1994: Infrastructure for Development.* Washington D.C.: World Bank.

B.D.Dhawan (1989). *The Big Dams: Claims and Counter Claims.* New Delhi: Commonwealth.

Biswas, A. and C. Tortajada (2001, November). Development and large dams: A global perspective. *International Journal of Water Resources Development*, 9–21.

Canning, D., M. Fay, and R. Perotti (1994). Infrastructure and growth. In M. Baldassarri and L. Paganetto (Eds.), *International Differences in Growth Rates: Market Globalization and Economic Areas.* New York: St. Martin's Press, Inc.

Deaton, A. (2003a). Adjusted indian poverty estimates for 1999-2000. *Economic and Political Weekly*, 322–326.

Deaton, A. (2003b). Prices and poverty in india, 1987-2000. *Economic and Political Weekly*, 362–368.

Dreze, J. and A. Mukherjee (1991). Labour contracts in rural india: Theory and evidence. In S. Chakravarty (Ed.), *The Balance between Industry and Agriculture in Economic Development.* New York.

Esfahani, H. and M. Ramirez (forthcoming). Institutions, Infrastructure and Economic growth. *Journal of Development Economics*.

Goldsmith, E. and N.Hildyard (1984). *The Social and Environmental Effects of Large Dams.* United Kingdom: Wadebridge Ecological Centre.

Gopalakrishnan, E. (2000, November). ICOLD Comment on report of world commission of dams. www.dams.org/report/reaction/icold$_{i}ndia.htm$.

Gramlich, E. M. (1994). Infrastructure Investment: A review essay. *Journal of Economic Literature 32(3)*, 1176–1196.

Jain, L., K.Sundaram, and S.D.Tendulkar (1988). Dimensions of rural poverty: An interregional profile. *Economic and Political Weekly November.*

J.Dreze, M.Samson, and S.Singh (1997). *The Dam and the Nation: Displacement and Resettlement in the Narmada Valley.* New Delhi: Oxford University Press.

McCully, P. (2001). *Silenced Rivers: The Ecology and Politics of Large Dams.* London: Zed Books.

Merrouche, O. (2004). The impact of dam construction on poverty and inequality: Evidence from indian states, 1960-1992. Mimeo, European Institute Florence.

Murthi, M., P.V.Srinivasan, and S.V.Subramanian (2001). Linking indian census with national sample survey. *Economic and Political Weekly*, 783–793.

Roy, A. (1999). *The Cost of Living.* The Modern Library.

R.Rangachari, N.Sengupta, R.Iyer, P.Banerji, and S.Singh (2000). *Large Dams: India's Experience.* World Commission of Dams.

Sharma, V. (1991). Health aspects of water resources projects. In C. B. of Irrigation and Power (Eds.), *Environmental Impact Assessment of Water Resources Projects.* New Delhi.

Singh, S. (2002). *Taming the Waters: The Political Economy of Large Dams in India.* New Delhi: Oxford University Press.

Tehri (1997). *Report of Expert Committee on Rehabilitation and Environmental Aspects: Tehri Hydro- Electric Project: Volume II.*

Thakkar, H. (2000). Assessment of irrigation in india. World Commission of Dams.

Thukral, E. (1992). *Big Dams, Displaced People: Rivers of Sorrow, Rivers of Change.* New Delhi: Sage.

Topalova, P. (2004, November). Factor immobility and regional effects of trade liberalization: Evidence on poverty and inequality from indian districts. Mimeo, MIT.

World Commission on Dams (2000). *Dams and Development: A New Framework for Decision-Making*. EarthScan Publications.

Table 1: Features of some Indian Dams

Dam	Year	District		Reservoir area	Dam Height	Irrigation			Submergence and displacement			
		Name	Area (sq. km)	(sq.km)	(meters)	Districts (number)	Potential (mill. ha)	Waterlogged	Land (acres)	Villages	Persons	Tribal (% of total)
Hirakud, Orissa	1957	Sambalpur	6698	743	61	4	3	0.06	1,82,592	285	1,10,000	33
Nagarjuna Sagar Andhra Pradesh	1969	Nalgonda	14,240	285	181	7	1.1	0.11	30,731	26	28,000	36
Ukai, Gujarat	1972	Surat	7,657	520	80.5	3	3.86	0.008	NA	172	52,000	19
Pong, Himachal Pradesh	1974	Kangra	5,739	360	132	NA	1.6	NA	71, 724	94	1,50,000	56

Source: Thukral 1992 and India Case Study, World Commission of Dams

Table 2: Descriptive Statistics

	1973	1999
A. Dams		
Number of dams in district	2.65	7.70
Number of dams upstream to district	4.10	13.15
B. Welfare		
Log (pce)	3.78	5.75
Headcount ratio	0.47	0.24
Poverty gap	0.27	0.05
Gini	0.28	0.26
C. Agriculture		
Log (agricultural wage)	1.23	1.62
% cultivated area that is irrigated	25.52	42.78
Log (total production)	10.80	10.61
Log (total yield)	4.81	4.73

Table 3: Means of Changes in Dam Construction, Poverty and Production by State Dam Incidence and District River Gradient

	River Gradient in District		
	High	Low	Difference
	(1)	(2)	(3)
Panel A: Dams constructed b/w 1973-1999 (per 100 dams)			
High dam construction states	0.154	0.076	0.078
	(0.019)	(0.013)	(0.023)
Low dam construction states	0.01	0.003	0.007
	(0.200)	(0.110)	(0.003)
Difference	0.144	0.073	0.071
	(0.017)	(0.013)	(0.023)
Panel B: Change in head count ratio b/w 1973-1999			
High dam construction states	-0.2315	-0.235	0.0035
	(0.011)	(0.011)	(0.017)
Low dam construction states	-0.231	-0.189	-0.042
	(1.570)	(0.019)	(0.025)
Difference	-0.0005	-0.046	0.0455
	(0.019)	(0.022)	(0.030)
Panel C: Change in log production b/w 1970-1986			
High dam construction states	0.124	0.311	-0.187
	(0.046)	(0.034)	(0.010)
Low dam construction states	0.201	0.059	0.142
	(0.012)	(0.113)	(0.022)
Difference	-0.077	0.252	-0.329
	(0.014)	(0.017)	(0.024)

Note:

1. A state is a high dam construction state if it had more than a hundred dams by 1999, and is a low dam construction state otherwise. A district is a low gradient district if more than 90% of area along river has gradient less than 1.5%, and is a high gradient district otherwise.
2. Standard errors in parantheses.

Table 4: Geography and Dam Construction

	Dams	
	Poverty sample	Wage sample
	(1)	(2)
Dams in state*(Fraction river slope of 1.5-3%)	0.151	0.191
	(0.064)	(0.070)
Dams in state*(Fraction river slope of 3-5%)	-0.255	-0.286
	(0.096)	(0.094)
Dams in state*(Fraction river slope of 6-10%)	0.088	0.066
	(0.123)	(0.123)
Dams in state*(Fraction river slope above 10%)	0.189	0.236
	(0.083)	(0.100)
F-test for slope along river	2.632	3.402
	[0.032]	[0.008]
Dams in state*River length	0.006	0.007
	(0.005)	(0.004)
Dams in state*(Fraction district slope of 1.5-3%)	0.147	0.159
	(0.075)	(0.082)
Dams in state*(Fraction district slope of 3-5%)	0.036	-0.008
	(0.137)	(0.140)
Dams in state*(Fraction district slope of 6-10%)	-0.200	-0.060
	(0.259)	(0.249)
Dams in state*(Fraction district slope above 10%)	0.165	0.000
	(0.237)	(0.221)
Dams in state*(Fraction district elevation between 250-500 metres)	0.013	0.009
	(0.010)	(0.010)
Dams in state*(Fraction district elevation between 500-1000 metres)	0.027	0.023
	(0.009)	(0.009)
Dams in state*(Fraction district elevation over 1000 metres)	-0.135	0.102
	(0.242)	(0.345)
Dams in state*District area (square kilometers)	0.0001	0.0002
	(0.0004)	(0.0004)
R-squared	0.97	0.97
N	1870	5896

Notes:

1.All regressions include district fixed effects and a full set of state*year interactions.

2.Standard errors clustered by district are reported in parentheses.

3.The poverty sample includes the years of 1973, 1983, 1987, 1993 and 1999. The wage sample includes years 1971-1994.

4. River length is per 1000 kms and District area per 10,000sq. kms. All coefficients are multipled by 100.

Table 5: First Stage of 2SLS Regression

	Poverty sample			Wage sample		
	Own district		Upstream	Own district		Upstream
Predicted dams	(1)	(2)	(3)	(4)	(5)	(6)
Own district	1.001	1.323	0.170	0.836	0.870	-0.221
	(0.183)	(0.243)	(0.303)	(0.271)	(0.318)	(0.320)
Upstream		0.002	0.985		0.026	0.888
		(0.031)	(0.072)		(0.035)	(0.093)
N	1855	1855	1855	6072	6072	6072

Notes:

1. All regressions include the elevation, slope along district, river length, district area and upstream variables specified in Table 32 as additional controls. Regressions also include district fixed effects and a full set of state*year interactions.

2. Standard errors clustered by district are reported in parentheses.

3. The poverty sample includes the years of 1973, 1983, 1987, 1993 and 1999. The wage sample includes years 1971-1994.

Table 6: Dam Construction and Overall Agricultural Outcomes

	Irrigated area				Cultivated area				HYV area	
	Gross		Net		Gross		Net		five crops	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A. OLS (all coefficients multiplied by 100)										
Dams:										
Own district	-0.092	-0.190	-0.132	-0.232	0.013	-0.022	0.026	-0.029	0.102	0.246
	(0.413)	(0.369)	(0.386)	(0.331)	(0.066)	(0.061)	(0.049)	(0.054)	(0.370)	(0.386)
Upstream		0.654		0.615		0.001		-0.058		0.599
		(0.203)		(0.208)		(0.038)		(0.036)		(0.273)
PANEL B. 2SLS (all coefficients multiplied by 100)										
Dams:										
Own district	0.832	0.705	0.820	0.649	-0.160	-0.289	0.026	-0.177	1.308	0.493
	(1.468)	(1.636)	(1.413)	(1.554)	(0.250)	(0.272)	(0.215)	(0.213)	(1.330)	(1.054)
Upstream		1.043		1.089		-0.048		-0.069		0.809
		(0.272)		(0.273)		(0.049)		(0.053)		(0.380)
N	6067	6067	6067	6067	6055	6055	6055	6055	4493	4493

Notes

1. All regressions include elevation, slope along district, river length and district area variables specified in Table 3 as additional controls. Regressions also include district fixed effects and a full set of state*year interactions.

2. Standard errors clustered by district are reported in parentheses.

3. Dependent variables are in logs. All coefficients are multiplied by 100. Irrigated and cultivated area regressions are for 1971-1994, the others are for 1971-1987. The crops included in production and yield regressions are wheat, rice, jowar, sugarcane

Table 7: Dam Construction and Overall Agricultural Outcomes

	Production		Yield		Fertilizer Use	
	nineteen crops				five crops	
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A. OLS (all coefficients multiplied by 100)						
Dams:						
Own district	-0.205	-0.255	-0.204	-0.244	0.383	0.597
	(0.208)	(0.224)	(0.202)	(0.209)	(0.296)	(0.311)
Upstream		0.218		0.169		0.231
		(0.121)		(0.106)		(0.227)
PANEL B. 2SLS (all coefficients multiplied by 100)						
Dams:						
Own district	0.085	-0.360	-0.503	-0.708	-1.026	-0.962
	(0.778)	(0.603)	(0.594)	(0.538)	(1.222)	(1.193)
Upstream		0.373		0.385		0.516
		(0.173)		(0.170)		(0.261)
N	4537	4537	4537	4537	4521	4521

Notes

1. All regressions include elevation, slope along district, river length and district area variables specified in Table 3 as additional controls. Regressions also include district fixed effects and a full set of state*year interactions.

2. Standard errors clustered by district are reported in parentheses.

3. Dependent variables are in logs. All coefficients are multiplied by 100. Regressions cover 1971-1987. The crops included in production and yield regressions are wheat, rice, jowar, sugarcane, groundnut, bajra, maize, gram, tur, other pulses, barley, tobacco, ragi, sesamum, cotton, potato, jute, soy, and sunflower. The HYV area includes HYV area for the crops rice, wheat, bajra, jowar and maize.

Table 8: Dam Construction and Cropwise Agricultural Outcomes

	Non-water intensive crops				Water intensive crops			
	All	Millet	Pulse	Wheat	All	Sugar	Cotton	Rice
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PANEL A. AREA CULTIVATED (all coefficients multiplied by 100)								
A.1 OLS								
Dams								
Own district	-0.53	0.16	-0.16	-0.27	0.14	0.72	1.99	-0.40
	(0.36)	(0.38)	(0.44)	(0.36)	(0.23)	(0.53)	(0.86)	(0.35)
Upstream	0.29	0.38	0.11	0.26	0.04	0.64	-0.02	0.33
	(0.15)	(0.17)	(0.22)	(0.18)	(0.14)	(0.37)	(0.42)	(0.21)
A.2. 2SLS								
Own district	-0.52	1.10	-1.22	-1.23	1.28	2.91	3.07	0.68
	(0.89)	(1.09)	(1.13)	(1.05)	(1.02)	(2.54)	(3.16)	(0.96)
Upstream	0.11	0.30	-0.18	0.34	0.41	1.38	0.19	0.79
	(0.18)	(0.24)	(0.30)	(0.21)	(0.22)	(0.44)	(0.50)	(0.29)
N	4530	4480	4464	4104	4489	4264	2790	4373
PANEL B. YIELD (all coefficients multiplied by 100)								
B.1. OLS								
Own district	-0.044	0.104	-0.136	0.226	-0.134	0.132	-0.397	-0.173
	(0.259)	(0.478)	(0.268)	(0.219)	(0.322)	(0.269)	(0.518)	(0.358)
Upstream	0.115	0.270	-0.085	0.142	0.247	0.064	-0.383	-0.196
	(0.134)	(0.182)	(0.086)	(0.103)	(0.207)	(0.105)	(0.254)	(0.138)
B.2. 2SLS								
Own district	-0.496	-0.567	0.280	0.173	-1.086	0.475	0.270	-1.634
	(0.612)	(0.828)	(0.502)	(0.876)	(0.926)	(0.838)	(1.371)	(0.977)
Upstream	0.238	0.471	-0.060	0.290	0.322	0.114	-0.309	-0.176
	(0.200)	(0.275)	(0.119)	(0.157)	(0.252)	(0.147)	(0.388)	(0.187)
N	4528	4476	4456	4073	4482	4252	2582	4359
PANEL C. PRODUCTION (all coefficients multiplied by 100)								
C.1. OLS								
Own district	-0.603	0.266	-0.343	-0.026	-0.015	0.719	1.345	-0.627
	(0.428)	(0.554)	(0.416)	(0.425)	(0.400)	(0.636)	(1.112)	(0.419)
Upstream	0.322	0.646	0.004	0.443	0.274	0.551	-0.242	0.154
	(0.198)	(0.235)	(0.208)	(0.219)	(0.198)	(0.394)	(0.517)	(0.215)
C.2 2SLS								
Own district	-1.003	0.565	-0.977	-1.054	0.159	4.192	2.469	-0.929
	(1.114)	(1.442)	(1.112)	(1.375)	(1.294)	(2.847)	(3.201)	(1.111)
Upstream	0.245	0.767	-0.247	0.672	0.720	1.335	0.140	0.630
	(0.249)	(0.336)	(0.290)	(0.292)	(0.235)	(0.444)	(0.722)	(0.318)
n	4531	4483	4456	4085	4497	4362	2587	4361

Notes:

1. All regressions include the elevation, slope along district, river length and district area variables specified in Table 2 as additional controls. All dependent variables are in logs. Regressions also include district fixed effects and a full set of state*year interactions.

2. Standard errors clustered by district are reported in parentheses.

3. Area, yield and production variables are in logs.

Table 9: Rural Poverty and Agricultural Wages

	Per capita expenditure		Head count ratio		Poverty gap		Gini coefficient		Agricultural wages	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A. OLS (all coefficients multiplied by 100)										
Dams										
Own district	-0.266	-0.316	0.267	0.277	0.083	0.081	0.006	-0.001	0.067	0.179
	(0.135)	(0.127)	(0.090)	(0.089)	(0.029)	(0.029)	(0.032)	(0.030)	(0.217)	(0.218)
Upstream		0.038		-0.051		-0.021		-0.005		0.207
		(0.053)		(0.036)		(0.012)		(0.013)		(0.098)
PANEL B. 2SLS (all coefficients are multiplied by 100)										
Dams										
Own district	-0.323	-0.767	0.476	0.786	0.142	0.234	-0.020	-0.020	-0.082	0.746
	(0.598)	(0.448)	(0.329)	(0.287)	(0.092)	(0.083)	(0.124)	(0.106)	(0.815)	(0.849)
Upstream		0.062		-0.093		-0.033		-0.007		0.221
		(0.070)		(0.048)		(0.016)		(0.019)		(0.151)
N	1799	1799	1799	1799	1799	1799	1794	1794	6072	6072

Notes:

1. All regressions include the elevation, slope along district, river length, district area and upstream variables specified in Table 3 as additional controls. Regressions also include district fixed effects and a full set of state*year interactions.

2. Standard errors are reported in parentheses. These are clustered by 1973 NSS region*year in columns (1)-(8), and by district in columns (9) and (10).

3. The poverty regressions include the years of 1973, 1983, 1987, 1993 and 1999. The wage regression includes years 1971-1994.

4. Per capita expenditure and agricultural wages are in logarithms.

Table 10: Non-linearities in the impact of dams: OLS estimates

	Headcount ratio		Poverty Gap		Production	
	(1)	(2)	(3)	(4)	(7)	(8)
PANEL A: NUMBER OF DAMS (all coefficients multiplied by 100)						
At least one dam	-0.029	-0.033	-0.008	-0.010	0.019	0.017
	(0.023)	(0.022)	(0.008)	(0.007)	(0.042)	(0.042)
Number of dams	0.265	0.274	0.083	0.079	-0.207	-0.273
	(0.089)	(0.090)	(0.029)	(0.030)	(0.208)	(0.222)
At least one dam upstream		-0.020		-0.012		0.016
		(0.024)		(0.009)		(0.066)
Number of dams upstream		-0.051		-0.021		0.207
		(0.036)		(0.012)		(0.117)
PANEL B: MEDIUM AND LARGE DAMS (all coefficients multiplied by 100)						
Number small	0.197	0.090	0.071	0.032	-0.233	-0.268
(less than 16 m)	(0.136)	(0.140)	(0.047)	(0.049)	(0.276)	(0.287)
Number medium	0.123	0.259	0.023	0.048	0.069	0.193
(16-30 m)	(0.227)	(0.222)	(0.072)	(0.074)	(0.489)	(0.587)
Number large	1.678	1.749	0.623	0.631	-0.035	0.056
(over 30m)	(0.814)	(0.835)	(0.283)	(0.283)	(1.148)	(1.344)
Number small dams		-0.030		0.021		-0.354
upstream		(0.108)		(0.040)		(0.223)
Number medium dams		-0.048		-0.045		0.785
upstream		(0.127)		(0.043)		(0.283)
Number high		0.262		-0.036		-0.320
upstream		(0.595)		(0.208)		(1.360)
N	1799	1799	1799	1799	4537	4537

Notes:

1. All regressions include the elevation, slope along district, river length, district area and upstream variables specified in Table 3 as additional controls. Regressions also include district fixed effects and a full set of state*year interactions.

2. Standard errors are reported in parentheses. These are clustered by 1973 NSS region*year in columns (1)-(6, and by district in c

3. "Atleast one dam" is a dummy which equals one if a district has atleast one dam. "Average dam is medium" is a dummy which equals one if average dam height in district is between 18.5 and 33.8 meters; "average dam is high" is a dummy which equals one if average dam height in the district exceeds 33.8 meters.

4. Per capita expenditure and agricultural wages are in logarithms.

Table 11 : Dams and Rainfall Shocks

PANEL A: POVERTY	Per capita expenditure			Headcount ratio			Poverty gap		
	OLS		2SLS	OLS		2SLS	OLS		2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rainshock	0.074	0.049	0.054	-0.060	-0.045	-0.052	-0.018	-0.009	-0.007
	(0.025)	(0.030)	(0.032)	(0.019)	(0.023)	(0.025)	(0.007)	(0.008)	(0.009)
Dams		-0.260	-0.677		0.232	0.720		0.059	0.195
		(0.129)	(0.451)		(0.093)	(0.282)		(0.030)	(0.081)
Dams*Rainshock		0.312	0.351		-0.244	-0.227		-0.131	-0.181
		(0.144)	(0.235)		(0.113)	(0.166)		(0.045)	(0.080)
Upstream Dams		0.026	0.027		-0.043	-0.069		-0.020	-0.024
		(0.055)	(0.074)		(0.038)	(0.051)		(0.012)	(0.016)
Upstream Dams*		-0.086	-0.157		0.063	0.116		0.015	0.035
Rainshock		(0.044)	(0.070)		(0.039)	(0.060)		(0.013)	(0.022)
N	1799	1799	1799	1799	1799	1799	1799	1799	1799
PANEL B: AGRICULTURE	Agricultural wages			Gross irrigated area			Total production		
Rainshock	-0.025	-0.046	-0.043	0.084	0.073	0.123	0.122	0.077	0.146
	(0.023)	(0.025)	(0.029)	(0.040)	(0.042)	(0.053)	(0.035)	(0.040)	(0.051)
Dams		0.188	0.753		-0.156	0.810		-0.140	-0.322
		(0.217)	(0.848)		(0.368)	(1.632)		(0.223)	(0.592)
Dams*Rainshock		0.317	0.286		0.752	0.042		1.212	0.236
		(0.156)	(0.257)		(0.260)	(0.306)		(0.343)	(0.333)
Upstream Dams		0.201	0.215		0.650	1.043		0.168	0.328
		(0.098)	(0.151)		(0.204)	(0.274)		(0.111)	(0.169)
Upstream Dams*		-0.022	-0.023		-0.180	-0.089		-0.333	-0.270
Rainshock		(0.043)	(0.055)		(0.077)	(0.076)		(0.103)	(0.097)
N	6072	6072	6072	6067	6067	6067	4537	4537	4537

Notes:

1. All regressions include the elevation, slope along district, river length and district area variables specified in Table 3 as additional controls. Regressions also include district fixed effects and a full set of state*year interactions.

2. Standard errors reported in parentheses. In panel A these are clustered by NSS region*year, while in panel B these are clusted by district.

3. Rainshock is the fractional deviation of the district's rainfall from the district mean (computed over 1973-1999). All agricultural variables are in logarithms. All coefficients multipled by 100.

Table 12: Dam construction and Neighborhood Effects: 2SLS estimates

	Rural welfare				Agriculture			
						Production		
	Per capita expenditure	Head count ratio	Poverty gap	Agricultural wage	Gross irrigated area	Total	Non-water intensive crops	Water intensive crops
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PANEL A. ALL NEIGHBORING DISTRICTS (all coefficients multiplied by 100)								
Dams								
Own district	-0.732	0.739	0.208	-0.309	2.238	-0.030	-0.544	0.762
	(0.541)	(0.327)	(0.100)	(0.890)	(1.540)	(0.707)	(1.037)	(1.403)
Upstream	0.064	-0.096	-0.033	0.201	1.034	0.339	0.182	0.721
	(0.075)	(0.050)	(0.016)	(0.160)	(0.267)	(0.169)	(0.256)	(0.215)
Downstream	-0.111	0.075	0.025	0.046	-0.449	0.007	0.969	-0.831
	(0.099)	(0.088)	(0.026)	(0.242)	(0.424)	(0.250)	(0.354)	(0.414)
Neighboring but not	0.031	-0.053	-0.020	-0.070	-0.137	-0.088	-0.359	0.276
upstream/downstream	(0.100)	(0.080)	(0.027)	(0.144)	(0.270)	(0.148)	(0.239)	(0.272)
PANEL B. NEIGHBORS OF NEIGHBORING DISTRICTS (all coefficients multiplied by 100)								
Dams								
Own district	-0.893	0.866	0.252	0.583	1.500	-0.461	-1.261	0.073
	(0.403)	(0.275)	(0.085)	(0.843)	(1.716)	(0.620)	(1.116)	(1.180)
Upstream	0.023	-0.054	-0.010	0.071	1.062	0.542	-0.232	0.819
	(0.125)	(0.105)	(0.035)	(0.232)	(0.445)	(0.280)	(0.415)	(0.382)
Upstream to upstream	0.006	-0.017	-0.012	0.112	0.018	-0.088	0.321	-0.094
districts	(0.064)	(0.060)	(0.021)	(0.101)	(0.263)	(0.122)	(0.203)	(0.188)
N	1799	1799	1799	6072	6067	4537	4531	4497

Notes:

1. All regressions include the elevation, slope along district, river length, district area and upstream and downstream variables specified in Table 3 as additional controls. Regressions also include district fixed effects and a full set of state*year interactions.

2. Standard errors are reported in parentheses. These are clustered by NSS region*year in columns (1)-(3), and by district in columns (4)-(8).

3. All agricultural outcomes, agricultural wages and per capita expenditure are in logs. All coefficients multiplied by 100.

Table 13: Census outcomes (population and public goods): 2SLS estimates

	Rural population	Rural migrants	Rural SC/ST population	Rural agricultural laborers	Rural cultivators
	(1)	(2)	(3)	(4)	(5)
PANEL A: POPULATION					
Dams					
Own district	7.364	-3.766	-21.840	5.072	8.789
	(14.950)	(12.915)	(16.630)	(4.537)	(6.680)
Upstream	0.563	1.685	2.682	0.194	-0.397
	(1.280)	(2.791)	(2.495)	(0.828)	(1.367)
N	952	947	947	947	947

PANEL B: PUBLIC GOODS

	Any education facility	Any medical facility	Any water facility	Any power	Any tarmac road
Own district	0.152	0.112	0.058	0.296	-0.105
	(0.206)	(0.328)	(0.063)	(0.352)	(0.235)
Upstream	0.016	0.021	-0.021	0.166	0.052
	(0.050)	(0.082)	(0.012)	(0.081)	(0.071)
N	849	849	849	850	850

Notes:

1. All regressions include the elevation, slope along district, river length, district area and upstream variables specified in Table 3 as additional controls. Regressions also include district fixed effects and a full set of state*year interactions.

2. Standard errors are reported in parentheses. These are clustered by district.

3. The regressions include 1971, 1981 and 1991. In panel A, the dependent variables refer to the male population. Rural population is reported as the percentage of district population, and all other dependent variables are the percentage of rural male population. In panel B, the dependent variable is the percentage of villages in the district with the public good facility. All coefficients mutiplied by 100.

3. Water facilities in column 3 do not include dams