ERF Working Papers Series

From Flow to Fortune? The Economic Impacts of Agricultural Dams

Hans R.A. Koster, Saeed Tajrishy,
Jos van Ommeren and Mohammad Vesal

Economic Research Forum
منتدى البحوث الاقتصادية

Working Paper No. 1719
August 2024

FROM FLOW TO FORTUNE? THE ECONOMIC IMPACTS OF AGRICULTURAL DAMS

Hans R.A. Koster, [*] Saeed Tajrishy, [†] Jos van Ommeren[‡] and Mohammad Vesal[§]

Working Paper No. 1719

August 2024

We thank seminar participants at VU and Birmingham University for useful insights.

This paper was originally presented during the ERF 30th Annual Conference on "Tragedies of Regional Conflicts and Promises of Peacebuilding: Responding to Disruptors and Enablers of MENA Development Pathway", April 21-23, 2024.

Send correspondence to:
Mohammad Vesal
Sharif University of Technology
m.vesal@sharif.edu

[*] Department of Spatial Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, email: h.koster@vu.nl. Hans is also research fellow at the Tinbergen Institute and affiliated to the Centre for Economic Policy Research.
[†] Department of Economics, Sharif University of Technology, Iran, email: saeed.tajrishy@gsme.sharif.edu
[‡] Department of Spatial Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, email:jos.van.ommeren@vu.nl. Jos is also research fellow at the Tinbergen Institute.
[§] Department of Economics, Sharif University of Technology, Iran, email: m.vesal@sharif.edu

First published in 2024 by
The Economic Research Forum (ERF)
21 Al-Sad Al-Aaly Street
Dokki, Giza
Egypt
www.erf.org.eg

Abstract

We examine the economic effects of agricultural dams for Iran. To estimate the benefits of having a larger dam, we introduce a novel triple difference identification strategy using information on rainfed and irrigated land. The benefits for landowners and agricultural workers are captured by a range of measures (e.g. land rent, profit, agricultural output, wages). We demonstrate that increasing water volume of dams that are upstream creates substantial benefits for landowners, but not for agricultural workers. About half of agricultural dams are profitable. There are substantial economies of scale in volume of dams, but as the benefits of volume are more strongly diminishing, so benefit to cost ratios is smaller for larger dams.

Keywords: Dams, Iran, Agriculture, Land rent, Infrastructure, Development
JEL Classifications: H54, O13, O20, Q15, Q19

ملخص

نقوم بدراسة الآثار الاقتصادية للسدود الزراعية على إيران. لتقدير فوائد وجود سد أكبر، ونقدم استراتيجية جديدة لتحديد الاختلافات الثلاثية باستخدام معلومات عن الأراضي التي تروى بواسطة الأمطار والمروية بواسطة مصادر مياة أخرى. وتستفيد استحقاقات ملاك الأراضي والعمال الزراعيين من مجموعة من التدابير (مثل إيجار الأراضي، والربح، والإنتاج الزراعي، والأجور). نثبت أن زيادة حجم المياه في السدود المنبثقة عن المنبع تخلق فوائد كبيرة لملاك الأراضي، ولكن ليس للعمال الزراعيين، فحوالي نصف السدود الزراعية مربحة. هناك وفورات كبيرة في حجم السدود، ولكن نظرًا لأن فوائد الحجم تتضاءل بقوة أكبر، فإن نسب الفائدة إلى التكلفة أصغر بالنسبة للسدود الأكبر.

1 Introduction

About 30% of irrigated agricultural land worldwide is fed by dams (World Commission on Dams 2000). Nevertheless, we know very little about their effectiveness (Jeuland 2020).[1] Hence, even a basic question, such as whether it is welfare improving to construct new irrigation dams to improve productivity in the agriculture sector is basically unanswered. The benefits of constructing agricultural dams have been contested as dams are associated with increases in malaria, the forced removal of population and environmental degradation (Mettetal 2019) and exacerbated water scarcity problems in non-agricultural sectors of the economy (Mesgaran et al. 2017). The ambiguous effects of agricultural dams sharply contrast to the results for *hydropower* dams, for which positive economic effects have been convincingly demonstrated (Lipscomb et al. 2013).

We focus on measuring the economic effects of agricultural dams for Iran, where there is a heated debate about the effects of dams. These dams affect local communities through several markets, including the land market through altering of land rents.

Land rent is, according to economic theory, one of the first-order indicators of local economic benefits, as efficiency gains in agriculture are thought to be largely captured by landowners, an idea developed in the 19th century by Ricardo (1809) and Von Thünen (1842). Using agricultural land rent as a proxy for welfare has become standard in economics and relies on the assumption of absentee landowners (Palmquist 1989). Agricultural land in Iran is typically not owned by absentee landowners who rent out the land to farmers, but owned by local farmers so efficiency gains may also end up in profits of farmers. We therefore supplement our estimates of the effects of dams on agricultural rent with effects on farmers' profits. We also examine the effects on other measures such as changes in irrigated and rainfed land size.[2]

We also examine the effect of dams on local wages, which is another indicator of increases in local economic benefits, and on changes in population, which captures migration. In competitive labour markets with migration, the effect of a local productivity increase on local wages are

[1] Note that Jeuland (2020) documents an impressive number of case studies of the effects of irrigation dams, but these studies are silent on the average effect we are interested in.

[2] We ignore issues related to health and the impacts on the local environment (Lewis et al. 2008, Provencher et al. 2008, Bohlen & Lewis 2009, Lipscomb et al. 2013). As these issues typically refer to additional costs induced by dams, it is plausible that we somewhat overestimate the net benefits of dams.

typically small, because of migration of workers, which increases labour supply. Rural labour markets in Iran are characterised by temporary employment with migrating workers, so, a priori, one expects minor increases in wages, and stronger effects on migration.

This paper contributes to a literature that aims to quantify the economic effects of *agricultural dams*. We employ a panel of highly spatially disaggregated data derived from village-level farmer surveys in Iran from 2000 to 2018 and administrative population data from 1976 to 2016. As emphasised by Strobl & Strobl (2011), finding those data over long periods for developing countries is challenging. Our analysis is at the rather detailed county level, which on average contains only about 10,000 households.

The paper aims to make 5 distinct contributions to the existing literature. First, we introduce a triple-difference method to address endogeneity issues using information on the difference in outcome measures (*e.g.*, in rent) between irrigated and rainfed lands within the same (small) area. In essence, we exploit that rainfall is the main water resource for rainfed agriculture, whereas stored water run-off from rivers, dams, lakes and aquifers is the main water source for irrigated agriculture.

We emphasise here that whether or not land is suitable for irrigation or rainfed is to a large extent exogenously determined by physical land features (type of soil and slope of land) and the quantity of rainfall (Worqlul et al. 2017).[3] Land owners with land suitable for irrigated agriculture face no concerns about the costs of regulated water, given its almost zero price. However, the challenge lies in the frequent scarcity of regulated water supply, especially during the summer. Owners of irrigated land will benefit from dams, whereas owners of rainfed land will not. Consequently, the main identifying assumption of the triple-differences method is that rainfed land is not *directly* affected by dams, whereas irrigated land, benefits from dams. We test this assumption by estimating standard difference-in-differences models where we show that outcome measures of rainfed land are not affected by the construction of dams in Table **??** in the Appendix. We will also discuss potential general equilibrium effects that may affects rainfed and irrigated land differently and the extent to which dams stimulate farmers to

[3] In line with that we will show that the quantity of irrigated land after the building of dams increases by about 10%, whereas there is no reduction in rainfed land, suggesting that rainfed land is rarely converted to irrigated land, and that the increase in irrigated land is almost completely coming from land not used for agriculture before the building of dams.

converse rainfed land into irrigated land.

Second, one contribution of the paper relates to the quality of our data that enable us to look at alternative and improved outcome measures, notably agricultural land rent, which is a common proxy for welfare, and information on farmers' income. Furthermore, as we have collected agricultural and population data at the county level, this enables us to provide evidence on the effect of dams at a more disaggregated level than previous studies, avoiding issues related to the use of large administrative areas. For example, one issue is that political factors play an important role in the placement of dams, which causes an omitted variable bias. By using a finer geographical unit at the county level, while controlling for district by year fixed effects, we address this issue as political factors important enough to influence the placement of dams are relevant at the district level, but not at the county level.

Third, it has been documented before that there is extreme variation in dam volume, also within countries. This also applies to Iran. For example, its largest dam has a volume of about 8 billion m^3, almost hundred times larger than the average volume. This motivates us to estimate the economic effects of dam volume instead of the effects of the number of dams, which is the focus in the extant literature.

Fourth, we distinguish between short-run and long-run estimates and demonstrate that approaches that use long differences (*e.g.*, 20 years) typically imply considerably higher estimates than those that are based on short-term differences. This makes it plausible that the long-run effects of dams are much more pronounced than suggested by the current estimates in the literature, which are essentially a mix of short and long-run estimates.

Our final contribution is that we improve upon the IV strategy introduced by Duflo & Pande (2007) and applied among others by Strobl & Strobl (2011), Lipscomb et al. (2013), Blanc & Strobl (2014) and Mettetal (2019). We show its limitations, at least in the context of Iran, which is one of the reasons we introduce an alternative strategy.

Related literature. Our study relates to a growing, but still small, economic literature that essentially starts with the seminal paper by Duflo & Pande (2007) for India. This paper introduced a novel instrumental variable approach employing information on river suitability, based on river gradient. They show that in districts located downstream from a dam, agricultural production slightly increases and vulnerability to rainfall shocks declines. While the contribution of this paper is unquestioned, this paper does have certain limitations that warrant consideration.

First, the study employs number of dams which is a rather crude measure of the amount of regulated water. Follow-up papers partially have addressed this issue by distinguishing between the effects of the number of small and large dams (see Strobl & Strobl 2011, Blanc & Strobl 2014).[4] An important finding of these papers is that the size of dams matters. For example, they make the plausible case that small dams increase cropland productivity downstream, but large dams reduce cropland productivity within their close proximity (Blanc & Strobl 2014). Moreover, large dams are more likely to have negative social, health and environmental consequences (Kitchens 2013). Mettetal (2019) shows that large dams in South Africa increase infant mortality, water pollution and reduce water availability. One interpretation of these papers is that the construction of small dams is more likely to be welfare improving than large ones, but this ignores economies of scale in building and ignores that large dams can play a potentially important role in enhancing the positive impact of local small dams because they can provide regulated water to small dams (Blanc & Strobl 2014).

Second, Duflo & Pande (2007) employ information at the level of districts in India, which are very large, so they are unlikely to capture the localised effects of smaller agricultural dams. In the study, upstream/downstream districts are defined based on administrative borders, rather than based on geographical features such as topographic elevation and river networks, which may create substantial measurement error. We will therefore follow the later literature, see Strobl & Strobl (2011), which addresses this issue.

Third, according to the engineering literature, river suitability for dam construction depends on *water flow accumulation*, which is defined by a combination of river width and river gradient (Andreadis et al. 2013). However, in the economic literature, river gradient is used to construct the instrumental variable Duflo & Pande (2007), Strobl & Strobl (2011), Lipscomb et al. (2013), Blanc & Strobl (2014) and Mettetal (2019). We improve the literature by using flow accumulation, and by showing that the instrument is indeed much stronger.

The plan for the remainder of the paper is as follows. In Section 2, we discuss the context of

[4] These studies still ignore important variation in the size of dams and also consider the threshold for belonging to the large or small dam category seemingly arbitrarily.

the study, followed by a discussion of the data in Section 3. Section 4 outlines the empirical methodology, while the results are reported in Section 5. Section 6 provides a back-of-the-envelope calculation to relate the overall benefits of dams to the costs. Section 7 concludes.

2 Dam construction and agriculture in Iran

In Iran, about 18% of workers are employed in the agricultural sector generating 13% of the national income. Iran aims to achieve food self-sufficiency. Its agricultural sector supplies about 90% of domestic food consumed, but at the same time uses about 92% of available freshwater, which is way above the global average of about 70% (Mesgaran et al. 2017, Ritchie & Roser 2017). Iran has a semi-arid climate, with many deserts, hence only 9.5% of Iran's land area is cultivated.[5] The majority of agriculture takes place in the west, north-west, and northern parts of the country where annual precipitation exceeds 250 mm. This is regarded as the minimum threshold for rainfed farming. However, irrigated agriculture is present in regions with precipitations as low as 200 mm per year, or even below 100 mm per year. Consequently, dams have become key for agriculture.

Irrigation is much more common in Iran than in other countries: more than half of its farmland is irrigated, about twice the global average (Portmann et al. 2010).[6] In Iran, pressurised irrigation is rather uncommon, is about 95% of agricultural land utilises gravity irrigation. After the 1979 revolution, promotion of self-sufficiency through cultivation of strategic staple crops, such as wheat, has led to hefty subsidies in the agricultural sector, which in turn increased the demand for water (Madani et al. 2016). Water is nearly free in Iran, also for the agricultural sector, and, therefore, water cost is not a limiting factor for agricultural activities. The only limitation is the unavailability of water. As a result, the country currently faces serious water challenges, including rising water demand, declining groundwater levels, deteriorating water quality, and increasing threats to the environment and various ecosystems.

Before the 1979 revolution, Iran had only 14 large dams. Since then, policies focused on increasing water supply for agricultural production, primarily through the construction of

[5] Annual precipitation is around 225 mm – far below the global average of 860 mm – and can be as low as 50 mm in deserts but exceeds 1500 mm in the north of Iran (Saemian et al. 2022).

[6] By comparison, in Sub-Saharan Africa it is less than 5% of farmland, while the corresponding figure is about 10% for Latin America, 25% for the Near East and North Africa, about 35% for East Asia and about 40% for South Asia (Wani et al. 2009).

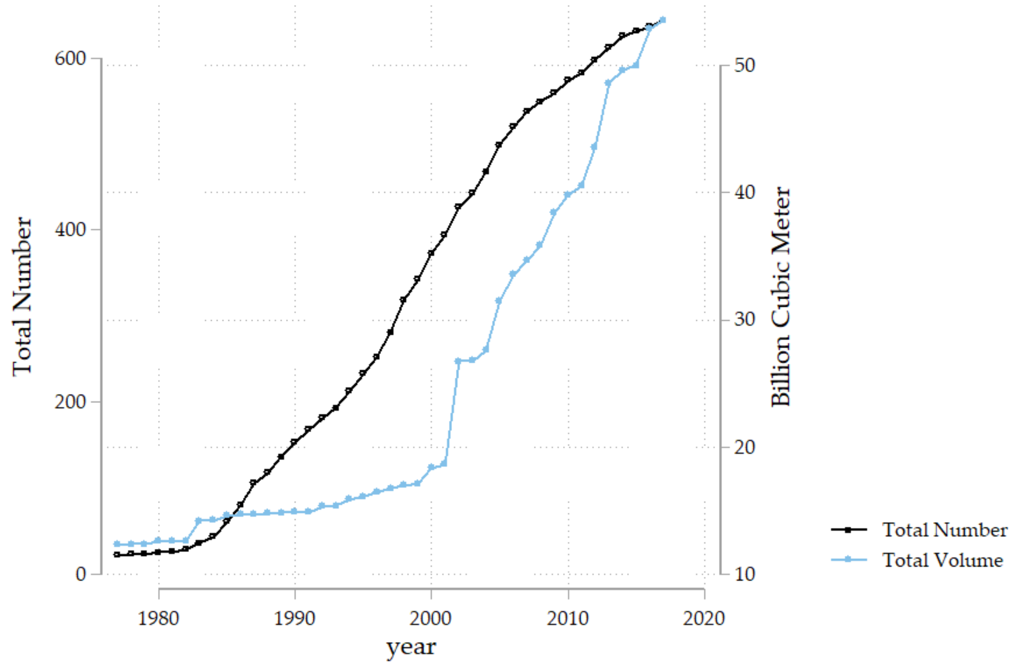

FIGURE 1: EVOLUTION OF DAMS IN IRAN DURING 1977-2017

dams (Yazdandoost 2016).[7] Iran is now the world's third leading country in terms of dam construction (Madani et al. 2016).

In Figure 1 we show the evolution of the number and volume of dams across Iran during 1976-2017. The number of dams has increased substantially, by about 660, since the 1980s. Until 2000, newly-constructed dams were rather small, but after 2000, mostly large dams were built tripling dam volume within a rather short period. The large majority, more than 80%, are (almost) entirely for agricultural purposes. Dam construction is far from over: there are still about 125 dams under construction and 500 dams under study to be constructed. In recent years, however, dam construction has slowed down, likely due to concerns about their negative impacts for example through environmental degradation.

Two ministries, *i.e.*, the Ministries of Energy and of Agricultural Jihad, are in charge of constructing dams: both ministries build about half of all dams. Agriculture stands as the primary reason behind the existence of nearly all dams and just a few dams do not have increasing agricultural yields as one of their main purposes. The Ministry of Agricultural Jihad builds smaller dams only for local agricultural use, with an average volume of 2 million m^3, while the Ministry of

[7]For the literature arguing that these policies have little concern for environmental impacts, see Tahbaz (2016).

Energy builds larger dams, with an average volume of 140 million m^3 for a range of purposes, including hydro-electricity, industry and drinking water, but usually also for agriculture. Both ministries are responsible for the construction of irrigation canals which might take longer than the construction of the dam itself.

3 Data

This paper combines hydrological and economic data to study the economic impact of dams. In the two following subsections, we discuss the hydrological, geographical and economic data.

3.1 Hydrological and Geographical Data

3.1.1 Dams

We have information on about 644 dams (out of 660 existing dams), which have a capacity of at least 30,000 m^3.[8] We have complete information about the purpose, opening year, volume and regulated water volume. Furthermore, we have (incomplete) information about height, irrigated area, lake area and agricultural water volume. We restrict our analysis to 596 dams reporting agriculture as one of their purposes. Almost all (about 94% on average) water is used for agriculture.[9] Figure 2 shows that dams are not evenly distributed across the country: most are in the north-western, western or eastern parts. There are also vast areas without dams, mostly in the central part of Iran, where rivers are non-existent.

Table 1 provides summary statistics. It shows that agriculture is not the only purpose. Flood control and drinking water provision are also common. For example, 24% of the dams also aim to improve flood control. Other purposes, such as hydro-electricity generation or supporting industry, do not play an important role. Dams have an average volume of 84 million m^3 and irrigate about 5 thousand hectares of downstream agricultural land. Variation in dam volume is extreme as its standard deviation is about 5 times the mean. The smallest dams have a volume of only 40 thousand m^3, whereas the largest have a volume of about 8 billion m^3. It takes on average 3.2 years to construct a dam. Small dams are obviously constructed much faster.

[8]This information is obtained through the Iran Water Resources Management Company within the Ministry of Energy. Some dams have missing information, which we complemented using public information available on internet.

[9]We know this percentage for the majority but not for all dams, so the exact percentage may be inaccurate.

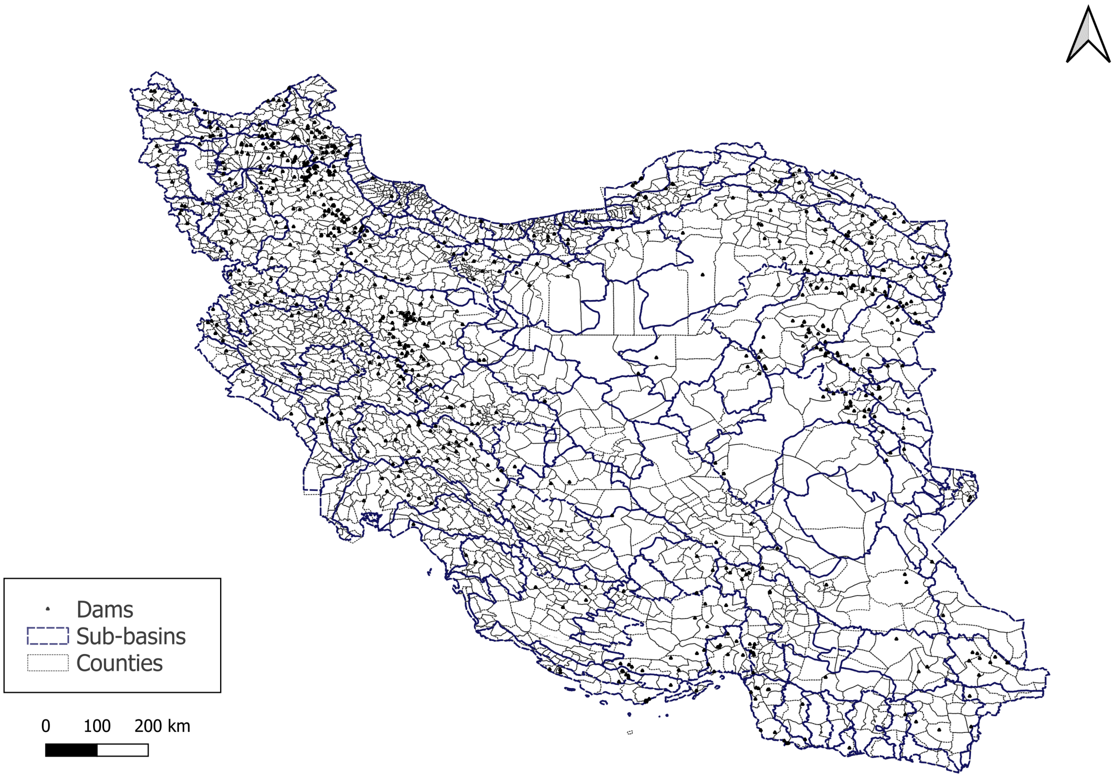

FIGURE 2: LOCATION OF IRRIGATION DAMS

Notes: This figure shows a map of Iran with dots indicating the location of irrigation dams. Lines depict polygons that delineate the boundary of counties.

3.1.2 Basins

We have collected geographical information about Iran's 30 major river basins, and 135 sub-basins, river kilometers, elevation (for cells of 30m by 30m), overall and river gradients, and flow accumulation using information from several sources (of Engineering 2012, JPL 2020, Andreadis et al. 2013).[10] Our econometric analysis later on will be at the level of counties (as defined for the year 2000), but we will supplement the estimates with analysis at the district level.[11] There are 2,194 counties with a median area of about 350 km^2, which is much (about 10 times) smaller than the areas defined in Duflo & Pande (2007) and follow-up studies. In Panel A of Table 2, we provide descriptive statistics on geography and hydrology characteristics at the level of counties. Counties are considerably smaller than sub-basins and it should be noted that sub-basins and

[10]The official sub-basin, as in of Engineering (2012), is defined to have one main river and has an average size of about 12000 km^2. Flow accumulation measures how much water flows in a geographical river cell of an area, see Figure A2 in the Appendix A.1.

[11]Each district in Iran consist of about 10 counties.

TABLE 1: DAMS SUMMARY STATISTICS

	Mean	Std. Dev.	Min	Max
Panel A. Additional purpose				
Flood control	0.24	0.48	0	1
Drinking water supply	0.17	0.37	0	1
Environment	0.13	0.32	0	1
Industry	0.07	0.25	0	1
Hydro-electrical	0.06	0.24	0	1
Touristic	0.02	0.12	0	1
Panel B. Dates				
Opening Year	1996.8	10.7	1951	2017
Construction period duration (*year*)	3.2	2.7	0	18
Panel C. Size and specifications				
Height (*m*)	27.1	28.4	3.6	205
Total irrigated Area (*hectare*)	5167.5	25496.7	6	364744
Lake area (*hectare*)	459.5	1719.2	0.35	21000
Dam Volume (*million* m^3)	90.1	476.9	0.04	7644.3
Regulated Water (*million* m^3)	65.9	335.8	0.04	4260
Agricultural Water (*million* m^3)	58.4	360.1	0.04	4260

Notes:This table shows summary statistics for 596 dams studying in this paper which are used for agricultural purpose. We do not have data on construction duration, total irrigated area and agricultural water for some of the dams in our sample.

counties or districts do not overlap since the latter are administrative or political units, while sub-basins refer to hydrological units. We identify how counties areas are interlinked in terms of water flow, *i.e.*, in terms of their upstream/downstream relationship using hydrological data following Strobl & Strobl (2011).[12] Throughout this study, we distinguish between upstream dams, downstream dams and 'own' dams, which are dams built in the own county.

As depicted in Figure 2, sub-basins vary greatly in shape and size. Most counties lie completely within a sub-basin which makes it easier to define upstream/downstream relationships.[13]

Panel B of Table 2 shows the *number* of own, upstream and downstream dams as well as the *volume* at the county level for the period from 2000 until 2017.[14] During this period, counties had on average 0.2 (own) dams, 6.3 upstream dams and 5.4 downstream dams. These data mask huge variation over time. In 1976, dams were almost absent and most counties (99%) had no dams.

In 2000, 12% had at least one dam and 4% at least two dams. Furthermore, 62%, have at least

[12]Hence we improve upon Duflo & Pande (2007) that use administrative areas to define these relationships, which is justified when catchment areas are contained within districts and one can classify districts as upstream/downstream to each other. This is generally not the case.

[13]Figure A1 shows an example of a sub-basin that contains several counties with rivers and a dam (shown by the triangle) upstream of counties 4, 9 and 10.

[14]The calculation of the volume of dams involves aggregating the volumes of individual dams dependent on whether they are upstream or downstream.

TABLE 2: DATA DESCRIPTION

		Mean	St. dev.	Min	Max
Panel A. Hydrology and geography					
Elevation (*m*)		1010.7	654.1	-16.4	2765.6
Slope (%)		5.8	4.9	1.0	30.1
River Length (*km*)		117.7	121.6	0.0	895.3
Area (km^2)		730.3	1319.6	11.1	20278.4
Distance to nearest big city (*km*)		93.3	62.6	2.6	422.7
Flow accumulation (km^2)		2190.8	18728.1	17.1	621447.8
Panel B. Dams					
Number	Own	0.2	0.5	0	8.1
	Upstream	6.3	11.1	0	66.1
	Downstream	5.4	11.3	0	87.8
Volume (*million* m^3)	Own	16.1	182.1	0	6219.4
	Upstream	298.2	1045.2	0	16054.4
	Downstream	562.5	1724.1	0	13860.8
Panel C. Agriculture and Population					
Share of irrigated land *(%)*		0.51	0.41	0	1
Irrigated	Rent	13.1	7.7	1.0	87.4
	Revenue	189.1	191.2	2.9	1761
	Profit	78.7	90.5	-72.4	1018
	Wage	452	70.2	299	837
Rainfed	Rent	3.9	4.3	0.1	51.4
	Revenue	122.1	189.1	0	5200
	Profit	45.3	91.2	-586	2097
	Wage	463	71.2	146	827
Population		9404.6	6917.4	187	81292
Rain (*mm/year*)		380.3	206.6	55.9	1183.3

Notes: We provide information on 2,194 counties. Panel A shows hydrological and geographical characteristics. In Panel B, we report the average of number and volume of own, upstream and downstream dams in each county spanning 2000 until 2018. Panel C shows agricultural variables, population and rain data. Agricultural rent, revenue, and profit are reported in million Rials per hectare and hourly wage is reported in thousand Rials and adjusted to 2018 prices. Agriculture data source is a survey at village level which is aggregated at county level resulting in an unbalanced panel sample between 2000 and 2017 for 2059 counties. Population data is from the Decennial Census from 1976 until 2016 which is available at the village level and aggregated at county level. For rain we use GSFC (2019).

one upstream dam By 2017, 20% of counties had at least one dam and 6% at least two dams. More importantly, in that year, the large majority of counties, 81%, have at least one upstream dam. The median distance of a county with its nearest upstream dam is about 84 km.

3.2 *Economic data*

We use 2 different datasets related to the farmer's expenditures and revenues, agricultural production and cultivated area, both obtained from the Ministry of Agricultural Jihad. The first dataset, summarized in Panel C of Table 2, is a farmer survey covering revenues and crop product costs in about 30 thousand villages across the country over 2000 to 2017.[15]

This rich dataset provides information on farmer costs and revenues for each product. We

[15] We do not have information on orchard products which cover 25 percent of cultivated area.

also know for each product whether it is rainfed or irrigated. Consequently, this allows us to calculate the rent (per m^2), wage, total labour costs, water costs, fertiliser costs, and farmer's revenues *separately for rainfed and irrigated land*.[16] We aggregate this information to 2,194 counties across 250 districts as we do not have information about the exact location of villages within counties.

In Panel C of Table 2, we show the descriptive statistics for the key outcome variables: land rent, labour wage, farmer profit and revenue, and population. As one expects, rents are substantially higher for irrigated than for rainfed land. On average, irrigated land rent is about 13 million Rials per m^2 (about $104 per m^2), whereas rainfed land rent is about 3.9 million Rials per m^2 (about $32 per m^2). Panel C also shows that about half of all land is irrigated. Furthermore, it appears that the wage for irrigated and rainfed land is almost the same. We supplement these data with census information on county level population for the years 1976, 1986, 1996, 2006, 2011, and 2016, so for a much longer period than the other data sources. A county hosts, on average, about 8 thousand people. Data on precipitation at county level is derived from NASA (GSFC 2019).

We do not have information about the size of the cultivated area or crop production at the county level, but we have this information at a more aggregated spatial level. The second dataset provides information on the size of the cultivated area and crop production *at the district level*, *i.e.*, across 225 districts from 1991 to 2015 (defined for the year 1991) with an average area of about 7,000 km^2. Combined with information on the crop price (which is district and year specific), we calculate the production value for 19 main crops separately for rainfed and irrigated products. This allows us to calculate for each year at the district level, the total cultivated area (in thousand hectares) and agricultural output (in million Rials).

4 Empirical framework

Our analysis exploits county panel data using variation over time in dam construction for identification of the causal effect of *upstream dam volume* on a range of outcomes in counties – land rent, profit (per unit of land), farmer revenue, population and wages, but with a focus

[16]We observe all of the farmer's costs including the amount of money paid to workers and the hours worked, so we calculate (hourly) labour wage for each farmer. In about 7% of counties, there is no information about agricultural costs and revenues due to the lack of agricultural activity.

on *rent* and *profit*.[17] Note that we use land rents rather than land prices, which is beneficial, because prices are forward looking, which complicates the analysis. In all analyses, we control for own and downstream dam volume, but we will see that controlling for these variables is not essential. Our preferred estimates rely on a triple-differences strategy, which applies a difference-in-differences (DID) strategy to the difference in outcomes between rainfed and irrigated land.

Before we discuss this strategy, we will discuss an alternative strategy, which boils down to a *standard DID strategy to outcomes of irrigated land*, for example the rent observed for irrigated land. Hence, we compare differences over time in several outcomes for irrigated land between counties that are treated by different volume of dams. Typically, as discussed extensively in the literature (see *e.g.* Duflo & Pande 2007, Strobl & Strobl 2011, Lipscomb et al. 2013, Blanc & Strobl 2014, Mettetal 2019), such a strategy would not be convincing, because it is subject to the concern that dam construction might be endogenous, as it is is plausible that governments favour the construction of dams at locations. For example, it might be the case that national governments tend to favour specific locations that exhibit stronger growth, introducing the issue that the outcomes observed for counties that receive dams may be on different trends.

However, this type of endogeneity issue can be easily addressed given our data as we employ information at a very spatially detailed level, i.e. at the level of counties, whereas the location of dams is politically decided at a much more aggregated level, i.e. the district level. Consequently, by including district-by-year fixed effects, we address endogeneity concerns in dam construction.[18]

There are also other reasons why several standard DID results shown by us are useful. These results will be used to test the identifying assumptions of the triple-differences strategy. Furthermore, for certain outcomes that do not refer to land, such as population, our triple-differences strategy is not feasible, so we rely on a standard DID (without distinguishing between rainfed and irrigated land).

[17] We will also examine the effects on agriculture output and the size of the irrigated land, but these effects are examined at a more aggregated, district, level.

[18] In line with that reasoning, our analyses obtained by the triple-differences strategy and by a standard DID strategy with district-by-year fixed effects yield almost identical results. Arguably, one may even prefer the latter estimation procedure, because it is somewhat more efficient i.e. provides smaller standard errors, as argued in a different context by Young (2022).

4.1 *Difference-in-differences estimation*

We now explain the DID strategy. Let $y^{\mathcal{I}}_{rdt}$ be one of the outcome variables in county r *for irrigated land* (denoted by a superscript $\mathcal{I}$) in district d in year t. The outcome variables refer to four different variables: agricultural land rent, farmers' revenues, profits and wages. We include three dam variables $D^{\mathcal{O}}_{rdt}$, $D^{\mathcal{U}}_{rdt}$ and $D^{\mathcal{D}}_{rdt}$, which refer to the volume of dams in the own county ($\mathcal{O}$), downstream ($\mathcal{D}$) of the county, or upstream ($\mathcal{U}$) of the county r, respectively.

Importantly, although we include three measures of dam volume, our main interest is in the effect of the volume of *upstream* dams, which is anticipated to have positive economic effects, whereas the effects of own and downstream dams are anticipated to be absent (or maybe even negative). Consequently, we include own and downstream dams as controls to prevent omitted variable bias.

Furthermore, we emphasise that we use the *volume* of dams, which is arguably a more accurate measure of the effect of the amount water regulated than the number of dams as is used in the extant literature (Duflo & Pande 2007, Strobl & Strobl 2011, Mettetal 2019, Blanc & Strobl 2014). Although this measure improves on the literature, arguably, it still ignores the role of distance between dam and county, which may create measurement error, therefore likely generates downwards biased estimates. We will therefore also use a weighted measure of upstream dam volume with weights equal to the inverse of distance. Typically, we find that the effects of the weighted measure are slightly more pronounced.

To estimate the effect of upstream dam volume, we estimate the following baseline specification:

$$\log y^{\mathcal{I}}_{rdt} = \beta^{\mathcal{I}}_1 + \beta^{\mathcal{I}}_2 \text{asinh}(D^{\mathcal{O}}_{rdt}) + \beta^{\mathcal{I}}_3 \text{asinh}(D^{\mathcal{U}}_{rdt}) + \beta^{\mathcal{I}}_4 \text{asinh}(D^{\mathcal{D}}_{rdt}) + \gamma^{\mathcal{I}}_r + \delta^{\mathcal{I}}_t + \epsilon^{\mathcal{I}}_{rdt}, \qquad (1)$$

where $\beta^{\mathcal{I}}_i$, with $i = 1, .., 4$, are parameters to be estimated.

We employ the inverse hyperbolic sine transformation (asinh) of the volume of upstream dams, because this variable is extremely right-skewed, but for some observations it is equal to zero. Particularly at the beginning of our observation period, there are quite some zeros, but less so at the end (only about 20% of counties have zero upstream dams). Because of the presence of zeroes, we apply the inverse hyperbolic sine as an approximation to the more common logarithmic transformation. For the same reasons, we also use the same transformation for upstream dams and for dams in the own county. We consider a wide range of robustness checks to the functional form assumption of the volume of upstream dams (see Table A2 of Appendix A.2). $\gamma_r^{\mathcal{I}}$ refers to a county fixed effect, which absorbs all time-invariant characteristics that jointly affect (upstream) dam volume and outcome variables in the county. $\delta_t^{\mathcal{I}}$ is a year fixed effect that controls for annual shocks. Finally, $\epsilon_{rdt}^{\mathcal{I}}$ is a county-year specific error term.

Placebo tests. We will also perform a range of placebo tests. For example, we estimate the apply the same difference-in-differences approach as in (1), but now we use information on outcome measures for *rainfed* land, $y_{rdt}^{\mathcal{R}}$, where the superscript $\mathcal{R}$ denotes that the outcome measure is observed for rainfed land:

$$\log y_{rdt}^{\mathcal{R}} = \beta_1^{\mathcal{R}} + \beta_2^{\mathcal{R}} \text{asinh}(D_{rdt}^{\mathcal{O}}) + \beta_3^{\mathcal{R}} \text{asinh}(D_{rdt}^{\mathcal{U}}) + \beta_4^{\mathcal{R}} \text{asinh}(D_{rdt}^{\mathcal{D}}) + \gamma_r^{\mathcal{R}} + \delta_d^{\mathcal{R}} + \epsilon_{rdt}^{\mathcal{R}}. \quad (2)$$

This placebo test implies that we aim to examine whether $\beta_2^{\mathcal{R}} = \beta_3^{\mathcal{R}} = \beta_4^{\mathcal{R}} = 0$. This test is important for two reasons. First, it can be interpreted as a placebo test of the above-discussed DID strategy that relies on estimation of equation (1). Second, it supports the use of a triple-differences approach, discussed in the next section. A finding of an absence of an effect of dams on rainfed production is a sufficient (but not necessary) condition for the absence of general equilibrium effects, which may invalidate our triple-difference approach. [19]

4.2 *Triple-differences approach*

The DID strategy relies on between-county differences in the volume of upstream dams over time, so it might still suffer from the endogeneity in dam construction. Arguably, when we control for district-by-year fixed effects, endogeneity issues are most likely absent. Nevertheless, although unlikely, we cannot exclude the possibility that, even within districts, counties that anticipate larger increases in agricultural productivity and population in the future (as it may take quite some time to build a dam) may be more likely to receive upstream dam investments, implying a spurious positive relationship between population increase, agricultural growth and dam building. Therefore, we consider here a triple-differences strategy (DDD), where the

[19] An example of a general equilibrium effect is the effect of dams on labour markets. The building of dams may affect the demand for labour for workers employed on land that is irrigated, which may increase wages and therefore indirectly increase cost for rainfed production.

additional difference comes from comparing outcomes for irrigated and rainfed agricultural land. Consequently, identification comes from that dam construction should have a differential effect on outcomes for irrigated and rainfed land.

We consider the following regression specification, which is based on the difference between equations (2) and (1):

$$\log y^{\mathcal{I}}_{rdt} - \log y^{\mathcal{R}}_{rdt} = \beta_1 + \beta_2 \text{asinh}(D^{\mathcal{O}}_{rdt}) + \beta_3 \text{asinh}(D^{\mathcal{U}}_{rdt}) + \beta_4 \text{asinh}(D^{\mathcal{D}}_{rdt}) + \gamma_r + \delta_t + \epsilon_{rdt}, \quad (3)$$

where the dependent variable refers to the difference between outcome variables of irrigated and rainfed land located in county r of district d in year t. So, for example, we use the difference in agricultural land rent between irrigated and rainfed land.

Here, all coefficients refer to the difference in coefficients identified by equations (2) and (1), so $\beta_i = \beta^{\mathcal{I}}_i - \beta^{\mathcal{R}}_i$ where $i = 1, .., 4$. Furthermore, $\epsilon_{rdt} = \epsilon^{\mathcal{I}}_{rdt} - \epsilon^{\mathcal{R}}_{rdt}$. Hence, taking the difference between the outcome variables of irrigated and rainfed land removes all *time-varying county-specific unobserved characteristics* that affect irrigated and rainfed land in the same way. The identification assumption now is that in the absence of treatment, the change over time in the difference in outcomes between rainfed and irrigated land within the same county would have been the same. We believe that this assumption is clearly non-restrictive.

Additional specifications and robustness. The above specifications implicitly assume that the effect of dams is immediate. If this is not the case, then it is plausible that these specifications provide underestimates of the long-run effect, which is particularly of interest. There are many reasons why the effect of dams might not be immediate. For example, even after construction of the dam, it may not be the case that the construction of irrigation system, including irrigation pipes, has been finished. For that reason, we also estimate long differences, i.e. we only keep the first and last observation.

The triple-difference strategy relies on the assumption that the general equilibrium effects are the same for rainfed and irrigated land *within a county*. As discussed above, we can test this using placebo tests. We also perform two additional tests. First, we will show that wages are the same for workers of rainfed and irrigated farms within the same county, so this makes any differential effects through the labour market extremely unlikely. Second, using data at the *district* level, we will show that dams do not affect the size of the *rainfed* land area. This is important, because if placement of dams has a strong effect on the conversion of rainfed land into irrigated land, then the DDD strategy would be biased because of employing a nonrandom sample of observations after treatment.[20]

We implement additional sensitivity checks for both the difference-in-differences (DID) and triple-difference (DDD) strategies. First, we introduce *basin-by-year* fixed effects to account for hydrological and geographical variations. Second, we control for *district-by-year* fixed effects to account for any endogeneity issues in the placement of dams. Third, we control for potential divergent time trends in irrigated and rainfed production using average rainfall and geographical indicators. Fourth, we incorporate a one-year lag of dam variables to address the issue that it takes time for effects of dams to materialise. Fifth, we weight the volume of dams by their distance to the county to account for potential dissipation effects. Alternatively, we impose a distance threshold of 50 km.

5 Results

5.1 Land rent and profits: main results

Table 3 shows the baseline results for the effect of the volume of upstream dams on the logarithm of agricultural land rent and the logarithm of profits.[21] So here, we only control for county and year fixed effects. Panel is based on annual panel data, while Panel B reports long-differences results.

Columns (1) and (2) report DID estimation results for irrigated land based on specification (1). Column (1) shows that an effect of about 0.023, which suggests that a 1% increase in the volume of upstream dams leads to 0.023% increase in the irrigated land rent, which is statistically significant at the 1% level. Column (2) shows that agricultural profits increase by 0.04% when the volume of upstream dams increases by 1%. The results also show that an increase in the

[20]For example, if rainfed land is more likely to be converted into irrigated land when the rent is low, then the triple-difference strategy is more likely to overestimate the causal effect of interest.

[21]We estimate the effect using the logarithm of profits, rather than using levels, to improve interpretation. The main disadvantage is that we exclude about 9% of observations with negative profits, which may cause selection bias. For alternative specifications, where we keep these observations, we find almost identical results. For example, when we estimate models using all observations while specifying profit in levels, the implied effects for the average dam are roughly the same.

TABLE 3: BASELINE RESULTS: RENTS AND PROFITS

	Irrigated land, DID		Rainfed land, DID		Triple differences	
Dep. var.	Rent *(log)*	Profit *(log)*	Rent *(log)*	Profit *(log)*	Rent *(log)*	Profit *(log)*
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Annual panel data						
Upstream volume of dams *(asinh)*	0.023***	0.051***	-0.014***	0.010	0.031***	0.029*
	(0.006)	(0.010)	(0.006)	(0.013)	(0.010)	(0.017)
Own volume of dams*(asinh)*	0.005	0.006	0.012	-0.024	0.027	-0.014
	(0.013)	(0.017)	(0.009)	(0.019)	(0.023)	(0.025)
Downstream volume of dams*(asinh)*	-0.001	-0.002	-0.009*	-0.010	0.014	0.006
	(0.007)	(0.010)	(0.005)	(0.010)	(0.010)	(0.017)
Number of observations	24,583	22,036	20,295	15,819	12,877	9,827
$\bar{R}^2$	0.807	0.535	0.850	0.584	0.652	0.552
Panel B: Long differences						
Upstream volume of dams*(asinh)*	0.034**	0.074**	-0.021	-0.045	0.055**	0.140**
	(0.014)	(0.030)	(0.015)	(0.044)	(0.025)	(0.060)
Own volume of dams*(asinh)*	0.016	-0.041	0.009	-0.063	0.043	-0.107
	(0.019)	(0.050)	(0.023)	(0.057)	(0.032)	(0.089)
Downstream volume of dams*(asinh)*	-0.007	0.006	-0.032***	0.021	0.021	0.016
	(0.009)	(0.020)	(0.009)	(0.031)	(0.019)	(0.048)
Number of observations	2,230	1,924	1,722	1,016	944	518
$\bar{R}^2$	0.880	0.709	0.898	0.736	0.705	0.581
Time fixed effects	Y	Y	Y	Y	Y	Y
County fixed effects	Y	Y	Y	Y	Y	Y

Notes: This table shows coefficient estimates from regressions of rent and profit on measures of dam volume. Columns (1) and (2) show the effect of volume of own, upstream, and downstream dams on agricultural land rent and agricultural profits of irrigated land. Columns (3) and (4) show the same for rainfed land. We include year and county fixed effects in all specifications. Columns (5) and (6) show DDD results based on the comparison between irrigated and rainfed land. Panel B uses the same specifications as in Panel A but only keeps observations from the first and last year of the data to find the long differencing effect. Standard errors are clustered at the county level and reported in parentheses. *, **, *** represent significance at 10%, 5%, and 1% level respectively.

volume of dams in the own county or downstream dams do not affect both outcome variables. This is consistent with the idea that counties are only affected by upstream dams (see Duflo & Pande 2007).

Columns (3) and (4) show estimates of the same specification (*i.e.*, the same dependent and independent variables), but now for rainfed land. It appears that the upstream dam volume has no impact on both outcome measures of rainfed land.[22] The latter finding makes the adoption of a triple differences (DDD) strategy more appealing, as it makes the absence of general equilibrium effects more plausible, as discussed earlier.

The last two columns of this table report triple-difference results based on specification (3), which deals with a range of endogeneity issues, and in particular the endogenous location

[22]We have also tested the hypothesis that the effects of all three volume measures (upstream, downstream, own) are simultaneously equal to 0. Here, again we do not find any evidence of any effect. For example, when we focus on rents, the corresponding p-value of the F-test is 0.70 indicating that we cannot reject the hypothesis that these effects are zero.

choice by governments. Here, we focus on a subset of observations, where rainfed and irrigated land are observed to be present within the same country in the same year, reducing the sample size by about a factor 2, increasing standard errors.[23] It appears that the results based on the triple difference estimation procedure are very similar to the standard DID. Column (5) shows that 1% increase in the volume of upstream dams leads to 0.031% increase in rent on irrigated land. Column (6) shows that the effect of volume on profit is about 0.030.

In Panel B of the same table, we move to the long-differences results, in which we only use first and last year observations (so a difference of 17 years). We find that the effects become somewhat larger: irrigated land rent and profit increase by 0.034% and 0.065%, respectively, for a 1% increase in the upstream dam volume. The DDD strategy also finds larger effects of about 0.055% and 0.140% for land rents and profits, respectively. Since it likely takes time for the effect of dams to materialise in downstream areas, it makes sense that the long-run effects are larger than those based on annual variation in the volume of dams. Own and downstream dam effects are, once again, zero in all long-differences specifications, providing more confidence in the upstream dam effects.

5.2 *Land rent and profits: sensitivity analysis*

To bolster confidence in the findings, we conduct various robustness checks. First, in all columns of Table 4, we introduce additional controls for rainfall and the interaction of time with a range of geographical factors, including river length, county elevation, slope, and area. Furthermore, for the DID analysis, we also incorporate district-by-year and basin-by-year fixed effects. District-by-year fixed effects allows rents and profits in different districts to have entirely different time trends. This is potentially important because, for example, the allocated budget to a district from the central government might vary over time. Once a district receives a higher budget it may invest at the same time in dams and provide agricultural subsidies (such as subsidised fertiliser). The inclusion of these fixed effects implies that we only exploit variation in the change of volume of upstream dams between counties *within a district* over time. The basin-by-year fixed effects capture different hydrological and geographical characteristics of basins which may change over time, and which may correlate with the placement of dams.

[23] Because we rely on a sample from a survey, it is likely that a substantial share of observations excluded by us also have both types of land in the county, but due to sampling error this is not observed by us.

Basins are hydrological units and not administrative units, so they do not overlap with district borders.

It appears that the results hardly change, independently of whether we use the annual data or apply long differences.[24] Notably, as anticipated, rainfall has a positive effect on profits of rainfed land, and therefore a negative effect for the DDD specifications. This makes sense as rainfall is beneficial for rainfed crops, while it is unlikely to significantly impact the yields of irrigated crops, as the price of irrigated water is very low.

One important concern is the definition of sub-basin, which determines how we define upstream and downstream areas. Wrongly classifying dams as upstream or downstream may result in substantial measurement error in the calculation of dam volume and therefore, typically downward, biased estimates. In Table A3 of Appendix A.2, we use an alternative definition for upstream and downstream.[25]

The second concern applies to the measurement of upstream dam volume, as we ignore the distance between dams and counties (this distance is constrained by that the dam must be in the same sub-basin). It is plausible that dams that are further away have a smaller effect than dams that are more nearby. To account for this, we use distance-weighted volume. Columns (1) and (2) in Table 5 show the estimation results for the DDD specification using these alternative explanatory variables. It appears that the results are robust with respect to this adjusted measure of dam volume. The oint estimates of the effects of rent appear to be somewhat larger. By contrast, the point estimate of the long-difference effect on profit is about the same as before, but due to large standard errors becomes statistically insignificant at conventional significance levels.

The third concern is the functional form of using the inverse hyperbolic sine transformation of the volume of upstream dams. Hence we have examined the robustness of the results regarding this functional form, where we have focused on the effects on rent, as shown in Table A2 of A.2.

[24] The same holds for similar sensitivity analyses not reported here. For example, for the triple difference analysis, even when we include district-by-year and basin-by-year, so we control for district-year-specific trends in outcome variables for rainfed and irrigated land *although they belong to same the county*, we find the same results, except for one estimated effect, i.e. the effect on profit when using annual data, where the standard error increases such that the effect becomes just statistically insignificant.

[25] We define sub-basins using river flow, as in Lehner & Grill (2013), to determine upstream/downstream relationships. This divides Iran into 70 hydrological basins with an average area of 22,900 km^2. The results remain similar.

TABLE 4: ROBUSTNESS: INCLUDING TRENDS AND CONTROLS

	Irrigated land (DID)		Rainfed land (DID)		Triple differences	
Dep. var.	Rent *(log)*	Profit *(log)*	Rent *(log)*	Profit *(log)*	Rent *(log)*	Profit *(log)*
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Annual panel data						
Upstream volume of dams *(asinh)*	0.013*	0.038***	0.001	0.014	0.030***	0.028**
	(0.007)	(0.011)	(0.006)	(0.020)	(0.010)	(0.014)
Own volume of dams *(asinh)*	-0.005	0.000	-0.006	-0.015	0.027	0.001
	(0.008)	(0.016)	(0.006)	(0.017)	(0.023)	(0.022)
Downstream volume of dams *(asinh)*	-0.003	-0.007	-0.002	0.016	0.013	0.007
	(0.008)	(0.011)	(0.005)	(0.017)	(0.011)	(0.013)
Rainfall *(log)*	0.061	0.121	-0.044	-0.044	-0.101*	-0.164**
	(0.113)	(0.127)	(0.146)	(0.508)	(0.053)	(0.080)
Number of observations	24,170	21,913	19,999	15,349	12,877	9,827
$\bar{R}^2$	0.886	0.560	0.916	0.678	0.656	0.570
Panel B: Long differences						
Upstream volume of dams *(asinh)*	0.037*	0.116*	-0.016	0.062	0.056**	0.096*
	(0.019)	(0.059)	(0.024)	(0.085)	(0.026)	(0.053)
Own volume of dams *(asinh)*	-0.004	0.007	-0.012	0.048	0.028	-0.114
	(0.021)	(0.071)	(0.017)	0.057	(0.042)	(0.077)
Downstream volume of dams *(asinh)*	0.003	0.026	0.014	0.032	0.019	0.018
	(0.019)	(0.061)	(0.017)	(0.084)	(0.024)	(0.044)
Rainfall *(log)*	0.533	0.900	-0.280	1.209	0.159	-0.225
	(0.460)	(0.963)	(1.005)	(6.536)	(0.222)	(0.253)
Number of observations	2,162	1,876	1,734	978	944	602
$\bar{R}^2$	0.906	0.674	0.932	0.675	0.712	0.703
Time fixed effects	N	N	N	N	Y	Y
County fixed effects	Y	Y	Y	Y	Y	Y
Basin-Year fixed effects	Y	Y	Y	Y	N	N
District-Year fixed effects	Y	Y	Y	Y	N	N
County-level controls	Y	Y	Y	Y	Y	Y

Notes: This table shows coefficient estimates from regressions of rent and profit on measures of dam volume. Columns (1) and (2) show the effect of volume of own, upstream, and downstream dams on agricultural land rent and agricultural profits of irrigated land. We include year and county fixed effects in all specifications. Columns (3) and (4) show DDD results based on the comparison between irrigated and rainfed land. Panel B uses the same specifications as in Panel A but only keeps observations from the first and last year of the data to find the long differencing effect. Standard errors are clustered at the county level and reported in parentheses. *, **, *** represent significance at 10%, 5%, and 1% level respectively.

Rather than employing the inverse hyperbolic sine transformation, we include two variables:a dummy variable for the presence of upstream dams (the extensive margin) and the logarithm of upstream volume given the presence of upstream dams (the intensive margin). It appears that the effect of the logarithm of upstream volume is close to, but somewhat higher, than the reported effect when using the inverse hyperbolic sine transformation. Furthermore, we provide evidence that the effect of the extensive margin is positive. In alternative specifications, we have used the logarithm of upstream volume for all observations, where we have replaced the zeros with a value equal to half of volume of the smallest dam, and we have excluded observations for which the upstream volume is zero. Again we find similar results.

TABLE 5: ROBUSTNESS: DISTANCE TO DAMS AND PROXIMITY TO CITIES

	Dam volume weighted by distance		Control for distance to city trends	
Dep. var.	Rent textit(log)	Profit *(log)*	Rent *(log)*	Profit *(log)*
	(1)	(2)	(3)	(4)
Panel A: Annual panel data				
Upstream volume of dams *(asinh)*	0.041**	0.044**	0.039***	0.049**
	(0.016)	(0.019)	(0.015)	(0.022)
Own volume of dams *(asinh)*	0.027	-0.004	0.026	-0.020
	(0.024)	(0.023)	(0.024)	(0.026)
Downstream volume of dams *(asinh)*	0.012	-0.001	0.019	-0.009
	(0.013)	(0.016)	(0.013)	(0.022)
Number of observations	12,197	9,342	12,197	9,342
$\bar{R}^2$	0.653	0.568	0.655	0.555
Panel B. Long differences				
Upstream volume of dams *(asinh)*	0.103***	0.0637	0.0940***	0.138*
	(0.0343)	(0.0754)	(0.0345)	(0.0808)
Own volume of dams *(asinh)*	0.0367	-0.0805	0.0345	-0.119
	(0.0424)	(0.0696)	(0.0415)	(0.0969)
Downstream volume of dams *(asinh)*	0.0313	0.00840	0.0507	-0.0217
	(0.0321)	(0.0563)	(0.0309)	(0.0710)
Number of observations	908	502	908	502
$\bar{R}^2$	0.720	0.601	0.732	0.589
Time fixed effects	Y	Y	Y	Y
County fixed effects	Y	Y	Y	Y
Distance to city trends	N	N	Y	Y

Notes: In columns (1) and (2) we divide volume of each upstream/downstream dam by its distance with the county and then add them up to find the weighted average effect of upstream/downstream dams. In columns (3) and (4), we control for distance with nearest big city trend as a robustness check. Panel B is the same as Panel A but just keep observations in the first and last year to find the long-difference effect. Standard errors are clustered at the counties and reported in parentheses. *, **, *** represent significance at 10%, 5%, and 1% level respectively.

Another concern is that the triple difference estimation strategy focuses only on those counties that have both irrigated and rainfed land, which may introduce a selection bias. As a robustness check, we have re-estimated the DiD for irrigated land using only observations for countries where both types of land are present. We find that the estimates are only slightly closer to the triple difference estimates, suggesting that the selection bias is not of major concern.

A final concern is that the increase in land rents may be due to higher demand for recreation and housing construction induced by dam construction. To see the importance of this alternative explanation, we add the interaction of distance to the nearest big city and year as an additional control variable in columns (3) and (4) of Table 5. This does not change the results and, therefore, we conclude that dams are unlikely to affect amenities or housing construction.

TABLE 6: EFFECT OF DAMS ON REVENUES, WAGES AND POPULATION

Dep. var.	DID				DDD	
	Irrigated revenue (log)	Rainfed revenue (log)	Irrigated wage (log)	Population (log)	Revenue (log)	Wage (log)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Annual panel data						
Upstream volume of dams *(asinh)*	0.032***	-0.009	0.0001	-0.002	0.023*	0.002
	(0.008)	(0.006)	(0.001)	(0.004)	(0.012)	(0.002)
Own volume of dams *(asinh)*	0.008	-0.006	-0.002	-0.006	-0.001	-0.003
	(0.016)	(0.009)	(0.002)	(0.008)	(0.016)	(0.003)
Downstream volume of dams *(asinh)*	0.001	-0.006	-0.002	-0.015***	0.020	0.000
	(0.009)	(0.006)	(0.002)	(0.004)	(0.013)	(0.001)
Number of observations	24,495	19,546	24,131	10,918	10,618	12,514
$\bar{R}^2$	0.768	0.820	0.988	0.804	0.622	0.178
Panel B. Long differences						
Upstream volume of dams *(asinh)*	0.051**	-0.041	0.002	0.032**	0.162***	-0.0002
	(0.020)	(0.028)	(0.009)	(0.014)	(0.049)	(0.006)
Own volume of dams *(asinh)*	-0.020	-0.042**	0.001	0.002	-0.038	0.0001
	(0.036)	(0.020)	(0.005)	(0.011)	(0.082)	(0.007)
Downstream volume of dams *(asinh)*	0.017	-0.001	-0.001	-0.044***	-0.019	0.001
	(0.018)	(0.019)	(0.008)	(0.015)	(0.036)	(0.004)
Number of observations	2,190	1,584	2,122	4,362	598	768
$\bar{R}^2$	0.835	0.895	0.995	0.600	0.695	0.48
Time fixed effects	Y	Y	Y	Y	Y	Y
County fixed effects	Y	Y	Y	Y	Y	Y
Basin-Year fixed effects	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y

Notes: This table shows the effect of own, upstream and downstream volume of dams on agricultural revenue, labour wage, and population. In columns (1) to (4), we see the difference in difference method and in columns (5) and (6), we report DDD results for agricultural revenue and labour wage. Panel B is the same as Panel A but just keep observations in the first and last year to find the long-difference effect. Standard errors are clustered at the counties and reported in parentheses. *, **, *** represent significance at 10%, 5%, and 1% level respectively.

5.3 *Wage, population and revenue*

Up to now, our main focus was on the effect of dams on agricultural land rent and profit, but dam construction may impact other outcomes. We now examine the effect on population and agricultural wages, as well as on agriculture revenue, which is a component of the profit analysed earlier, see Table 6.

As explained above, one cannot apply the DDD strategy for population, so the inclusion of district by year fixed effects is desirable here to address endogeneity issues. Another important difference is that we have data for population over a much longer period, such that the long-difference analysis refers to a 40 year difference. As before, Panel A shows the short-run results while Panel B shows the long-run results.

Columns (1) and (4) show that upstream dams lead to an increase in agricultural revenue, consistent with our effects for profits. The results in columns (2) and (5) show that dams do not have any detectable effect on agricultural wages. This is in line with the extremely competitive market in Iran for unskilled agricultural workers that migrate between counties, which in turn erodes any local increases in wages. The absence of a wage effect due to migration between counties is in line with the results for population: dam construction in the downstream of a county reduces population. This effect is particularly strong in the long run. We find now that in the long-run a 1% increase in the volume of regulated water downstream reduces population by 0.044% (significant at the 1% level), while it increases by 0.032% when upstream dam volume increases by 1%.[26] This strongly suggests that migration occurs from counties upstream of dams to those in the downstream.

5.4 *Agricultural output, irrigated land area: district-level results*

The above analysis – at the level of county – is silent on two important effects of dams, that are also unexplored in the extant literature. First, it is silent to what extent dams increase the size of irrigated land area and whether this reduces the size of the rainfed land area. Second, it is also silent to what extent dams increase agricultural output.

Note that our previous results on farmer revenues suggest, but does not demonstrate, that output increases. This is due to 2 underlying reasons. First, the effects of dams on output and revenue will differ when dam construction affects local prices of agricultural products. For example, if dam construction reduces local prices of agricultural products, then the impact of new dams on output will exceed the impact on revenue. Second, more fundamental, in the county analysis, we have examined the impact on revenues *per farmer*, which ignores the change in number of farmers.

We are able to estimate both effects – on land size and output– using an additional dataset at a more aggregated, i.e. district, level using the DiD as well as the DDD methodology introduced above. At this level distinguishing between upstream and downstream dams is rather ambiguous, since districts are large, so they are typically in more than one sub-basin, which makes it frequently impossible to define whether or not they are upstream or downstream

[26]When focusing on a shorter long-term difference, e.g. from 1996 until 2016 (instead of 1976 until 2016), the effect of downstream dam is the same but the effect of upstream dam become statistically insignificant although still positive.

of dams. So, in this analysis, we focus on the effect of volume of dams located in the own district, and do not aim to disentangle the effects of other dams, which is consistent with the literature (Duflo & Pande 2007).[27]

Columns (1) and (2) of Table 7 show that dam volume has a positive effect on the size of the irrigated area as well as on agricultural output. The long difference analysis suggests that doubling of own dam volume leads to a $7 - 8\%$ increase in the size of irrigated area and agricultural output.[28] In the long difference analysis of columns (3) and (4), it is weakly suggested that rainfed land is converted into irrigated land after increasing dam volume, but the annual data analysis does not suggest this at all.[29] To take into account that placement of dams might be endogenous, we implement the DDD approach as reported in columns (5) and (6). This analysis also supports the conclusion that output of irrigated land increases relative to the output of rainfed land after placement of dams. As an aside, note that the effects of rain on output and area size are much stronger for rainfed land than for irrigated land, which supports our distinction between these two types of lands.[30]

5.5 An IV approach

In the economic literature on dams, previous studies employed an IV approach using an instrument that is based on *non-linear* interaction effects of gradient differences of rivers (i.e., the steepness of the river) and the predicted local number of dams constructed in a specific year. The idea behind this non-linearity is that rivers have to be sufficiently steep, but not too steep, for the building of dams. In these studies, it appeared that the instrument was not strong enough (in the first stage) to estimate the effect on dam volume (see Duflo & Pande 2007, Strobl & Strobl 2011), which is the main reasons that these studies focus on the effect of the number of dams rather than on dam volume. In the current study, when we use this instrument, we have the same issue, so the instrument is too weak for reliable inference (the F-statistics that measure the strength of the instrument are below 5).

[27] For consistency, although the presence of zeros is rare in this context, we continue to apply the hyperbolic sine transformation of water volume.

[28] We have used here that log(2x) - log(x) = log(2) = 0.7.

[29] In the cost benefit analysis later on, we will make the conservative assumption that new irrigated land is converted rainfed land.

[30] Note that the effect of rain on the size of the irrigated area is small but not zero. We hypothesise that this is due to that a small proportion of irrigated area partially relies on rain or, more likely, that water from dams is more abundant during periods with more rain.

TABLE 7: LAND AREA AND AGRICULTURAL OUTPUT

	Irrigated land, DID		Rainfed land, DID		Triple differences	
Dep. var.	Area *(log)*	Output *(log)*	Area *(log)*	Output *(log)*	Area *(log)*	Output *(log)*
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Annual panel data						
Upstream volume of dams *(asinh)*	0.028*	0.041**	0.021	-0.013	0.010	0.062**
	(0.015)	(0.017)	(0.024)	(0.020)	(0.026)	(0.024)
Own volume of dams*(asinh)*	0.037*	0.045**	-0.007	-0.014	0.042	0.055
	(0.019)	(0.020)	(0.026)	(0.024)	(0.034)	(0.034)
Downstream volume of dams*(asinh)*	0.020	0.020	-0.011	-0.001	0.031	0.013
	(0.014)	(0.016)	(0.021)	(0.020)	(0.020)	(0.020)
Rainfall *(log)*	0.165***	0.215***	0.688***	0.988***	-0.548***	-0.813***
	(0.045)	(0.051)	(0.152)	(0.158)	(0.162)	(0.171)
Number of observations	3,858	3,855	2,919	2,893	2,894	2,865
$\bar{R}^2$	0.820	0.810	0.857	0.858	0.827	0.838
Panel B: Long differences						
Upstream volume of dams*(asinh)*	0.078**	0.110**	-0.085	-0.077	0.131	0.131
	(0.037)	(0.053)	(0.081)	(0.076)	(0.084)	(0.089)
Own volume of dams*(asinh)*	0.081**	0.115**	-0.097	-0.024	0.197	0.145
	(0.040)	(0.051)	(0.141)	(0.126)	(0.150)	(0.142)
Downstream volume of dams*(asinh)*	-0.048	0.063	-0.089	-0.095	0.092	0.092
	(0.037)	(0.043)	(0.067)	(0.062)	(0.063)	(0.061)
Rainfall *(log)*	0.321	0.472	0.379	0.940	-0.576	-0.986
	(0.238)	(0.340)	(0.529)	(0.574)	(0.634)	(0.648)
Number of observations	222	222	156	156	154	154
$\bar{R}^2$	0.692	0.591	0.714	0.755	0.716	0.745
Time fixed effects	Y	Y	Y	Y	Y	Y
District fixed effects	Y	Y	Y	Y	Y	Y

Notes: This table shows coefficient estimates from regressions of Agricultural area and product at district level on measures of dam volume. Columns (1) and (2) show the effect of volume of own, upstream, and downstream dams on agricultural area and agricultural output of irrigated land. Columns (3) and (4) show the same for rainfed land. We include year and district fixed effects in all specifications. Columns (5) and (6) show DDD results based on the comparison between irrigated and rainfed land. Panel B uses the same specifications as in Panel A but only keeps observations from the first and last year of the data to find the long differencing effect. Standard errors are clustered at the district level and reported in parentheses. *, **, *** represent significance at 10%, 5%, and 1% level respectively.

We have therefore constructed a more suitable, i.e. stronger, instrument, relying on the engineering/hydrology literature. Here, it is indicated that *flow accumulation*, and not river gradient, is used by engineers to decide dam volume (Andreadis et al. 2013). Flow accumulation aims to capture how much water flows through a river (so it takes the width of a river into account, and not only the gradient). We measure flow accumulation for the whole of Iran and use this (rather than river gradient) to construct an improved instrument, see Appendix A.3 for details. We now find that the improved IV strategy based on flow accumulation provides enough statistical power in our first stage to estimate the effect of volume of dams (with F-statistics which measure the strength of the instrument of above 12). In the second stage, in line with our other findings, we find that upstream volume has a positive effect on rents, profits and revenue. Unfortunately, the corresponding standard errors are so large that the point estimates are essentially non-informative. Consequently, despite the large number of dams constructed in Iran, even an improved IV approach using flow accumulation cannot be usefully applied to investigate the effect of dam volume. We speculate that such our improved approach is still useful for the largest countries in the world or applications which use information for whole continents. For details, we refer to Appendix A.3.

6 Cost-benefit analysis

This section provides a cost-benefit analysis (CBA) of agricultural dams, where we emphasise the role of dam size. To do so, first we focus on the benefits, then on the costs, and then we provide the cost-benefits analysis. As far as we know, this analysis is novel, as it is the world's first ex-post analysis based a large number of dams. We focus here on 248 dams that are only, or predominantly, used for agricultural purposes and have been built after the year 2001.

6.1 Benefit analysis

To calculate the benefits, we take into account that dams have been built in a certain order. When calculating the marginal benefit of a dam we take into account that a new dam affects the benefits of all other dams present. Our estimates of the dam benefits will be based on our long-difference estimates of changes in agricultural profits, land rents and cultivated agricultural land induced by the construction of dams.

To calculate the beneficial effect through increased profits and rents, we adopt a four-step approach which calculates the marginal benefit of each dam. First, for each county, we calculate the increase in upstream volume resulting from the construction of a new dam. Second, we employ information from the Agricultural Census 2014 for each county to compute the total downstream agricultural area for each dam. We make the assumption that the cultivated area remains unchanged due to dam construction, for which we have provided evidence in Table 7. Third, we multiply our land rent elasticity estimate (0.05, as shown in Table 3) by the province-level average land rent in that year to determine the total annual benefit in terms of land rent from the new dam. We deflate benefits by the yearly consumer price index (with 1997 as the base year). Additionally, we assume a 100-year economic lifespan and apply a discount factor of 0.98 to find the net present value of benefits for each dam.[31] We follow a similar procedure to calculate the benefits arising from higher profits, using the long-run profit elasticity of 0.14 for volume.

To calculate the beneficial effect of increased land size that is irrigated, we use the results at the district level and apply again a four-step approach. First, for each county, we calculate the increase in upstream volume resulting from the construction of a new dam. Second, we employ information from the Agricultural Census 2014 for each county to compute the total downstream agricultural area for each dam. Third, we use the irrigated area elasticity estimate (0.08, as shown in Table 7) to determine the total irrigated area increase from the new dam. Finally, we multiply the irrigated area size increase with the marginal benefit of land size. To obtain the marginal benefit, we make the conservative assumption that all new irrigated land is converted from rainfed land (rather than from non-cultivated land), which justifies the use of the difference in average rent levels per square metre (as well as profit levels) between irrigated and non-irrigated land as a monetary measure of marginal benefit of irrigated land size. Additionally, the same as before, we assume a 100-year economic lifespan and apply a discount factor of 0.98 to find the net present value of benefits for each dam.

One expects that volume of a dam is the main predictor of it's benefits. In Figure 3, we show the effect of log dam volume on log benefits, where we control for year dummies and (time-averaged) rainfall (without these controls the figure looks very similar). It appears that there

[31] Most likely, we underestimate the economic lifespan of dams by assuming a 100- year lifespan. Dams are constructed to last way longer. However, mechanical components that move have to be replaced 50 years.

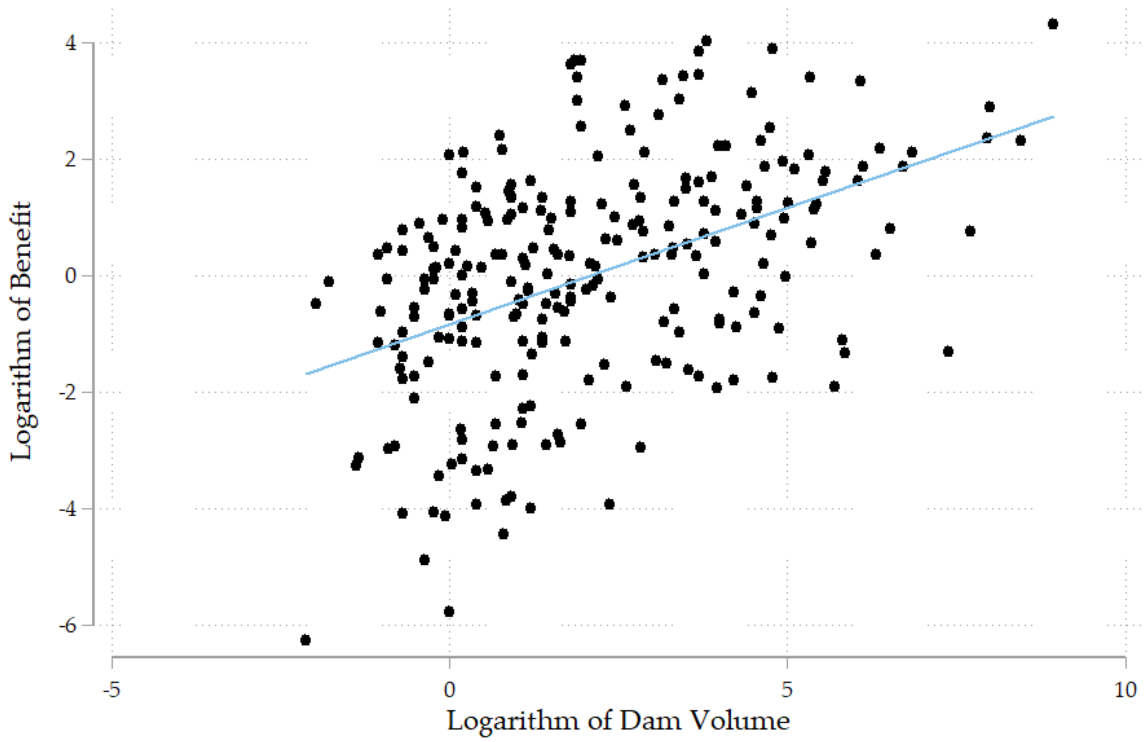

FIGURE 3: DAM BENEFIT AND VOLUME OF DAMS

Notes: Each point shows log of a dam's benefit versus log of its volume while controlling for rainfall and year fixed effects. Benefits are deflated by the consumer price index. The slope is 0.40. Number of observations is 248.

are strong diminishing returns to scale: the volume elasticity of benefits is positive and equal to 0.40, far below one.

6.2 Cost analysis

We will now investigate the effect of dam volume on their costs, and in particular the construction costs, which arguably are the main costs of dam construction. Information about these costs comes from Ministry of Energy. We deflate costs by the yearly consumer price index (with 1997 as the base year). For 116 dams out of the 248 dams we focus on, we observe (construction) costs. For another 98 dams, we also observe dam volume and construction costs. We will include these dams in this subsection only.

Economies of scale arguments are typically key in the debates about the size of dams. In Figure 4, we show the effect of log dam volume on log costs, where we control for year dummies and (time-averaged) rainfall (including year fixed effects is useful as construction costs and volume

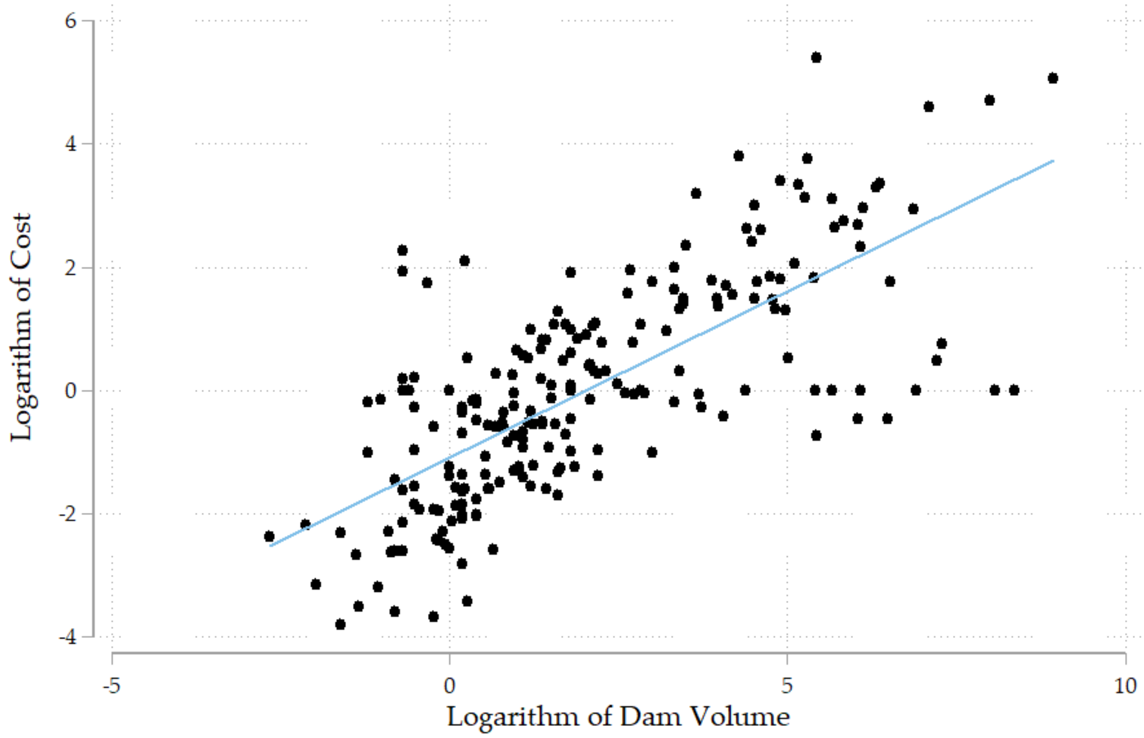

FIGURE 4: CONSTRUCTION COSTS AND VOLUME OF DAMS

Notes: Each point shows log of a dam's construction cost versus log of its volume after controlling for rainfall and year fixed effects. Costs are deflated by the consumer price index. The slope is 0.54. Number of observations is 214.

of dams both vary over time). It appears that the volume elasticity of costs is 0.54, far below one, implying substantial economies of scale in costs.

6.3 Benefits and costs analysis

The above two subsections have provided us with information of the effects of volume on agricultural benefits and on total costs for dams that are primarily built for agricultural purposes. Nevertheless, for a sound cost benefit analysis of dams, we wish to take into account that most dams are not pure agricultural, so, by construction, by using total costs, we will overestimate the costs attributed to agricultural purposes if we ignore this issue. We deal with this issue in two ways.

First, we focus on 208 (out of 248) dams for which we observe the share of volume which is allocated to agricultural water.

We then continue to attribute the construction costs for agricultural purposes based on this share. This can be interpreted as an accounting exercise, but it also has an economic rationale. Given the assumption that a rational planner maximises the profitability of a dam by allocating the share of water to agricultural or non-agricultural purposes, then this share would be chosen such that the marginal benefits of agricultural purposes would be equal to the marginal benefits of the non-agricultural purposes. According to this logic, the total benefit to cost ratio would be roughly the same as the agricultural benefit to attributed-to-agriculture cost ratio, where the approximation would be more accurate for large values of the share.

Table 8 report the costs and benefits for the 248 dams in our sample.[32] Panel A shows the costs. The benefits in panel B reflect the above-described procedure. We report benefit minus cost and benefit divided by cost as two measures. Both mean and median NPV of dams are negative (*i.e.* -33 and -0.4 billion Rials). Looking at the benefit to cost ratio we observe a large mean of 3.8 but the median shows benefits add up only to 86% of costs. One important point here is the high dispersion of the CBA distribution. Figure 5 shows the scatter plot of log of benefits versus log of costs. There are 132 dams that have a negative NPV. On the other hand 116 dams have a positive NPV. Figure **??** shows the histogram of benefit to cost ratio. Most dams have a negative NPV suggesting an overall loss. A few dams have very large benefits as they serve a very large area. The sizeable heterogeneity in NPV of dams underlines the importance of dam placement and design. By construction, dams that regulate water for large areas of fertile land are the most profitable, while those built with limited downstream areas create a net loss.

As above, Figure 6 shows the relationship between dam volume and benefit to cost ratios. It shows that there are substantial diminishing net returns to scale (where net return is defined as the difference between log benefits and log costs), demonstrating that larger dams have smaller cost to benefit ratios.

Panels C and D in Table 8 present two robustness checks designed to consider the diminished benefit for areas farther from a dam. **First, in Panel C we weight downstream areas by the inverse of distance from the upstream dam. **The first exercise shows that benefits decreased but the results are not materially influenced.** Second, we assume no benefits for areas further than 50 km from the dam. We show in Panel C and D reduction in NPVs and respectively 205

[32] For 131 dams, costs are missing, so we rely on imputed costs that are obtained using the estimates from the previous subsection. Combined with information on volume of these 131 dams, one can predict these costs.

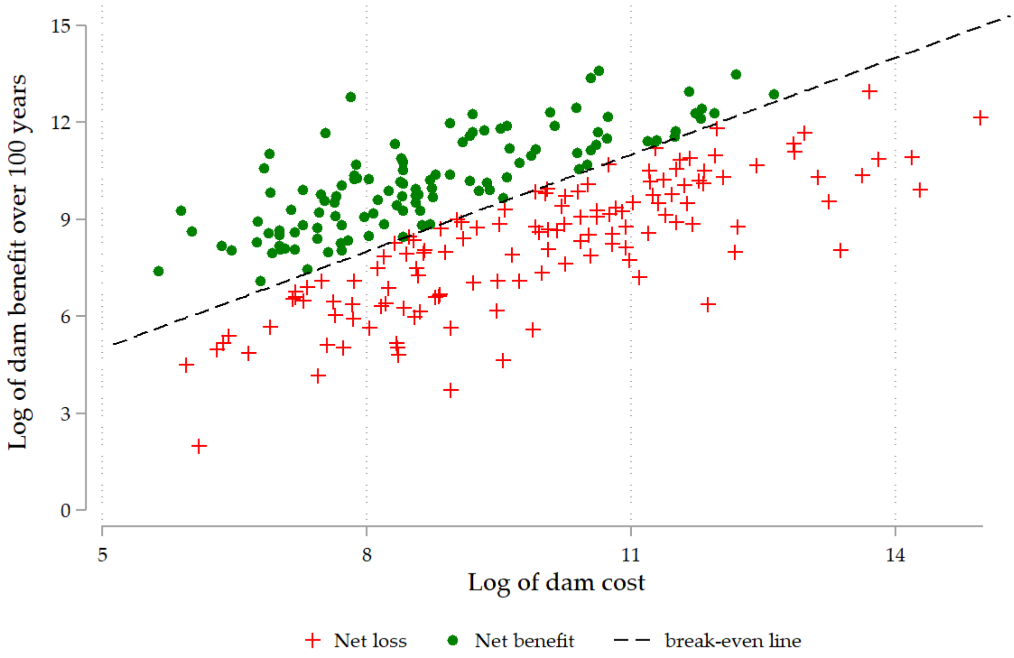

FIGURE 5: SCATTER PLOT OF LOG OF CALCULATED BENEFITS VERSUS LOG OF CALCULATED COSTS OF EACH DAM

Notes: Each point shows a dam built during our sample period. Red plus and green dots show dams with respectively negative and positive NPVs. The dashed line is the break-even point. The vertical axis shows the log of benefits accruing due to higher land rent and agricultural profits in the downstream of the dam over the course of 100 years, discounted by a 0.98 discount factor. The horizontal axis shows the log of predicted costs of the same dam.

and 194 dams show a negative NPV.**

6.4 *Sensitivity analysis*

Our cost benefit analysis is based on a number of assumptions. We discuss here how adjusting these assumptions affects our main conclusion.

First, the lifespan of dams is much larger than most other material investments, which implies that, by design, results are sensitive regarding the chosen discount rate. Given our finding that on average dams are not profitable, one may wonder to what our results change when the future is less, or, rather extremely, not, discounted. So we repeat the analysis assuming that the discount rate is 0.99 and 1.00. We then find that if the discount rate is 0.99 then the benefits increase by a factor of 1.5, so average benefits are still only 83% of costs, but in this case a bit more than half of dams (53 percent) are profitable. In the extreme case of no discounting, then we find that average benefits (101 Billion Rials) are 31% above costs, and the majority of dams

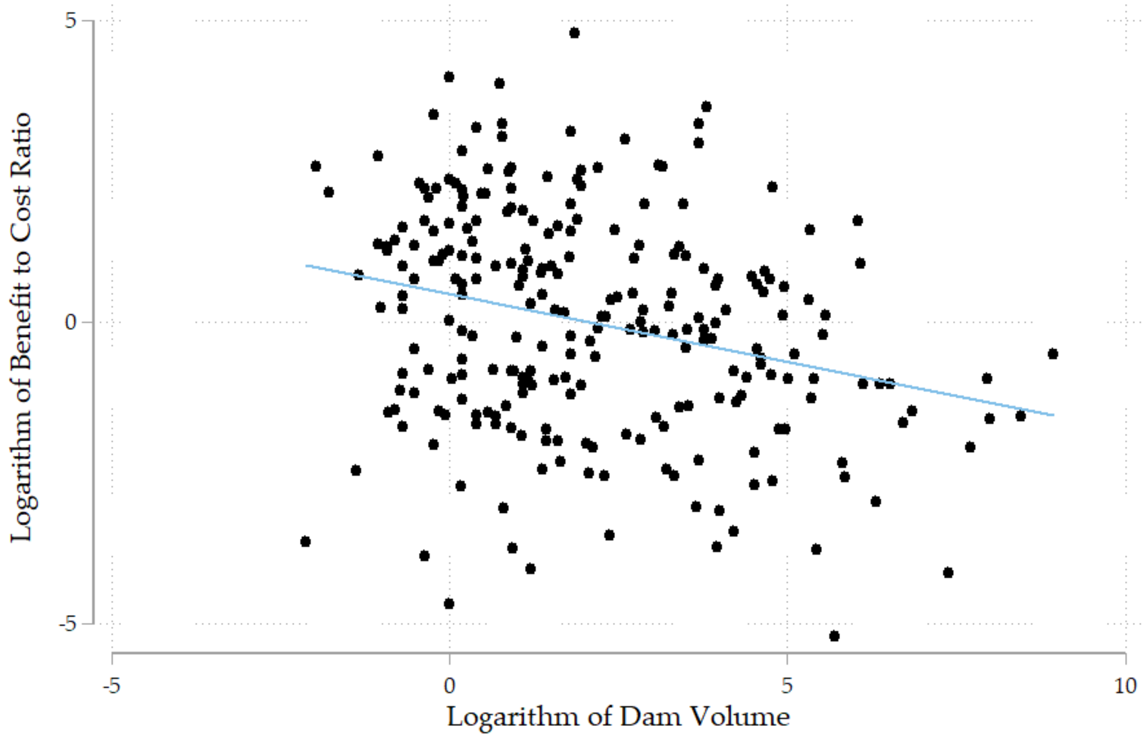

FIGURE 6: BENEFIT TO COST RATIO AND VOLUME OF DAMS

Notes: Each point shows log of a dam's benefit to cost ratio versus log of its volume after controlling for rainfall and year fixed effects. The slope is -0.23 and number of observations is 248.

(62 percent) are profitable.

Second, for a substantial share of dams (208 out of 248 dams built in our study period), we know how much percent of the adjusted water is used for agriculture. By using this ration, we adjust the cost which is attributed to agriculture and estimate the cost-benefit ratios for just agriculture. We find that average benefits are 5.8 time larger than average costs, and about half of dams (103 out of 208 dams) are profitable. Furthermore about one-third of dams are exclusively used for agricultural purposes. Subsequently, we repeat the analysis for those dams only. It appears now that 41 out of 85 pure-agricultural dams are profitable and the average benefit is 4 times bigger than cost. These results can be seen in Table 9.

Third, our estimated effects may be interpreted as average effects, but one may wonder to what extent these results are valid for subsamples of the population of dams. We remark here that the estimates are not precise enough to divide our overall population in two subsamples. We therefore repeat the analysis by excluding the 25% of counties that are located in (extremely)

TABLE 8: SUMMARY STATISTICS OF CBA UNDER VARIOUS SCENARIOS

	unit	N	mean	p10	p25	median	p75	p90
Panel A. Costs								
Cost	Billion Rials	248	77	1.1	2.6	8.5	46.1	134.2
Panel B. Unweighted dam volume (Long-Run)								
Benefit	Billion Rials	248	43.8	0.5	2.9	10.3	36.1	117.4
Benefit - Cost	Billion Rials	248	-33	-87.1	-15.5	-0.4	12.1	56.6
Benefit / Cost	-	248	3.8	0.06	0.19	0.86	3.43	9.65
Panel C. Unweighted dam volume (Short-Run)								
Benefit	Billion Rials	248	11.5	0.1	0.8	2.7	10.0	30.0
Benefit - Cost	Billion Rials	248	-65.5	-110.5	-35.1	-4.2	-0.2	5.5
Benefit / Cost	-	248	1.06	0.01	0.05	0.22	0.90	2.68
Panel D. Distance-weighted dam volume								
Benefit, distance weighted	Billion Rials	248	16.2	0.2	0.7	2.7	11.5	51.8
Benefit - Cost, distance weighted	Billion Rials	248	-60.1	-99.3	-28.6	-3.7	-0.5	2.9
Benefit / Cost, distance weighted	-	248	0.71	0.02	0.08	0.26	0.75	1.76
Panel E. Dam volume <50 km								
Benefit	Billion Rials	232	9.1	0.1	0.4	2.1	9.2	24.6
Benefit - Cost	Billion Rials	232	-71.1	-110.4	-36.4	-4.7	-0.7	2.7
Benefit / Cost	-	232	0.64	0.01	0.04	0.16	0.62	1.86

TABLE 9: SENSITIVITY ANALYSIS OF BOST-BENEFIT ANALYSIS

Elasticity	Number of Dams	Percent of Profitable Dams	Average Profit (Billion Rials)	Median Benefit/Cost	Note
Annual	248	24%	-65	0.23	All dams
Long differences	248	45%	-33	0.86	All dams
Long differences	85	49%	11	0.91	Pure Agricultural
Long differences	208	50%	7.8	0.96	Agricultural-Adjusted Cost

Notes: In first row, we use annual elasticities reported in Panel A of Table 3 and in second to fourth row, we use long difference elasticities reported in Panel B of Table 3 to calculate benefit of dams. In third row we report the cost-benefit analysis for 85 dams that are used just for agriculture. In the last row, we use the share of agricultural water use for 208 dams that we have this information and adjust the cost which is attributed to agriculture and estimate the cost-benefit ratio for agricultural use.

dry areas. .We also repeat the analysis by excluding the largest 25% dams.

7 Conclusions

In this paper, we study the economic effects of agricultural dams. We focus on Iran for which we have a wide variety of economic measures, including locally-measured agricultural land rent and profit. We demonstrate that the water volume of upstream dams moderately increases agricultural land rents and profits. Reassuringly, we do not find any effects for downstream dams. Standard difference in differences and triple-difference estimation strategies yield fairly similar results. We also show that long-differences estimates are several times larger, which is consistent with the idea that building supporting infrastructure, such as irrigation canals, takes time.

The triple difference strategy is a novel addition to the literature and relies on a comparison between irrigated and rainfed land. The robustness of results and the range of falsification test confirm that underlying identifying assumptions are valid. The introduction of a new strategy is highly desirable, as the more well-known shift share strategy that uses river gradient as an instrument suffers from a weak first-stage when applied. We also improve on this strategy relying on the engineering/hydrology literature by using water accumulation as an instrument, which provides indeed a strong instrument, but the second-stage confidence intervals of the estimates are still non-informative. The triple difference strategy also needs a different set of assumptions that can be tested.

The set of outcome variables used in this paper are new as well. We concentrate on agricultural land rent, farmer profits and wages paid to agricultural workers, but also show estimates for firms' revenue, agricultural output, irrigated land area and population. Land rent and farmer profits are probably the most relevant variables. Given efficient rental land markets, rent reflects the maximum value that could be extracted from land. For land owners, reported profits measures the added value to landowners. Our finding that rents and profits increase after dam construction shows that regulation of water resulted from dam construction, increase water security for farmers. We do not find any impact on wages, which may reflect the excess labour supply of unskilled labour in Iran. Our results for the effects of dams on population imply that dams induce local migration such that workers benefit from the improved agricultural opportunities created by dams.

We also conducted a back-of-envelope calculation of the costs and benefits of agricultural dams, where we zoom in on the role of dam size. We show that about half of all dams are profitable, and the aggregated profits, i.e. benefits minus costs, are close to 0. Furthermore, are substantial economies of scale in dam volume, but that there are also strong, even stronger, diminishing returns to dam volume. Therefore, a novel, but we believe important, result is that larger dams have lower benefits to cost ratios. Hence, the economic case to build more agricultural dams is sound, but the economic case to build large agricultural dams is rather weak.

References

Andreadis, K. M., Schumann, G. J.-P. & Pavelsky, T. (2013), 'A simple global river bankfull width and depth database: Data and Analysis Note', *Water Resources Research* **49**(10), 7164–7168.

Blanc, E. & Strobl, E. (2014), 'Is Small Better? A Comparison of the Effect of Large and Small Dams on Cropland Productivity in South Africa', *The World Bank Economic Review* **28**(3), 545–576.

Bohlen, C. & Lewis, L. Y. (2009), 'Examining the economic impacts of hydropower dams on property values using GIS', *Journal of Environmental Management* **90**, S258–S269.

Duflo, E. & Pande, R. (2007), 'Dams', *The Quarterly Journal of Economics* **122**(2), 601–646.

GSFC, P. P. S. P. A. N. (2019), 'GPM IMERG Final Precipitation L3 1 month 0.1 degree x 0.1 degree V06'.

Jeuland, M. (2020), 'The economics of dams', *Oxford Review of Economic Policy* **36**(1), 45–68.

JPL, N. (2020), 'NASADEM Merged DEM Global 1 arc second V001'.

Kitchens, C. (2013), 'A Dam Problem: TVA's Fight Against Malaria, 1926–1951', *The Journal of Economic History* **73**(3), 694–724.

Lehner, B. & Grill, G. (2013), 'Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems', *Hydrological Processes* **27**(15), 2171–2186.

Lewis, L. Y., Bohlen, C. & Wilson, S. (2008), 'Dams, Dam Removal, And River Restoration: A Hedonic Property Value Analysis', *Contemporary Economic Policy* **26**(2), 175–186.

Lipscomb, M., Mobarak, A. M. & Barham, T. (2013), 'Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil', *American Economic Journal: Applied Economics* **5**(2), 200–231.

Madani, K., AghaKouchak, A. & Mirchi, A. (2016), 'Iran's Socio-economic Drought: Challenges of a Water-Bankrupt Nation', *Iranian Studies* **49**(6), 997–1016.

Mesgaran, M. B., Madani, K., Hashemi, H. & Azadi, P. (2017), 'Iran's Land Suitability for Agriculture', *Scientific Reports* **7**(1), 7670.

Mettetal, E. (2019), 'Irrigation dams, water and infant mortality: Evidence from South Africa', *Journal of Development Economics* **138**, 17–40.

of Engineering, B. (2012), 'Guidelines and criteria for classification and coding basins in iran'.

Palmquist, R. B. (1989), 'Land as a Differentiated Factor of Production: A Hedonic Model and Its Implications for Welfare Measurement', *Land Economics* **65**(1), 23.

Portmann, F. T., Siebert, S. & Döll, P. (2010), 'MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling', *Global Biogeochemical Cycles* **24**(1), n/a–n/a.

Provencher, B., Sarakinos, H. & Meyer, T. (2008), 'Does Small Dam Removal Affect Local Property Values? An Empirical Analysis', *Contemporary Economic Policy* **26**(2), 187–197.

Ricardo, D. (1809), From the Principles of Political Economy and Taxation, *in* 'Readings in the Economics of the Division of Labor: The Classical Tradition', pp. 127–130.

Ritchie, H. & Roser, M. (2017), 'Water use and stress', *Our World in Data* . https://ourworldindata.org/water-use-stress.

Saemian, P., Tourian, M. J., AghaKouchak, A., Madani, K. & Sneeuw, N. (2022), 'How much water did Iran lose over the last two decades?', *Journal of Hydrology: Regional Studies* **41**, 101095.

Strobl, E. & Strobl, R. O. (2011), 'The distributional impact of large dams: Evidence from cropland productivity in Africa', *Journal of Development Economics* **96**(2), 432–450.

Tahbaz, M. (2016), 'Environmental Challenges in Today's Iran', *Iranian Studies* **49**(6), 943–961.

Von Thünen, J. H. (1842), *Der isolirte Staat in Beziehung auf Landwirthschaft und Nationalökonomie: Untersuchungen über den Einfluß, den die Getreidepreise, der Reichthum des Bodens und die Abgaben auf den Ackerbau ausüben*, Vol. 1, Leopold.

Wani, S. P., Rockström, J. & Oweis, T. Y. (2009), *Rainfed agriculture: unlocking the potential*, number 7 *in* 'Comprehensive assessment of water management in agriculture series', CABI, Wallingford.

World Commission on Dams (2000), *Dams and development: A new framework for decision-making: The report of the world commission on dams*, Earthscan.

Worqlul, A. W., Jeong, J., Dile, Y. T., Osorio, J., Schmitter, P., Gerik, T., Srinivasan, R. & Clark, N. (2017), 'Assessing potential land suitable for surface irrigation using groundwater in ethiopia', *Applied Geography* **85**, 1–13.

Yazdandoost, F. (2016), 'Dams, Drought and Water Shortage in Today's Iran', *Iranian Studies* **49**(6), 1017–1028.

Young, A. (2022), 'Consistency without inference: Instrumental variables in practical application', *European Economic Review* **147**, 104112.

Appendix

TABLE A1: AGRICULTURAL DATA AT DISTRICT LEVEL

Variable	Number of Observations (end-beginning)	Average (Standard Deviation) Total	Beginning	End	Min	Max
Area Under Cultivation (Thousand Hectares)						
Total	229-229	41959 (50136)	32247 (51533)	43907 (52160)	0	357301
Irrigated	229-229	19303 (22461)	13919 (21144)	19658 (22631)	0	149262
Rainfed	229-229	22656 (38889)	18328 (39604)	24249 (40507)	0	281885
Value of Production (Million Rials, Real)						
Total	229-229	5276356 (7158854)	4242352 (6831440)	6757302 (8233091)	0	$8.19 * 10^7$
Irrigated	229-137	4469903 (6534909)	5998252 (7463598)	5626409 (7320456)	0	$7.48 * 10^7$
Rainfed	229-229	878311 (1813207)	653892 (1500061)	1131195 (1988047)	0	$1.92 * 10^7$
Yield (Million Rials per Hectare)						
Total	227-138	189.49 (177.65)	219.93 (210.62)	217.57 (193.81)	0	2244.22
Irrigated	223-137	254.92 (214.72)	275.09 (189.92)	308.32 (234.05)	0	6446.64
Rainfed	184-100	50.1 (62.59)	35.74 (18.52)	68.38 (88.08)	0	1166.9

Notes: Beginning Period is 1991 and End period 2015. Data source is at district and crop level which is aggregated all crops in each district.

In this appendix, we first present hydrological information examples, i.e. how we delineate sub-basins and calculate flow accumulation. Second, we provide information about our sensitivity analyses. Third, we provide information on underlying methodology and results of the instrumental variable approach, where we discussed the standard instrument river flow as well as the improved instrument flow accumulation.

A.1 Hydrological information examples

We provide examples of how we delineate sub-basins and calculate flow accumulation. Figure A1 shows an example of one sub-basin (purple borders) which contains 10 counties and one dam. We define downstream using elevation of counties and river flow direction. The dam (triangle mark) is located in county 5 and is in the upstream of counties 4, 9 and 10. County 4 is downstream of counties 1, 2 and 3, because of river flow direction caused by elevation differences.

Figure A2 shows how we calculate flow accumulation of each cell in a hypothetical 6×6 grid. The left matrix shows the direction of water flow based on land gradient. Flow accumulation in each cell is then calculated by the cumulative count of all arrows that lead to the given cell. The resulting flow accumulation numbers are reported in the right matrix.

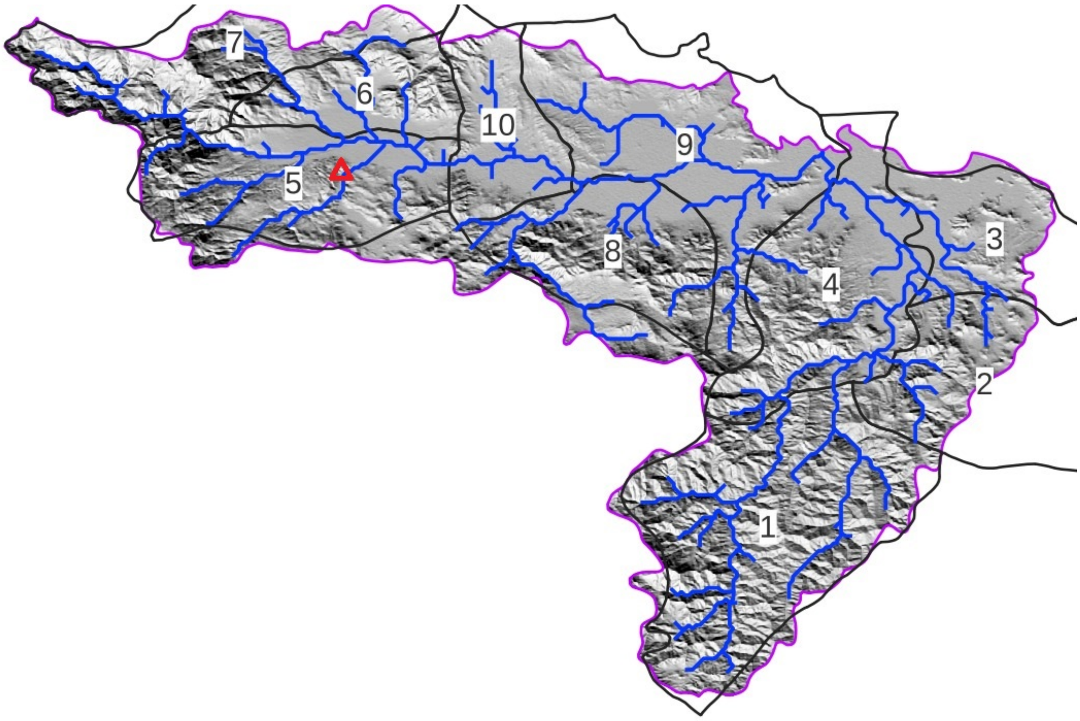

FIGURE A1: A SAMPLE SUB-BASIN BORDER, COUNTIES WITHIN IT, RIVER NETWORK AND A SAMPLE DAM

Notes: Border of the sub-basin is depicted in purple. Border of 10 counties overlapping with this sub-basin are shown in black. The triangle is a dam.

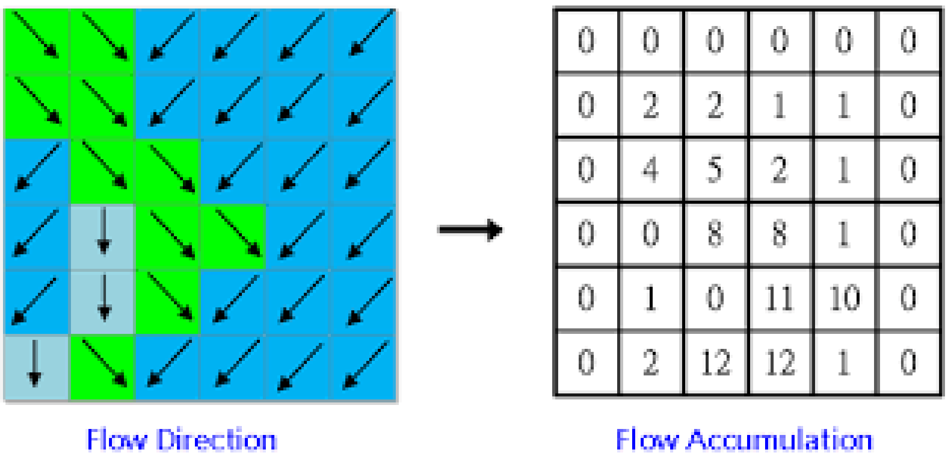

FIGURE A2: FLOW ACCUMULATION DEFINITION

Notes: This figure shows a schematic of flow accumulation calculation. The left figure shows water direction between adjacent cells and by using these directions, the right figure shows the flow accumulation (number of cells) to a given cell.

A.2 Robustness checks

We conduct several robustness checks. First, we look into the importance of changes in the functional form of the effects of upstream dam volume on agricultural rent. Column (1) and (2) of Table A2 include a dummy variable for the presence of upstream dams (the extensive margin) and interact this with the (standardised) logarithm of upstream volume, where the standardisation refers to that we subtract the logarithm of the average volume of upstream. This standardisation implies that marginal effect of the extensive margin when you build a dam with the average water value, is equal to the coefficient of dummy variable for the presence of upstream dam. Both columns suggest that the effects of the extensive as well as of the intensive margin are positive, supporting the use of the inverse hyperbolic sine transformation.

TABLE A2: ROBUSTNESS CHECK ON FUNCTIONAL FORM

Dep. var.: log (rent)	Interact log volume with dummy		volume=0.2, if no upstream dam		Counties with upstream dam	
	DID (1)	DDD (2)	DID (3)	DDD (4)	DID (5)	DDD (6)
Upstream	0.060	0.302***				
	(0.044)	(0.057)				
Upstream × log (vol. upstream)	0.017*	0.047***				
	(0.008)	(0.011)				
log (volume upstream)			0.014***	0.029***	0.006	0.031*
			(0.004)	(0.005)	(0.011)	(0.017)
Observations	24,303	13,258	24,583	13,258	18,494	9,737
$\bar{R}^2$	0.885	0.450	0.807	0.450	0.834	0.444

Notes: Columns (1) and (2) show the effect of upstream dam dummy and its interaction with logarithm of upstream dam volume on land rent using DiD approach for irrigated land in column (1) and DDD approach in column (2). In columns (3) and (4), we set upstream dam volume of counties without any upstream dam equal to 0.2. In last two columns, we exclude counties without upstream dams and use logarithm. We include year and county fixed effects in all columns. Standard errors are corrected for county clusters and reported in parentheses. *, **, *** represent significance at 10%, 5%, and 1% level respectively.

Columns (3) and (4) use the logarithm of volume upstream and set the value of upstream volume equal to 0.2 when there are no upstream dams. Columns (5) and (6) only includes counties that have upstream dams, so the number of observations is substantially reduced. Note that all specification support the results of the inverse sine hyperbolic transformation. Column (5) suggests that the effect of the inverse sine hyperbolic transformation is identified using information of the extensive margin.

Second, one important concern is the definition of sub-basin, which determines how we define upstream and downstream areas. Wrongly classifying dams as upstream or downstream may result in substantial measurement error in the calculation of dam volume and therefore,

TABLE A3: ALTERNATIVE DEFINITION FOR UP/DOWNSTREAM

Dep. var.	Rent *(log)* (1)	Profit *(log)* (2)	Revenue *(log)* (3)	Population *(log)* (4)
Upstream volume of dams *(asinh)*	0.028***	0.030**	0.020**	0.018***
	(0.008)	(0.013)	(0.009)	(0.005)
Own volume of dams *(asinh)*	0.001	0.011	0.004	-0.006
	(0.013)	(0.021)	(0.016)	(0.008)
Downstream volume of dams *(asinh)*	0.001	0.007	0.004	-0.014***
	(0.006)	(0.013)	(0.00879)	(0.003)
Observations	24,583	22,200	24,495	10,948
$\bar{R}^2$	0.807	0.616	0.768	0.753

Notes: This table shows the effect of volume of own, upstream and downstream dams on logarithm of land rent, farmer's profit and revenue, and population by using the second definition of upstream and downstream which is based on 6-digit HYDRO1K sub-basins. All regressions include time and county fixed effects. Standard errors are corrected for county clusters and reported in parentheses. *, **, *** represent significance at 10%, 5%, and 1% level respectively.

typically downward, biased estimates. Therefore, we use alternative sub-basin definitions (6-digit HYDRO1K) to designate upstream and downstream dams. Table A3 shows the estimation results for several variables of interest. The results seem very similar to our main results and show a statistically significant coefficient for the upstream variable while own and downstream coefficients are statistically insignificant for rent, profit, and revenue.

A.3 IV approach

In the extant literature on dams, studies employ an IV approach using an instrument that is based on *non-linear* interaction effects of gradient differences of rivers (i.e., the steepness of the river) and the predicted local number of dams constructed in a specific year. The idea behind this non-linearity is that rivers have to be sufficiently steep, but not too steep, for the building of dams.

We will show that these Duflo & Pande (2007) instrumental variables are not very suitable for our case. Then we propose a different instrumental strategy and compare results with our approaches.

We create four variables to capture the river gradient in counties. These variables capture the fraction of total rivers' length that fall into [0,3), [3,6), [6,9), and [9,∞) percent gradients. We take the first as the reference category and include the rest as instruments in a cross-sectional evaluation of their suitability. Columns (1) and (3) of Table A4 show the *first-stage* estimation results together with F-stats, respectively for number and volume of dams using cross-section variation only of dams built in 2017. While the three instruments are statistically significant, their F-stats are not large enough to minimize the risk of weak instruments. Columns (2) and (4) of Table A4 follow the literature more closely, where we use all panel data, but also include basin-by-year fixed effects.

To create a more suitable instrument, in the spirit of shift-share instrument, we multiply the initial share of dams in each water basin with the actual number of dams in the whole country in each year to construct the predicted number of dams for each basin, we then interact this with the river gradient variables to construct the instruments, as is common in the literature (Duflo & Pande (2007), Strobl & Strobl (2011), Blanc & Strobl (2014), Mettetal (2019)). In column (4) a similar strategy is used but volume is used instead of number of dams. The results of the first stage regressions are worse in these columns. Ignoring the weak first stage results, we find large estimated coefficients that are insignificant in the second stage (results not shown). These suggest that the usual instrument in the literature is not suitable in our case and probably more generally in fine-level studies like ours.

To construct a more suitable instrument, we follow the engineering literature to use county flow accumulation measures as predictors of dams (Andreadis et al. (2013)). Flow accumulation measures how much water is expected to flow through each grid cell given land gradient and upstream cells. For an example of how flow accumulation is calculated see appendix A.1. We restrict attention to cells in a county through which a river flows and compute the average of flow accumulation in these cells. The logic is that larger dams can be built in counties with higher flow accumulation levels. Hence, the first stage regression becomes:

$$\ln D_{rbt} = \beta_1 + \beta_2 FA * \ln \bar{D}_{bt} + \beta_3 Z_{rbt} * \ln \bar{D}_{bt} + \gamma_r + \alpha_{bt} + \epsilon_{rbt} \tag{A.1}$$

where D_{rbt} is the IHS of dam volume in county r and water basin b in year t. $\bar{D}_{bt}$ is the predicted volume of dam incidence in the basin. This is constructed by multiplying total volume of dam construction in Iran with the initial fraction of dam volume in the basin in 2000. Use of predicted, rather than actual, dam volume ensures that the measure is exogenous with respect to the dam volume in the county.

To generate instruments for D_{rbt}, $D^{\mathcal{U}}{}_{rbt}$, and $D^{\mathcal{D}}{}_{rbt}$, we use the same method as Duflo & Pande (2007). We use parameters from A.1 to predict the volume of dams per county $\hat{D}_{rbt}$. For upstream counties, the predicted volume of dams, $\hat{D}^{\mathcal{U}}_{rbt}$, is the sum of predicted values from A.1 for all upstream counties (it equals zero if the county has no upstream county). We do the same for downstream counties. Then, we estimate 1 using $\hat{D}_{rbt}$, $\hat{D}^{\mathcal{U}}_{rbt}$, $\hat{D}^{\mathcal{D}}_{rbt}$, $\hat{Z}_{rbt}$, $\hat{Z}^{\mathcal{U}}_{rbt}$, and $\hat{Z}^{\mathcal{D}}_{rbt}$ as instruments.

Table A5 shows the first and second stage results for the flow accumulation instrument. The first stage show F-stat improvements for both cross-sectional (column 1) and panel versions of the instrument (column 2). The flow accumulation is a much stronger predictor of dam volume than river gradient because from an engineering perspective it is the amount of water that could be regulated that matters and not just river gradient. For the second stage we observe mostly insignificant results reflecting very imprecisely estimated coefficients. However, the coefficients have expected signs and are large. Therefore, we believe that our differencing strategies are better and more precise compared to instrumental variables estimation.

TABLE A4: FIRST STAGE RESULTS USING DUFLO & PANDE (2007) INSTRUMENTAL VARIABLES

Dep. var.	Number of Dams		Volume of Dams	
	Cross Section	Panel	Cross Section	Panel
	(1)	(2)	(3)	(4)
Fraction river Gradient	-0.0621**	-0.146	-0.112**	-0.225
3-6%	(0.0271)	(0.120)	(0.0483)	(0.326)
Fraction river Gradient	0.189***	0.860***	0.357***	2.045***
6-9%	(0.0450)	(0.259)	(0.0974)	(0.744)
Fraction river Gradient	-0.181***	-0.572***	-0.217**	-1.206
above 9%	(0.0468)	(0.180)	(0.101)	(0.870)
F-Test	7.57	5.28	4.68	3.21
p-value	[0.000]	[0.001]	[0.003]	[0.02]
Observations	2,173	39,114	2,173	39,114
$\bar{R}^2$	0.124	0.954	0.052	0.838
Fixed effects	Sub-basin	county	Sub-basin	county
Basin-Year FE	N	Y	N	Y

Notes: This table shows first stage regression results for instruments similar to Duflo & Pande (2007). The dependent variable in columns (1) and (2) is the number of dams, while it is the volume of dams in columns (3) and (4). Columns (1) and (3) use the cross-sectional sample of dams in 2017 with cross-sectional geographical instruments, including the fraction of river gradient between 3-6 percent, between 6-9 percent, and above 9 percent. The omitted reference gradient category is 0-3 percent. Columns (2) and (4) use the whole sample with a time varying instrument derived by multiplying predicted number (columns 1 and 2) or volume (columns 3 and 4) of dams in the basin by the cross-sectional instruments. Panel regressions include county, basin-by-year fixed effects, annual rainfall and linear trends interacted with geographical controls including elevation, overall gradient, county area, and river length. Cross-sectional regressions include similar controls for the year of the cross-section. Basin fixed effects are used here. Standard errors are corrected for county clusters and reported in parentheses. *, **, *** represent significance at 10%, 5%, and 1% levels respectively.

TABLE A5: FIRST AND SECOND STAGE RESULTS USING FLOW ACCUMULATION INSTRUMENT

Dep. var.	Volume of Dams		Irrigated rent		
	Cross Section	Panel	Baseline	+Geo. Controls	+ Basin-Year FE
	(1)	(2)	(3)	(4)	(5)
Flow Accumulation	0.0121***	0.0409***			
	(0.0034)	(0.0108)			
Own			0.288	0.305	0.393
			(0.222)	(0.221)	(0.294)
Upstream			0.250*	0.230	0.150
			(0.140)	(0.148)	(0.129)
Downstream			-0.053	-0.058	-0.078
			(0.044)	(0.045)	(0.060)
F-test	12.66	14.3			
p-value	[0.000]	[0.000]			
Observations	2,173	39,240	24,560	24,505	24,505
Time FE	N	Y	Y	Y	Y
Fixed effects	basin	county	county	county	county
Basin-Year FE	N	Y	N	N	Y
Geography controls	Y	Y	N	Y	Y

Notes: Column (1) and (2) show the first stage results of regressing IHS of dam volume in each county on flow accumulation. Column (1) follows a cross-sectional estimation for the year 2017. Column (2) interacts flow accumulation by the predicted volume of dams in the basin. Columns (3) to (5) show the second stage. Column (3) we does not control for geographical trend variables and basin-by-year fixed effects. Column (4) controls for geographical trends and column (5) controls for both geographical trends and basin-by-year fixed effects. The first stage corresponding to column (5) is shown in column (2). All panel regressions include year fixed effects. Standard errors are corrected for county clusters and reported in parentheses. *, **, *** represent significance at 10%, 5%, and 1% levels respectively.