Adapting to Disasters: Evidence from Drought in Brazil

Guilherme DePaula* Behzad Jeddi† David Keiser‡

February 7, 2022

Abstract

Can societies adapt to more frequent and intense weather disasters? We estimate the value and the limit of disaster adaptation in Northeast Brazil, a semi-arid region with a long history of droughts. The main contribution of this paper is the identification of the effects of drought frequency and reservoir construction on economic and population outcomes. We create a dataset that replicates the network of rivers and monitoring stations in Brazil and we use within state-biome variation in extreme river flow and river gradient to identify the impact of drought frequency and the adaptation value of reservoir placements. We find that the adaptation value of reservoirs is large but declines rapidly with drought frequency. Adding a reservoir offsets the impact of an increase in drought frequency on GDP per square kilometer by 27% and on population density by 26%. However, the contribution of a reservoir declines to zero at a frequency of 3 droughts per decade. Reservoirs expand subsistence agriculture, increasing vulnerability, but also accelerate rural-urban migration, improving resilience.

Keywords: Disaster adaptation; Migration; Drought, Reservoirs; Brazil

JEL Codes: Q54; O18; N56; N96

*Assistant Professor, Department of Economics and Center for Agricultural and Rural Development (CARD), Iowa State University. gdepaula@iastate.edu

†Graduate Student, Department of Economics, Iowa State University. bjeddi@iastate.edu

‡Professor, Department of Resource Economics, University of Massachusetts Amherst. dkeiser@umass.edu

1 Introduction

Can societies adapt to more frequent and intense weather disasters? Climate change can lead to more frequent and intense extreme weather events (IPCC, 2012) and adapting to changes in disaster frequency and intensity may be harder than adapting to slow changes in temperature and precipitation (Olmstead 2014; Hansen et al. 2011; Zaveri et al. 2021; Annex 2012; Botzen et al. 2020; Coronese et al. 2019; Van Aalst 2006; Douville et al. 2021). In this paper, we estimate the value and the limit of disaster adaptation in Northeast Brazil (NEB), a semi-arid region with a long history of droughts.

NEB has a population of 57 million people in an area of 600 thousand square miles, close to the size of Alaska. Over 60% of the NEB is semi-arid land with a pattern of irregular rainfall concentrated in a three-month rain season. In such an environment, even moderate changes in rainfall could lead to droughts. After the "Great Drought" of 1887, which killed 500 thousand people in NEB, the Brazilian emperor approved the construction of the first large reservoir (Marengo 2010; Marengo et al. 2017). Since then, over 300 large public reservoirs have been built in NEB. The great drought also started the first wave of migrants from NEB as 250 thousand northeasters left NEB to work on rubber plantations in the Amazon (Villa 2000). By 2000, 9.67 million northeasters were living in other regions of Brazil, or 17 percent of the people born in NEB (IBGE, 2000).

The main contribution of this paper is the identification of the effects of drought frequency and reservoir construction on economic and population outcomes. We construct two sets of instrumental variables to address the endogeneity of drought reporting and the endogeneity of reservoir placement.

The reporting of disasters such as droughts may increase in more vulnerable locations where the potential damage is higher. Empirical models using observed disaster frequency would likely overstate the damage from droughts. We address this endogeneity problem by creating a counting model that explains disaster occurrence using extreme flow and rainfall events. We use a similar approach as Keiser and Shapiro (2019) to build a dataset that replicates the river network and the network of flow and rainfall monitoring stations in NEB. The shapefiles and measurements are sourced from the Brazilian National Water and Sanitation Agency (ANA) (ÁGUAS-ANA 2019). We use daily records of flow and rainfall to identify 3- and 20-year flow and rainfall events across the river network and predict drought occurrences.

The placement of reservoirs is influenced by political and economic factors. Estimates based on observed reservoirs would likely bias the benefits of reservoirs. We follow the approach of Duflo and Pande (2007) to instrument the number of reservoirs constructed using variation in river gradient. We use remote sensing data on surface elevation and slope from the Brazilian National Institute of Space Research (INPE) to compute the fraction of

different slope classes within a municipality: flat, moderate, steep, and very steep gradients (ESPACIAIS-INPE 2019).

We construct a dataset combining a panel with outcomes and instrumental variables with reservoir and disaster data. Our reservoir dataset comes from the Brazilian National Department of Works Against Drought (FEDERAL 2005) and has data for 388 large reservoirs. We use disaster reports from the Integrated Disaster Information System (S2ID) (Civil 2015). There are 35,586 short droughts and 10,784 long droughts reported in NEB. Our sample has a total of 280 microregions, 3,032 municipalities, and 25 state-biome combinations. Our outcome data is a panel with annual GDP, population, and agricultural production measurements based on surveys from the Brazilian Institute of Geography and Statistics (Populacional 2010; Agropecuário 1996)). GDP data is available from 2002 to 2019. Population data is available from 1872 to 2020. Data on agricultural outcomes are available from 1974 to 2020. The microregion-level panels with GDP, population, and agricultural outcomes have 4,158, 9,492, and 10,566 observations, respectively.

Our empirical model is a log-linear regression equation with instrumental variables for drought frequency and number of reservoirs. The dependent variables are the log of GDP per square kilometer and the log of population per square kilometer. The parameters of interest are the effect of an increase in disaster frequency without reservoirs and the mediating effect of reservoirs. We identify drought frequency effects using within state-biome exogenous variation in extreme river flow events in the NEB river network and we identify reservoir effects using within state-biome exogenous variation in river gradients.

We have four main results.

First, we find that the exponential function form for disaster damage fits well the economic damages from an increase in drought frequency in Brazil. This result is consistent with Martin and Pindyck (2015) specification of disasters as consumption shocks with Poisson arrival rates. The exponential damage function departs from common smooth specifications for the damages of changes in mean temperature (Mendelsohn et al. 1994; Deschênes and Greenstone 2007). The impact of changes in mean temperature tends to become nonlinear at high temperatures or with large temperature changes (Schlenker and Roberts 2008). By contrast, the impact of small changes in drought frequency is higher in places with no droughts, for example, close to the border of the drought polygon in Brazil.

The adaptation value of reservoirs is large but declines rapidly with drought frequency. Adding a reservoir offsets the impact of an increase in drought frequency on GDP per square kilometer by 27% and on population density by 26%. However, the contribution of a reservoir declines to zero at a frequency of 3 long droughts per decade. We also find a distinct pattern for the effects of drought frequency and reservoirs on productivity and population density. In places with few droughts, the drought frequency effect on productivity and population

density are comparable after accounting for adaptation. However, in locations with higher drought frequencies, further increases in frequency affect more productivity than population density.

The places most vulnerable to increases in drought frequency are those that currently have no droughts or reservoirs. In these locations, an additional drought every 10 years reduces GDP per square kilometer by approximately 47% and population density by 27%.

Finally, We examine the linkage between drought adaptation and migration in a developing economy. We find mixed evidence to support FURTADO (1959) hypothesis for the retention of potential NEB migrants in drought-prone locations. Reservoirs preserve and expand subsistence agriculture in NEB as Furtado predicted. However, reservoirs also facilitate urbanization and rural-urban migration in NEB. These results highlight the linkages between disaster adaptation and migration policies in developing countries(Cattaneo and Peri 2016; Boustan et al. 2012; Zaveri et al. 2021).

The remaining of the paper is organized as follows. We first extend the model of Martin and Pindyck (2015) to derive a damage function for disasters and an equation for the value of disaster adaptation. We describe the historical context in section 3 and the dataset in section 4. We present our empirical analysis and main results in section 5. We extend our analysis to investigate the trade-off between adaptation and migration in section 6. Section 7 concludes. The construction of the dataset and descriptive statistics are documented in the appendix.

2 A Damage Function for Disasters

The starting point of our analysis of disaster adaptation is a damage function that explains the economics losses of an increase in disaster frequency and severity. We start with a simple model for disaster damages following the work of Martin and Pindyck (2015) and Barro (2009). Disasters are modeled as permanent shocks to consumption and GDP. Without disasters, GDP per capita grows at a constant rate g. Each disaster event then permanently reduces log of GDP per capita by a random amount ϕ. The log of GDP per capita at time t for an economy with $Q(n)$ disasters is:

$$logY_t = logY_0 + gt + \sum_{n=1}^{Q(n)} \phi_n \tag{1}$$

Disasters are defined by two parameters: frequency and severity. We assume that the number of disasters $Q(n)$ follows a Poisson process. The frequency of the disasters is then the Poisson arrival rate λ. λ measures the average number of disasters in a fixed period of time. For example, for a 100-year period of analysis, an arrival rate equal to 0.05 (0.5) implies that a disaster occurs once every 20 (2) years.

The severity of a disaster event i, ϕ_i, is the economic damage of the disaster. The severity parameter ϕ is approximately the fraction of GDP per capita reduced by the disaster. We assume that the disaster severity is an independent and identically distributed random variable with a power distribution function (Martin and Pindyck 2015; Barro 2009). Define $z_i = e^{-\phi_i}$ as the fraction of GDP per capita remaining after disaster i ($0 < z_i < 1$). We assume that z_i has the power distribution $\beta_i z_i^{\beta_i - 1}$. The expected value of z_i is $\frac{\beta_i}{\beta_i + 1}$ and therefore a large value of β implies a small expected disaster damage. The most damaging disasters will have small values of β.

Disaster adaptation is a reduction in disaster frequency. For example, governments can build infrastructure such as water reservoirs or seawalls to reduce the probability of damaging droughts and floods. A complete adaptation would mean reducing the expected arrival rate of a disaster to zero. Partial adaptation would reduce the arrival rate by a fraction p. The expected disaster arrival rate with adaptation is then $(1-p)\lambda$ where p captures the level of disaster adaptation. Large investments in adaptation would result in a large value for p.

We can now express an economic measure of welfare as a function of disaster frequency, severity, and adaptation [1]. The expected value for the discounted stream of GDP per capita is:

$$W(\lambda, \beta, p) = E \int_0^\infty Y_t e^{-\delta t} dt = \frac{Y_0}{\delta - g + \frac{1-p}{1+\beta}\lambda} \tag{2}$$

The denominator can be interpreted as a discounted rate adjusted for economic growth and disaster impact. The term $\frac{1-p}{1+\beta}\lambda$ is the effective disaster frequency after accounting for adaptation and disaster severity. The effective frequency decreases with a high level of adaptation, high p, and with a low disaster severity, high β.

Equation 2 is the damage function of disasters. Figure 1 shows the damage function for disasters with high and low severity as a function of disaster frequency. The solid lines represent the damage without adaptation, $p = 0$. The dashed and dotted lines are the damage functions with 50% and 80% adaptation respectively ($p = 0.5$ and $p = 0.8$).

A striking feature of the disaster damage function is the nonlinear relation with respect to frequency, severity, and adaptation. For low severity disasters, the damage function is a convex function of frequency and damages tends to increase slowly (Figure 1 - Panel A). In this case, the value of adaptation, the vertical distance between the solid and the dashed or dotted lines, increases with frequency. By contrast, with high severity disasters, damages grow rapidly with frequency and the value of adaptation is a concave function of frequency (Figure 1 - Panel B). In the high severity case, the value of adaptation is very high at low frequencies but decreases at medium to high frequencies.

[1] We follow Martin (2008) and Martin and Pindyck (2015) in using the cumulant-generating function (CGF) to model the consumption and output process. The CGF is the log of the moment-generating function: $k_t(\theta) = logEY_t^\theta$. Assuming that output is a Levy process, the CGF is $k_t(\theta) = g + \lambda(Ee^{-\theta\Phi} - 1)$ and we can express output as $EY_t = e^{k_t(1)t}$. The integration in Equation (2) then follows by using the CGF transformation of the output process: $E\int_0^\infty Y_t e^{-\delta t} dt = E\int_0^\infty e^{k_t(1)t} e^{-\delta t} dt$.

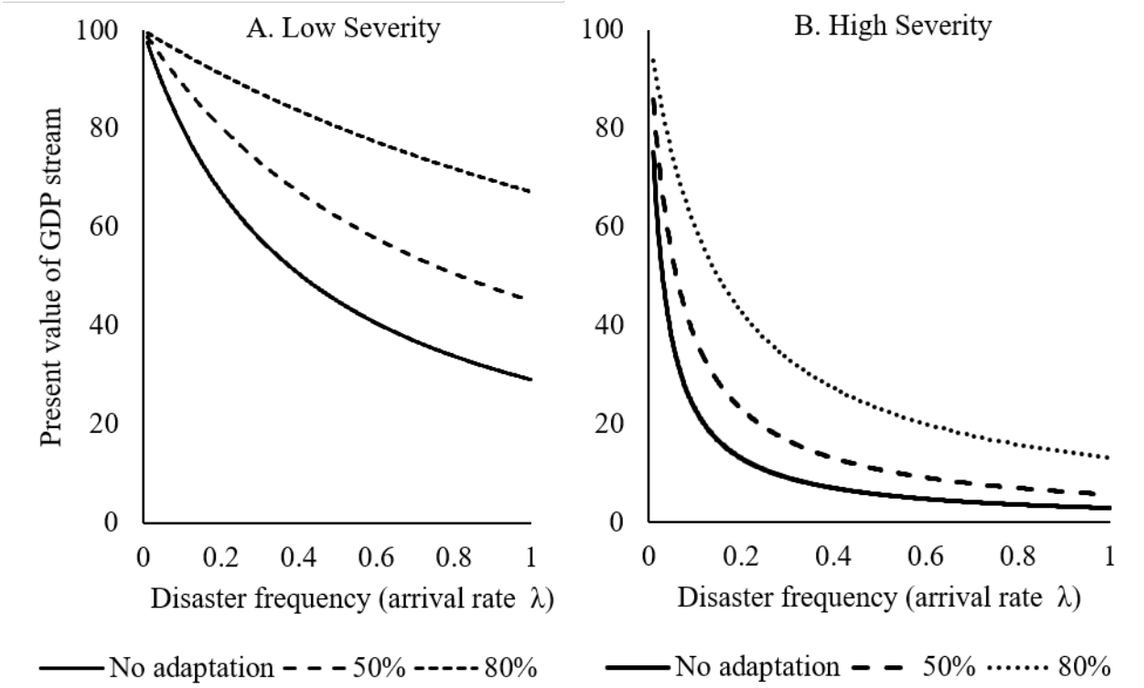


Figure 1: Disaster Damage function

More formally, we can express the value of adaptation as the cross derivative of the welfare function with respect to frequency and adaptation, $W_{\lambda.p}$:

$$W_{\lambda,p} = \frac{1}{1+\beta}W^2 + 2(\frac{1-p}{1+\beta})^2 W^3 \tag{3}$$

Equation (3) measures the marginal benefit of adaptation or the welfare gain from reducing the frequency of disaster by a small amount. The value of adaptation is always positive but decreases fast with a cubic function of welfare. There are therefore limits to adaptation in the case of disasters that are not related to the actual costs of adaptation investments but to the catastrophic nature of severe disasters. For example, it is possible to build an infrastructure of water reservoirs and channels that protects populations and economies from the sporadic long drought. However, as the frequency of long drought increases the economic impacts in sectors like agriculture becomes so large that the benefit of an additional reservoir is very small.

3 Drought Adaptation in Northeast Brazil

The northeast of Brazil (NEB) has a long history of droughts and drought-adaptation policies. Figure 2 shows the spatial distribution of droughts in Brazil. NEB is the dark red northeast corner of Brazil in the map, a region with 57 million people in an area of 600 thousand square miles, close to the size of Alaska. Over 60 percent of the NEB is classified as semi-arid land and is called the "drought polygon" for its polygon shape and high frequency of droughts [2]. This semi-arid land is distinct for a pattern of irregular rainfall concentrated

[2]NEB has 9 states: Maranhao, Piaui, Ceara, Rio Grande do Norte, Paraiba, Pernambuci, Alagoas, Sergipe, and Bahia. The drought polygon extends south of NEB and covers part of the state of Minas Gerais.The Brazilian government determines which municipalities are part of the drought polygon using three criteria: (1) average annual rainfall under 800 millimeters, (2) aridity index under 0.5, and (3) drought

in a three-month rain season, a network of rivers that drains towards the Atlantic ocean and is dry for six months of the year, shallow soils, and a shrubland vegetation called Caatinga. In such an environment, even moderate changes in rainfall could lead to droughts. Over 130 severe droughts have been documented in the NEB since the 16th century (Marengo et al. 2017) [3].

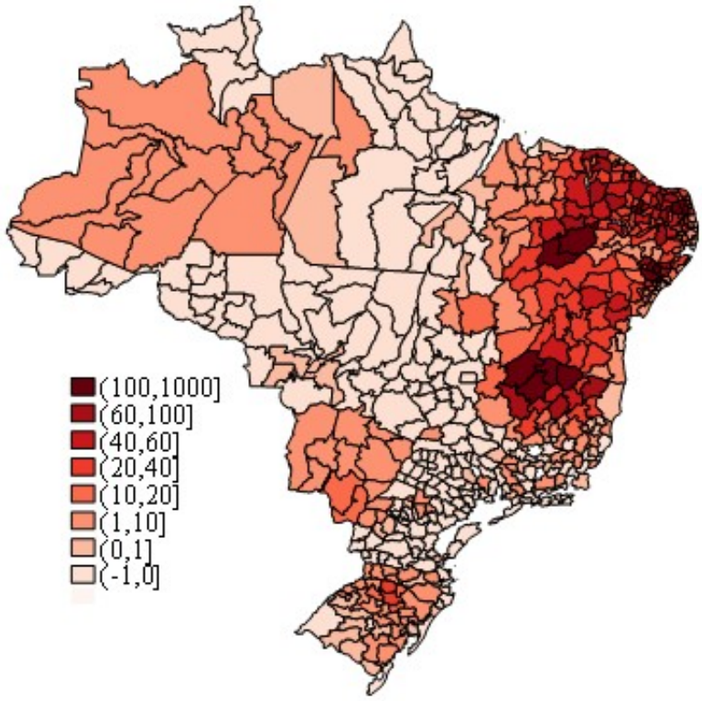


Figure 2: Number of Long Droughts in Brazil

The Brazilian government intervened to mitigate the impacts of droughts in NEB with long-term and short-term policies. The most important long-term policy was the construction of large water reservoirs. These large water reservoirs are dams build across rivers to store water during the rain season to ensure water supply to population and agriculture during the dry season and long droughts. After the "Great Drought" of 1887, which killed up to 500 thousand people in NEB (Marengo 2010; Marengo et al. 2017), the brazilian emperor Don Pedro II approved the construction of the first large reservoir in NEB, the Cedro reservoir. The Cedro reservoir was completed in 1906 with the capacity to store 100 million cubic meters of water. In 1909, the Brazilian government created the first institution for the planning and construction of a hydraulic infrastructure in NEB, the "Inspetoria de Obras Contra as Seca" (IOCS). In 1949, IOCS became the current institution, the "Departmento Nacional de Obras Contra Seca" (DNOCS). Figure 3 shows the time series of large public reservoirs built in NEB. Reservoir construction tended to accelerate after the occurrence of severe droughts and slow down during periods of normal rainfall. By 2005, DNOCS had constructed 323 reservoirs in NEB with the capacity to store 27 billion metric cubics of water (FEDERAL 2005).

risk over 60 percent based on drought occurrence from 1970 to 1990. The aridity index is the ratio of annual precipitation to annual potential evapotranspiration. There are currently 1,133 municipalities in the drought polygon, including 85 municipalities outside of NEB, in the neighboring state of Minas Gerais.

[3]The first documented drought in NEB happened in 1533. This first drought was reported by Jesuit priest Fernao Cardim (CARDIM 1925). Cardim reported that thousands of indigenous people migrated to the coast to escape the drought.

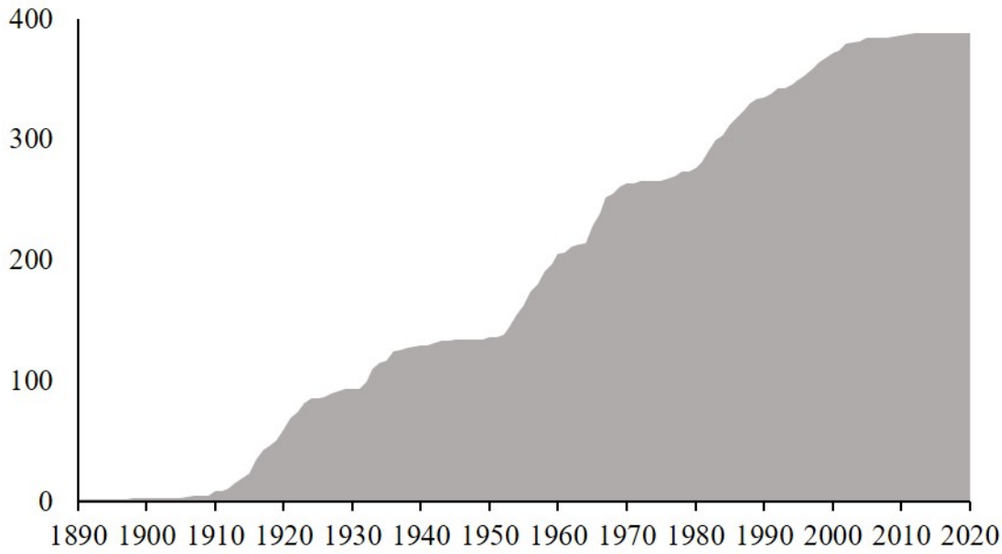


Figure 3: Number of Public Reservoirs Constructed in NEB

The short-term policies are a variety of emergency assistance programs. One of the first such programs were the work fronts designed to employ millions of farmers affected by droughts (Campos and de Carvalho Studart 2008). For example, a work front created after the 1979-83 drought generated 3 million temporary jobs in the NEB (De Nys et al. 2016). Another policy still used in NEB is the water truck program, designed to transport water from large reservoirs to small private water wells in the semi-arid using thousands of water trucks (Assunção and Livingstone 1993). More recently, there has been a series of transfer programs designed to supplement the income of poor families in NEB. The Drought Stipend program (Bolsa Estiagem) assists family farmers with a monthly income of two minimum wages, The Pronaf Semiarido provides subsidized credit for family farmers, and the Garantia-Safra offers crop insurance for extreme drought events (Marengo 2010; Marengo et al. 2017, 2021) [4]. These emergency programs have been criticized in Brazil for enabling an agricultural subsistence economy in NEB that is vulnerable to droughts. The influential NEB economist Celso Furtado argued that short-term interventions have contributed to fixate the low-income population in the semi-arid, increasing the population density of the region and preserving a subsistence agriculture sector considerd the most vulnerable to droughts (FURTADO 1959) [5].

Many northeasters have migrated to cities within and outside NEB to escape from droughts. The great drought of 1887 started the first wave of migrants from NEB. About 250,000 northeasters left NEB after the great drought to work on rubber plantations in the Amazon region (Villa 2000). A few decades later, the industrialization of the southeast region of Brazil spurred new waves of migration from NEB. This industrialization process started

[4]There is an extensive literature reviewing the history of drought relief programs in Brazil. See Marengo (2010), Campos and de Carvalho Studart (2008), and Fonseca et al. (2014) for a review of this literature.

[5]A striking feature of the drought polygon is its relatively high population living under extreme climate and economic conditions. In 2019, the polygon's population was 27.7 million people, almost half the NEB population, and the polygon's GDP was 72.6 billion USD, or 37 percent of the NEB GDP (IBGE, 2020). The polygon's GDP per capita of 2,621 USD is the same as Zimbabwe's, which has the 33rd lowest GDP per capita in Africa.

in the 1930s and lasted for 50 years. The Department of Immigration and Colonization of Sao Paulo (DTCI) recorded 924 thousand migrants from NEB from 1950 to 1959 (De Nys et al. 2016). The peak of migration from NEB to Sao Paulo happened during the 1951-1953 drought when 450 thousands northeasters were registered by DTCI. In 2000, there were 9.67 million northeasters living in other regions of Brazil, or 17 percent of the total number of people born in NEB. 70 percent of these NEB migrants lived in the southeast region of Brazil in 2000 (IBGE, 2000). An even larger migration happened within the NEB as part of the urbanization of the semi-arid (Lima and Magalhães (2019); De Nys et al. (2016)), 2014; (Ojima and Fusco 2014). Rural workers moved into growing mid-size cities in the semi-arid looking for urban jobs. According to the census of 2000, 65.5 percent of the northeasters that left their residency moved to another municipality in NEB and 54.4 percent moved to a municipality within the same state of residency (Ojima and Fusco 2014; do Nascimento and de Oliveira 2015; Assunção and Livingstone 1993).

4 Data

Our main dataset is a panel with annual measurements of GDP, population, and agricultural production. The outcome variables are GDP per capita, GDP per area, population density, rural and urban population, agricultural value per area, agriculture value per employee, and cattle stock. All the outcome variables are from surveys from the Brazilian Institute of Geography and Statistics (IBGE). The number of years of data available varies by outcome variable. IBGE has GDP data at the municipality level from 2002 to 2019 (de Geografia and de Contas Nacionais 2005). Population data also at the municipality level is available from two IBGE surveys. The first survey is the population census with data for every 10-year census survey, starting in 1872 and ending with the last census in 2010 (Populacional 2010). IBGE also has annual estimates of municipality population from 1992 to 2019 (**?**). The agricultural outcome variables are available from the Municipal Livestock Production survey (PPM) and the Municipal Agricultural Production survey (PAM) for the period from 1974 to 2020 (IBGE 2020). See Appendix A for a description of each variable and summary statistics.

We then match the DNOCs reservoir data with the information from the reservoir tracking system (SAR) from the National Water Agency (ÁGUAS-ANA 2019) and from the World Register of Dams database maintained by the International Commission on Large Dams (ICOLD) (on Large Dams 1979). We identify 388 large reservoirs built in NEB and neighboring states. Figure 3 shows a time series of reservoir construction in NEB.

We obtained the disaster data from the Integrated Disaster Information System (S2ID) created in 2011 by the Brazilian National Secretariat for Civil Protection and Defense (SEDEC) and the Center of Studies and Research in Engineering and Civil Defense at the Federal University of Santa Catarina (CEPED). We merged two S2ID reports: the S2ID

management reports on disaster damages since 2013 (S2ID, 2021a) and the digital archives files on disaster reports since 1900 (S2ID, 2021b). Both reports include data on federal disaster declarations, damage reports, and SEDEC disaster responses. Disaster records have the type, date, and location of the disaster. We gathered information on four types of disasters: (1) short drought ("estiagem") identified by a delay over 15 days from the start of the rainy season and average monthly precipitation under 60 percent of historic averages; (2) long droughts ("secas") identified by a sustained depletion of water reserves; (3) flash floods ("enxurrada") defined as intense and rapid flow of water resulting from intense and concentrated extreme rainfall; (4) floods ("inundacao") defined as overflow of rivers and streams. There are 35,586 short droughts, 10,784 long droughts, 2,825 flash floods, and 3,914 floods reports for NEB and neighboring states. We calculated disaster frequencies by first aggregating disaster reports by year and microregion to avoid double counting disasters. Multiple reports of a type of disaster in the same microregion and year were counted as one disaster. We then calculated frequencies for each microregion and disaster type using disaster reports after 1980 [6].

Our preferred sample covers NEB plus three neighboring states: Minas Gerais, Goias, and Tocantins. The north region of Minas Gerais is also semi-arid land and is part of the drought polygon. Tocantins and Goias also have dry winter seasons with propensity to droughts. There were 172 droughts reported for both states. The inclusion of the three neighboring states helps the estimation of damage functions by increasing the variation in drought frequency and economic outcomes. Our preferred unit of observation is a microregion. There are 280 microregions and 3,032 municipalities in our sample. A microregion is on average 10 times larger than a municipality. The average area of a municipality in our sample is 909 square kilometers whereas the average area of a microregion is 9,887 square kilometers, about the size of Puerto Rico [7]. The choice for a larger unit of observation mitigates double counting of disasters and facilitates the matching of the extreme events across the river network with disaster occurrences.

We also built two novel datasets to create instrumental variables for disaster frequency and reservoir. We built a dataset replicating the river network in NEB and integrated it with flow and rainfall data to compute extreme flows and rainfall measures that explain the frequency of disasters at NEB. We used remote sensing data to construct a dataset for the slope of the NEB terrain to explain the placement of reservoirs. We explain the construction of these datasets in the empirical section.

[6] There is a clear change in the number of disasters reported before and after 1980 likely due to misreporting in earlier decades.

[7] smallest municipality has 41 squared miles and is located in Minas Gerais and the largest municipality has 16,404 squared miles and is located in Bahia. The smallest microregion is the metropolitan region of Fortaleza in the NEB coast. The largest microregion is in the southeast of Goias, close to the western border of NEB.

A Simple Decomposition for the Reservoir Effect

Table 1 shows a decomposition of the effect of reservoirs on GDP per capita. Adding a reservoir in a drought-prone region could improve land productivity and reduce migration by stabilizing water supply. We can disentangle these two effects by decomposing the log of GDP per capita:

$$log(\frac{GDP}{population}) = log(\frac{GDP}{area}) - log(\frac{population}{area}) \tag{4}$$

The ratios $\frac{GDP}{area}$ and $\frac{Population}{area}$ measure land productivity and population density. Table 1 shows the GDP per capita decomposition for microregions in the drought polygon with high and medium frequency of long droughts. The high frequency group has the microregions with reported long droughts above the median for the polygon. The medium frequency has the remaining microregions. The group with high frequency microregions has a total of 5,021 long droughts reported and 250 reservoirs and the medium frequency group has 1,485 long droughts and 135 reservoirs. The frequency of droughts in the high frequency group is almost three times higher than in the medium groups as the two groups have similar land area. We further breakdown each frequency group into microregions with and without reservoirs. We then aggregate GDP, area, and population for each group to calculate the ratios in Table 1.

Table 1: Decomposition of the Reservoir Effect.

	Log GDP per capita	Log GDP per area	Log population per area
High frequency	2.50	5.91	3.41
no reservoir	2.54	5.32	2.78
with reservoir	2.50	6.04	3.55
reservoir effect	-0.04	0.73	0.77
Medium frequency	2.37	5.51	3.14
no reservoir	2.51	5.27	2.76
with reservoir	2.30	5.67	3.37
reservoir effect	-0.21	0.41	0.61
Differential effect	0.17	0.32	0.15

Note: The table shows summary statistics for microregions within the drought polygon grouped by the frequency of reported long droughts and the existence of reservoirs in the microregion. All values are logs of each ratio of aggregated measures of GDP, population, and area. The GDP and population data are gathered from IBGE surveys (IBGE,; IBGE,). The number of reservoirs are based on DNOCS (2005) and ANA (2021) reports. The number of long drougths are based on S2ID reports (S2ID, 2021). The reservoir effect is estimated as the difference between the statistics of groups with and without reservoirs. The Differential effect is the difference in difference between the groups with high and medium frequencies.

The summary statistics in Table 1 are puzzling. The region with three times higher frequency of droughts is more productive and more populous. The GDP per capita in the microregions with higher frequency of droughts is 0.13 log points or about 14 percent higher than in locations with medium frequency. The high frequency group has higher

land productivity (50 percent) and higher population density (31 percent). One possible explanation is reverse causation. There could be more monitoring and reporting of droughts in locations that are more vulnerable in terms of economic and human impacts. There could also be more investments in adaptation in the most vulnerable locations where the return to adaptation is higher. The breakdown of GDP per capita by frequency and reservoirs in column 1 (Table 1) does not add much clarity. GDP per capita is similar in places with high frequency with and without reservoirs and is 0.21 log points or about 23 percent higher in low frequency regions without reservoir.

The decomposition by land productivity and population density (columns 2 and 3) is more informative. High frequency places with reservoir have twice the land productivity and population density than high frequency places without reservoir. This result explains the small reservoir effect on GDP per capita as a higher population density reduces the positive contribution of higher land productivity to GDP per capita. In medium frequency locations the reservoir effect is also positive but about half the magnitude than in high frequency locations. In these locations the reservoir effect on population density are higher than in land productivity resulting in a negative reservoir effect on GDP per capita. The difference in differences statistics are also more intuitive. These results suggest that reservoir are more valuable in locations with more droughts and that the land productivity effect is larger that the migration effect. However, these statistics should be interpreted with caution as we can not disentangle the direction of causation. Such positive reservoir effects could just reflect the concentration of adaptation investments in the most populous and productive regions. In the following empirical analysis we create two instrumental variables to identify the contributions of drought severity and drought adaptation on productivity and population density.

5 Empirical Analysis

Our baseline model for the effect of drought frequency and reservoirs on economic and population outcomes is:

$$logY_{mt} = \mu + \alpha F_m + \beta R_{mt} + \theta F_m R_{mt} + \delta_{sbt} + \epsilon_{mt} \tag{5}$$

The dependent variables are log of GDP per square kilometer and log of population per square kilometer at microregion m and year t. The explanatory variables are the long drought frequency at microregion m, F_m, and the cumulative number of reservoirs constructed at microregiom m and year t, R_{mt}. The parameter δ_{sbt} captures the state-biome-year fixed

effects and controls for linear trends, year specific shocks to productivity and population, and state-biome unobserved productivity factors such as local policies and climate, as well as the interaction of year and local effects. Our sample has 10 states, 4 biomes, and 25 state-biome combinations[8]. The parameters of interest are α, the effect of an increase in disaster frequency if there were no reservoirs constructed, and θ, the mediating effect of reservoirs on the damage of disaster frequencies. Both F_m and R_{mt} are endogenous variables.

The Endogeneity of Drought Reporting

The reporting of disasters such as droughts and floods may increase with potential damages. For example, the monitoring capability may be better in more vulnerable places. Also, many reports are emergency declarations by local governments and these emergency declarations are approved by state and federal agencies based on an assessment of reported damages. Once the emergency declaration is approved, the local government can operate under a special regulatory framework to accelerate the acquisition and deployment of resources for emergency response and social assistance. For example, NEB municipalities under emergency condition can access additional federal funding and credit lines to finance the recovery process. The observed disaster frequency would likely be higher in locations that are more productive and have higher potential damages. Our empirical model would then overestimate the damage from disasters because of this positive correlation between the omitted potential damage and both the disaster frequency and the outcome variable.

We propose to solve the endogeneity of disaster reporting by creating a counting model of disasters that explains disaster occurrence using extreme flow and rainfall events. The idea is to separate disaster occurrences that are caused by extreme weather from occurrences that may have been over reported based on variation in potential damages. Long droughts are hydrology droughts resulting from extreme river flows. The propensity for droughts in a location thus depends on the characteristics of the local river network. We use the spatial variation in the river network in combination with the occurrence of extreme river flows and rainfall event to predict droughts. The identification assumption is that the interaction of extreme weather events with the features of the local river network is uncorrelated with the variation in potential damages within a state-biome region.

We use a similar approach as Keiser and Shapiro (2019) to build a dataset that replicates the river network, the river flows, and the network of flow and rainfall monitoring stations in NEB. The shapefiles and the flow and rainfall measurements are sourced from the Brazilian National Water and Sanitation Agency (ANA). Figure 3 shows a map for our sample data.

[8]The 10 states are Maranhao, Piaui, Ceara, Rio Grande do Norte, Paraiba, Pernambuco, Sergipe, Bahia, Tocantins, Goias, and Minas Gerais. The four biomes are Caating, Atlantic Forest, Cerrado, and Amazon.

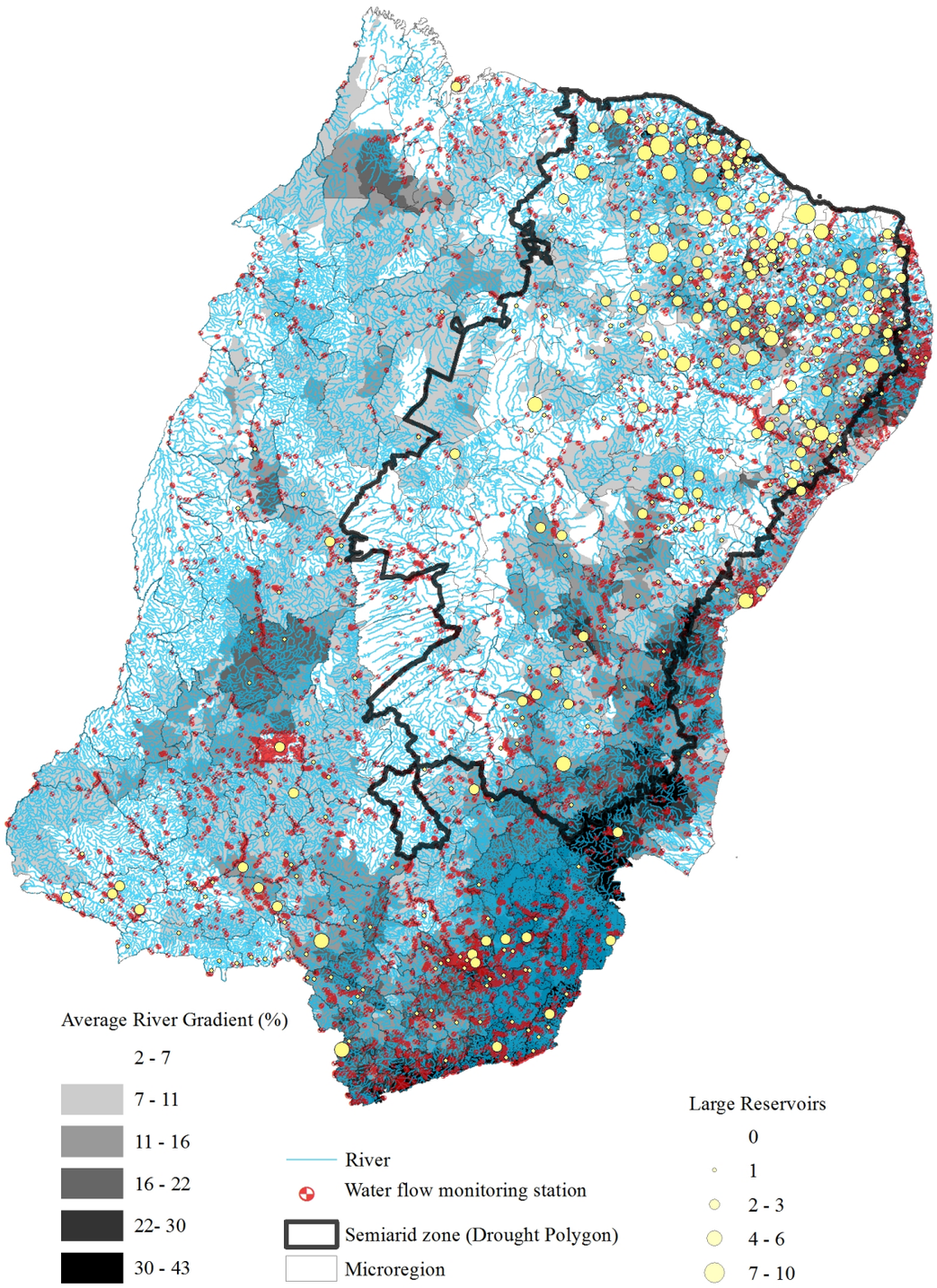


Figure 4: Network of Rivers, Stations, and Reservoirs in NEB

Note: Map of NEB and neighboring states of Tocantins, Goias, and Minas Gerais, with the network of main rivers, flow and rainfall stations, and reservoirs. The unit of observation is a microregion, represented in the figure by the thin black boundary lines. There are 280 microregions, 388 reservoirs, 1,575 flow stations, and 5,324 rain stations in our sample. The shaded areas represent different levels of river gradient computed using remote sensing data from INPE (2020). The river network shape file, the and monitoring stations are from ANA (2020) and the reservoirs are from DNOCS(2005), IOCS(2021) and SNISB(2021). NEB has an exoreic river network that drains towards the Atlantic Coast in the east (left to right in the map). Most rivers are dry for up to sevens months of the year.

The map covers 280 microregions in NEB and the three neighboring states of Tocantins, Goias, and Minas Gerais. Our sample has 1,575 flow stations with 14.8 million daily flow measurements from 1911 to 2020 and 5,324 rainfall stations with 55.6 million daily rainfall measurements. We use the flow and rainfall data to identify 3- and 20-year extreme flow

and rainfall events across the river network and predict drought occurrences. We compute the distribution of flow and rainfall during the high rain and dry seasons for each station and use the river drainage network to link upstream stations to each microregion[9].

Table 2 shows the Negative Binomial model for the number of drought reports as a function of extreme rainfall and flow events. The explanatory variables are dummies for events as extreme as the frequency specified. For example, the 3-year low flow dummy equals 1 if there is an event at least as extreme as a 3-year low river flow in region m and year$t$. The coefficients can be interpreted as percentage changes in the number of droughts and are additive. 3-year and 20-year low flow events increase the expected number of long droughts by 46% and 89.5% (46% + 43.5%), respectively. These low flow events have a significant but much smaller effect on short droughts. By contrast, the effects of low rainfall extreme events is higher on shorter droughts. These results are consistent with the definition of long droughts as hydrological droughts and short droughts as meteorological droughts [10].

We use the Negative Binomial model to predict drought occurrences and estimate drought frequency for each microregion. We restrict our sample to drought reports after 1980 to mitigate measurement error in historical reports and we control for variation in microregion area. All specifications have year and mesoregion fixed effects. There are 9,571 and 3,790 predicted short and long droughts after 1980, about 6% more than observed droughts (9,022 and 3,325). The difference between predicted and observed varies across the distribution of reported droughts. The Negative Binomial model overpredicts droughts in places with less reported droughts. We compute a Pearson statistic for the relative difference between predicted and observed droughts. Table 2 shows the sum of Pearson statistics across all observations. We find a larger difference between predicted and observed short droughts, likely due to a higher uncertainty in assessing damages of short droughts. We then regress the Pearson statistic on the number of reported droughts and find a positive correlation between the number of reported droughts and the difference between predicted and observed measures. This result suggests overreporting in locations with frequent droughts [11].

[9]The three-month high flow season in our sample data is January, February, and March. The three-month low flow season is July, August, and September. The three-month high rainfall season is February, March, and April and the three month dry season is July, August, and September. See Marengo et al. (2021) for a description of the rain season in NEB.

[10]The NB model is a flexible version of the Poisson regression that allows for the variance to exceed the mean (Overdispersion). The Poisson model assumes equal mean and variance. The NB model functional form is similar to the Poisson: $E[y|x] = exp(\beta X)$. The variance differs by an overdispersion parameter α : $var[y|x, \alpha] = (1+\alpha)E[y|x]$. We report the estimated overdispersion parameter in table 2. The overdispersion parameter is significant, indicating overdispersion, particularly in the long drought regression.

[11]The Pearson statistic is a relative measure of the difference between predicted and observed values. $Pearson = (observed - predicted)^2/Predicted$. See Cameron and Trivedi (2005) for a description of Poisson and Negative Binomial regressions.

Table 2: Weather Extremes and Droughts.

	Number of reported droughts	
	Long Drought	Short Drought
	(1)	(2)
River flow as extreme as	0.460***	0.251***
3-year low	(0.094)	(0.055)
River flow as extreme as	0.435***	0.292***
20-year low	(0.143)	(0.069)
Rainfall as extreme as	0.026	0.369***
3-year low	(0.092)	(0.058)
Rainfall as extreme as	0.227**	0.223***
20-year low	(0.096)	(0.059)
Log of microregion area	0.638***	0.570***
	(0.083)	(0.056)
Year and mesoregion fixed effects	Yes	Yes
Observations	9,034	9,034
Wald test for extreme flow and rainfall	50.80	92.80
p-value	[0.000]	[0.000]
Overdispertion parameter	2.513***	1.163*
Sum of Pearson statistic	13,859	16,110
OLS Pearson on reported droughts	2.937***	1.057***

Note: The table shows the results for a Negative Binomial Model (NB) of drought occurrence using a pooled dataset for the 280 NEB microregions with annual observations from 1980 to 2020. The dependent variable is the number of reported droughts by microregion-year. The dummy variables identify extreme events during the rain season. All specifications have year and mesoregion fixed effects. Standard errors are clustered at the microregion level. The variable log area control for variation in the size of microregions. The table shows the Chi2 statistic and p-value for a Wald test of the joint significance of the four dummy variables of extreme events. The table reports the sum of Pearson statistics for the two samples and the coefficient of the OLS regression of the Pearson statistic on the number of reported droughts at each microregion-year. *, **, and *** denotes significant level at 10%, 5%, and 1% respectively.

The Endogeneity of Reservoir Placement

The placement of large public reservoirs is likely influenced by economic outcomes and political connections. States that are richer or have more potential economic losses from droughts are more likely to build reservoirs. In this case, empirical estimates of reservoir effects would overestimate the benefits of reservoirs. Also, large public reservoirs are at least partially funded and managed by federal agencies such as DNOCs. The choice of reservoir location could then be influenced by political connections, which would tend to underestimate reservoir benefits.

We follow the approach of Duflo and Pande (2007) to instrument the number of reservoirs constructed using variation in river gradient. The idea is based on the engineering literature. As Duflo and Pande explain: "Low (but nonzero) river gradient areas are more suitable for

irrigation dams while very steep river gradient areas are suitable for hydroelectric dams. Regions where the river gradient is either flat or somewhat steep are the least likely to receive dams". In the case of NEB, most large reservoirs were built for water supply to population and agriculture during droughts. Reservoir placement and construction in NEB is managed by DNOCs, a federal agency created to manage infrastructure investments for drought adaptation. However, as in the case of India, there are multipurpose reservoirs in NEB that are also designed for electricity generation.

To create the river gradient variables for NEB we integrate remote sensing measurements of slope and elevation from the National Institute for Space Research (INPE, 2021) with river and catchment areas shape files from IBGE (2021). For each municipality, we first exclude catchment areas without rivers and then we classify surface gradients for the catchment areas with rivers into four groups: flat (0-1.5%), moderate (1.5-3%), steep (3-6%), and very steep (above 6%). We then construct the variable fraction river gradient for each municipality by dividing the area classified for each gradient group by the total area of the municipality catchments with rivers.

We estimate the following panel regression for the number of reservoirs constructed [12]

$$R_{ist} = \lambda_0 + \sum_{k=2}^{4} \lambda_{2k}(RG_{ki} * \overline{D_{st}}) + \lambda_3(\mathbf{M}_i * \overline{D_{st}}) + \sum_{k=2}^{4} \lambda_{3k}(RG_{ki} * l_t) + \nu_i + \eta_{st} + \omega_{ist} \quad (6)$$

where R_{ist} is the number of reservoirs constructed in municipality i, state s, and year t. The time-invariant river gradient variables RG_{ki} are interacted with $\overline{D_{st}}$, the predicted reservoir incidence in state s and year t, and with l_t, dummy variables for each year. The predicted reservoir incidence is calculated by multiplying the total number of reservoirs built in year t by the share of reservoirs constructed in state s in year 1950. The geographic controls, $\mathbf{M}_i$, river length and municipality area, are also interacted with predicted reservoir incidence. ν_i are municipality fixed effects, and η_{st} are state-year interactions. We use three three sources of variation to predict reservoirs: withing state variation in river gradient, variation in reservoir construction across years, and variation in the contribution of each state to reservoir construction.

Table 3 shows our results for NEB and comparable results from Duflo and Pande (2007) analysis in India. The coefficients in Table 3 are for the interaction between river gradient and predicted reservoir incidence, λ_{2k}. The baseline river gradient group is flat (0-15%). We estimate the NEB model for the period of 1950 to 2020. In both models a moderate and a very steep river gradient increase the number of reservoirs, consistent with the engineering

[12]The regression model in equation **??** is the same as Equation (2) **?**

guidance. Also, in both models the effect of moderate river gradient is almost three times higher than the effect of very steep river gradient, reflecting a predominance of water supply and irrigation reservoirs. The effect of very steep river gradient captures the construction of multipurpose reservoirs that are designed for water supply and electricity generation. The F-test for the joint significance of the river gradient interactions with predicted reservoir incident shows the power of the instrumental variables. We use the predicted number of reservoirs using the model in Equation 6, $\widehat{R_{ist}}$, as instrumental variables for the observed number of constructed reservoirs. $\widehat{R_{ist}}$ represents the number of reservoirs predicted based on gegraphical features of municipalities within a state and excludes the effect of unobserved political influence and economic outcomes.

Table 3: Geography and Reservoir Construction.

	Number of Reservoirs	
	NEB	India (Duflo and Pande, 2007)
	(1)	(2)
Fraction river gradient 1.5-3%	0.538***	0.153***
	(0.104)	(0.040)
Fraction river gradient 3-6%	-0.024	-0.191***
	(0.050)	(0.065)
Fraction river gradient above 6%	0.189***	0.075**
	(0.033)	(0.031)
Geography controls	Yes	Yes
State*year and river gradient*year interactions	Yes	Yes
Fixed effects	Municipality	District
Observations	19,320	1,855
F-test for river gradient	51.68	6.372
p-value	[0.000]	[0.000]

Note: The table show results for the OLS regression of the number of constructed reservoirs on fractions of river gradient at 1.5-3%, 3-6%, and above 6%. The coefficients reported are for the interactions between river gradient and predicted reservoir incidence, the parameter λ_{2k} in Equation 6. Specification (2) results are from Duflo and Pande (2007) analysis in india (Table II, poverty sample, Column (2)). The baseline group are locations with river gradient lower than 1.5%. Geographic controls are river length and municipality or district area. River length and area are interacted with the river gradient fraction variables. Specification (2) also includes controls for district elevation. All specifications include a full set of state*year interactions and river gradient*year interactions. Standard errors in parenthesis are clustered at the state level for model (1) and at the NSS region*year level for model (2). The NEB model includes all years of data available, from 1850 to 2020. The India model,, poverty sample, includes years 1973, 1983, 1987, and 1999. *, **, and *** denotes significant level at 10%, 5%, and 1% respectively.

Main Results

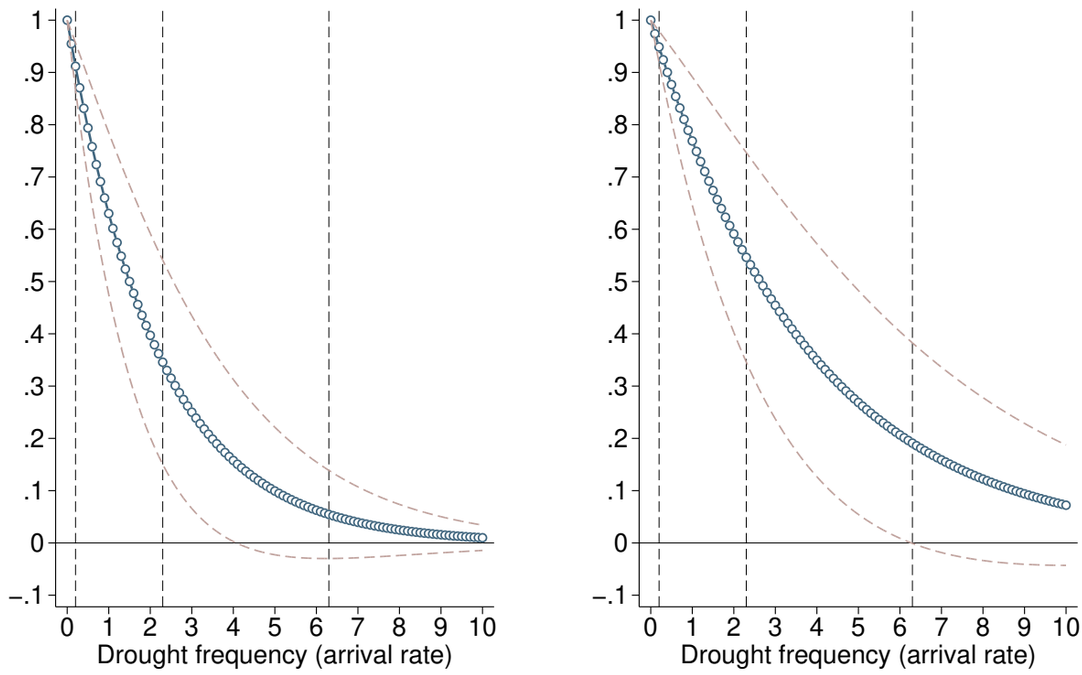
Table 4 reports ordinary least squares (OLS) and instrumental variable (IV) estimates for the coefficients of interest, α, the effect of an increase in disaster frequency without reservoirs, and θ, the mediating effect of reservoirs. Columns (1) and (6) display the strong correlation between reported droughts and the two dependent variables, GDP per square kilometer and Population density. The unit for the frequency variables is occurrences per 10-year period. Adding one long drought every 10 years reduces GDP per square kilometer by approximately 60% and population density by 33%. However, this drought frequency effect is halved when we replace reported by predicted drought frequency (columns (2) and (7)). This large difference is consistent with under reporting of long droughts in the least vulnerable places. The drought frequency effects reported in columns (1), (2), (6), and (7) incorporate the effect of reservoirs. Columns (3)-(5) and (8)-(9) reports the drought frequency effect after controlling for reservoirs. The drought frequency effect increases as it now captures the impact of an increase in drought frequency without reservoirs. Adding one long drought every 10 years in places without reservoirs reduces GDP per square kilometer by approximately 47% and population density by 27%.

Figure 5 plots the counterfactual estimates for the impact of drought frequency on GDP and population density if there were no reservoirs. Figure 5 (a) is a damage function for drought frequency and figure 5 (b) captures a migration outcome measured by a reduction in population density. The values plotted in figure 5 are the relative levels of the outcome variables. Inverting the empirical model in equation 5 and setting the number of reservoirs to zero, $y_{mt} = K_{mt}e^{\alpha F_m}$, where $K_{mt} = e^{\mu+\delta_{sbt}+\epsilon_{mt}}$ is the level of the outcome if frequency and the number of reservoirs equal zero. The values plotted in figure 5 are the expectation of y_{mt}/K_{mt}, $E(y_{mt}/K_{mt}) = E(e^{\alpha F_m})$. For example, an increase in drought frequency from zero to 1 reduces GDP per square kilometer by 47% (Table 4, column (4)). The GDP per square kilometer for a disaster frequency equal to 1 is then 0.53 in Figure 5 (a), reflecting the damages from a higher drought frequency. The vertical dashed lines represent the median, 95 quantile, and 99 quantile of disaster frequency in our sample. We estimate the values plotted in Figure 5 using our preferred model, the IV model with state*biome*year interactions (Columns (4) and (9)).

The two striking features of the estimated damage and migration functions are the rapid decline in places with low drought frequency and the faster decline in GDP per square kilometer compared to the decline in population density. We find robust estimates for the frequency effect using the exponential log-linear function form [13]. The log-linear functional

[13]We also estimated models using quantiles of disaster frequency as explanatory variables and we find

form is consistent with Martin and Pindyck (2015) specification of disasters as consumption shocks with Poisson arrival rates and is more sensitive than common smooth specifications for the damages of changes in mean temperature (Mendelsohn et al. 1994; Deschênes and Greenstone 2007). The impact of changes in mean temperatures tends to become nonlinear at high temperatures or with large changes in mean temperature (Schlenker and Roberts 2008). By contrast, the impact of small changes in drought frequency is higher in places with no droughts, for example, close to the border of the drought polygon in Brazil.



(a) GDP per square kilometer

(b) Population density

Figure 5: Drought Damage and Migration without Reservoirs

Note: Graphs show the counterfactual estimates for GDP per square kilometer and Population density without reservoirs based on the coefficients of Predicted Frequency in model 5 in Table 4. The plotted values are the level predictions for model 5 and the reference value is the GDP per square kilometer and Population density with zero drought frequency. The vertical dashed lines represent the median, 95 quantile, and 99 quantile of long drought frequency in our sample. The unit for the frequency variables is occurrences per 10-year period. For example, a frequency of 1 means 1 drought every 10 years.

The drought frequency effect on GDP per square kilometer is almost twice the effect on population density (Table 4 columns (4) and (9)). Figure 5 shows the faster decline in the GDP per square kilometer function. A region with 2 long droughts per decade and without reservoirs would have a 62% reduction in GDP per square kilometer and a 39% reduction in population density. There could be different explanations for this differentiated effect. There could be short term emergency measures that supplement the income of potential migrants and reduce the incentive to leave. There could also be constraints to migration as the alternative to look for new jobs in urban areas or different regions could be too risky. The emergency assistance and the constraints to migration may be particularly important if there

a similar exponential functional form. Our estimates for models with short droughts using the exponential functional form are also similar and robust.

were a large population of subsistence farmers at the margin for potential migration. We further investigate alternative explanations for the differentiated productivity and migration effect in section 6.

Reservoirs partially reduce the impact of droughts. Table 4 shows the estimated coefficients for the interaction between reservoirs and predicted frequency. To interpret the reservoir effect across drought frequencies we compute the contribution of adding one reservoir for a doubling of the drought frequency. This statistic can be easily computed in the log-linear model using the ratio of the impact of doubling the drought frequency with and without a reservoir:

$$E[\frac{LnY(2F, R=1) - LnY(F, R=1)]}{LnY(2F, R=0) - LnY(F, R=0)} = 1 + \frac{\theta}{\alpha} \tag{7}$$

where α is the coefficient of predicted frequency and θ is the coefficient of the interaction term. The ratio $\frac{\theta}{\alpha}$ captures the reservoir effect with a doubling of the drought frequency. Adding a reservoir offsets the impact of an increase in drought frequency on GDP per square kilometer by 27% and on population density by 26% (Table 4 columns (4) and (9)). The overall contribution of reservoirs could be larger than 27% as some locations have multiple reservoirs. There are one or more large public reservoirs in 40% of the microregions in our sample (112 out of 279 microregions) and more than 2 reservoirs in 30% of the sample. Our IV estimate is larger than the OLS estimate for the reservoir effect. This negative bias is consistent with political influence on reservoir placements. It is possible that multiple reservoirs were placed in locations with a lower potential benefit because of political connections [14].

The offsetting reservoir effect is significant but it declines rapidly with an exponential damage function. We compute the marginal reservoir effect relative to the level of the outcomes at zero frequency and zero reservoirs to assess the declining contribution of reservoirs. The mediating value of reservoirs is captured by the cross-derivative of the outcome with respect to frequency and reservoirs, as in Equation 3. The cross-derivative with the log-linear function assessed at zero reservoirs is $\frac{\partial^2 Y_{mt}}{\partial F \, \partial R} \frac{1}{K_{mt}} = (\theta + \alpha(\beta + \theta F))e^{\alpha F_m}$, where K_{mt} is the level of the outcome if frequency and the number of reservoirs equal zero.

[14] The coefficient for the reservoir variable represents the effect of adding a reservoir in locations with zero drought frequency. The estimate for the reservoir effect in locations without droughts should be intercept with caution as 50% of our sample do not have reservoirs. This negative intercept shift is a feature of the nonlinear log-linear functional form as locations that already had a reservoir to start with are less sensitive to changes in frequency. In this case, the exponential functions becomes more linear and the intercept shifts down.

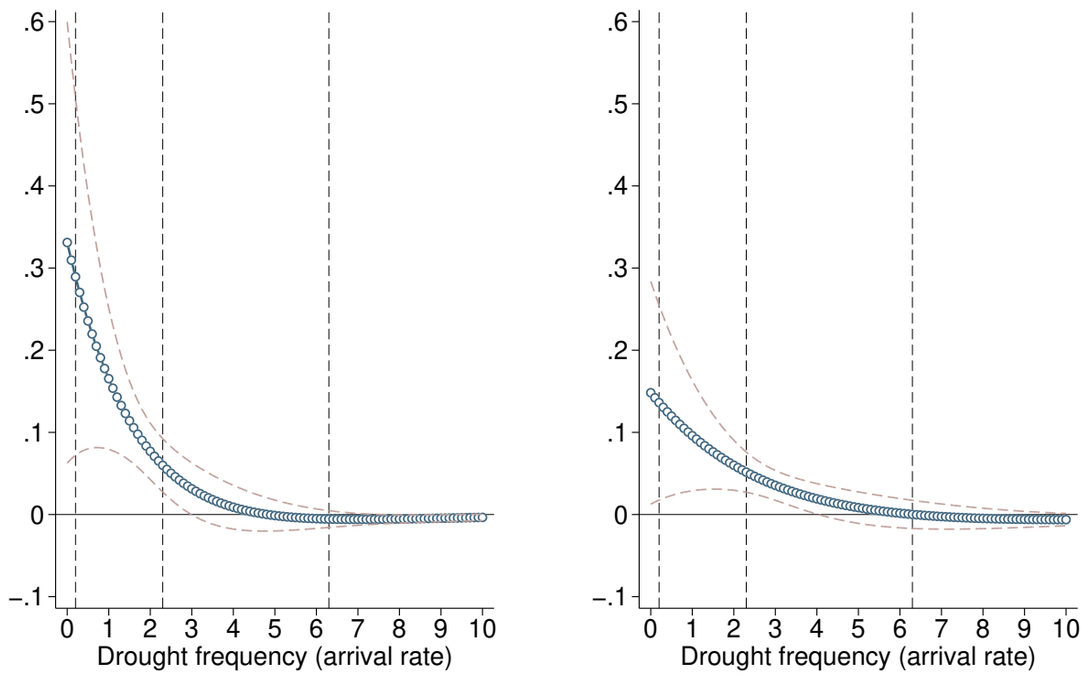
Table 4: The Impact of Drought Frequency and Reservoirs.

	Log GDP per square kilometer					Log population per square kilometer				
	OLS	OLS	OLS	IV	IV	OLS	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Reported frequency	-0.60***					-0.33***				
	(0.11)					(0.08)				
Predicted frequency		-0.28***	-0.35***	-0.47***	-0.46***		-0.19***	-0.23***	-0.27***	-0.26***
		(0.06)	(0.07)	(0.11)	(0.13)		(0.04)	(0.05)	(0.07)	(0.08)
Reservoirs			-0.12***	-0.35***	-0.40***			-0.08***	-0.18**	-0.23***
			(0.04)	(0.11)	(0.10)			(0.03)	(0.08)	(0.07)
Reservoirs*predicted frequency			0.06**	0.13**	0.15**			0.04**	0.07*	0.09**
			(0.03)	(0.06)	(0.06)			(0.02)	(0.04)	(0.04)
Reservoir effect with doubling of drought frequency (θ/α)			0.171	0.27	0.326			0.17	0.259	0.346
State*biome*year interactions	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
State*year and biome*year interactions	No	No	No	No	Yes	No	No	No	No	Yes
Observations	4,986	4,986	4,986	3,438	3,438	8,586	8,586	8,586	5,920	5,920
R^2	0.61	0.59	0.60	0.52	0.43	0.61	0.61	0.62	0.60	0.51
First stage F-statistic				9.06	10.00				8.43	9.02

Note: Table 4 presents OLS and IV estimates for Equation 5. We compute predicted frequency by microregion using variation in weather extremes and river network (Model (1) in Table 2). The instrumental variables are the predicted number of reservoirs, $\widehat{R_{ist}}$, and the interaction between predicted reservoirs and predicted frequency. The unit for the frequency variables is occurrences per 10-year period. For example, a frequency of 1 means 1 drought every 10 years. All specifications include a full set of state*year interactions and river gradient*year interactions. The reservoir effect for a doubling of drought frequency is estimated as the ration of the coefficients of the interaction Reservoirs*predicted frequency (θ) and Predicted frequency (α). All standard errors in parenthesis are clustered at the microregion level. *, **, and *** denotes significant level at 10%, 5%, and 1% respectively.

Figure 6 plots the reservoir effects for the two outcome variables using models (5) and (10) in table 4. The expected offsetting effect of adding the first reservoir in locations with zero frequency is 33% for GDP per square kilometer and 35% for population density. The estimates at zero frequency are significant but noisy as there are not many reservoir in places without droughts. The estimates are more precise in the upper half of the sample, represented in the figure by the interval between the median and 95 quantile vertical dashed lines. The contribution of the first reservoirs declines to zero at a frequency of 3 long droughts per decade.

The reservoir effects plotted in Figure 6 are relative to outcomes at zero drought frequency. We find again a distinct pattern for the productivity and population density outcomes, similar to the results plotted in Figure 5. Both the impact of higher frequencies and the reservoir effects are larger for GDP per square kilometer than for population density. For small increases in frequency the higher reservoir effect compensates for the larger loss in productivity and the declines in productivity and population density are comparable when we consider reservoir adaptation. However, in locations with higher drought frequencies, further increases in frequency will tend to affect more productivity than population density.



(a) GDP per square kilometer (b) Population density

Figure 6: Adaptation Value: Reservoirs

Note: Graph shows the reservoir effect on the level GDP per square kilometer and Population density without reservoirs based on model 5 in Table 4. Each value in the graph represents the fraction of the outcome variable that would be recovered with the addition of one reservoir. The reference value is the level of the outcome variable with zero drought frequency and no reservoirs. The vertical dashed lines represent the median, 95 quantile, and 99 quantile of long drought frequency in our sample. The unit for the frequency variables is occurrences per 10-year period. For example, a frequency of 1 means 1 drought every 10 years.

6 Drought Adaptation and Migration

We examine in this section the linkage between drought adaptation and migration in a developing economy. Our starting point is the FURTADO (1959) analysis of the development of NEB [15]. The crux of Furtado's assessment of drought impacts in NEB is the vulnerability of the subsistence agriculture predominant in the region (FURTADO 1959)). Furtado argues that long-term and short term adaptation measures failed to change the dependency of the NEB economy on subsistence farming. Furtado identifies three main economic activities in the semi-arid section of NEB: extensive ranching, xerophyte crops such as arboreous cotton production, and subsistence farming. Furtado then argues that long-term adaptation measures such as the construction of large reservoirs mostly benefited ranching as there was not an expansion of irrigated agriculture. Also, the short-term policies designed to provide emergency supplies and income did not change the structure of the farming sector [16]. Furtado then concludes that these adaptation measures tended to fixate subsistence farmers in NEB and increase their vulnerability to food production shocks resulting from long droughts. In the remaining of the text we refer to Furtado's arguments on the fixation of the NEB population as Furtado's hypothesis.

We use our empirical model for the effects of long drought frequency and reservoirs to test Furtado's hypothesis. Ideally, we would use a time-series of data on NEB workers and migrants associated to each of the economic activities identified by Furtado. Unfortunately, migration data is only available at the state level and disaggregated labor data by agricultural activities surveyed in the agricultural census is confidential. Instead, we constructed a microregion-level dataset with agriculture and population outcomes using agricultural census and population census surveys from IBGE. We construct measures of agricultural outcomes using data from five agricultural censuses: 1975, 1985, 1995, 2007, and 2017 (IBGE, 2017). We use population measures from 13 population census surveys: 1872, 1880, 1890, 1910, 1920, 1930, 1950, 1960, 1970, 1991, 2000, 2010, and 2020. We estimate population measures for the decades of 1900, 1940, and 1980 using the data from the census surveys immediately before and after. Finally, we create outcome variables for rural and urban population using data from 9 census surveys: 1940, 1950, 1960, 1970, 1980, 1991, 2000, and 2010 (IBGE, 2010).

[15] Celso Furtado was an influential Brazilian economist that studied the economic development of NEB. In addition to his academic work, Furtado designed the Superintendency for the Development of Northeast (Sudene), a government agency intended to create development policies for NEB. Furtado directed Sudene from 1956 to 1961 and later served as minister of Planning and minister of Culture in Brazil.

[16] Furtado's plan for restructuring the economy of the semi-arid NEB, the hinterland, included the integration of the agriculture sector to export channels, incentives to industrialization, and the planned migration of NEB farmers to western frontier of NEB, the state of Maranhao, where the climate is more favorable to agriculture production (FURTADO 1959).

Table 5 shows our estimates for the drought frequency and reservoir effects on agricultural outcomes. All estimates are based on the IV model with state-biome-year interactions. The results are consistent with Furtado's assessment of the impact of reservoirs. An increase in drought frequency reduces cropland but adding a reservoir offsets about 30% of the impact. Adding multiple reservoirs then further preserves cropland. The agricultural census separates annual crops such as rice and beans, which are typical of subsistence farming, from perennial crops, which include the arboreous cotton produced in NEB. We find different effects on annual and perennial cropland. The drought frequency effect on annual crops is almost three times the effect on perennial crop, consistent with Furtado's emphasis on the vulnerability of subsistence farming. Naturally, xerophyte crops such as cotton are more resistant to drought conditions. The reservoir effect on Annual cropland is also higher. Adding a reservoir offsets 32% of the impact of drought frequency on Annual cropland. One possible explanation for this large effect of reservoirs would be an expansion of irrigated crop land. However, we find that drought frequency or reservoirs have no effects on irrigated cropland. According to the 2017 agricultural census survey, only 10% of farms in NEB used irrigation and only 12% of total NEB cropland was irrigated. In the drought polygon about 10% use irrigation and only 9% of the cropland is irrigated.

The agricultural census separately tracks family and hired farm labor. Family labor is the number of workers in the farm that are related to the farm manager and is more common in subsistence farming. Family labor includes male and female workers 14-year-old or older. Hired labor is the number of farm workers not related to the farm manager. It includes permanent and temporary workers. We find a clear shift in labor composition as the drought frequency increases and with the addition of reservoirs. An increase in drought frequency shifts labor shares towards family labor. Adding a reservoir offsets about 25% of the change in labor shares. Adding reservoirs expands temporary crop farming with an increase in cropland and in hired labor. The reduction in hired labor with more droughts is consistent with migration to urban centers or other regions. In this case, the reservoir effect in labor composition also supports Furtado's argument of labor fixation in subsistence agriculture.

Table 5: The Impact of Drought Frequency and Reservoirs on Agriculture.

	Cropland share	Annual crop land share	Perennial crop land share	Family labor share	Hired labor share	Irrigated land share	Pasture land share	Log cattle stock
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted frequency	-0.035***	-0.028***	-0.011**	0.038***	-0.039***	0.004	-0.002	-0.155***
	(0.008)	(0.006)	(0.005)	(0.009)	(0.009)	(0.007)	(0.015)	(0.050)
Reservoirs	-0.025**	-0.017**	-0.012**	0.023**	-0.027***	-0.011	0.018	-0.073
	(0.010)	(0.008)	(0.006)	(0.010)	(0.009)	(0.009)	(0.013)	(0.045)
Reservoirs*predicted frequency	0.010**	0.009***	-0.002	-0.009**	0.010**	0.000	0.002	0.046
	(0.004)	(0.003)	(0.002)	(0.004)	(0.004)	(0.003)	(0.006)	(0.034)
Reservoir effect with doubling of drought frequency (θ/α):								
IV	-0.297***	-0.319***	-0.219*	-0.243***	-0.281***	0.115	-0.761	-0.297*
OLS	-0.177**	-0.206**	-0.118	-0.099	-0.138**	0.214	-0.329*	-0.0739
State*Biome*year interactions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,106	953	762	953	950	191	952	952
First stage F-statistic	6.54	7.36	6.54	7.36	7.28	8.68	7.35	7.35

Note: Table 5 presents IV estimates for Equation 5 with agricultural outcomes. The dependent variables are cropland share (crop area divided by total land area), annual cropland share (annual cropland divided by total cropland), perennial land share (perennial cropland area divided by total cropland area), family labor share (number of farm workers related to the farm manager divided by total number of farm workers), hired labor share (number of permanent and temporary workers not related to farm manager divided by total number of farm workers), irrigated land share (irrigated land area divided by total crop area), pasture land share (natural and artificial pasture land divided by total land area), and log of cattle stock (number of cattle heads divided by total pasture area). All models are estimated using agricultural census surveys for 5 decades: 1975, 1985, 1995, 2006, and 2017. The unit of observation for all models is a microregion. All standard errors are clustered at the microregion level. *, **, and *** denotes significant level at 10%, 5%, and 1% respectively.

Finally, we investigate the effects of drought frequency and reservoir on ranching. The agricultural census tracks natural and artificial pastureland as well as the number of cattle heads. We do not find an effect on pasture land. Both drought frequency and reservoir effects are statistically and economically insignificant. This results is expected as most ranching in NEB is extensive and uses natural pasture area. Artificial pastures are costly and sensitive to drought. However we find a very significant, statistically and economically, drought effect on the cattle stock, the number of cattle heads divided by total pastureland. An increase in drought frequency reduces cattle stock by 0.155 log points. About 30% of this reduction is offset by adding a reservoir. Adding multiple reservoirs could them have a significant effect on preserving the ranching activity. Furtado argues that commercial ranching increases the demand for labor and incentives workers to stay in the semi-arid region (FURTADO 1959)). The large effect of drought frequency on ranching is also consistent with Furtado's hypothesis.

Droughts, Reservoirs, and Population Growth

How many people would have left NEB if not for the Reservoirs? We use the long time series of population census surveys to estimate a counterfactual population change. The outcome variables are population growth rates, approximated by first-differences in log of population. Table 6 summarizes our results. We estimate models with state-biome fixed effects and models with microregion fixed-effects. The river gradient instrumental variable has more explanatory power with the long-time series and using the within-microregion variation. However we can not estimate the effect of time-invariant drought frequency using a microregion fixed effects model. We find consistent estimates with the two models. Standard errors tend to be higher with the microregion fixed effects model but the magnitude of the drought frequency and reservoir effects is similar in the two models.

In 1872, there were 6.8 million people living in NEB. 38% of the NEB population lived in the drought polygon at the time, about 2.5 million people. In 2020, the NEB population increased to 87.3 million, a 17% growth rate per decade, and the drought polygon population increased to 27.8 million, a 15% growth rate per decade. Table 6 presents our estimates for the drought frequency and reservoir effects on population growth in NEB. We estimate that adding one long drought in places without reservoirs reduces the population growth rate by 2%, or about 9 million people in NEB if we apply the rate to the 1872 population. This approximation is not too far from the number of northeastern living in other states according to the 2000 population census, 9.37 million. Naturally, many of these potential migrants would not be reported in the 2000 census and also many NEB migrants could have migrated anyway as they pursued better opportunities in the growing southern economy.

Table 6: The Impact of Drought Frequency and Reservoirs on Population Growth.

	Δ log population		Δ log rural population		Δ log urban population		Δ urban-rural growth	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted frequency	-0.021***		0.006		-0.020*		-0.026**	
	(0.006)		(0.010)		(0.010)		(0.008)	
Reservoirs	-0.016**	-0.026**	-0.000	0.042	-0.008	-0.067***	-0.004	-0.109***
	(0.007)	(0.010)	(0.005)	(0.030)	(0.007)	(0.019)	(0.007)	(0.037)
Reservoirs*predicted frequency	0.008**	0.010**	-0.002	0.002	0.009*	0.017*	0.009*	0.015
	(0.004)	(0.005)	(0.005)	(0.009)	(0.005)	(0.010)	(0.005)	(0.015)
Reservoir effect with doubling of drought frequency (θ/α)	-0.402***		0.071		-0.475**		-0.352**	
	(0.126)		(1.009)		(0.218)		(0.148)	
State*year interactions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	StatexBiome	Microregion	StatexBiome	Microregion	StatexBiome	Microregion	StatexBiome	Microregion
Observations	2,591	2,591	1,325	1,325	1,325	1,325	1,325	1,325
First stage F-statistic	6.48	19.7	6.33	23.83	6.33	23.83	6.33	23.83

Note: Table 6 presents IV estimates for Equation 5 with population growth outcomes. The dependent variables are the first differences (Δ) of log population, log rural population, log urban population, and the difference between urban and rural population growth (urban-rural growth difference - URGD). Columns (1) and (2) use census data for 16 decades (1872,1880,1890,1900,1910,1920,1930,1940,1950,1960,1970,1980,1991,2000,2010,2020). Columns (3) to (8) used census data for 10 decades (1940,1950,1960,1970,1980,1991,2000,2010). The unit of observation for all models is a microregion. The reservoir effect for a doubling of drought frequency is estimated as the ration of the coefficients of the interaction Reservoirs*predicted frequency (θ) and Predicted frequency (α). The instrumental variables are the predicted number of reservoirs, $\widehat{R_{ist}}$, and the interaction between predicted reservoirs and predicted frequency. The unit for the frequency variables is occurrences per 10-year period. For example, a frequency of 1 means 1 drought every 10 years. All standard errors in parenthesis are clustered at the microregion level. *, **, and *** denotes significant level at 10%, 5%, and 1% respectively.

We estimate that the offsetting reservoir effect on population growth is about 40%. This offsetting effect implies that about 3.7 million potential migrants would have stayed in NEB because of the adaptation benefits of water reservoirs. This number would increase if we were to account for multiple reservoirs in a microregion. If we were to consider only the drought polygon, the number of potential migrants that stayed would be 1.4 million. If we were to consider only the top three quintiles of drought frequency within the drought polygon, the number of potential migrants that would have stayed would be 578 thousand people. This number would likely be higher as the places with more droughts also have multiple reservoirs. We estimate that the counterfactual number of potential migrants that stayed because of reservoirs would likely be in the range of 0.5 to 3.7 million people.

Finally, we assess the impact of drought frequency and reservoirs on urbanization. Poor countries tend to urbanize as they develop and the economy and the population shift from the agricultural sector to industry and services. in 1940, the NEB rural population was 16.8 million people and the urban population was 5.2 million people. The share of urban population in NEB was only 24%. However, in the following decades, the NEB urban population grew at a rate of 42% per decade whereas the rural population in NEB grew at a rate of 1% per decade. In 2010, the rural population was 16.8 million and the urban population was 62 million. The share of urban population grew to 77% in 2010. This urbanization process in NEB is consistent with the analysis of do Nascimento and de Oliveira (2015) and Ojima and Fusco (2015) as they describe the expansion of mid-size cities within the drought polygon in the last decades.

We estimate that adding one drought per decade in places without reservoirs reduces the growth rate of urban population by 2% and the differential urban-rural growth rate by 2.6%. To the extent that urbanization is part of the development process, this results suggests a slow down in development. For example, the urbanization in the most developed southeast region of Brazil was faster than in the rest of the country. However, we find a large reservoir effect on urbanization. We estimate that adding a reservoir offsets 48% of the impact of drought frequency on urban population and 35% of the impact on the differential urban-rural growth. This result suggests that reservoirs were important to stabilize the water supply in urban areas and supported the expansion of the mid-size cities in NEB. As multiple reservoirs are placed in locations with high drought frequency, a large part of the urbanization effect was likely offset by reservoirs. This benefit from reservoirs was not considered in Furtado's hypothesis. Although reservoirs preserved subsistence agriculture in places with high drought frequency, reservoirs also seem to have facilitated urbanization and labor movement away from agriculture and to urban jobs.

7 Conclusion

Our results highlight the linkages between climate adaptation and migration policies in developing countries. The optimal climate mitigation policy depends on the shape of the damage function and the value of adaptation. We find that the damage from severe disasters grows rapidly in areas with low disaster frequency and the value of adaptation is positive but limited. In places with frequent disasters, adaptation investments have diminishing effects on productivity but significant effects on net migration, potentially increasing the population vulnerable to climate change.

We find mixed evidence to support (FURTADO (1959)) hypothesis for the retention of potential NEB migrants in drought-prone places. Reservoirs preserve and expand subsistence agriculture in NEB as Furtado predicted. However, reservoirs also supported urbanization and rural-urban migration in NEB. A natural extension of the analysis of potential migration would be to assess the relative benefits of migration from the NEB. It could be that staying in the NEB would be better anyway because of the lack of good jobs for a largely uneducated agricultural work force. In this case, the offsetting effect of reservoirs would be a positive effect on the welfare of potential migrants. Unfortunately, we do not have the data on migrant outcomes for a migration welfare analysis. However, Furtado's reasoning is insightful on potential welfare effects. Furtado argues that the drought adaptation measures in NEB did not enhance welfare because they did not change the structure of the NEB economy. To the extent that subsistence farming is the economic activity most vulnerable to droughts, the retention of potential migrants in subsistence agriculture increased their vulnerability to further changes in the severity and frequency of these extreme events. Furtado proposed a planned migration to the western border of NEB where the distribution of rainfall is more regular.

References

Agropecuário, I. C. (1996). Censo agropecuário 2006.

ÁGUAS-ANA, A. N. D. (2019). Sistema nacional de informações sobre recursos hídricos (snirh).

Annex, I. (2012). Managing the risks of extreme events and disasters to advance climate change adaptation. Sciences, 10:97–104.

Assunção, L. M. and Livingstone, I. (1993). Desenvolvimento inadequado: construção de açudes e secas no sertão do nordeste. Revista Brasileira de Economia, 47(3):425–448.

Barro, R. J. (2009). Rare disasters, asset prices, and welfare costs. American Economic Review, 99(1):243–64.

Botzen, W. W., Deschenes, O., and Sanders, M. (2020). The economic impacts of natural disasters: A review of models and empirical studies. Review of Environmental Economics and Policy.

Boustan, L. P., Kahn, M. E., and Rhode, P. W. (2012). Moving to higher ground: Migration response to natural disasters in the early twentieth century. American Economic Review, 102(3):238–44.

Cameron, A. C. and Trivedi, P. K. (2005). Microeconometrics: methods and applications. Cambridge university press.

Campos, J. N. B. and de Carvalho Studart, T. M. (2008). Drought and water policies in northeast brazil: backgrounds and rationale. Water policy, 10(5):425–438.

CARDIM, F. (1925). Sj tratados da terra e gente do brasil.

Cattaneo, C. and Peri, G. (2016). The migration response to increasing temperatures. Journal of development economics, 122:127–146.

Civil, B. D. (2015). Sistema integrado de informações sobre desastres.

Coronese, M., Lamperti, F., Keller, K., Chiaromonte, F., and Roventini, A. (2019). Evidence for sharp increase in the economic damages of extreme natural disasters. Proceedings of the National Academy of Sciences, 116(43):21450–21455.

de Geografia, I. B. and de Contas Nacionais, E. C. (2005). Produto interno bruto dos municípios. IBGE.

De Nys, E., Engle, N., and Magalhães, A. R. (2016). Drought in Brazil: proactive management and policy. CRC Press.

Deschênes, O. and Greenstone, M. (2007). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. American economic review, 97(1):354–385.

do Nascimento, T. C. L. and de Oliveira, H. C. G. (2015). Análise das migrações intrarregionais no semiárido setentrional1. Migrações Nordestinas no Século 21: Um Panorama Recente, page 113.

Douville, H., Raghavan, K., Renwick, J., Allan, R., Arias, P., Barlow, M., Cerezo-Mota, R., Cherchi, A., Gan, T., Gergis, J., et al. (2021). Weather and climate extreme events in a changing climate.

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

ESPACIAIS-INPE, I. N. D. P. (2019). Divisão de geração de imagens (didgi).

FEDERAL, S. P. (2005). Relatório 2005.

Fonseca, J. E., Carneiro, M., Pena, J. L., Colosimo, E. A., Silva, N. B. d., Costa, A. G. F. d., Moreira, L. E., Cairncross, S., and Heller, L. (2014). Reducing occurrence of giardia duodenalis in children living in semiarid regions: impact of a large scale rainwater harvesting initiative. PLoS neglected tropical diseases, 8(6):e2943.

FURTADO, C. (1959). Gtdn uma política de desenvolvimento econômico para o nordeste. Rio de Janeiro, Presidên.

Hansen, Z. K., Libecap, G. D., and Lowe, S. E. (2011). 9. Climate Variability and Water Infrastructure. University of Chicago Press.

IBGE, D. D. P. (2020). Pesquisa pecuária municipal. Produção aquícola no período de, 1.

Keiser, D. A. and Shapiro, J. S. (2019). Consequences of the clean water act and the demand for water quality. The Quarterly Journal of Economics, 134(1):349–396.

Lima, J. R. d. and Magalhães, A. R. (2019). Secas no nordeste: registros históricos das catástrofes econômicas e humanas do século 16 ao século 21. Parcerias Estratégicas, 23(46):191–212.

Marengo, J. A. (2010). Vulnerabilidade, impactos e adaptação à mudança do clima no semi-árido do brasil. Parcerias estratégicas, 13(27):149–176.

Marengo, J. A., Galdos, M. V., Challinor, A., Cunha, A. P., Marin, F. R., Vianna, M. d. S., Alvala, R. C., Alves, L. M., Moraes, O. L., and Bender, F. (2021). Drought in northeast brazil: A review of agricultural and policy adaptation options for food security. Climate Resilience and Sustainability.

Marengo, J. A., Torres, R. R., and Alves, L. M. (2017). Drought in northeast brazil—past, present, and future. Theoretical and Applied Climatology, 129(3):1189–1200.

Martin, I. W. (2008). Disasters and the welfare cost of uncertainty. American Economic Review, 98(2):74–78.

Martin, I. W. and Pindyck, R. S. (2015). Averting catastrophes: The strange economics of scylla and charybdis. American Economic Review, 105(10):2947–85.

Mendelsohn, R., Nordhaus, W. D., and Shaw, D. (1994). The impact of global warming on agriculture: a ricardian analysis. The American economic review, pages 753–771.

Ojima, R. and Fusco, W. (2014). Migrações nordestinas no século 21. São Paulo: Blucher.

Ojima, R. and Fusco, W. (2015). Migrações nordestinas no século 21: um panorama recente. Editora Blucher.

Olmstead, S. M. (2014). Climate change adaptation and water resource management: A review of the literature. Energy Economics, 46:500–509.

on Large Dams, I. C. (1979). World Register of Dams: Registre Mondial Des Barrages: Deuxieme Mise a Jour, 31 Decembre 1977: Second Updating, December 1977. International Commission on Large Dams.

Populacional, C. (2010). Instituto brasileiro de geografia e estatística (ibge)(29 de novembro de 2010). Página visitada em, 11.

Schlenker, W. and Roberts, M. J. (2008). Estimating the impact of climate change on crop yields: The importance of nonlinear temperature effects. Technical report, National Bureau of Economic Research.

Van Aalst, M. K. (2006). The impacts of climate change on the risk of natural disasters. Disasters, 30(1):5–18.

Villa, M. A. (2000). Vida e morte no sertão: histórias das secas no Nordeste nos séculos XIX e XX. Ática.

Zaveri, E., Russ, J., Khan, A., Damania, R., Borgomeo, E., and Jägerskog, A. (2021). Going with the flow.