The University of Chicago

Fragmented Futures: Understanding the Role of Spatial Boundaries on Groundwater in India

by

Shamil Shirish Khedgikar

August 2023

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts Program in the Social Sciences

Faculty Advisor : Dr. Pedro Amaral

Preceptor : Dr. Yan Xu

1 Abstract

This research provides the first spatially explicit understanding of the effects of *de-jure* administrative boundaries on groundwater extraction in India. By discovering a link between the spatial configurations of settlements, administrative boundaries, and reservoirs of groundwater, we develop a new methodological approach for understanding the drivers of groundwater scarcity, as measured through estimates of volumetric extraction and persistence of dry wells over time. To contextualize the problem, we first undertake a qualitative review of India's past and present policies on water governance and identify some structural deficiencies and their persistence across larger systems of urban and rural governance. We find two critical historical linkages in India's Water-Food-Energy nexus that can be attributed for driving over-extraction of groundwater over time. We formulate the persistence of these linkages as a larger problem of fragmentation - lack of coordination between contiguous administrative units sharing common aquifer systems and misdirected investments due to federated governance structures exacerbating this problem. We posit that the misdirected investments arise due to the "out of sight" characteristic of groundwater, leading to a difference between actual and perceived scarcity of water. To test this explanation empirically using models in spatial econometrics, we develop a new class of spatial weights called RelWeights that allow for interaction between resource and administrative boundaries through irregular 2-dimensional contiguity, a unique spatial characteristic of non-overlapping administrative and resource boundaries. We then compare model outcomes using RelWeights in comparison to Queen Weights in an ensemble of spatial econometric models (SAR, SEM, Spatial Durbin Model, Spatial Lag & Error Probit) to identify drivers of over-extraction, finding strong evidence against the normative policies of groundwater welfarism adopted by the state and a distinct set of drivers for real and perceived scarcity of groundwater. Finally, we study the possibility of creating alternative clusters for the effective regulation of groundwater while synergistically improving the sustainable use of groundwater through a series of counterfactual simulations.

2 Introduction

Global estimates indicate that approximately 97% or more of the Earth's freshwater supply comes from groundwater (Rodell et al., 2009). A globally ubiquitous resource, it is also notoriously prone to the tragedy of commons (Ostrom, 1990) - a phenomenon that arises due to the non-excludable and rivalrous nature of the resource. The increase in anthropogenic consumption of groundwater has been a growing concern globally for almost three decades, however, the estimates on the extent of this consumption and the resultant effects on water levels have only been established more recently. This has primarily been attributed to the increased availability of high-resolution and high-frequency geospatial data due to advances in geoinformatics (Tapley et al., 2004), the ability of coupling remotely sensed data with on-ground monitoring of groundwater levels time and space (Blakeslee et al., 2020; Bhanja et al., 2017), and advances, and the increasing frequency of crises of water being felt across the globe. The most significant of these estimates suggests that groundwater is effectively a finite resource, with only 6% of the total volume being capable of replenishment over a span of 50 years (Gleeson et al., 2016). Recent evidence also shows that over-extraction of groundwater could possibly be a major contributor to sea-level rise, as extensive extraction of groundwater is a very recent phenomena that could have potentially disrupted the global water-cycle (Wada et al., 2016).

Within this global context of groundwater, India is placed rather precariously and is considered to be on the brink of a groundwater crisis (Subhadra, 2015). India's groundwater levels are declining rapidly and the per capita availability of groundwater stands below 200 cubic meters as of 2020 (Bhanja et al., 2017). As a first-order effect, rapid urbanization in India, coupled with a shift in the primary source of rural irrigation towards groundwater continues to deplete sub-surface water doubly across the country. According to the Aqueduct rankings by WRI, India is now classified as a highly stressed country.[1]. It ranks 4^{th} in the world in Drought Risk Index and 13^{th} under the Baseline Water Stress Index (Hofste et al., 2019). These metrics are likely to worsen over time, given the increase in the uncertainty of rainfall coupled with an overall decrease in its magnitude accelerating groundwater depletion.

Responses or "solutions" to the myriad of present and looming problems of groundwater scarcity have started to emerge in both academic literature and practice over the last decade. They range from empirically assessing the effectiveness of water reallocation from rural to urban areas (Garrick et al., 2019) to a call for radical paradigm shifts in approaching groundwater problems to fundamentally alter the human-nature and human-human dynamics in groundwater systems through deep radical interventions (Morrison et al., 2022). However,

[1] See https://www.wri.org/applications/aqueduct/country rankings/

the substantial lag between the state of academic research and policy in developing countries like India means that a precursor to effective policy is the development of pragmatic, methodological innovations that promise a visibility of comparative statics and dynamics of groundwater resources as an integral part of a complex, interconnected system (McPhearson et al., 2022).

To address this in our research, we contextualize the problem of groundwater over-extraction and scarcity in India to be inherently "wicked" - that is defying traditional linear problem-solving approaches and requiring iterative, adaptive, non-linear methods for diagnosis and inference (Rittel and Webber, 1973). We rely on three intersecting bodies of literature in our research - economic theory for understanding the tragedy of commons for groundwater being a Common-Pool Resource (CPRs) (Ostrom, 1990, Ostrom (2008)), dimensions of Integrated Water Resource Management (IWRM) (Savenije and Van der Zaag, 2008), and a framework for assessing the problem of over-extraction of groundwater and the resultant scarcity inspired by the new science of Socio-Hydrology (Sivapalan et al., 2012)[2].

This formulation is based on four unique characteristics of groundwater we discuss in our research - First, the "interplay of spatial and temporal variations of groundwater at the local level, which are often the determining factors for water scarcity, lead to complex systems dynamics" (Savenije, 2000) cited in Sivapalan et al., 2012, pp. 1273). Second, other than the non-excludability and rivalrous characteristics of CPRs (Ostrom, 1990), groundwater (or generally all underground CPRs) also additionally suffer from the problem of "out of sight, out of mind" (Taylor and Shamsudduha, 2022) as its loss is realized only through dried-up tube wells and bore wells *observed* by a user over a limited time and space. Thus, the local *perception of scarcity* of groundwater over space and time may vary from the *actual scarcity* of the resource. Third, the extraction and replenishment of groundwater is usually an outcome of interactions between hydrological (surface and sub-surface flows, rainfall), geological (rock systems with varying permeability), ecological (forests and ecosystems) political (communities, cities, and state), and sociological (consumption patterns) systems which makes it difficult to measure, regulate, and forecast (Dangar et al., 2021). Finally, there is a lack of theoretical consensus on how scarcity of groundwater affects its consumption. There is ample evidence indicating both scarcity leading to an increase in efficiency of groundwater use as well as an exacerbation of the scarcity as a result of competitive over-extraction (Varghese et al., 2013).

We observe that *spatial interdependence* and *spatial heterogeneity* underpin each of these four characteristics. Both the extent and patterns of use of the resource and the agency of governance are spatially explicit in nature. Spatial patterns of settlements have always co-

[2]We explicitly situate this problem within the larger literature in Section 2

evolved with access to surface water (Rodriguez-Iturbe et al., 2009), which in turn has shaped the prevalence of groundwater through sub-surface flows. When we consider aggregate units of governance with nesting such as wards/municipalities, cities, districts, states and so on, these patterns persist and become more complex with number and size of the governance units, due to the heterogeneity in the local availability of groundwater over space, and its use and regulation by institutions. While the spatial extents of groundwater stay the same over long periods of time[3], it is the spatial extents of administration which are modifiable, systemic elements that offer the opportunity for inducing non-linear, positive or negative impacts on groundwater extraction that can be estimated empirically.

Through this research, we aim to further the understanding of one such key modifiable element that largely remains unexplored in literature - the role of *de-jure* administrative boundaries of governance. There are the units at which all state infrastructure for water supply, programs and welfare schemes including investments for irrigation and agriculture inputs are implemented. Methodologically, we employ and extend the theory and methods developed in Spatial Econometrics as they are well suited to the asymmetric and interdependent spatial relations between groundwater reserves and units of governance, the importance of explanatory factors located outside units of interest affecting groundwater use over a distance (also known as 'space distant explanatory factors'), and the need to explicitly model topology given the co-evolutionary dynamics between settlements and groundwater resources (Paelinck and Klaassen, 1979, pp. 5-11 as cited in Anselin, 2010, pp. 4).

To the best of our knowledge, no such attempt has been made so far in literature across different disciplines, and especially in the case of India - a critical gap which this research attempts to fill. The significance of the study is also reinforced by the ubiquitous nature of administrative and political units and the potential of generalizing methods developed in this research to improve the global landscape of groundwater. Despite the scope of the empirical work undertaken here being limited to India, the methodological developments and approaches are generalizable across all geographies.

2.1 Research Questions and Outline

Policies related to groundwater regulation in India have evolved over time, and their historical context provides insights into the challenges and successes faced in managing this vital resource. Therefore, we begin by trying to understand (1) What are some key historical

[3]The spatial extents of aquifers can change over short-term, i.e years to decades due to seasonal factors and anthropogenic activities like pumping and land-use changes. However, significant changes only occur over long-term (century to millennia) primarily caused by geological processes, climate variations, and significant alterations to the geography due to anthropogenic activities such as deforestation and mining

policy decisions that have contributed to the current state of groundwater in India? Next, we focus on (2) What are the lacunae in present policy ensemble for regulation of water at the urban and rural levels that aid the over-extraction of groundwater? In Section 3 situate our study within the larger literature on groundwater and governance in India and undertake a qualitative analysis of how policies in India's Water-Food-Energy nexus lead to the problem of overextraction. We identify two historical linkages as drivers of the problem of present day over-extraction of groundwater and outline the current structure for regulating groundwater in India. In Section 4, we delve deeper and focus on the institutional linkages and operational deficiencies India's governance structure with respect to groundwater. We discuss measurement of scarcity and its effects on driving groundwater investment, and a possibility of mismatch leading to exacerbation of scarcity. We then establish the thesis for empirical inquiry based on the possibility that actual scarcity of water may be different than the perception of scarcity observed locally through dried up and dysfunctional wells.

We then build upon our observations to inform the quantitative aspects of our paper. First, we try to understand (4) How can the two-way interaction between administrative boundaries and sub-basins be captured in a spatial econometric framework to understand the drivers of groundwater in India? Next, we investigate whether our thesis can be validated empirically, i.e (5) Whether the *perceived scarcity* of water, as observed from seasonal and perennial dry wells, distinct from the *actual scarcity*, as estimated from measurements of annual drafts? Finally, we investigate (6) Whether differences in *settlement patterns of urban and rural areas in terms of their physiographic locations*, *accessibility to surface-water* and *transportation networks*, *water requirements for agriculture*, and public *investments* in *physical infrastructure* for groundwater extraction and *financial welfare* significantly explain the differences in *perceived* and *actual* scarcity of groundwater? These questions are addressed at length Sections 6-11. We start with a preliminary analysis of resource and administrative boundaries in Section 5, followed by a theoretical model and simulation experiment that establishes key claims driving our empirics in Sections 6. Section 7 outlines the methodology and sources for data collection for a spatial cross-section. In Section 8 we undertake Exploratory Spatial Data Analysis (ESDA) on the created cross-section. The concept and theory behind a new class of Spatial Weights - RelWeights is developed and applied to structural forms of externalities in Section 9. In Section 10 we outline the model specifications for our empirics and compare models based on residual spatial correlation, Lagrange Multiplier tests and model selection based on the Akaike Information Criterion. This is followed by a discussion and the implications of our findings for creating alternative spatial clusters.

3 Context

We provide a brief overview of the current state of literature related to groundwater in Economic Theory, IWRM, and Socio-Hydrology and situating our research within it. Broadly, economic theory emphasizes a focus on economic efficiency and market mechanisms, IWRM adopts a multidimensional approach to water management, and Socio-Hydrology examines the feedback between human actions and hydrological processes.

3.1 Situating the Problem of Groundwater Extraction

1. *Economic Theory and Groundwater:* From an economic perspective, issues related to overextraction of natural resources such as groundwater arise from inadequate property rights enforcement. When property rights are unclear, private decision-makers overlook social benefits and costs in their actions, leading to over-exploitation or under-provision of public goods (Libecap, 2009, pp.130-132). Due to this public nature of groundwater, typically there is no inherent incentive for an individual or a group of individuals to conserve it. Hence, the cost of its over-extraction is borne by all the participants extracting groundwater, often allowing anyone to extract groundwater without individually bearing its cost. This susceptibility to free-riding which leads to the over-extraction of groundwater is known as the Tragedy of Commons. As a phenomenon, it was first discussed by Thomas Hobbes in Leviathan in the context of the necessity of an external enforcer, explicated in the context of inevitability of the destruction of natural resources by Garret Hardin (Hardin, 1968), and further formalized as a special case of a multi-party Nash Equilibrium generalizable to all Common Pool Resources (CPRs) by Elinor Ostrom (Ostrom, 1990, p. 1-28). Methodologically, econometric methods enable us to empirically investigate the extent and magnitude of externalities of groundwater extraction and provide us with a framework for understanding groundwater over-extraction and scarcity. We use these ideas in Section 6 to develop a two-unit-one-resource model which also informs our empirics.

However, while economic theory proposes that command and control approach can eliminate the Tragedy of Commons, this has rarely been observed in practice (Libecap, 2009, pp. 131) due to a variety of non-economic factors. A purely economic analysis of supply and demand of groundwater fails to acknowledge the non-economic externalities, environmental degradation, and the interconnectedness and feedback mechanisms that exists in hydrological and social systems.

2. *The IWRM perspective:* IWRM takes a holistic approach to managing water resources, considering the interactions between social, economic, environmental, and institutional factors to achieve sustainable water management. IWRM emerged in response to the

recognition that traditional sectoral and disciplinary approaches to water management informed only by economic or sociological theory were inadequate to address the complexities in studying groundwater. As outlined in (Savenije and Van der Zaag, 2008, pp.291) there are four dimensions in IWRM - (1) The *water resources*, or the natural dimension, taking the entire hydrological cycle into account, including stock and flows, as well as water quantity and water quality, (2) the water users, or the human dimension, all economic interests and stakeholders, the *spatial scale* including spatial distribution of water resources and uses and heterogeneity of spatial scales at which water is being managed, i.e. individual user, user groups (e.g. user boards), watershed, catchment, (international) basin; and the institutional arrangements that exist at these various scales, and (4) the *temporal scale*, taking into account the temporal variation in availability of and demand for water resources, but also the physical infrastructure that allows effective matching of water demand and supply over time. Figure 1 shows the first three dimensions of IWRM through an intuitive illustration.

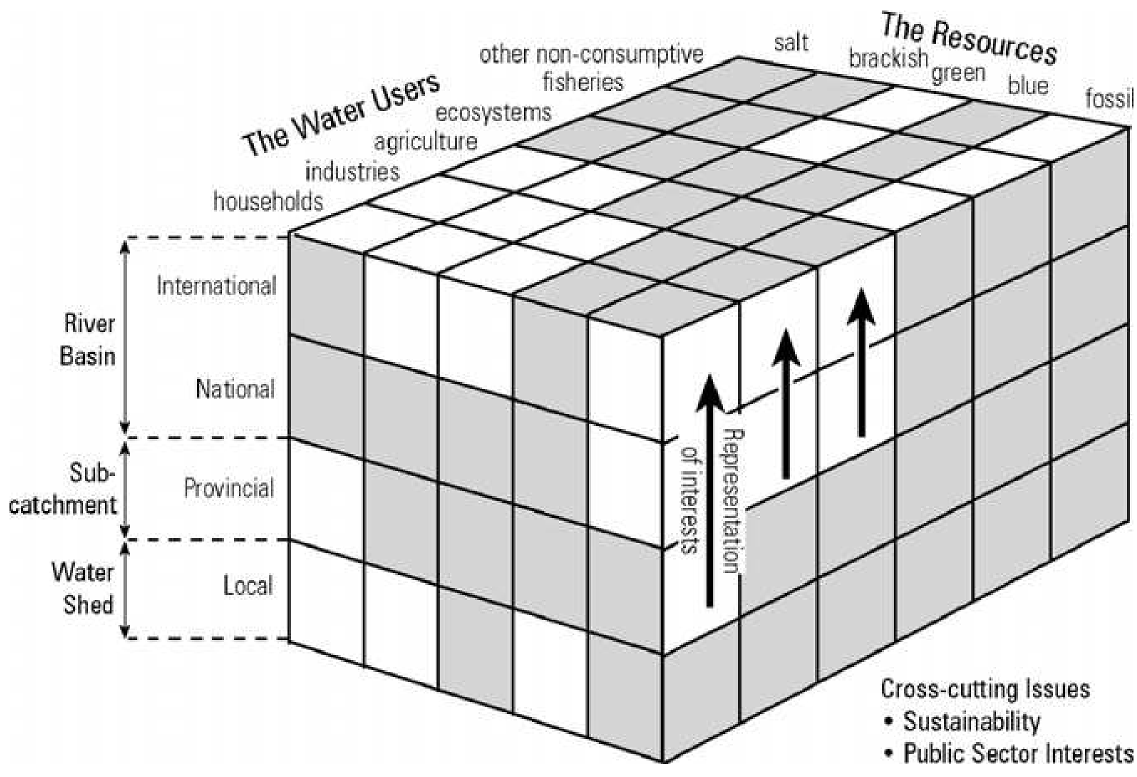


Figure 1: Visualizing the 3 Dimensions of IWRM (Source: Savenije and Van der Zaag, 2008)

The IWRM framework takes cognizance of the spatial and temporal scales of water use. In particular, the explicit acknowledgment is that spatial scales of "hydrological boundaries seldom concur with administrative boundaries. River basins are very appropriate units for operational management but present problems for institutions which have a different spatial logic. At all times, two-way interaction should take place between the institutional levels."

(Savenije and Van der Zaag, 2008, pp. 291). Another important aspect of IWRM is that it offers guidance for an appropriate scale of analysis - in our research we focus on the second 'plane' in Figure 1, thereby limiting our analysis to the hydrological units of Sub-Basins for a provincial analysis, focusing on households, agriculture, and ecosystems in terms of the Water Users, and considering precipitation over land partitions into runoff via surface water and groundwater (blue water) and evapotranspiration (green water). By it's very formulation, the IWRM framework enables an interdisciplinary discourse on groundwater. But IWRM is also self-limiting because it is focused on operational and managerial aspects of water systems, which makes the framework highly contextual and does not yield itself to empirical generalization.

3. *Socio-Hydrology:* As outlined in the (Sivapalan et al., 2014, pp.226), "Socio-hydrology has three goals: (1) analyze multi-scale, space-time patterns and dynamics of interacting social and hydrological processes, and interpret them in terms of the underlying structural features of biophysical and human systems and their interactions; (2) explain and interpret socio-hydrological responses in terms of outcomes relevant to human well-being, and discern possible future scenarios of their evolution; and (3) understand the meaning and value of water as a culturally, politically, and economically embodied resource necessary to human life, and do so in a manner that explicitly accounts for biophysical and human interactions." Thus, while the aim of IWRM is to control or manage the water system to reach desired outcomes for society and the environment, "the focus of socio-hydrology is on observing, understanding and predicting future trajectories of co-evolution of coupled human-water systems."(Sivapalan et al., 2012, pp.1271).

Socio-Hydrology consolidates contextual empirical evidence gathered from IWRM projects into a scientific body of knowledge. The focus on spatial and temporal which is emphasized in IWRM is augmented with a focus on real or estimated water use, cognizance of past and present human interactions across space, and feedback loops. We build upon these ideas through a spatial econometric lens, which provides us with the perfect set of tools to understand the feedback across space and time, interaction between hydrological (natural systems) and governance (social systems), and prediction and simulation of effects of change in the system.

3.2 Groundwater as a Regulated Resource

Groundwater typically exists within the bounds of permeable or non-permeable geological formations formed over long time periods of time. These formations, known as aquifers[4], in-

[4]Aquifers are often also considered as sub-basins within shallower formations as Surficial aquifers (the water table) generally mimic surface-water watersheds, which aggregate to form sub-basins and their flow

teract with neighboring formations, percolating rainwater, and sub-surface flows from rivers, based on varying levels of permeability that are determined by the properties of geological structures which bound them. When groundwater is extracted through vertical drilling, several complex hydro-geological processes are triggered, and the groundwater-surfacewater interaction system looses its equilibrium. The effect of this exogenous shock propagates across the sub-surface flows, and results in the development of a depression in the water-table[5]. The only natural way this water-table rises back to it's original level over relatively shorter periods of time is through heavy rainfall.

Therefore, a given aquifer may undergo recharge at some variable rate, but if the rate of water extraction continues to exceed the rate of recharge, the groundwater will be eventually exhausted. This rate deficit creates scenarios where demand exceeds the availability of water across space and time, and is experienced as water scarcity. Water scarcity typically pertains to a particular region or population. In the context of regulation, measuring water scarcity helps identify areas where there is a significant imbalance between water supply and demand, highlighting the need for effective management strategies. However, as an underground resource, this actual scarcity is hard to measure, particularly because of a lack of volumetric understanding of the geological formations. Thus, water table levels or amount of water in dug-wells or tubewells are typically used as an approximation of scarcity for any given location.

Increasing rates of urbanization and the resultant growing demand for food and water needed for its sustenance, an increase in the variation and intensity of rainfall due to climate change, and a long-term decrease in the cost of groundwater extraction are forcing us to rethink groundwater as a finite resource in terms of its capacity at a reservoir level, which will continue to be replenished at increasingly slower, irregular rates. Beyond these simple first-order effects, there are also more complex natural and artificially induced second-order effects that may affect groundwater replenishment rates and increase instances of water scarcity. If interventions to tackle scarcity are not executed with urgency, regions are confronted with a rapid decline in the availability of water per capita and eventually, a complete depletion of groundwater. While the effects of complete depletion are yet to be observed and documented

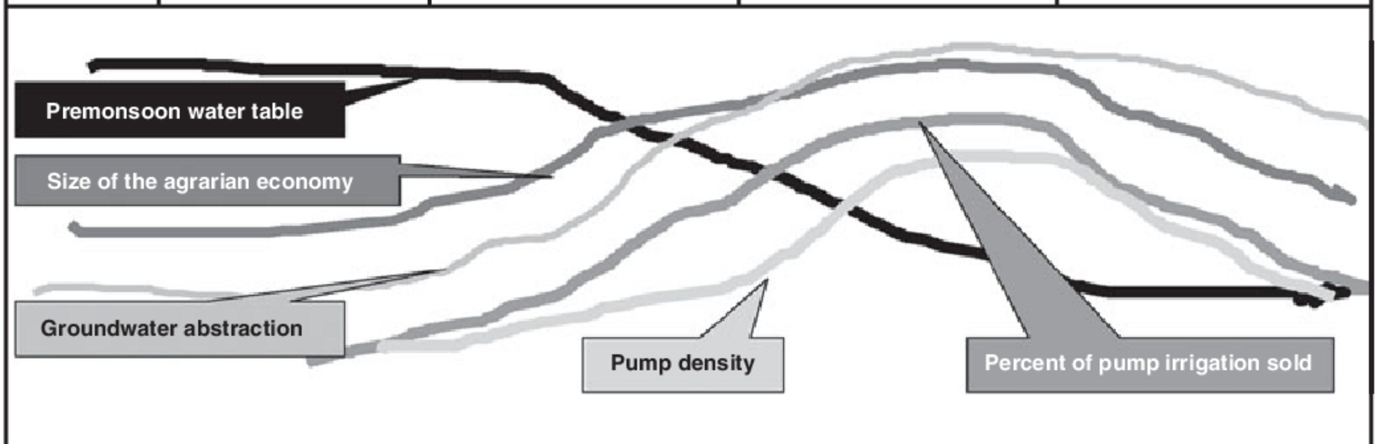
usually does not cross surface boundaries.

[5] Water-table is a common term used in lieu of a piezometric surface, which is an imaginary surface at the subsurface flow level that represents the pressure and potential energy of the groundwater in an aquifer. It is the level to which water will naturally rise in a well or borehole that taps into the groundwater. The shape of the piezometric surface is influenced by various factors, including the topography of the land, the geology of the aquifer, the rate of recharge (infiltration of water from the surface), and the rate of groundwater discharge (through natural springs or human-made wells). Changes in the piezometric surface can impact the availability of groundwater for wells and can also affect natural ecosystems and surface water bodies connected to the aquifer. See (Edwards and Guilfoos, 2021) for a detailed hydrological assessment of effects of groundwater extraction.

extensively, studies indicate that the effects will be most severe in urban areas and will require substantial redistribution of existing water resources at a regional level (LaVanchy et al., 2021).

As the continuous extraction of water grows, the effect and dependence on the resource grows with it. We can identify four typological stages of groundwater 'development' that are specific to South Asia. Figure 2 summarizes these stages in terms of their characteristics and interventions. We also see some interesting indicator trends that go with it. India's groundwater development now stands between Stage 3 and Stage 4, and strategies that would prolong the arrival of Stage 4 are to be emphasized. In the next section, we analyze the historical policy decisions that brought about this accelerated progression in groundwater development in just a matter of 50 years.

	Stage 1	Stage 2	Stage 3	Stage 4
Stages	The rise of Green Revolution and tube well technologies	Groundwater-based agrarian boom	Early symptoms of groundwater overdraft and degradation	Decline of the groundwater socio-ecology with immiserizing impacts



Examples	North Bengal, North Bihar, Nepal Terai, Orissa	Eastern Uttar Pradesh, western Godavari, central and South Gujarat	Haryana, Punjab, western Uttar Pradesh, central Tamilnadu	North Gujarat, coastal Tamilnadu, coastal Saurashtra, southern Rajasthan
Characteristics	Subsistence agriculture; protective irrigation traditional crops; concentrated rural poverty; traditional water-lifting devices using human and animal power	Skewed ownership of tube wells; access to pump irrigation prized; rise of primitive pump irrigation 'exchange' institutions; decline of traditional water-lifting technologies; rapid growth in agrarian income and employment	Crop diversification; permanent decline in water tables. The groundwater-based 'bubble economy' continues booming, but tensions between economy and ecology surface as pumping costs soar and water market become oppressive; private and social costs of groundwater use part ways	The 'bubble' bursts; agricultural growth declines; pauperization of the poor is accompanied by depopulation of entire clusters of villages; water quality problems assume serious proportions; the 'smart' begin moving out long before the crisis deepens; the poor get hit the hardest
Interventions	Targeted subsidy on pump capital; public tube well programmes; electricity subsidies and flat tariff	Subsidies continue; institutional credit for wells and pumps; donors augment resources for pump capital; NGOs promote small farmer irrigation as a livelihood programme	Subsidies, credit, donor and NGO support continue apace; licensing, siting norms and zoning system are created but are weakly enforced; groundwater irrigations emerge as a huge, powerful vote bank that political leaders cannot ignore	Subsidies, credit and donor support reluctantly go; NGOs and donors assume conservationist posture; zoning restrictions begin to get enforced with frequent pre-election relaxations; water imports begin for domestic needs; variety of public- and NGO-sponsored ameliorative actions start

Figure 2: Stages and Indicators of Groundwater Development (Adopted version from Shah, 2007a,b)

3.3 Groundwater in India - Policy Pathways and Responses

Future estimations of water scarcity in India show that the magnitude of global groundwater scarcity will be disproportionately clustered in India and will coincide with the largest centers of urban agglomerations - most notably Delhi, Mumbai, and Bengaluru (He et al., 2021). India's National Commission on Integrated Water Resources Development (NCIWRD) expects growing urban areas to drive up India's water demand by 843 Billion Cubic Meter between 2010 and 2025. This corresponds to a 18.7% jump with respect to the current consumptive demand in a short span of 15 years (Verma and Phansalkar, 2007).

Taking cognizance of the academic responses and the pre-existing "contours of governance" (Kulkarni et al., 2015, p. 172) of groundwater in India, a conscious effort towards policy reform has been outlined in India's forthcoming National Water Policy 2022[6]. The policy calls for the need for strategic paradigm shifts and policy instruments to encourage conservation and prolong the depletion of groundwater. One motivating example that responds to the policy provides the first quantification of water savings for potentially bridging water demand in urban areas through crop substitution in adjoining rural catchments around cities and towns (Singh et al., 2020). Along similar lines, more recent empirical work provides primary evidence that competition arising from administrative fragmentation of groundwater resources is significant in India and leads to the over-extraction of groundwater as observed from depleted groundwater wells (Bhogale and Khedgikar, 2022), albeit in a spatially implicit manner. These types of studies represent explicit problem inspired approaches that take into account the complex interactions between systems of resources, governance, and individual actions (Sivapalan et al., 2014).

Like most developing, democratic countries, the burden of regulating groundwater in India is considered as the *de-jure* responsibility of the state. However, the limited (or gradually growing) capacity of states and the expansive extent of groundwater development typically leads to a situation of unenforced regulation, operational paralysis or misdirected policies (Bosch et al., 2021). The latter is more commonly observed in India, where promotion of agricultural interests by using unregulated, subsidized access to groundwater as a strategic instrument in the 'vote-bank' politics typically seen in parliamentary elections (Mukherji and Shah, 2005; Shah, 2010).

From a legal perspective, regulation of groundwater is typically linked with common property rights, developed and enforced either through formal institutions or in an informal manner locally by users of the groundwater (Tiboris, 2019). Yet, while robust property rights are established and enforced throughout most countries, the same cannot be said

[6] Also known as the Mihir Shah Committee with all authors in (Kulkarni et al., 2015) serving as executive committee members.

for resource rights, especially groundwater. The problem is not so much about the state capacity, availability of policy instruments or doctrines of ownership as it is about the extent to which laws on property rights fail to reconcile with the concept of groundwater regulation. For example, if we consider the Indian Easements Act 1882 - a fundamental act that defines non-possesory rights over land, it recognizes the right of the landowner to collect and use groundwater on their property. However, this Act fails to act as an instrument to ensure effective use of groundwater over different type of properties or land use[7].

3.4 Historical Policy Linkages and the Water-Food-Energy Nexus

The dependence of agriculture and urban areas on groundwater in India has been consistently on the rise and since the mid-1960s. The need for water above and beyond the natural surface water sources and monsoons was triggered in India by the introduction of high-yielding varieties (HYV) of wheat and rice during the Green Revolution (1960s), with an aim of simultaneously tackling drought and famine. As irrigation was necessary for the productivity-enhancing effects of varieties, an extensive network of dams and canals was proposed first under India's central planning regime.

However, by the mid-1970s, technology improvements in digging deeper tubewells and borewells, coupled with a decentralization of India's electricity grids enabled the capitalization of groundwater for enhanced agricultural productivity (Mukherji, 2022; Palit and Bandyopadhyay, 2017). As a result, the burden of infrastructure development for canals and dams was lessened, but the effects of unregulated extraction of groundwater for three decades are now being experienced by the Indian population in the form of acute and chronic scarcity. This free, unregulated use of groundwater has also spilled over into the consumption baskets in India over time, leading to price distortions in markets for less-water intensive crops, and a steady increase in virtual exports of water (Nishad and Kumar, 2022) due to the indirect incentivization of producing of water-intensive cash crops. We can trace this issue of groundwater scarcity to two linkages created across systems of critical agriculture inputs & outputs - *policies for management of crop price* and *electricity tariffs*.

The Minimum Security Price in India [8] - which was introduced by the Government of India around the same time as the HYVs. The intent behind the MSP system was to develop

[7]In Subsection (3.4), we provide an overview of other acts that enable or inform policies for groundwater in India

[8]The Minimum Support Price (MSP) is the rate at which the government of India purchases crops from farmers. The MSP is calculated as at least one-and-a-half times the farmers' cost of production. The government sets the MSP twice a year for 24 commodities. The MSP is a form of government intervention to insure farmers against a steep decline in the prices of their goods and to help them prevent losses. When the market price falls below the declared MSP, the government would purchase the entire quantity from the farmers at MSP.

a pan-India solution to simultaneously tackle the occurrences of droughts and famines, after the Green Revolution had allowed operationalization of distribution channels for inputs (seeds, fertilizers, and primarily surface water irrigation through canals).

In order to ensure that farmers had a social security net in case of crop failure, the MSP was initially established for the most 'successful' HYV products of the Green Revolution – Wheat and Rice but quickly extended to other crops such as maize, sugarcane, and cotton. Simultaneously, the central and state governments also established output channels (marketing and procurement channels through mandis, fair price distribution shops enabled with the issuing of 'Ration Cards', and distribution and storage infrastructure in the form of large warehouses) as a brute-force strategy for livelihood and food security. The MSP system has undergone several changes over the years. Most notably:

- In 1972, the government established the Central Agricultural Prices Commission (CAPC) to advise on MSPs and ensure that the prices paid to farmers were remunerative. In 1985, the government introduced a state bonus on MSP to incentivize farmers to increase production.
- In 2004, the M.S. Swaminathan Committee was set up to review the MSP system and suggest measures to improve the income of farmers. The Committee recommended that the MSP should be at least 50% more than the cost of production, which has been accepted by the government[9].
- In 2014, the Shanta Kumar Committee was set up to review the procurement policy of food grains and the management of food stocks. The Committee recommended several measures to revamp the MSP system and make it more efficient.
- Finally, in 2017, the Ashok Dalwai Committee was set up to review the MSP system and suggest measures to double farmers' income. The Committee recommended that the MSP should be based on comprehensive cost factors, including labor and management costs.

A result of the MSP policy and multiple its revisions has created an inextricable linking of the production and consumption patterns across the country, leading to a decrease in the prices of wheat and rice nationally, an increase in consumption of rice and wheat in favor of traditional crops such as millets, while farmers moving away from traditional crop choices in order to minimize their risk.

[9]See Report Committee available at https://agricoop.nic.in/sites/default/files/NCF3%20%281%29.pdf

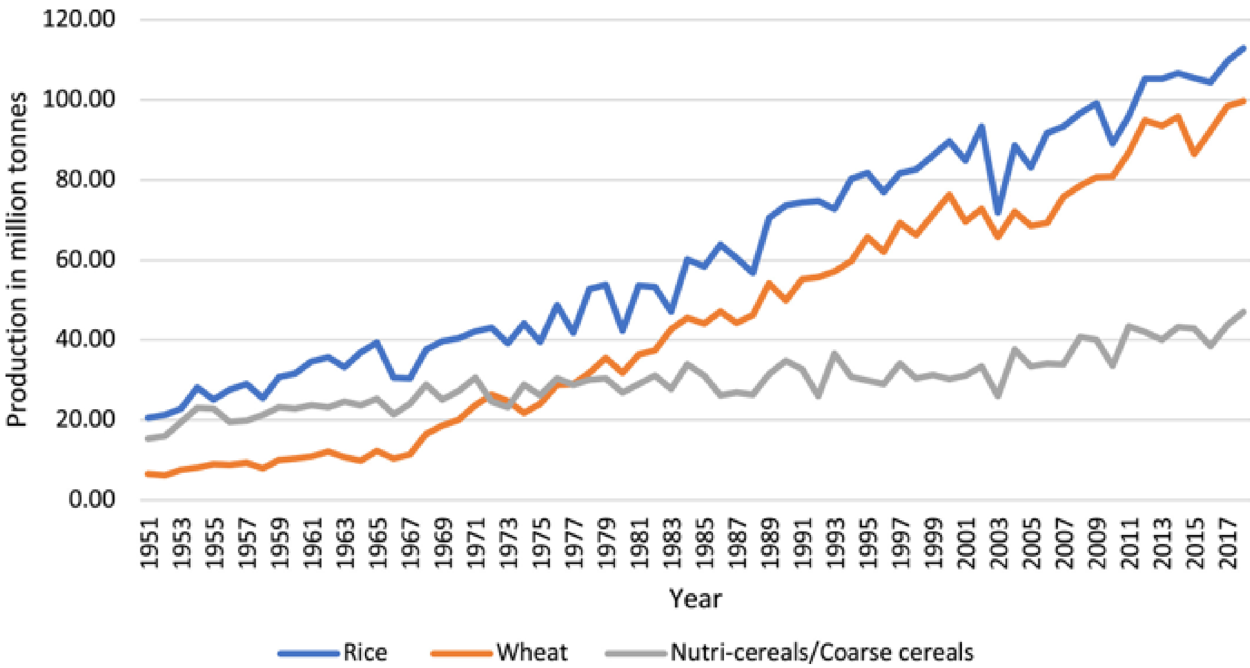


Figure 3: Cereal Production in India 1950-2017 (Source: Mukherji, 2022)

The second, more complex form of linkage coincided with advances in groundwater extraction technology and the operationalization of rural electrification, which started in the late 1960s. Key historical policy decisions that facilitated this were:

- Based on the recommendations of the Reserve Bank of India, the Rural Electrification Corporation Limited (RECL) was set up in 1969 to promote pump electrification (Palit and Bandyopadhyay, 2017). By the early 1970s, there were more than a million wells with electric pumps in the country. Eventually, the lack of robust central-grid infrastructure and the expansive nature of India's settlement patterns meant that the electricity grids and their management had to be decentralized.

- By 1980 State Electricity Boards were formed, which strengthened local techno-political linkages, spurring the growth of the cooperative movements for cash crops such as sugarcane. These movements created strong, politically and industrially backed lobbies that kept the cost of water extraction low (Shah, 1993).

- Finally, when India liberalized its economy in 1992, these linkages strengthened further as India's exports of water-intensive cash crops grew rapidly and the political economy of water-intensive crops deepened.

These reforms led to a link between subsidized, flat electricity tariffs and groundwater extraction, leading to a complete loss of incentive for the government to understand the *spillover of distorted price dynamics into the consumption basket.*

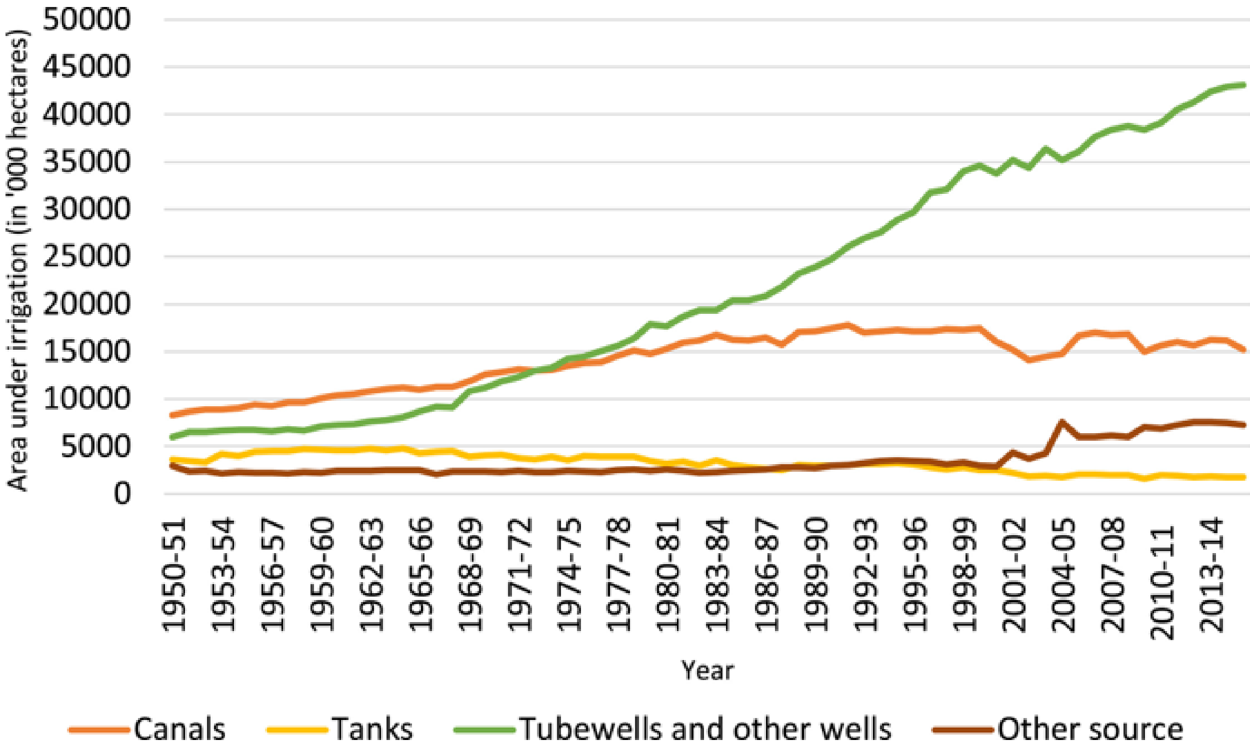


Figure 4: Net Irrigated Area by Type of Irrigation 1950-2012 (Source: Mukherji, 2022)

We posit here that these linkages and their effects across geographies and institutions, which shape India's Water-Energy-Food nexus, are centered around the phenomena of scarcity. Understanding the spatial and temporal dynamics of variables that may drive scarcity and thus becomes an important consideration that should be analyzed further.

For example, consider the components of the *Cost of Cultivation*[10] that are currently used to determine the MSP by the central government - we see that there is no component that accounts for the variation , which have now been found to be substantial for the same crops in different parts of the country. From an economic perspective, it has been shown that change to consumption baskets in order to normalize the distorted prices would require a significant investment materialized as a loss to the Exchequer before the long-term benefits of changes in consumption basket are visible in nutritional outcomes (Chhatre et al., 2019).

Currently, the policy ensemble for water regulation and the key decisions related to water

[10]The cost of cultivation of crops in India is the sum of the variable costs and the fixed costs. The variable costs include the cost of seed, manure, fertilizers, irrigation, and labor. The fixed costs include the value of hired human labor, hired bullock labor, owned bullock labor, owned machinery labor, hired machinery charges, value of seed, value of insecticides and pesticides, and value of manure. The total cost of cultivation is the sum of the variable costs and the fixed costs.

continue to be driven by the Central government of India, with very little autonomy given to the States for decision-making. Thus, despite water being a State Subject as per the Indian Constitution, the central government has become the *de-facto* regulator of water use. The current landscape of water regulation in India is an ensemble of one directive policy with overarching guidance and non-statutory guidelines - the National Water Policy, missions and programs like the Jal Jeevan Mission (JJM) for drinking water supply, Pradhan Mantri Krishi Sinchan Yojana (PMKSY) for irrigation, Atal Mission for Rejuvenation & Urban Transformation for urban water services, and several campaigns such as SahiFasal Campaign (Correct Crop Campaign), Jal Shakti Abhiyan (Water Empowerment Program), Catch the Rain, and Water Heroes for increasing awareness about water problems and conservation [11]. The way this ensemble operates is highly hierarchical and top-down, with key decisions regarding public *investments* in *physical infrastructure* for groundwater extraction and indirect *financial welfare* determined almost entirely at the central level.

At a policy level, two problems in this ensemble are apparent. The first problem is rooted in the conflicting legislative assignment of water as a state subject for efficient use, versus the judicial interpretation of the right to water being an implicit extension of constitutional rights[12]. The second problem is that due to the expansive, federated nature of government in India, urban and rural governance silos that originate at the central level are devolved to the local level. Moreover, since the legislative frameworks for urban and rural local governments in India are distinct, governed by the 73^{rd} and 74^{th} Amendments to the Constitution[13], there is no policy instrument, act or even an ad-hoc framework for implementing checks and balances to regulate the dynamics of coupled water extraction at the District level.

Other than the Indian Easements Act discussed in the last section, regulation of groundwater is also loosely governed by 4 other acts. These are

1. *The Water (Prevention and Control of Pollution) Act, 1974:* Although primarily

[11] See https://pib.gov.in/PressReleseDetail.aspx?PRID=1810525 for an updated legislative Press Release on the landscape of India's water regulation.

[12] In India, water is a State subject, but the provisions have historically been a matter of conflict themselves. The primary entry in the Constitution relating to water is Entry 17, which is in the State List. It brings water including water supplies, irrigation and canals, drainage and embankments, water storage, and water power under the purview of the state list. But there is also a provision that enables the Union to deal with Inter-State rivers if Parliament legislates it to be in the public interest, via Entry 56 in the Union List. Since water conflicts primarily arose on inter-state surface-water disputes starting in the 1950s, no specific exception has been made in differentiating groundwater and surface-water within the Union list

[13] The 73rd Amendment of the Indian Constitution, also known as the Panchayati Raj Act, was passed in 1992 to empower local self-government institutions in rural areas. It granted constitutional status to Panchayats and provided them with the authority to govern and make decisions on local matters. The 74th Amendment, known as the Nagarpalika Act, was passed in 1992 as well, aiming to strengthen urban local bodies. It granted constitutional status to Municipalities and enabled them to perform functions related to urban planning, governance, and service delivery. It emphasized citizen participation in decision-making and aimed to improve urban governance and development.

focused on surface water, this act also includes provisions to prevent and control water pollution from wells, tube wells, and other sources that may impact groundwater quality.

2. *The Model Groundwater (Regulation and Control of Development and Management) Bill, 2011:* The central government drafted this model bill, encouraging states to adopt and enact it. It provides a legal framework for the regulation and control of groundwater development and management. The bill empowers state governments to formulate rules for the registration of wells and boreholes, assessment of groundwater availability, and restrictions on the drilling of new wells.

3. *The Environment (Protection) Act, 1986:* This act enables the central government to take measures to protect and improve the quality of the environment. Under this act, the Central Ground Water Authority (CGWA) was established to regulate and control the development and management of groundwater resources at the national level.

4. *State-Level Groundwater Regulations (Different Years):* In addition to the central-level regulations, each state in India may have its own specific policies and regulations related to groundwater. These may include restrictions on drilling new wells, pumping groundwater, and implementing water conservation and recharge measures.

Based on this discussion, we argue that these problems manifest as water stress at the District level, where urban water rights and use remains de-coupled from rural water rights and use in the adjoining areas. Moreover, the allocation of water rights continues to follow a traditional riparian system at the local level, and water use thus while the land use might change over time, the cognizance that underlying water reserves remain the same is lacking. Since the uses of water are distinct for urban and rural areas, the de-coupling results in a negative spillover between urban and rural areas (or surrounding rural areas with different dominant crops), which needs to be further understood through a theoretical insight. The most significant effect of these spillovers is seen when scarcity of water faced in urban areas is a result of subsidization and consequent over-extraction of water from growing water intensive crops that fetch higher market values globally but ultimately lead to virtual exports of water. From an urban perspective, the coincidental contiguity with agricultural areas that cultivate water intensive crops causing a reduction in the rate of urbanization and limiting economic development and social growth in the long run remains a major challenge that remains unaddressed.

The negative spillovers have an impact beyond just the apparent scarcity of groundwater. There is substantial evidence to show that the variety of nutritional deficits in the Indian population are linked to a disproportionate consumption of rice and wheat which is indirectly a consequence of India's consumption basket (Davis et al., 2018). Since consumption baskets in India are driven by distribution of grains through the Public Distribution Systems, the floor of Minimum Support Prices creates price distortions indirectly by affecting the crop choice. If farmers are guaranteed a MSP for a particular crop that is water-intensive, they are more likely to over-extract water relying on subsidized electricity rates based on the MSPs declared by the Government as a risk-averse option. Wheat, Rice, and Sugarcane are some of the key crops that fall within this category. This difference has been mapped with the help of the WATNEEDS model (Chiarelli et al., 2020) for Blue & Green water requirements in Figure 5.

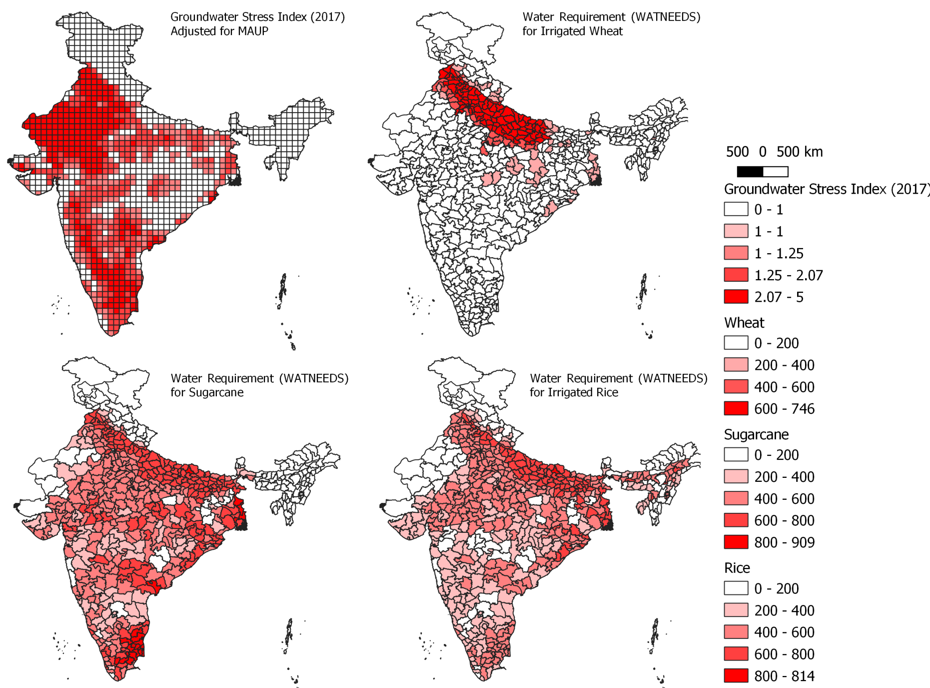


Figure 5: Groundwater Stress and Average Total Water Requirements (mm) based on WATNEEDS model

Having outlined the problems at the policy level, we now focus on the issues in operational aspects and institutional linkages that contribute to the water stress.

4 Operational Issues & Water Scarcity

In India, the relationship between the central government, state governments, and local governments is defined by the principles of federalism. The country follows a federal system of governance where power and authority are divided between the central government and the state governments. We begin with a brief overview of the the different tiers of government in India and the overarching legislative framework that guides policies of water regulation in a top-down manner.

1. *Central Government:* The Central government, also known as the Union government, is responsible for matters that affect the entire country or multiple states. It has jurisdiction over subjects listed in the Union List and Concurrent List of the Constitution of India, such as defense, foreign affairs, currency, communication, and interstate commerce. The central government is headed by the Prime Minister and exercises executive, legislative, and financial powers at the national level.

2. *State Governments:* India is a union of states, and each state has its own government headed by a Chief Minister. State governments have jurisdiction over subjects listed in the State List of the Constitution, such as police, public health, local government, agriculture, and irrigation. They have the authority to make laws and policies on these subjects within their respective states. State governments also have residual powers, meaning they can legislate on matters not specifically listed in the Union List or Concurrent List.

3. *Local Governments:* Local governments in India are referred to as Panchayats in rural areas and Municipalities or Municipal Corporations in urban areas. They are responsible for local administration, governance, and service delivery at the grassroots level. Local governments have limited powers and functions, including provision of basic amenities like water supply, sanitation, roads, health, education, and local-level planning and development. The 73rd and 74th Amendments to the Constitution in 1992 strengthened the role of local governments and ensured their constitutional status.

The relationship between the central government, state governments, and local governments is characterized by a division of powers and responsibilities characteristic of Cooperative Federalism. Figure 6 shows this relationship and major linkages between the different tiers of the government. We intend to limit our focus to the district level (boxed out in Figure 6) as it is the key *spatial, administrative*, and *fiscal* unit of interest our research.

At the District level, the operational problems can be characterized along three interdependent dimensions - *Fragmentation of Authority, Inadequate Local Capacity, Cold Spots in Transitional Areas, and Lack of Monitoring Infrastructure.* Figures 6 & 7 act as a reference for the discussion on each of these dimensions.

1. *Fragmentation of authority*: The division of powers between different levels of government can lead to a fragmented approach to water regulation. Water management involves multiple aspects, including allocation, conservation, pollution control, and infrastructure development. When different levels of government have overlapping or unclear responsibilities, it can result in gaps, duplication of efforts, and lack of coordination, leading to ineffective regulation. Just at the state and District level, Apex bodies, State Advisory bodies, Parastatal agencies, and Urban and Rural local government departments (along with the *Zilla Parishad* or the District council) act are agencies across with authority is fragmented. How-

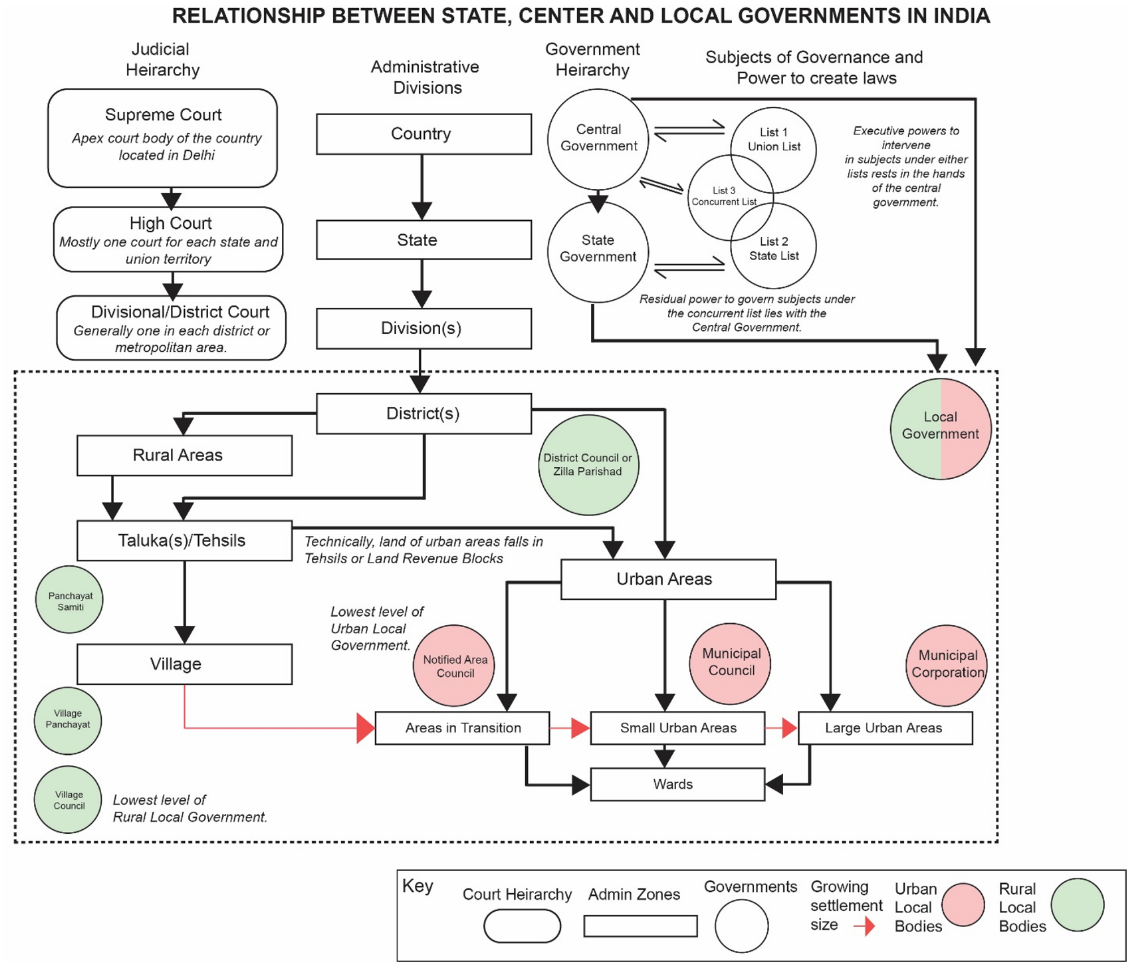


Figure 6: Relationship between State, Center, and Local Government in India

ever, despite the presence of these different agencies at an operational level, the key decisions are still made at the central level as outlined previously.

2. *Inadequate Local Capacity*: The reason why key decisions with respect to water regulation are made at the central level is that most urban and rural governments at the District level face capacity constraints in terms of technical expertise, human resources, and a lack of financial autonomy for implementing effective water regulation measures. These gaps in capacity lead to delays, inefficiencies, and inadequate enforcement of regulations if the responsibility for regulation is devolved locally.

3. *Transitional Cold Spots* - The transitional cold spots arise due to a combination of a lag between the notification of transitional areas from rural to urban and a failure in establishment of operational Metropolitan and District Planning Committees (MPCs & DPC). In the majority of states, DPCs are not fully functional. In a few states, they are not even properly constituted, whereas, in the majority of states, they are constituted

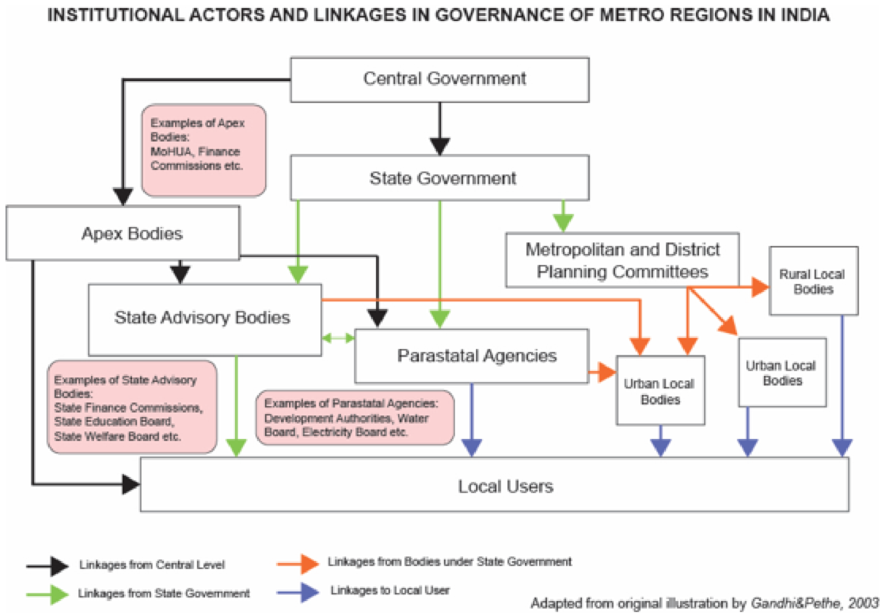


Figure 7: *De-Jure* Institutional Actors and Linkages at the Urban District Centers of India

but not in constitutionally desired ways. Ministers serve as DPC Chairperson in several states, including Chhattisgarh, Gujarat, Madhya Pradesh, Orissa, Maharashtra, Punjab, and Himachal Pradesh. As a result, the participatory nature of the planning process in DPCis severely hampered.

4. *Lack of Monitoring Infrastructure*: Perhaps the most important amongst the operational issues from the point of view of this research is the lack of comprehensive and up-to-date groundwater monitoring systems across India, leading to insufficient data on groundwater levels, quality, and recharge rates. Due to the lack of local capacity, the monitoring of groundwater through well levels is still undertaken in most states [14] through a quarterly primary survey of wells that is conducted by the Central Groundwater Regulatory Board (CGWB). The Central Groundwater Regulatory Board is the de-facto regulator of all groundwater across the country. Constituted under the Ministry of Jal Shakti, Government of India, the CGWB is responsible for groundwater management and regulation in the country. The CGWB operates under the overall administrative control of the Department of Water Resources, River Development, and Ganga Rejuvenation. The primary legal framework for the CGWB's activities is provided by the Water (Prevention and Control of Pollution) Act, 1974, and the Environment (Protection) Act, 1986. These acts empower the CGWB to regulate and control groundwater extraction and protect groundwater resources

[14]The only exception to this is the newly restructured state of Andhra Pradesh, where a program called APGRACE in collaboration with an MIT funded Vassar Labs has implemented state of the art IoT devices for real-time monitoring of groundwater wells. See https://www.vassarlabs.com/cms/portfolio-item/apgrace/ for more information.

from pollution and degradation. The CGWB also works in close coordination with state groundwater departments and various other agencies involved in water resources management. One of the key regulatory tasks of the CGWB is to monitor wells that are typically constructed as open wells or tube wells. As of 2017, there are 23209 monitoring wells (also known as Groundwater Monitoring Stations) across the country, which have been mapped in Figure 8.

Given the dimensions of operational deficiencies outlined above, we argue here that there is a possibility of misdirected investments that promote water intensive practices in districts that are already low on groundwater. Investigating this argument becomes a central thesis of this research, because qualitatively, the operational deficiencies discussed here make the possibility of a misdirected welfarism driven by political or perceptual guidance more tenable across rural areas. Two definitive audits reports on the state of groundwater in India highlight these issue in much greater detail (Garduño et al., 2011, pp. xiv and Comptroller and Auditor General of India, 2021, pp. i-xii,21). The salient points from both the reports that support our thesis are summarized below:

1. "CGWA, CGWB, and SGBs are severely handicapped by their chronic under-staffing and lack of coordination with a large number of other government agencies impacting the resource. severely handicapped by their chronic understaffing and lack of coordination with a large number of other government agencies impacting the resource."

2. "Institutional technical capacity at the national and state levels is generally weak, but there is a core base of high-level professionals and the knowledge base and potential of research organizations, government institutions and universities is well-developed."

3. "Given the millions of water wells in India—mainly drilled, operated, and maintained by private users—and the scant institutional enforcement capacity, we cannot currently expect users to contribute toward groundwater management costs. There is, however, an outstanding issue that merits discussion: the groundwater-energy nexus and one possible way forward through upscaling Gujarat's promising Jyotigram Scheme."

4. "Groundwater agencies in the states are not adequately equipped to take up these roles. These agencies are located at relatively low levels in the state hierarchy and tend to have much less clout than their counterpart departments focused on, for example, irrigation or water supply. In many cases, there is no dedicated state groundwater agency."

5. "There were numerous cases in which conditions stipulated in the NOCs were violated.

Despite the widespread violations, CGWA issued (2013-18) show cause notices to only 99 project proponents."

6. "In 18 of the 28 states of India, groundwater legislation has been either partially implemented (4) or not implemented at all (14)."

While the CGWB is responsible for monitoring the wells, the financial outlays for welfare schemes primarily directed at farmers is mostly undertaken by the locally Department of Agriculture - which unlike the CGWB has a very strong local presence due to the expansive operationalization of HVYs that was undertaken from the 1960s (Bhogale and Khedgikar, 2022). If local observations of scarcity are not based on a scientific understanding of groundwater levels, then the phenomena and represents another type of, albeit institutionally driven negative spillover that exacerbates the scarcity of groundwater. Along similar lines, as urban areas grow within a district, these areas become the dominant intensive users of groundwater. Therefore, the percentage of urban population may also drive the difference between real and apparent scarcity of water. Lastly, presence of other water river bodies or a large surface-water reservoir such as a lake in districts might also affect the perception of water scarcity as typically,.

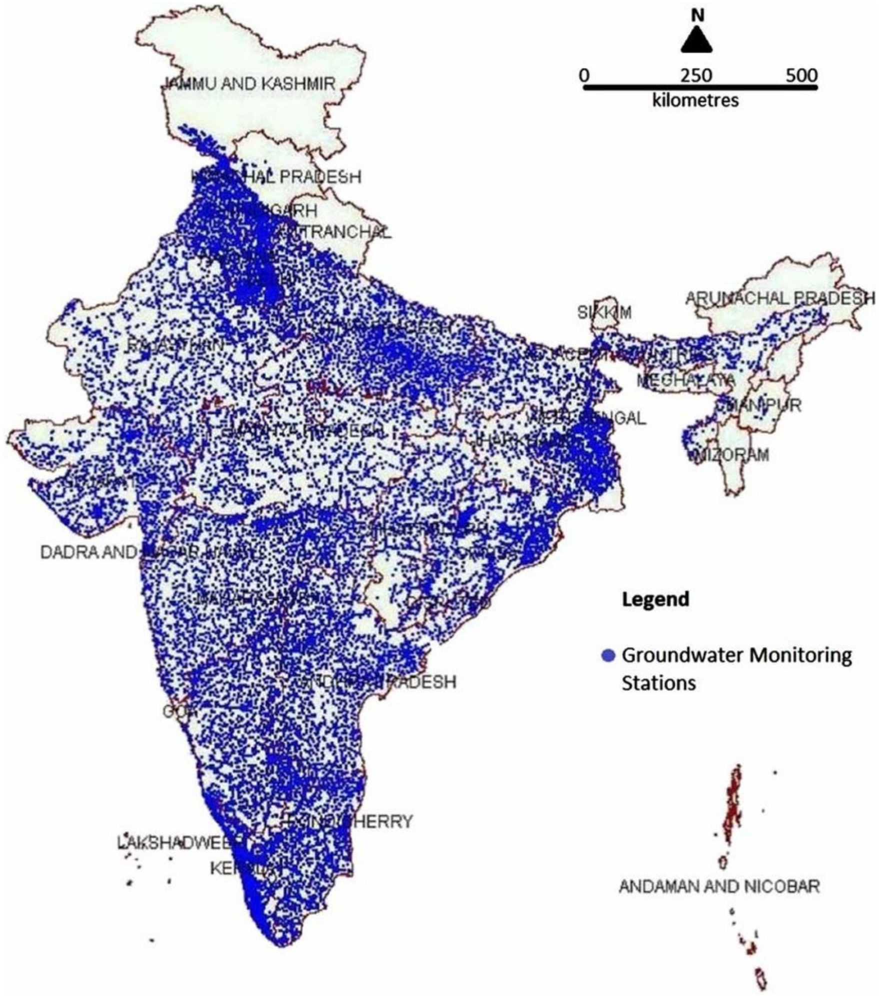


Figure 8: CGWB Monitoring Wells in India

5 Districts and Sub-Basins: The Importance of Spatial Boundaries

It is important to emphasize the distinction between *actual scarcity* and *perceived scarcity* discussed in the previous section in regards to the effects of competitive extraction of groundwater in India because a failure to understand this distinction can lead to over-extraction due to misdirected investments. Specifically, due to the complex hydrogeological structures that form aquifers (approximated in this study as sub-basins), it is often possible that a well that is used to monitor the level of water at a particular location might inaccurately represent the true volume of water in the underlying aquifer over a larger area. Across a complex network of government agencies like is the case discussed for India, if some financial investments are determined on the basis of the *perceived scarcity* of water, as observed from seasonal and perennial dry wells, it might be lead to inefficient allocation of funds and consequential over-extraction of water, which can only be understood through *actual scarcity*.

Focusing on districts for subsequent analysis also follows from the previous section, since districts are primary agencies that have autonomy on the welfare investments at the agricultural and domestic level (primarily rural).

5.1 LISA for Sub-basins and Districts (Local Moran's I)

First, we observe that for 99999 permutations, the pseudo p-values are extremely low in the case of Districts, and relatively higher for sub-basins, but in both cases allows us to overwhelmingly reject the null of Spatial Randomness. Next, we observe some very interesting patterns of clustering in both districts and sub-basins which are interpreted below and illustrated in Figure 9. Note that in each case, we have used the Queen contiguity to generate the row-standardized spatial weights matrix.

- For districts, we look at the area as the variable. The high-high clusters are seen in regions that were integrated into India after independence, and existed as autonomous Princely states. The high-high clusters are seen in Kashmir, Rajasthan and AP/Telangana region, which were all one or more erstwhile princely states.
- Next, we observe that the low-low clusters are found in Punjab and Haryana, and Himachal Pradesh and Uttarakhand, which are states that essentially underwent splits in 2000. The contiguous portion of Delhi, which was also designated as NCR is another low-low cluster.
- For sub-basins, we look at the number of fractions as a variable. We see that the Indo-Gangetic Plane has a large, contiguous high-high cluster of fraction count. This is the most fertile part of India, through which the rivers Ganga and Yamuna flow.
- Lastly, we observe that there is very little overlap between the district and sub-basin clustering, indicating that contiguous non-clustered districts fragment large sub-basins, whereas the overlap of larger districts and less fragmented sub-basins indicates is only found along the northern Himalayas and Rajasthan (i.e along boundaries of the geography).

Thus, we realize that the boundaries between districts and sub-basins do not overlap. In terms of their area, the sub-basins are fractionalized more in regions where districts are smaller. Moreover, district areas also show a significant amount of clustering, in different regions. Thus, we now need to focus on the selection of a measure for fragmentation that captures this relationship.

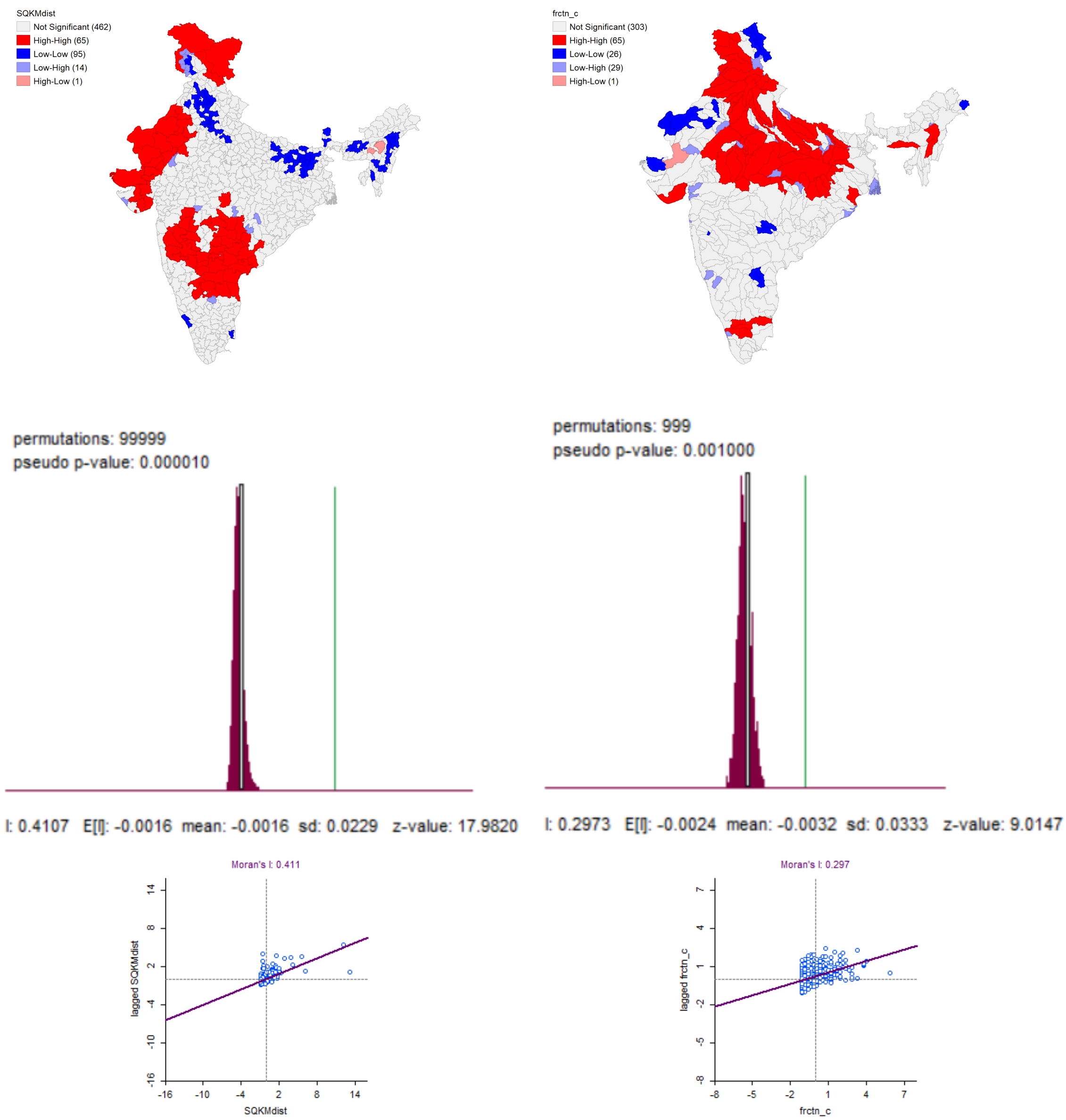


Figure 9: Clustering in Districts (left) and Sub-Basins (right) by Area and Fraction Count (Local Moran's I)

5.2 Indices of Fragmentation

We briefly review the indices shortlisted for the analysis and their trends over time, to justify the choice of the I-HHI as an index for measuring fragmentation.

1. Fractions per Sub-basin - Total fragments of a sub-basin due to overlapping Districts, given by n, where n is the total number of fractions.

2. Mean Area per Fraction - The average of the area of each fragment. This is given by

A_{sub}/n, where A_{sub} is the Area of the sub-basin and n is the total number of fractions.

3. Mean Proportion of Fractions - Mean at the sub-basin level. This is given by A_{frac}/A_{sub}, where A_{frac} is the Area of the fraction of each District-Subasin.

4. Weighted Gini Coefficient - Measures the inequality of distribution of fragmentation. Higher the Fragmentation, Higher the Gini Coefficient. The coefficient is given as $G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j |x_i - x_j|}{2\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j x_j}$, where n is the number of observations, x_i and x_j represent the values of the variable of interest for observations i and j respectively, w_i and w_j are the corresponding weights assigned to observations i and j.

5. Inverse HHI - Interpreted as the Concentration of Fragmentation in a sub-basin or the probability that any two districts fragment the same sub-basin. Higher the inverse HHI, higher the concentration of districts fragmenting a given sub-basin. An alternative explanation of this is also that the inverse HHI is lower in the cases where there is a dominant district in terms of its overlap with the sub-basin (i.e more concentration of 'area' within a sub-basin) whereas it is higher in the case of highly fragmented sub-basin with no single dominant district overlapping with a sub-basin. Mathematically, the formula for the inverse HHI is given as $I_i = 1 - \sum_{k=1}^{n} x_{ik}^2$, where x_{ik} is the fractional area.

6. Theil's L Index - Measures the redundancy or lack of diversity in fragmentation configurations. The numerical result is in terms of negative entropy so that a higher number indicates more order that is further away from the 'ideal' of maximum disorder. Mathematically, it is given by $T_L = \frac{1}{N}\sum_{i=1}^{N} \ln\left(\frac{\mu}{x_i}\right)$ where $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$ and x_i is the fractional area.

7. Krugman Index - Measures the average of Relative Deviation from the Mean of a fragmentation configuration. It is given by $K = \frac{x_{ij}}{\frac{x_i + x_j}{2}}$, where x_{ij} is the fraction created by district i and sub-basin j with areas x_i and x_j respectively.

8. Januszewski's Coefficient - A well-known index of land fragmentation. A value of 0 indicates the worst possible land fragmentation situation while a value of 1 indicates no land fragmentation. The Januszewski coefficient is given as $JI = \frac{\sqrt{\sum_{a=1}^{N} X_i}}{\sum_{a=1}^{N} \sqrt{X_i}}$

We compare these indices for India over time, using different boundaries for districts and keeping the sub-basin boundaries the same. The density of the indices from 1960-2011, calculated at decadal intervals indicates that there are some subtle changes in densities due to an increase in number of administrative units and creation of new states. This can be attributed to the 'deepening' of democratic institutions within India as a result of increasing population over time. As regions urbanize, their classification within the Indian system changes. Figure 10 summarizes these densities, whereas Figure 11 gives an insight into the cross sectional trends in district configurations in India.

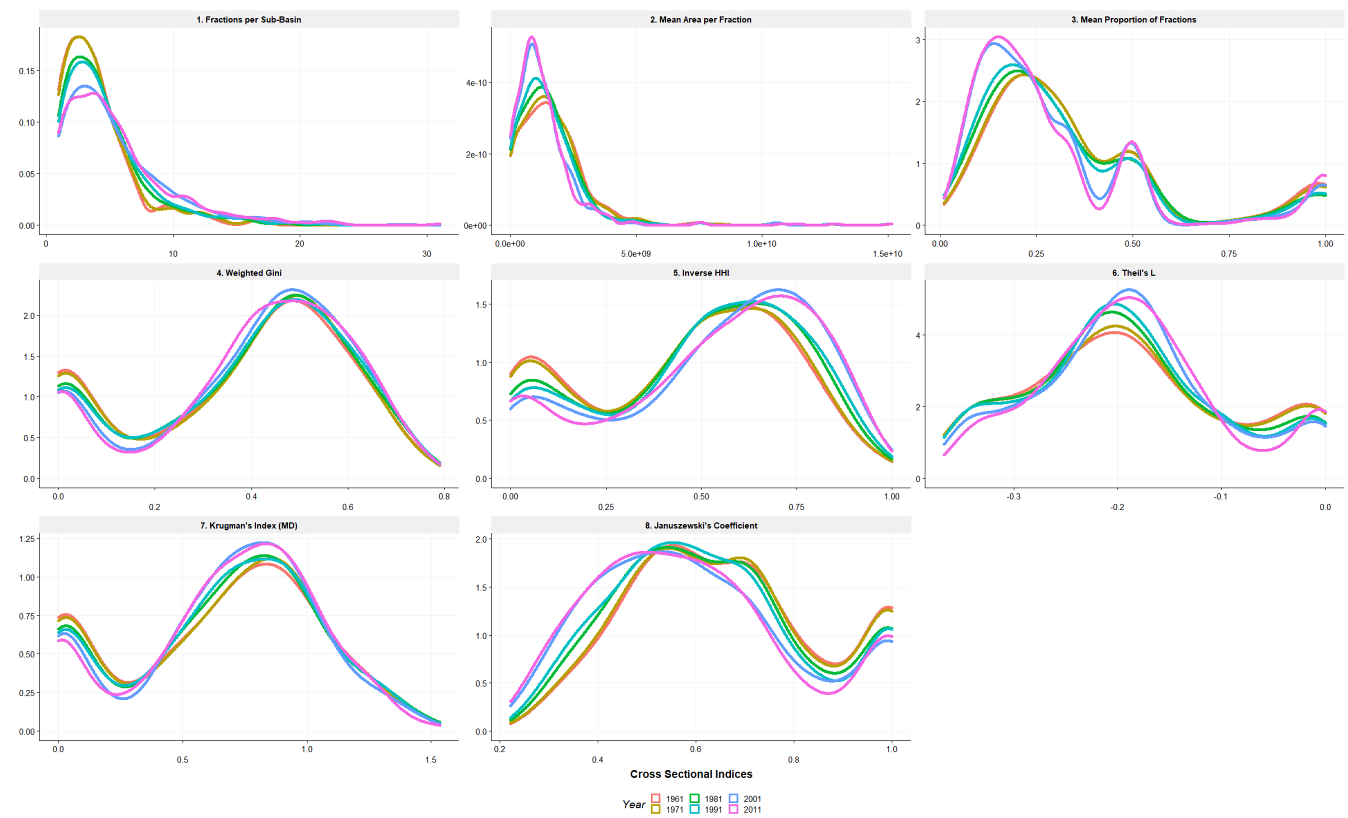


Figure 10: Cross Sectional Trends in Densities of fragmentation of Sub-Basins by District (1961-2011)

We see that the fragmentation has changed over the years as a consequence of the changing shape of districts but constant shapes of the Sub-basins. The densities also reflect the change as a result of a significant delimitation of districts that was undertaken for India in 2001. The number of districts are growing, the mean district area is falling, indicative of growing fragmentation and based on the theoretical model, indicating a trend towards over-extraction of groundwater.

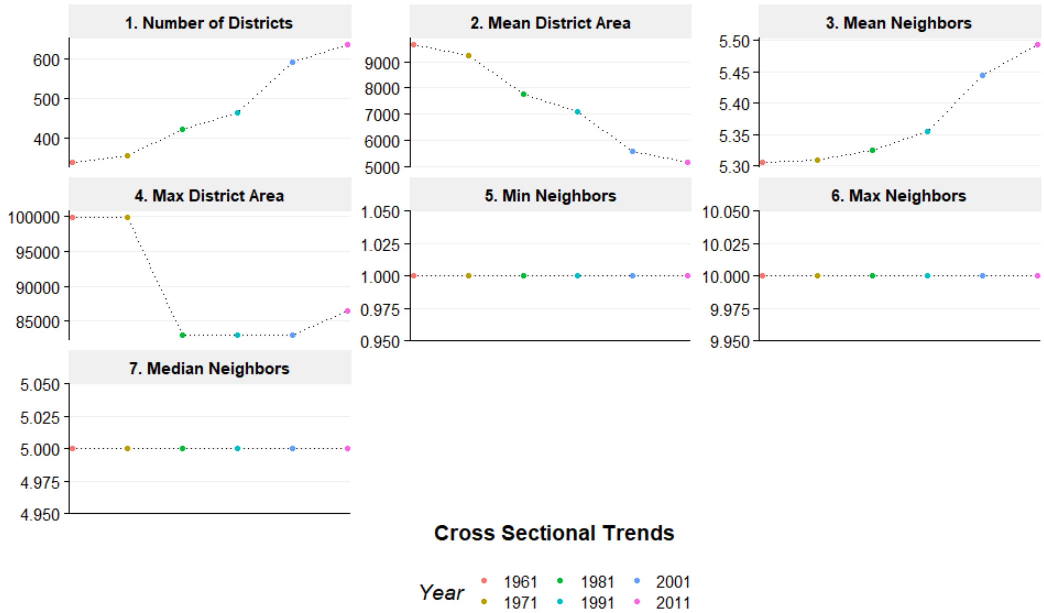


Figure 11: Cross Sectional Trends in District Configuration (1961-2011)

6 Theoretical Model: Externalities in Groundwater Extraction

With a measure of fragmentation in place, we now turn to a theoretical model wherein the dynamic of fragmentation and the spillover of externalities due to unregulated groundwater extraction can be understood as an economic equilibrium.

We develop the initial ideas for the theoretical model based on ideas in (Bhogale and Khedgikar, 2022) to motivate the idea of externalities spilling over across space. This is done through setting up a simple prisoner's dilemma model. We can assume one water-resource with a permeable aquifer system and two overlapping districts: i, capturing share α of the resource, and j, capturing share $(1-\alpha)$. There are two time periods- the agricultural cycle $(t=1)$, when both districts extract water for agriculture denoted by $w_{\{i,j\}}$, and the summer $(t=2)$, when districts consume the remaining water. Both districts can equally access the stock of water available in period 1, denoted by S^1 proportional to their overlap with the resource. The degree of permeability of the water-bearing layer between both districts is measured by $k \in (0,1]$. Here, k determines the magnitude of externalities depending on relative water extraction decisions. The externalities $(k(.))$ are reflected in the period 2 stock of water given by:

$$S^2_{i|w_j} = \alpha S^1 - w_i - k((1-\alpha)w_j - \alpha w_i)$$
$$S^2_{j|w_i} = (1-\alpha)S^1 - w_j + k((1-\alpha)w_j - \alpha w_i)$$

Note that the externality for i is negative if j's extraction $w_j > w_i$, and vice-versa. The returns to water in both periods are concave ($R(w_i)$ and $\pi(S_i^2)$) and costs of extraction in period 1 are convexly proportional to their own extraction $C(w_i)$.

In order to look at comparative statistics for equilibrium extraction with respect to the level of resource capture, we fix the number of districts to 2. We assume that i and j have information on the following components: the stock of water in period 2 ($S^2_{\{i,j\}}$), their own extraction $w_{\{i,j\}}$, and their share of resource-capture (α) is observable to them. They may infer k over time from $\frac{\partial S^2}{\partial w_{\{i,j\}}}$ by experimenting with their own extraction, and through information from experts and engineers in the district administration to study water-levels and movements. They may or may not directly observe the other district's extraction level, but this can be inferred under the current functional form of S^2. Additionally, we assume a linear relationship between water extraction and the externality on the stock of water. This assumption enables districts to learn about the other district's extraction level from their experience of other observable parameters.

Claim 1: Extraction is higher in a decentralized regime relative to the social planner

The social planner's objective function maximizes the sum of the net returns for i and j districts over both time periods, internalizing externalities across boundaries.

$$w_i^{SP} : \max_{w_i, w_j} \Big[R(w_i) + R(w_j) - C(w_i) - C(w_j) + \delta\pi\big[\alpha S^1 - w_i - k((1-\alpha)w_j - \alpha w_i)\big] + \delta\pi\big[(1-\alpha)S^1 - w_j + k((1-\alpha)w_i - \alpha w_j)\big]\Big]$$

The first-best equilibrium extraction levels reflect the equality of the marginal cost (MC^{SP}) and marginal revenue (MR^{SP}) for i and j under the social planner's regime.

$$\implies w_i^{SP} : \big[\frac{\partial R(w_i)}{\partial w_i} - \delta(1-\alpha k)\frac{\partial \pi}{\partial w_i}\big(S_i^2\big) - \delta(1-\alpha)k\frac{\partial \pi}{\partial w_i}\big(S_j^2\big)\big] = \frac{\partial C(w_i)}{\partial w_i}$$

$$\text{and, } w_j^{SP} : \big[\frac{\partial R(w_j)}{\partial w_j} - \delta(1-(1-\alpha)k)\frac{\partial \pi}{\partial w_j}\big(S_j^2\big) - \delta\alpha k\frac{\partial \pi}{\partial w_j}\big(S_i^2\big)\big] = \frac{\partial C(w_j)}{\partial w_j} \tag{1}$$

In contrast, under a decentralized regime, each district maximizes their own returns, agnostic to the externality, over both periods. The equilibrium extraction is:

$$\implies w_i^{Dec} : \frac{\partial R(w_i)}{\partial w_i} - \delta(1 - \alpha k)\frac{\partial \pi}{\partial w_i}\big(S_i^2\big) = \frac{\partial C(w_i)}{\partial w_i}$$

$$\text{similarly, } w_j^{Dec} : \frac{\partial R(w_j)}{\partial w_j} - \delta(1 - (1-\alpha)k)\frac{\partial \pi}{\partial w_j}\big(S_j^2\big) = \frac{\partial C(w_j)}{\partial w_j} \tag{2}$$

where MR^{Dec} is the marginal revenue and MC^{Dec} is the marginal cost under the decentralized regime. Comparing these, $MR^{Dec} > MR^{SP}$, as $\delta > 0$, $k \in (0, 1]$, and $\frac{\partial \pi}{\partial w_j}\big(S_i^2\big) > 0$ implying there are uninternalized externalities under the decentralized regime that lead to inflated returns to extraction. However, the marginal cost function is unchanged $MC^{Dec} = MC^{SP}$. Given, concave $R(.)$ and $\pi(.)$ function and convex $C(.)$ function, holding cost constant if w is such that $MR = MC$, then,

$$w^{Dec} > w^{SP}$$

Claim 2: Share of overlap on the resource influences extraction levels in decentralization
Here, we look at the comparative statistics for extraction with increasing resource-capture (α). Assuming the parameters as $k = 0.5$; $\delta = 0.8$; $S^1 = 10$ (based on experimental calibration), imposing a functional form on concave returns, $R(w) = \log(w)$ and $\pi(S^2) = \log(S^2)$, with convex costs $C(w) = \frac{w^2}{10}$, and adding constraints $w_i + w_j \leq S^1$, $w_i \geq 0$, we look at comparative statistics for total equilibrium extraction with respect to α.

For sake of simplicity, this can be thought of as the central line, or the "spatial boundary" between the two district systems to be moving (see Figure 12.), resulting in different extraction statics. Intuitively it is easy to see that as α increases towards a more equitable split between the two districts, total equilibrium extraction on the resource goes up. This is aligned with increased competition for the resource leading to increased extraction. However, as α continues to go up, the total equilibrium extraction falls. This is aligned with lower competition leading to less water-intensive practices.

While this model has been developed for a 2-units-1-resource scenario, it is easy to see that it is generalizable to n-units-k-resource scenario, where the resource and administrative boundaries are independently contiguous. However, deriving first order conditions for an (n, k) type of model is not trivial and beyond the scope of this research. Based on this theoretical model, we focus on the empirics, outlining the variables of interest for the creation of a spatial cross-section in the next section.

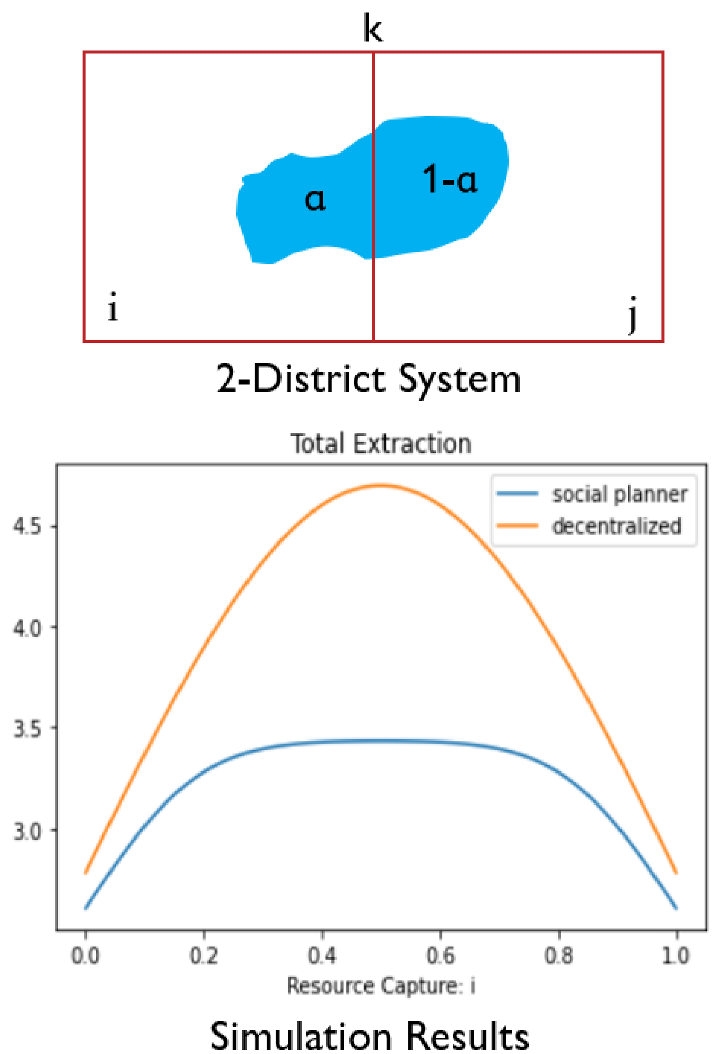


Figure 12: Illustration and Simulation results

7 Construction of the Spatial Cross-Section

In order to develop a cross-section of data that would allow us to answer the research questions, we merge datasets on the the shapefile of 2017 scraped from BharatMaps[15], combining it with variations in India's Local Government Directory[16]. Next, we use the Central Groundwater Board's assessment of Groundwater draft statistics derived from the National Compilation on Dynamic Ground Water Resources of India (2017) [17]. The statistics reviewed include:

[15]https://bharatmaps.gov.in/map.aspx
[16]https://lgdirectory.gov.in/
[17]http://cgwb.gov.in/Dynamic-GW-Resources.html

1. *Net Annual Ground Water Availability* - The net annual ground water availability for future use is obtained by deducting the allocation for domestic use and current extraction for Irrigation and Industrial uses from the Annual extractable Ground Water Recharge. It is calculated separately for non-command areas and command areas.

2. *Annual Groundwater Draft for Irrigation and Agricultural, and Domestic Use* - These measurements are derived based on the Groundwater Extraction Committee recommendations for computing the ground water draft. The norm used for computing ground water draft is the unit draft. The unit draft can be computed by field studies. This method involves selecting representative abstraction structures of aquifers and calculating the discharge from that particular type of structure and collecting the information on how many hours of pumping is being done in various seasons and number of such days during each season based on random sampling. The Unit Draft during a particular season is then computed using the following equation:
   Unit Draft = Discharge (m^3hr) x Pumping (hrs) x No. of days

   This Unit draft is differentiated based on the use of the water in samples where the water is observed, namely for domestic use or agriculture and irrigation purposes.

3. *Annual Replenishable Groundwater Resource* - This is estimated as a proportion of draft that is withdrawn overtime from permeable aquifers that are capable of recharging through sub-surface flows and aggregate percolation.

4. *Total Natural Discharge* - It is measured as the volumetric flow rate of groundwater through an aquifer. Total groundwater discharge, as reported through a specified area for India is measured through an ensemble of techniques, primarily relying on the groundwater flow equation (also known as Darcy's Law).

There is a considerable geographical variation in the availability of groundwater as well as the administrative and political units that fragment an underlying groundwater resource across them. The capacity of regulating groundwater and the quantum of groundwater extraction is substantially different given factors such as the morphology of urban and rural areas, agricultural productivity, and the presence (or absence) of a surface water body within a unit. Given the innumerable points of simultaneous extraction of groundwater across different households in villages, towns, and cities, farms in rural areas, and pre-existing infrastructure of supply of water which lie within these political and administrative boundaries, we need to first define and understand extraction as a 'construct' and then relate it to a variable of measurement.

We cannot expect to account for each drop extracted from a groundwater reservoir, but we can estimate the extent of extraction through a carefully proposed quasi-experiment. We now consider each of the empirical questions in greater detail, along with the variety of data available to help answer them.

These CGWB statistics are zonally aggregated to the District levels (sums and averages) as measures of actual scarcity. For perceived scarcity, we rely on the Groundwater wells data scraped from 1997 to 2017 and available through supplementary data link listed in (Hora et al., 2019), reducing the temporal panel to a mean centered cross-section at 2017 for the *perceived* measurement of scarcity, where a higher frequency of value of 0 indicates an increasing perception of scarcity, which does not necessarily reflect the actual scarcity. The mean centering is necessary to eliminate the bias of observation, where dysfunctional wells may be counted as dry wells due to the nature of permeability of the underlying aquifer.

Next, we scrape the annual Area, Production, and Yield (APY) statistics for India at the sub-district level from the Ministry of Agriculture and Cooperation's Crop Production and Statistics Information System [18]. This data is cleaned and aggregated for all districts throughout the country. The next set of variables are derived from a combination of multiple datasets on features of rivers and reservoirs available on the WRIS portal of India[19]. we calculate the length of Rivers (in thousand sq.kms or mil. Ha) and Reservoir Areas larger than 10,000 Ha as a ratio of the total district area. It is important to note that the lengths of rivers are restricted to stream orders (or Strahler Numbers[20]) less than 4 (counted from upstream), meaning that the 4 highest orders of drainage by flow capacity are considered to be significant. Next, we clean and aggregate the financial investments for schemes on drinking water supply (Jal Jeevan Mission or *JJM*[21]), and irrigation (Pradhan Mantri Kisaan Sinchan Yojana or *PMKSY*[22]) to the district level. Given the complexity of the financial outlays for the schemes at the household level for JJM and local government level for PMKSY, we convert it into a simple ratio of public water supply achieved at the district level (aggregated from habitation level), and financial outlay aggregated at the District Level. Next, we use data from the Pradhan Mantri Gram Sadak Yojana or *PMGSY*[23] released under the Open Government Data license for information on Count of Habitations per District and Marginal Population per District.

[18] https://www.aps.dac.gov.in/APY/Public_Report1.aspx

[19] https://indiawris.gov.in/wris/#/Geoviewer

[20] Strahler number or Horton–Strahler number of a mathematical tree is a numerical measure of its branching complexity. These numbers were first developed in hydrology, as a way of measuring the complexity of rivers and streams

[21] https://ejalshakti.gov.in/jjmreport/JJMIndia.aspx

[22] https://pmksy.gov.in/dashboardpage.aspx

[23] http://geosadak-pmgsy.nic.in/OpenData

Finally, we calculate fragmentation indices shortlisted on the basis of a comprehensive literature review of metrics of two-way spatial association and fragmentation in Locational Economics and Regional Science. A simplistic method of assessing the proportion of groundwater available can be through the proportion of areas. However, this measure is inaccurate because extraction of groundwater per unit area is actually determined by the activities on a unit rather than its extent - that is, a spatially extensive process that needs to be reduced to a spatially intensive measure. This has been widely studied in the context of industrial concentrations as well as in literature on economic and locational theory. The most common measures that applied are the "Gini Coefficient, the Herfindahl-Hirschman index (HHI) (Laine, 1995), the Maurel-Sedillot (MS) index (Maurel and Sédillot, 1999), and more recently the decomposable Theil index (TI)(Bickenbach et al., 2013)" Van Egeraat et al., 2018, p. 2. Other indices include the Krugman index and the Januszewski's coefficient, reviewed for previously conducted analysis (Bhogale and Khedgikar, 2022). Overall, in situations of resource fragmentation, the inverse of the Herfindahl-Hirschman index, also called Generality (Jaffe and Trajtenberg, 2002; Hall, 2005) has generally proven to be an effective measure. Summary statistics for the calculated indices and other variables in the cross section are shown in Table 1 below:

Table 1: Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Annual GWD Draft (Total)	663	36850	43890	0	6428	53558	374631
Annual GWD Draft (Ag. & Irrigation)	663	32912	42160	0	3669	47926	368502
Annual GWD Draft (Domestic)	663	3938	3847	0	1442	5196	34021
****Perceived Scarcity (Dry Wells '97 to '17)***	663	0.49	0.5	0	0	1	1
Total Production ('17)	663	0.11	0.23	0	0.0031	0.11	2.5
Total Annual Yeild('17)	663	6.5	5.7	0	1.2	10	32
Percentage of Urban Area	663	11	16	0	3.7	12	100
Length of River in District (S.O $\leq$ 4)	663	0.91	0.95	0	0.27	1.3	13
Reservoir Area per District	663	0.43	1.1	0	0	0.39	15
Irrigation Investments (mil. RS)	663	0.95	1.7	0	0.14	1	16
Ratio of Public Water Supply	663	0.39	0.3	0	0.13	0.62	1
Count of Habitations per District	663	994	821	1	439	1383	8698
Marginal Population Ratio per Dist.	663	0.38	0.26	0.0014	0.21	0.47	1
Count of River Streams (S.O $\leq$ 4)	663	1753	2470	0	133	2396	35535
Inverse HHI Index	663	0.4	0.24	0	0.18	0.58	0.95
Weighted Gini Coefficient	663	0.41	0.21	0	0.29	0.56	0.82
Theil's L Index	663	-0.18	0.1	-0.36	-0.26	-0.11	0
Krugman Coefficient	663	0.71	0.38	0	0.49	0.97	1.6
Januszewski Coefficient	663	0.68	0.17	0.3	0.56	0.78	1

8 ESDA

8.1 Cross-Section Variables

We discuss the variables outlined in Table 1 further in this section in an attempt to offer some insights that underscore the rationale, importance of the research question, and point out to evidence of spatially explicit mismatch between water consumption investments in agriculture, and consequent development.

The implications of these observations are consistent with our theoretical understanding of 'splillover' effects as outlined in the model, but more importantly they reinforce the need for a way to reduce or inherit the interaction effect of lattices across 2-dimensions into a single dimension in order to more accurately represent the extent of spillover. By creating a new method of capturing this contiguity within the weights specification, we can represent the true extent of the fragmentation of sub-basins by districts, or a comparable unit of administrative region with a unit of sub-unit of the Basin (Eg. Basins, Watersheds, or Microwatersheds).

8.2 Eliminating the Bias of Dysfunctional Wells

Eliminating the bias in observations due to dysfunctional wells is not trivial, especially in this cross-sectional setting. However, we can infer the perception of scarcity over time based on the prevalence of dry wells within a given region (in this case, districts). Given that well observations from CGWB are available as a time-series (4 observations per year are Pre-Monsoon, Monsoon, Post-Monsoon *Kharif* and Post-Monsoon *Rabi*[24]), if the number of dysfunctional wells are constant over time and it is the number of dry wells that are changing, it is easy to find the difference $x_t - x_{dys}$, where x_t is the number of wells at any cross-section instance, and x_{dys} are the total number of dysfunctional wells. However, in this case, the latter is not known to us. To overcome this problem, we first aggregate zeroes across each season's x_t and center the values around the global mean (i.e $x_t - \bar{x}_t$). Next we divide this value by the total number of wells in each district w_t^{tot}, to get a measure that approximates the *perceived* scarcity, and is approximately normally distributed around the mean. The effect of this transformation on the density is seen in Figure 13.

This transformation also allows us to treat *perception of scarcity* as a latent variable (specifically for Probit models) while retaining the balance in the dataset, as seen in the

[24]Kharif and Rabi are the two agricultural seasons of India. The two post-monsoon observations refer to the southwest and northeast monsoons in India which are the two monsoon seasons each leading Kharif and Rabi sowing.

parallel Coordinate plot at the Bottom of Figure 14, which is a critical necessity for effectiveness of Probit Models. The LISA clusters in Figure 14 also provide validation that the Actual and Perceived scarcity clusters are actually different.

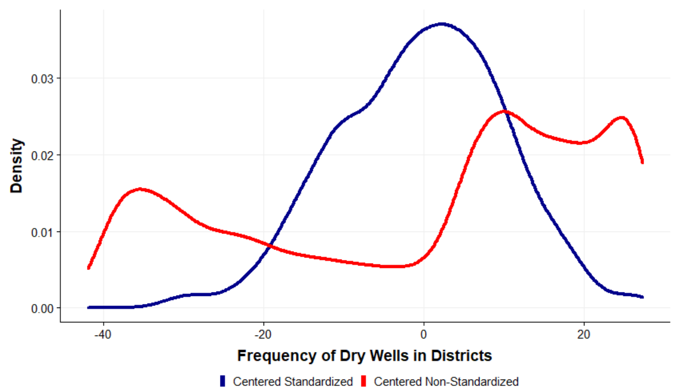


Figure 13: Effect of Standardization on density of dry wells

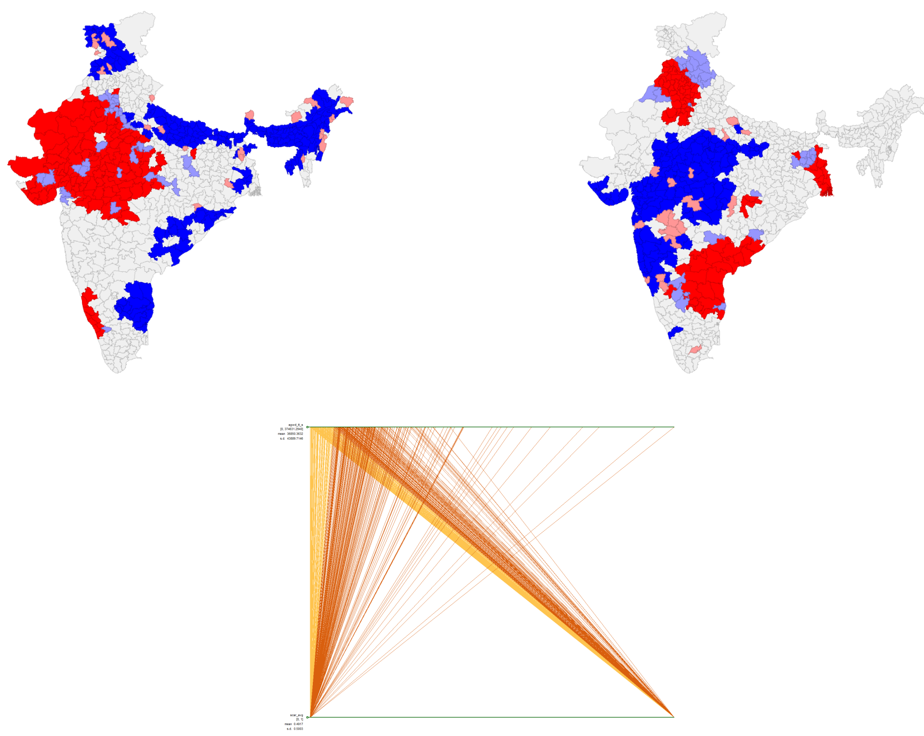


Figure 14: (Top-Left):Local Moran's I LISA for Actual Scarcity of Water (Total Annual GWD Draft 2017), (Top-Right):Perceived Scarcity of Water(Mean Centered Frequency of Dry Wells), (Botton:) Parallel Coordinate Plot between Actual Scarcity(2 Quantiles: High and Low) versus Perceived Scarcity(Binary)

9 RelWeights: Basic Principles and Advantages

The idea for RelWeights comes from two important ideas discussed for 1-layer examples in Anselin, 1988, pp. 18-21. First, with regards to configurations of lattice structures (for centroids of polygons) or tessellations over space (for points) that are regular, we note that "the determination of contiguity is not unique." This allows creation of the family of contiguity weights with Queen, Rook and Bishop criterion. Secondly, as argued in (Anselin, 1980, 1982, 1984), we note that a judicious choice of spatial weights is important to relate it to the underlying spatial relationship being studied.

The RelWeights grows out of these ideas in an attempt to develop contiguity that extend non-uniformly along different directions, based on a overlayed contiguity structure. For example, in the illustration below, we see that the relationship between Layer 1 (A,B,C) and overlaying Layer 2 (W,X,Y,Z) is a many-to-many relationship. In functional programming, this is also called *HasAndBelongsToMany* relationship.

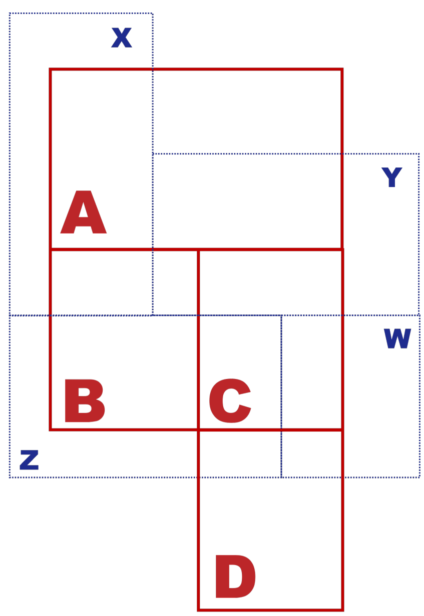


Figure 15: Interaction between two layers

Note that in the above lattice structure in Figure 15, D is not related to A if a first order contiguity is considered for *Layer 1* (Red). However, if we consider an inherited relationship based on first order contiguity of *Layer 2*(Blue), we see that A and D are indeed connected via Y and W. This is really the essence of RelWeights. If we construct a binary relationship matrix, we can naturally see that matrices based on the RelWeights contiguity would be much denser than say the Queen contiguity and as a result, the contiguity histogram has higher variance.

One of the biggest advantages of this kind of this formulation of the Weights matrix specification is that it allows us to model irregular orders of contiguity that may be borrowed through lattice structures (or regimes) that ***do not coincide*** *with the original layer's queen contiguity. This is very important because currently, all regime specifications and models in spatial econometrics are defined as discrete units based on a single layer which means any form of specification for two-layers needs to be evaluated at the unit size and contiguity of* ***Tr(W)*** *one of the layers, and values need to be imputed at .*

The RelWeights specification allows us to capture the neighborhood relationship between districts and the overlayed (but underlying!) sub-basins. I primarily construct these as row-standardized weights, capturing the effect of fragmentation through the regressor. Figure 16 shows the effect it has on the neighborhood linkage density, where we see an increase in the variance and the mean, as well as skewness is added to the linkage density. Note that the highlighted bins are for n=5 for Queen (Left), and the corresponding denser linkages distributed across in the RelWeights specification (Right). Figure 17 shows the effects of the RelWeight contiguity, and how it enables us to capture *irregular lags* across space.

The mathematical intuition behind RelWeights is quite straightforward, which allows us to experiment with it quite easily. Let **W** be an $n \times k$ matrix derived from the intersection of any two layers with the same extent α & β (Note that α & β here are just two layers similar to the illustrative figure). The problem is that we need to inherit the relationship between the lattice structure of α into β, or α into β, with n polygons of α an k polygons of β. In this case when polygons (or points) completely overlap with each other or are perfect subsets of each other, the relationship reduces to Block Weights[25]. However, when they do then:

R = WW' - [WW']I then represents this relationship as a weights matrix that is $n \times n$, and row-standardized. **Tr([WW']I)** represents the total fragments created through the intersections of the layers. The 'excess' inherited relationship can also be inferred through **R - W***, where **W*** represents the standard $n \times n$ row-standardized Queen contiguity matrix. In order to create such "first-order" RelWeights, we define a *tab2relweights* function for this in R which we outline in the form of a pseudo-code below:

[25] Block weights assign weights to groups or clusters of spatial units based on their proximity or contiguity. These weights represent the strength of the spatial relationship between the blocks. The purpose of block weights is to capture the spatial dependence and connectivity among groups of spatial units rather than individual observations.

```
require(sf)
require(spdep)
tab2relweights <- function(rel_layer,inh_layer) {
	#Generate IDs
	rel_layer$id_x <- seq.int(nrow(rel_layer))
	inh_layer$id_y <- seq.int(nrow(inh_layer))

	#Intersection using st_intersect
	int <- st_intersect(rel_layer,inh_layer)

	#Order intersection based on rel_layer
	int <- int[order(as.numeric(int$id_x)),]

	#Create W and W`
	table_weights <- table(int$id_x,int$id_y)
	table_weights_t <- t(table_weights)

	#Create R = WW` - WW`I
	WW_t <- table_weights%*%table_weights_t
	WW_tI <- diag(diag(WW_t))
	RelWeights <- WW_t - WW_tI

	#Use pmin for converting cell-values > 1 to 1
	RW4spdep <- pmin(RelWeights, 1)
	gal <- mat2listw(final4spdep, style = "W")

	#Return
	return(gal)
}
#Writing GAL file
write.nb.gal(gal$neighbours, "RelWeights.gal", oldstyle=TRUE, shpfile = tempfile)
```

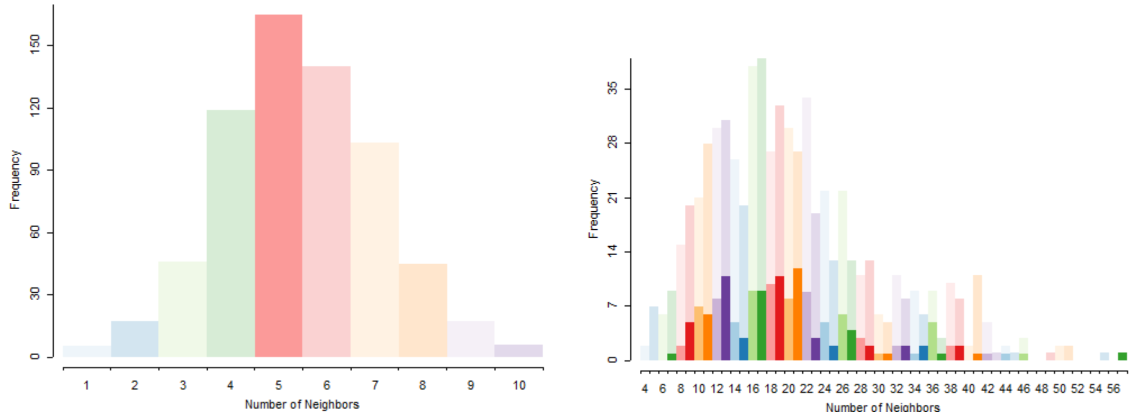


Figure 16: Histogram of Neighbor Connectivity for Queen(Left) and RelWeights(Right) contiguity

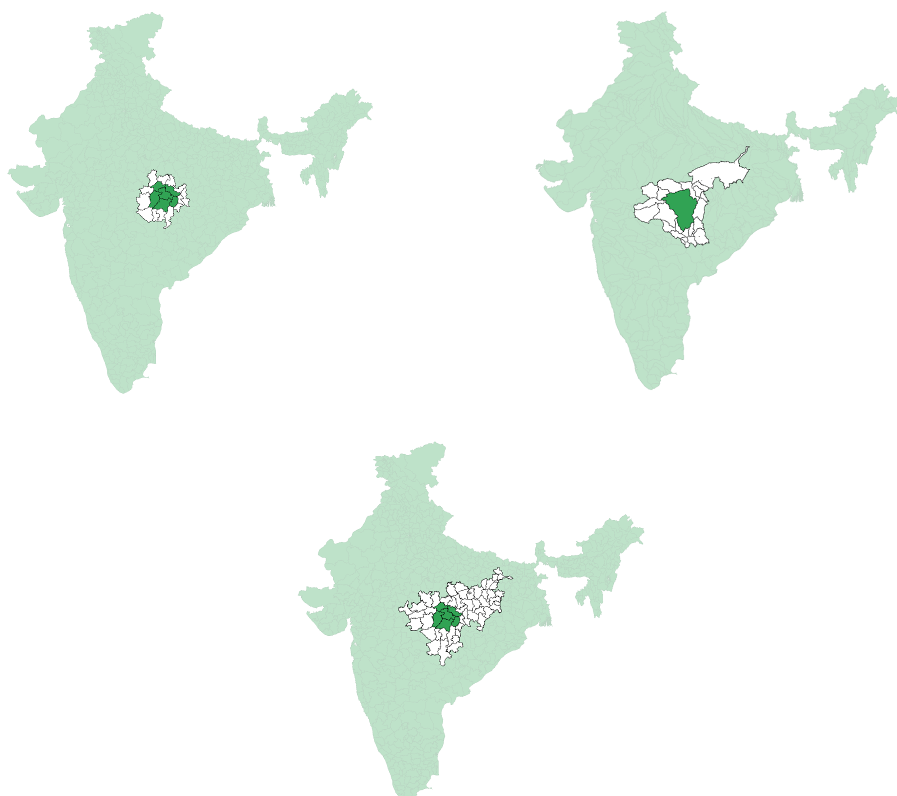


Figure 17: (Top-Left): Queen Contiguity for a district and neighboring units, (Top-Right): Queen Contiguity for a Sub-Basin, (Bottom): Relweights for District with Sub-Basins as inherited layer

The effect of RelWeights specification can be seen by cross-linking queen-contiguity with the RelWeights and cross-linking the connectivity Maps in GeoDa. Figure 17 is an illustration of this for our current use case, where the regular orders are captured using the Queen Weights and irregular orders are captured through RelWeights. Notice how the shape of the queen specification for the Sub-Basins is inherited in the RelWeights while simultaneously extending the contiguity order irregularly on all sides based on the structure of the sub-basins. When we compare the connectivity maps generated by the Queen's Weights and RelWeights, we see that the inheritence of the graph of the overlayed layer of sub-basins on underlying layers of districts results in a much denser graph. This is shown below in Figure 18.

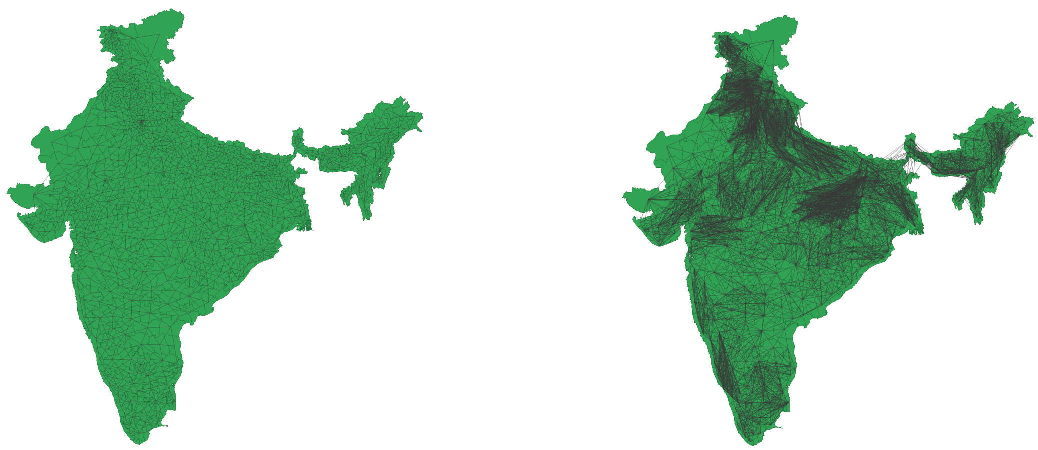


Figure 18: Connectivity Graph based on Queen Contiguity (Left) and RelWeights Contiguity (Right)

9.1 Effect of RelWeights on calculation of Local Moran's I

When calculating LISA statistics based on Local Moran's I, the order of contiguity determines the spatial weights used to quantify the strength of spatial relationships between neighboring units. The spatial weights matrix assigns weights to each unit's neighbors based on their proximity. The higher the order of contiguity, the more neighbors are considered, leading to a broader definition of spatial relationships.

Formally, the Local Moran's I is given as:

$$I_i = \frac{\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S^2 \sum_{j=1}^{n} w_{ij}}$$

where I_i is the Local Moran's I value for observation i, w_{ij} represents the spatial weight between observations i and j, x_i and x_j are the values of the variable of interest for observations i and j respectively, $\bar{x}$ is the mean value of the variable of interest across all observations, and S^2 is the variance of the variable of interest. We can now use the RelWeights in comparison to the Queen Weights one any one of the variables to test the comparative effect of the new boundaries.

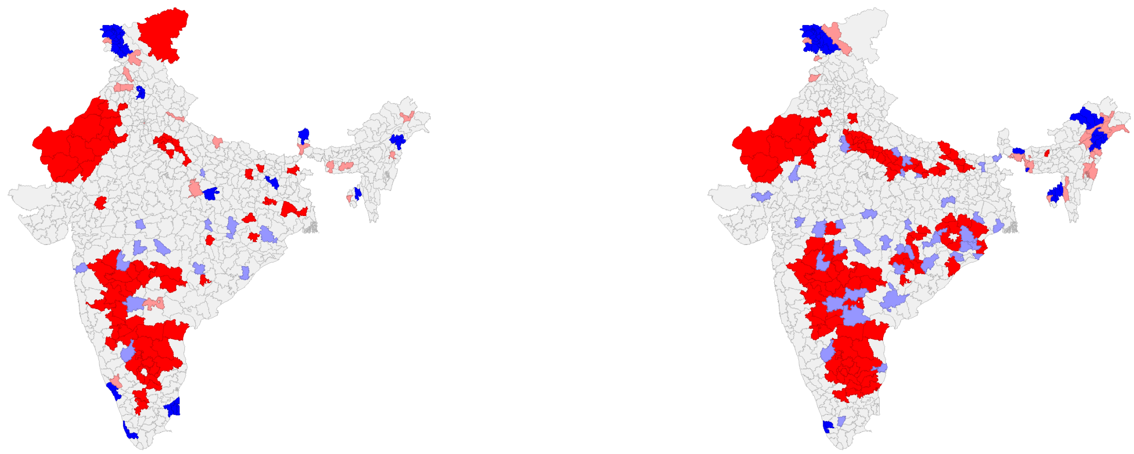


Figure 19: LISA Clusters based on Queen Contiguity (Left) and RelWeights Contiguity (Right)

The LISA clusters on Moran's I reveal a very interesting pattern. We see that the clustering around the 'true' extent of inherited contiguity (RelWeights) are much larger that the simple first order lags (Queen Weights). Second, we observe that the clustering is much more broader - implying that the variance in the Moran's I plot is much higher. Naturally, this effect is also seen in the increase of Low-High and High-Low clusters, with an increase in their significance (only the cluster maps are shown here as this is a Univariate Local Moran's I LISA).

It is important to note that by increasing the order of the contiguity,

Figure 20: Moran's I Plot for Clusters based on Queen Contiguity (Left) and RelWeights Contiguity (Right)

10 Specifications

Our point of departure for outlining the specifications is based on the taxonomies of local and global externalities as outlined in (Anselin, 2003). These taxonomies determine the different models seen in Spatial Econometric Literature in order to study the local and global effects of spillovers.

	Local Externalities	*Global Externalities*
u	$y = X\beta + u + \gamma W u$	$y = \lambda W y + X\beta - \lambda W X \beta + u$
X	$y = X\beta + W X \rho + u$	$y = \rho W y + X\beta + u - \rho W u$
Both	$y = X\beta + W X \rho + u + \gamma W u$	$y = (\rho + \lambda) W y - \rho\lambda W^2 y + X\beta - \lambda W X \beta + u - \rho W u$
	$y = X\beta + W X \rho + u + \rho W u$	$y = \rho W y + X\beta + u$

Figure 21: Taxonomy of Structural Forms

The independent variables for our empirics are the measures of *actual* and *perceived* scarcity, and the comparisons of effects of the Queen Contiguity and the RelWeights Contiguity. First, for *actual scarcity*, the dependent variables are:

1. Annual GWD Draft (Total Draft)
2. Annual GWD Draft (Ag & Irrigation Draft)
3. Annual GWD Draft (Domestic Draft)

Next, for *perceived scarcity* We treat the previously calculated Centered Standardized measure of Frequency of Dry wells as a latent variable (which is unobserved) and the perception of scarcity being high if value is > 0 and low if value is < 0. More formally,

$$Y_i = \begin{cases} 1 & \text{if} \frac{x_t - \bar{x}_t}{w_t^{tot}} > 0 \\ 0 & \text{if} \frac{x_t - \bar{x}_t}{w_t^{tot}} < 0 \end{cases}$$

where $\frac{x_t - \bar{x}_t}{w_{tot}}$ is the unobserved perception of scarcity (latent variable) derived from the frequency of wells. $\bar{x}_t$ is the global mean of dry wells per district (43) and x_t is frequency of dry wells in each t and w_t^{tot} is the total number of wells in the district t.

We consider 6 major models for the analysis - Non-Spatial OLS regression, Spatial Autoregressive Model, Spatial Error Model, Spatial Durbin Model (with Independent Lags) and the Spatial Probit Model (Spatial Lag Probit and Spatial Error Probit).

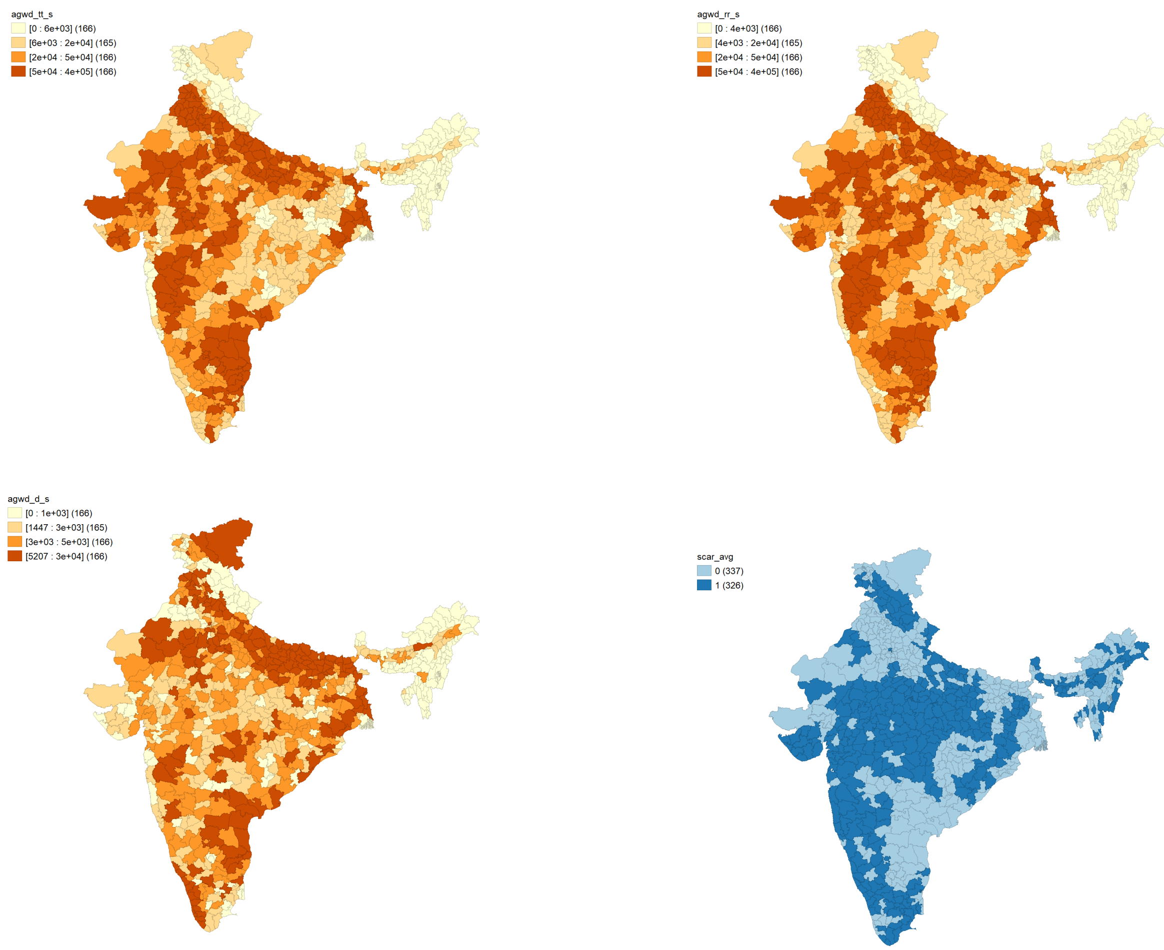


Figure 22: Comparison of Dependent Variables: Quantiles for actual scarcity variables (Top Left: Total Draft), (Top Right: Ag & Irrigation Draft), (Bottom Left: Domestic Draft), and Unique values for perceived scarcity variable (Bottom Right)

1. OLS Regression - The OLS regression is used to model relationship between a dependent variable and one or more independent variables by minimizing the sum of squared differences between observed and predicted values. In its Matrix form, it is given by:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon \tag{3}$$

where Y is the vector of dependent variables, X is the matrix of independent variables (including a column of ones for the intercept term), β is the vector of coefficients, ε is the vector of error terms.

2. *Spatial Autoregressive Model* - In a SAR model, the dependent variable of interest is regressed on both its own lagged values and the lagged values of neighboring observations. The inclusion of spatial lag terms captures the influence of neighboring observations on the

target variable, acknowledging spatial autocorrelation and spatial dependence in the data. The basic form of a SAR model can be written as:

$$\mathbf{Y} = \rho\mathbf{W}\mathbf{Y} + \mathbf{X}\beta + \epsilon \tag{4}$$

Where $\mathbf{Y}$ is a vector of the dependent variable, ρ is the spatial auto-regressive coefficient, representing the strength of spatial dependence, Wy is the spatial lag term, computed as the weighted average of neighboring $\mathbf{Y}$ values, $\mathbf{X}$ is a matrix of independent variables, β is a vector of coefficients for the independent variables and ϵ is the error term.

3. *Spatial Error Model* - The SEM is a spatial econometric model accounts for spatial dependence in the error term of a regression model. It is used when the residuals of a traditional regression model exhibit spatial autocorrelation, indicating that nearby observations are more likely to have similar error terms. The basic form of SEM model is given by:

$$\mathbf{Y} = \mathbf{X}\beta + \lambda\mathbf{W}\varepsilon + \varepsilon$$

where Y is the dependent variable, $\mathbf{X}$ is a matrix of independent variables (exogenous variables), β is a vector of coefficients for the independent variables, $\mathbf{W}$ is the spatial weights matrix, ε is the error term, λ is the spatial autoregressive coefficient for the error term.

4.*Spatial Durbin Model* - The Spatial Durbin Model is a spatial econometric model that extends the traditional linear regression framework by incorporating spatial dependence in both the dependent variable and the independent variables. It allows for analyzing the impact of both spatial lagged dependent variables and spatial lagged independent variables on the dependent variable of interest. The Spatial Durbin model is given by:

$$\mathbf{Y} = \rho\mathbf{W}\mathbf{Y} + \lambda\mathbf{W}\mathbf{X} + \mathbf{X}\beta + \varepsilon$$

where $\mathbf{Y}$ is the dependent variable, ρ is the spatial autoregressive coefficient for the dependent variable, $\mathbf{W}$ is the spatial weights matrix, λ is the spatial autoregressive coefficient for the independent variables, $\mathbf{X}$ is a matrix of independent variables (exogenous variables), β is a vector of coefficients for the independent variables, ε is the error term.

5. *Spatial Probit Regression* - The Spatial Probit model is used to analyze binary or categorical data with spatial dependence. It estimates the probability of an event occurring based on independent variables while accounting for spatial autocorrelation. Types used in this analysis are Spatial Autoregressive Probit, Spatial Error Probit. In it's most general form, the Spatial Probit Model can be written as:

$$Y_i = \begin{cases} 1 & \text{if } \Phi(\rho WY + X_i\beta + \varepsilon_i) > 0.5 \\ 0 & \text{otherwise} \end{cases}$$

where Y_i is the binary dependent variable for observation i, Φ is the cumulative distribution function of the standard normal distribution, ρ is the spatial autoregressive coefficient for the dependent variable, W is the spatial weights matrix, specifying the spatial relationships between observations, Y is a vector of the dependent variable, X_i represents the values of the independent variables for observation i, β is a vector of coefficients for the independent variables, and ε_i is the error term for observation i.

Depending upon the nature of the spatial process, further modifications can be made such as the *Autoregressive* Spatial Probit Model which is given as:

$$Y_i = \begin{cases} 1 & \text{if } \Phi(\rho WY + \lambda WX + X_i\beta + \varepsilon_i) > 0.5 \\ 0 & \text{otherwise} \end{cases}$$

where Y_i is the binary dependent variable for observation i, Φ is the cumulative distribution function of the standard normal distribution, ρ is the spatial autoregressive coefficient for the dependent variable, W is the spatial weights matrix, specifying the spatial relationships between observations, Y is a vector of the dependent variable, λ is the spatial autoregressive coefficient for the independent variables, X is a matrix of independent variables (exogenous variables), β is a vector of coefficients for the independent variables, X_i represents the values of the independent variables for observation i, and ε_i is the error term for observation i.

Similarly, the Error Spatial Probit Model is given as:

$$Y_i = \begin{cases} 1 & \text{if } \Phi(X_i\beta + \lambda W\varepsilon + \varepsilon_i) > 0.5 \\ 0 & \text{otherwise} \end{cases}$$

where Y_i is the binary dependent variable for observation i, Φ is the cumulative distribution function of the standard normal distribution, X_i represents the values of the independent variables for observation i, β is a vector of coefficients for the independent variables, λ is the spatial autoregressive coefficient for the error term, W is the spatial weights matrix, specifying the spatial relationships between observations, ε is the spatially correlated error term, ε_i is the individual-level error term for observation i.

For determining the independent variables, we use a stepwise strategy which is outlined along with the explicit specification in the next section.

11 Discussion

As seen in Table 2, we begin by performing stepwise backward and forward specifications for Actual Scarcity across all the three dependent variables. We observe that as expected, Total production, Total Yeild, length of river, ratio of public water supply, marginal population, and the I-HHI are overall the signficant determinants of actual scarcity. This provides us a rough guidance for finetuning the independent variables we will use to evaluate the spatial econometric model. Overall, we find strong signficance for all the critical variables, giving some preliminary validation about the setup and the determinants of over-extraction of water, and consequently water scarcity.

Next, we conduct a series of Lagrange Multiplier Tests using both the Queen and RelWeights contiguity matrices. LM (Lagrange Multiplier) tests are commonly used in spatial econometrics to test for the presence of spatial dependence or misspecification in regression models. LM tests help assess the adequacy of the model and the presence of spatial autocorrelation in the residuals. The LM tests help us determine the best-fit models while adjusting the scale of the variables based on their distribution. The independent variables, i.e β (and spatial lags of variables for Spatial Durbin model) used in LM Tests are based on the following specifications:

- *sqrt(Draft)* (Total, Ag & Domestic)) = β_0 + β_1 Inverse HHI Index + β_2 Total Production (mil Ha) + β_3Urban Area (%) + β_4River Length (th sq.km) + β_5Reservoir Area (%) + β_6Irrigation Investment (mil Rs.) + β_7Drinking Water Coverage (%)

The results show that in all cases, we overwhelmingly reject the Null of No Spatial Dependence. However, we see that for the same dependent variable, the use of RelWeights substantially increases the test statistic as well as the level of signficance in Error Models, indicative of error models capturing the spillover of local externalities better than the lag models. Even in the case of Robust tests, we find that there is atleast a 5x increase in the LM test statistic against the same specification and same dependent variables.

When comparing between models, the SARMA model performs the best, however, since the implementation of SARMA model is not trvial in R - we go with the next best models based on the Robust test statistics. We find a very interesting observation here. In the case of the Queen Weights, the Robust LM Lag Tests perform overwhelmingly better than the Robust LM Error Tests. However, in the case of RelWeights, we find that the test statistics are much higher in the case of the Robust LM Error Tests than the Robust LM Lag Tests.

From Tables 5 to 7, we look at the SAR, SEM and Spatial Durbin specifications independently for each of the dependent variables of actual scarcity. While all the coefficients are mostly significant, the most important difference is seen between the Total, Ag & Irrigation versus Domenstic drafts. For the Total and Ag & Irrigation drafts, we find that the SAM model with Queen weights performs much better in terms of AIC, whereas for the domestic draft we see that the error model with the Queen weights performs much better.

Lastly, we look at Spatial Probit Models. We find that only a partial set of variables that drive the *actual* scarcity of water also significantly predict the *perception* of scarcity of water, as inferred from our binary variable. Only index of fragmentation, the ratio of public water supply, and the length of river are highly significant variables. We see that there is a very good accuracy of more than 65% across all the models, and the Spatial Lag Probit models perform the best in terms of prediction.

For all the models, we do not see any significant pattern of clustering in the residuals. We do not report them here but include them in our supplementary documentation, along with the effects of clustering for counterfactual districts based on the newly created RelWeights.

12 Conclusion

Through this work, we have successfully managed to formulate a model which effectively explains drivers of groundwater with extremely high significance. We have also shown that the drivers of perceived scarcity are different from the drivers of actual scarcity of water. By creating a new class of spatial weights matrix, we have generalized the concept of popularly used Block Weights, when the 'blocks', in our case Sub-Basins, are not perfect super-sets of the units of analysis, i.e the Districts. The contiguity structure generated by RelWeights shows us how just the local contiguity assigned through Queen Weights do not capture the expansive extent of sub-surface flows across sub-basins. While the effect of increasing orders of contiguity irregularly on each side, which may be seen as introducing heteroskedasticity in the model, we see that the effect is minimal and in cases of Durbin specifications yeilds a lower AIC.

This method of introducing irregular spatial lags helps in identifying additional significant spatially constrained clusters which can be used further to generate counterfactual aggregations, which may be the first step towards an algorithmic approach to delineation of water districts for effective management. Studying these counterfactual districts and their resultant groundwater fragmentation can provide valuable information for effective and sustainable groundwater management, ecosystem conservation, and equitable development planning. It enables policymakers, researchers, and communities to make informed decisions that are crucial for the long-term well-being of both people and the environment.

13 Tables

Table 2: OLS Stepwise (Backward &Forward) Specifications for *Actual Scarcity*

	Dependent variable:		
	sqrt(Total Draft)	sqrt(Ag & Irrigation Draft)	sqrt(Domestic Draft)
	(1)	(2)	(3)
Total Production (mil Tonnes)	0.013***	0.013***	0.001***
	(0.001)	(0.001)	(0.0004)
Total Yeild	3.400***	3.526***	0.799***
	(0.588)	(0.597)	(0.183)
Urban Area (%)			0.243***
			(0.062)
River Length (th sq.km)	45.160***	45.090***	13.121***
	(9.013)	(9.163)	(2.755)
Irrigation Investment (mil Rs.)	2.085***	1.945***	0.559***
	(0.529)	(0.537)	(0.160)
Drinking Water Coverage (%)	44.729***	50.461***	
	(10.524)	(10.699)	
Count of Habitations per Dist.	0.020***	0.019***	0.009***
	(0.004)	(0.004)	(0.001)
Marginal Population Ratio per Dist.	−81.548***	−74.561***	−27.946***
	(12.470)	(12.678)	(3.759)
Length of River in District (S.O $\leq$ 4)	−0.018***	−0.018***	−0.005***
	(0.003)	(0.003)	(0.001)
Inverse HHI Index	49.765**	52.209**	
	(21.335)	(21.691)	
Theil's L Index	−71.161	−76.593	
	(46.713)	(47.493)	
Weighted Gini Coefficient			−7.233
			(4.883)
Januszewski Coefficient			−19.388***
			(6.353)
Constant	60.446***	40.137***	56.689***
	(11.585)	(11.779)	(6.474)
Observations	663	663	663
R^2	0.473	0.467	0.364
Adjusted R^2	0.464	0.459	0.354
Residual Std. Error (df = 652)	77.407	78.698	23.833
F Statistic (df = 10; 652)	58.419***	57.153***	37.245***

*p<0.1; **p<0.05; ***p<0.01

Table 3: LM Tests with Queen Weights

	Lagrange multiplier diagnostics for spatial dependence		
	Total Draft	Ag & Irrigation Draft	Domestic Draft
LMerr	572.11	602.35	334.33
	(1, p < 2.2e-16)	(1, p < 2.2e-16)	(1, p < 2.2e-16)
RLMerr	23.167	25.755	10.147
	(1, p = 1.486e-06)	(1, p = 3.876e-07)	(1, p < 0.001445)
LMlag	621.29	651.7	335.99
	(1,p < 2.2e-16)	(1,p < 2.2e-16)	(1,p < 2.2e-16)
RLMlag	72.345	75.102	11.806
	(1, p < 2.2e-16)	(1, p < 2.2e-16)	(1, p < 0.0005903)
SARMA	644.46	677.45	346.13
	(2, p < 2.2e-16)	(2, p < 2.2e-16)	(2, p < 2.2e-16)

Note: The specification for the LM tests are outlined in section 11

Table 4: LM Tests with RelWeights

	Lagrange multiplier diagnostics for spatial dependence		
	Total Draft	Ag & Irrigation Draft	Domestic Draft
LMerr	711.15	761.38	440.87
	(1, p < 2.2e-16)	(1, p < 2.2e-16)	(1, p < 2.2e-16)
RLMerr	140.11	148.38	54.194
	(1, p < 2.2e-16)	(1, p < 2.2e-16)	(1, p = 1.816e-13)
LMlag	633.43,	679.87	404.05
	(1,p < 2.2e-16)	(1,p < 2.2e-16)	(1,p < 2.2e-16)
RLMlag	62.399	66.872	17.38
	(1, p < 2.2e-16)	(1, p = 3.331e-16)	(1, p = 3.061e-05)
SARMA	773.54	828.25	458.24
	(2, p < 2.2e-16)	(2, p < 2.2e-16)	(2, p < 2.2e-16)

Note: The specification for the LM tests are outlined in section 11

Table 5: Comparison of Models and Weights for Total Draft

	Dependent variable:					
	sqrt(Total Draft)					
	Spatial Autoregressive		*Spatial Error*		*Spatial Durbin*	
	(1-Q)	(2-R)	(3-Q)	(4-R)	(5-Q)	(6-R)
Inverse HHI Index	61.913***	74.921***	54.979***	69.565***	54.776***	68.842***
	(9.145)	(11.782)	(9.844)	(12.211)	(9.746)	(12.258)
Total Production (mil Tonnes)	0.009***	0.014***	0.009***	0.016***	0.009***	0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Urban Area (%)	−0.168	−0.389**	−0.231	−0.487**	−0.166	−0.480**
	(0.134)	(0.174)	(0.190)	(0.209)	(0.194)	(0.213)
River Length (th sq.km)	5.015**	0.034	16.414***	3.401	16.943***	4.206
	(2.422)	(3.127)	(3.067)	(3.671)	(3.047)	(3.733)
Reservoir Area (%)	−2.271	−2.677	0.821	−1.734	0.616	−1.768
	(1.865)	(2.411)	(1.863)	(2.464)	(1.871)	(2.470)
Irrigation Investment (mil Rs.)	1.334***	1.453***	2.207***	1.352**	2.347***	1.335**
	(0.368)	(0.472)	(0.528)	(0.637)	(0.525)	(0.646)
Drinking Water Coverage (%)	20.749***	28.304***	18.107**	31.121***	19.369**	30.220***
	(7.135)	(9.198)	(8.376)	(10.583)	(8.281)	(10.634)
lag.Inverse HHI Index					−0.168	30.867
					(17.528)	(35.219)
lag.Total Production (mil Tonnes)					−0.001	−0.010***
					(0.002)	(0.003)
lag.Urban Area (%)					−0.104	0.409
					(0.279)	(0.414)
lag.River Length (th sq.km)					−25.226***	−11.895
					(4.861)	(9.548)
lag.Reservoir Area (%)					−2.722	−2.673
					(4.129)	(10.557)
lag.Irrigation Investment (mil Rs.)					−1.221*	−0.163
					(0.653)	(0.934)
lag.Drinking Water Coverage (%)					7.539	−4.083
					(12.111)	(20.598)
Constant	−20.793***	−19.874**	77.854***	97.227***	−6.854	−21.355
	(6.313)	(8.764)	(14.707)	(18.150)	(8.411)	(14.981)
Observations	663	663	663	663	663	663
Log Likelihood	−3,633.338	−3,767.548	−3,632.791	−3,764.645	−3,610.949	−3,759.444
σ^2	2,888.583	4,834.885	2,741.400	4,687.824	2,664.903	4,673.134
Akaike Inf. Crit.	7,286.677	7,555.096	7,285.583	7,549.290	7,255.898	7,552.889
Wald Test (df = 1)	1,018.455***	369.824***	1,441.765***	534.201***	898.999***	320.645***
LR Test (df = 1)	535.512***	267.092***	536.606***	272.898***	481.870***	195.789***

*p<0.1; **p<0.05; ***p<0.01

Table 6: Comparison of Models and Weights for Ag & Irrigation Draft

	Dependent variable:					
	sqrt(Ag & Irrigation Draft)					
	Spatial Autoregressive		*Spatial Error*		*Spatial Durbin*	
	(1-Q)	(2-R)	(3-Q)	(4-R)	(5-Q)	(6-R)
Inverse HHI Index	62.993***	76.944***	57.537***	71.732***	9.647**	11.994***
	(8.965)	(11.729)	(9.637)	(12.145)	(3.898)	(4.108)
Total Production (mil Tonnes)	0.009***	0.014***	0.009***	0.016***	0.001***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.0004)	(0.0004)
Urban Area (%)	−0.350***	−0.609***	−0.563***	−0.810***	0.302***	0.249***
	(0.131)	(0.173)	(0.186)	(0.208)	(0.078)	(0.071)
River Length (th sq.km)	3.813	−1.294	14.343***	1.535	6.913***	4.071***
	(2.371)	(3.112)	(3.006)	(3.653)	(1.219)	(1.251)
Reservoir Area (%)	−2.429	−2.864	0.511	−2.040	0.757	0.372
	(1.827)	(2.400)	(1.823)	(2.451)	(0.748)	(0.828)
Irrigation Investment (mil Rs.)	1.208***	1.314***	1.945***	1.027	0.979***	0.701***
	(0.359)	(0.468)	(0.518)	(0.634)	(0.210)	(0.216)
Drinking Water Coverage (%)	22.913***	31.561***	22.453***	34.726***	−1.795	0.176
	(7.006)	(9.156)	(8.200)	(10.529)	(3.309)	(3.564)
lag.Inverse HHI Index					10.820	21.651*
					(6.957)	(11.738)
lag.Total Production (mil Tonnes)					0.001	−0.002*
					(0.001)	(0.001)
lag.Urban Area (%)					−0.325***	−0.318**
					(0.112)	(0.139)
lag.River Length (th sq.km)					−9.703***	−7.962**
					(1.940)	(3.198)
lag.Reservoir Area (%)					0.044	0.784
					(1.651)	(3.539)
lag.Irrigation Investment (mil Rs.)					−0.579**	−0.383
					(0.260)	(0.311)
lag.Drinking Water Coverage (%)					4.830	2.141
					(4.759)	(6.830)
Constant	−22.600***	−22.293***	69.583***	90.095***	6.515*	0.386
	(6.139)	(8.582)	(15.090)	(18.620)	(3.474)	(5.107)
Observations	663	663	663	663	663	663
Log Likelihood	−3,621.775	−3,765.061	−3,621.201	−3,761.566	−2,980.660	−3,031.675
σ^2	2,773.389	4,791.548	2,628.543	4,636.566	426.319	525.053
Akaike Inf. Crit.	7,263.551	7,550.123	7,262.402	7,543.132	5,995.320	6,097.350
Wald Test (df = 1)	1,116.487***	391.178***	1,594.683***	573.821***	304.822***	212.609***
LR Test (df = 1)	569.131***	282.559***	570.280***	289.550***	239.130***	139.632***

*p<0.1; **p<0.05; ***p<0.01

Table 7: Comparison of Models and Weights for Domestic Draft

	Dependent variable:					
	sqrt(Domestic Draft)					
	Spatial Autoregressive		*Spatial Error*		*Spatial Durbin*	
	(1-Q)	(2-R)	(3-Q)	(4-R)	(5-Q)	(6-R)
Inverse HHI Index	14.376***	15.015***	11.149***	12.826***	57.497***	71.482***
	(3.607)	(3.935)	(3.926)	(4.094)	(9.559)	(12.193)
Total Production (mil Tonnes)	0.002***	0.002***	0.002***	0.003***	0.009***	0.016***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.001)	(0.001)
Urban Area (%)	0.149***	0.142**	0.262***	0.228***	−0.506***	−0.808***
	(0.054)	(0.059)	(0.072)	(0.070)	(0.190)	(0.212)
River Length (th sq.km)	2.690***	2.001*	6.036***	3.574***	14.772***	2.197
	(0.964)	(1.049)	(1.196)	(1.224)	(2.988)	(3.713)
Reservoir Area (%)	−0.058	−0.036	0.630	0.382	0.290	−2.111
	(0.741)	(0.808)	(0.749)	(0.826)	(1.835)	(2.457)
Irrigation Investment (mil Rs.)	0.497***	0.444***	0.925***	0.703***	2.080***	1.009
	(0.146)	(0.161)	(0.203)	(0.212)	(0.515)	(0.642)
Drinking Water Coverage (%)	−0.408	−1.258	−1.429	0.060	23.775***	33.910***
	(2.796)	(3.056)	(3.324)	(3.535)	(8.122)	(10.577)
lag.Inverse HHI Index					−6.406	27.373
					(17.181)	(35.068)
lag.Total Production (mil Tonnes)					−0.001	−0.010***
					(0.002)	(0.003)
lag.Urban Area (%)					0.211	0.790*
					(0.274)	(0.412)
lag.River Length (th sq.km)					−22.524***	−8.228
					(4.769)	(9.494)
lag.Reservoir Area (%)					−2.496	−3.746
					(4.051)	(10.502)
lag.Irrigation Investment (mil Rs.)					−1.106*	0.037
					(0.639)	(0.925)
lag.Drinking Water Coverage (%)					1.054	−10.583
					(11.875)	(20.473)
Constant	3.771	1.154	31.658***	35.211***	−8.054	−24.355
	(2.715)	(3.159)	(3.861)	(4.950)	(8.249)	(14.927)
Observations	663	663	663	663	663	663
Log Likelihood	−3,001.936	−3,041.022	−2,995.471	−3,036.386	−3,600.994	−3,756.283
σ^2	455.516	542.271	437.614	527.882	2,564.434	4,623.267
Akaike Inf. Crit.	6,023.873	6,102.043	6,010.942	6,092.771	7,235.987	7,546.567
Wald Test (df = 1)	332.817***	237.253***	403.434***	307.853***	1,011.192***	338.109***
LR Test (df = 1)	248.717***	170.546***	261.648***	179.818***	512.740***	204.952***

*p<0.1; **p<0.05; ***p<0.01

Table 8: *Perceived Scarcity Comparison for Spatial Lag Probit and Spatial Error Probit*

	Dependent variable:			
	Perceived Scarcity			
	Spatial Lag		Spatial Error	
	(1-Q)	(2-R)	(3-Q)	(4-R)
(Intercept)	0.2318	0.2244**	0.0935*	0.0673**
	(2.3547)	(3.8910)	(3.8005)	(4.3938)
Inverse HHI Index	-0.7939**	-0.7804***	-0.8385**	-0.7422**
	(7.5649)	(11.4235)	(6.0285)	(10.6852)
Total Production (mil Tonnes)	0.0000	0.0000	0.0000	0.0000**
	(0.1989)	(0.0997)	(1.0483)	(4.8921)
Urban Area (%)	-0.0066*	-0.0069**	-0.0140**	-0.0134***
	(2.7198)	(9.0754)	(8.2134)	(15.5922)
River Length (th sq.km)	0.0000***	0.0000***	0.0000**	0.0000***
	(14.9473)	(14.4039)	(9.9508)	(11.7333)
Reservoir Area (%)	0.1821**	0.1768**	0.1706**	0.1422**
	(5.5679)	(5.5498)	(8.3404)	(6.9824)
Irrigation Investment (mil Rs.)	0.0000*	0.0000*	0.0000**	0.0000**
	(3.4964)	(3.6589)	(7.8348)	(9.2247)
Drinking Water Coverage (%)	-0.6311***	-0.6497***	-0.8043**	-0.7187***
	(10.9294)	(15.2405)	(9.7767)	(14.4897)
Constant(ρ)	0.6376***	0.5613***	0.6282***	0.6543***
	(104.9145)	(57.5344)	(88.2613)	(58.3788)
Observations	663	663	663	663
No. of Covariates	8	8	8	8
Estimation Method	conditional	conditional	full-link	full-link
DGP	SAR	SAR	SEM	SEM
Order of iW	6	6	6	6
Accuracy	0.6983409	0.6862745	0.6576169	0.6576169
Sensitivity	0.7269939	0.6993865	0.6472393	0.6349693
Specificity	0.6706231	0.6735905	0.6676558	0.6795252
Positive Predicted Value	0.6810345	0.6745562	0.6532508	0.6571429
Negative Predicted Value	0.7174603	0.6984615	0.6617647	0.658046

*p<0.1; **p<0.05; ***p<0.01

References

Anselin, Luc, "Estimation of Spatial Autoregressive Structures; A Contribution to Spatial Econometrics," 1980.

— , "A note on small sample properties of estimators in a first-order spatial autoregressive model," *Environment and Planning A*, 1982, *14* (8), 1023–1030.

— , "Specification tests on the structure of interaction in spatial econometric models," in "Papers of the Regional Science Association," Vol. 54 Springer 1984, pp. 165–182.

— , *Spatial econometrics: methods and models*, Vol. 4, Springer Science & Business Media, 1988.

— , "Spatial externalities, spatial multipliers, and spatial econometrics," *International regional science review*, 2003, *26* (2), 153–166.

— , "Thirty years of spatial econometrics," *Papers in regional science*, 2010, *89* (1), 3–25.

Bhanja, Soumendra N, Abhijit Mukherjee, Matthew Rodell, Yoshihide Wada, Siddhartha Chattopadhyay, Isabella Velicogna, Kishore Pangaluru, and James S Famiglietti, "Groundwater rejuvenation in parts of India influenced by water-policy change implementation," *Scientific reports*, 2017, *7* (1), 1–7.

Bhogale, Shweta and Shamil Khedgikar, "Run on the Reservoir: Evidence on Administrative Competition for Groundwater in India," in "North Eastern Universities Development Conference 2022 - Working Paper" November 2022.

Bickenbach, Frank, Eckhardt Bode, and Christiane Krieger-Boden, "Closing the gap between absolute and relative measures of localization, concentration or specialization," *Papers in Regional Science*, 2013, *92* (3), 465–479.

Blakeslee, David, Ram Fishman, and Veena Srinivasan, "Way down in the hole: Adaptation to long-term water loss in rural India," *American Economic Review*, 2020, *110* (1), 200–224.

Bosch, Hilmer J, Joyeeta Gupta, and Hebe Verrest, "A water property right inventory of 60 countries," *Review of European, Comparative & International Environmental Law*, 2021, *30* (2), 263–274.

Chhatre, A, M Sinha, and S Khedgikar, "Promotion of cultivation and consumption of nutri-cereals in scheduled areas of Maharashtra," 2019. Publisher: Indian School of Business.

Chiarelli, Davide Danilo, Corrado Passera, Lorenzo Rosa, Kyle Frankel Davis, Paolo D'Odorico, and Maria Cristina Rulli, "The green and blue crop water requirement WATNEEDS model and its global gridded outputs," *Scientific data*, 2020, *7* (1), 1–9. Publisher: Nature Publishing Group.

Comptroller and Auditor General of India, "Report of the Comptroller and Auditor General of India on Ground Water Management and Regulation," Technical Report 9, CAG, Delhi 2021.

Dangar, Swarup, Akarsh Asoka, and Vimal Mishra, "Causes and implications of groundwater depletion in India: A review," *Journal of Hydrology*, 2021, *596*, 126103. Publisher: Elsevier.

Davis, Kyle Frankel, Davide Danilo Chiarelli, Maria Cristina Rulli, Ashwini

Chhatre, Brian Richter, Deepti Singh, and Ruth DeFries, "Alternative cereals can improve water use and nutrient supply in India," *Science advances*, 2018, *4* (7), eaao1108.

Edwards, Eric C and Todd Guilfoos, "The economics of groundwater governance institutions across the globe," *Applied Economic Perspectives and Policy*, 2021, *43* (4), 1571–1594.

Egeraat, Chris Van, Edgar Morgenroth, Rutger Kroes, Declan Curran, and Justin Gleeson, "A measure for identifying substantial geographic concentrations," *Papers in Regional Science*, 2018, *97* (2), 281–300.

Garduño, Héctor, Saleem Romani, Buba Sengupta, Albert Tuinhof, and Richard Davis, "India groundwater governance case study," 2011.

Garrick, Dustin, Lucia De Stefano, Winston Yu, Isabel Jorgensen, Erin O'Donnell, Laura Turley, Ismael Aguilar-Barajas, Xiaoping Dai, Renata de Souza Leão, Bharat Punjabi et al., "Rural water for thirsty cities: A systematic review of water reallocation from rural to urban regions," *Environmental Research Letters*, 2019, *14* (4), 043003.

Gleeson, Tom, Kevin M Befus, Scott Jasechko, Elco Luijendijk, and M Bayani Cardenas, "The global volume and distribution of modern groundwater," *Nature Geoscience*, 2016, *9* (2), 161–167. Publisher: Nature Publishing Group.

Hall, Bronwyn, "A note on the bias in herfindahl-type measures based on count data," *Revue d'économie industrielle*, 2005, *110* (1), 149–156.

Hardin, Garrett, "The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality.," *Science*, 1968, *162* (3859), 1243–1248.

He, Chunyang, Zhifeng Liu, Jianguo Wu, Xinhao Pan, Zihang Fang, Jingwei Li, and Brett A Bryan, "Future global urban water scarcity and potential solutions," *Nature Communications*, 2021, *12* (1), 4667.

Hofste, Rutger Willem, Samantha Kuzma, Sara Walker, Edwin H Sutanudjaja, Marc FP Bierkens, Marijn JM Kuijper, Marta Faneca Sanchez, Rens Van Beek, Yoshihide Wada, Sandra Galvis Rodríguez et al., "Aqueduct 3.0: Updated decision-relevant global water risk indicators," *World Resources Institute*, 2019, *784.*

Hora, Tejasvi, Veena Srinivasan, and Nandita B Basu, "The groundwater recovery paradox in South India," *Geophysical Research Letters*, 2019, *46* (16), 9602–9611. Publisher: Wiley Online Library.

Jaffe, Adam B and Manuel Trajtenberg, *Patents, citations, and innovations: A window on the knowledge economy*, MIT press, 2002.

Kulkarni, Himanshu, Mihir Shah, and PS Vijay Shankar, "Shaping the contours of groundwater governance in India," *Journal of Hydrology: Regional Studies*, 2015, *4*, 172–192.

Laine, Charles R, "The Herfindahl-Hirschman index: a concentration measure taking the consumer's point of view," *The Antitrust Bulletin*, 1995, *40* (2), 423–432.

LaVanchy, G Thomas, James K Adamson, and Michael W Kerwin, "Integrating groundwater for water security in Cape Town, South Africa," in "Global Groundwater," Elsevier, 2021, pp. 439–449.

Libecap, Gary D, "The tragedy of the commons: property rights and markets as solutions to resource and environmental problems," *Australian Journal of Agricultural and Resource Economics*, 2009, *53* (1), 129–144.

Maurel, Françoise and Béatrice Sédillot, "A measure of the geographic concentration in French manufacturing industries," *Regional Science and Urban Economics*, 1999, *29* (5), 575–604.

McPhearson, Timon, Elizabeth M Cook, Marta Berbes-Blazquez, Chingwen Cheng, Nancy B Grimm, Erik Andersson, Olga Barbosa, David G Chandler, Heejun Chang, Mikhail V Chester et al., "A social-ecological-technological systems framework for urban ecosystem services," *One Earth*, 2022, *5* (5), 505–518.

Morrison, Tiffany H, W Neil Adger, Arun Agrawal, Katrina Brown, Matthew J Hornsey, Terry P Hughes, Meha Jain, Maria Carmen Lemos, Lucy Holmes McHugh, Saffron O'Neill et al., "Radical interventions for climate-impacted systems," *Nature Climate Change*, 2022, pp. 1–7.

Mukherji, Aditi, "Sustainable groundwater management in India needs a water-energy-food nexus approach," *Applied Economic Perspectives and Policy*, 2022, *44* (1), 394–410.

_ **and Tushaar Shah**, "Groundwater socio-ecology and governance: a review of institutions and policies in selected countries," *Hydrogeology Journal*, 2005, *13*, 328–345.

Nishad, Shiv Narayan and Naresh Kumar, "Virtual water trade and its implications on water sustainability," *Water Supply*, 2022, *22* (2), 1704–1715.

Ostrom, Elinor, *Governing the commons: The evolution of institutions for collective action*, Cambridge university press, 1990.

_ , "The challenge of common-pool resources," *Environment: Science and Policy for Sustainable Development*, 2008, *50* (4), 8–21.

Paelinck, Jean H. P. and Leo H. Klaassen, *Spatial econometrics / Jean H. P. Paelinck, Leo H. Klaassen, with the assistance of J.-P. Ancot, A. C. P. Verster, Sj. Wagenaar*, Saxon House Farnborough, Eng, 1979.

Palit, Debajit and Kaushik Ranjan Bandyopadhyay, "Rural electricity access in India in retrospect: A critical rumination," *Energy Policy*, 2017, *109*, 109–120.

Rittel, Horst WJ and Melvin M Webber, "Dilemmas in a general theory of planning," *Policy sciences*, 1973, *4* (2), 155–169.

Rodell, Matthew, Isabella Velicogna, and James S Famiglietti, "Satellite-based estimates of groundwater depletion in India," *Nature*, 2009, *460* (7258), 999–1002. Publisher: Nature Publishing Group.

Rodriguez-Iturbe, Ignacio, Rachata Muneepeerakul, Enrico Bertuzzo, Simon A Levin, and Andrea Rinaldo, "River networks as ecological corridors: A complex systems perspective for integrating hydrologic, geomorphologic, and ecologic dynamics," *Water Resources Research*, 2009, *45* (1).

Savenije, Hubert HG, "Water scarcity indicators; the deception of the numbers," *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 2000, *25* (3), 199–204.

_ **and Pieter Van der Zaag**, "Integrated water resources management: Concepts and issues," *Physics and Chemistry of the Earth, Parts A/B/C*, 2008, *33* (5), 290–297.

Shah, Tushaar, "Ground water markets and Irrigation Development," 1993.

— , "The groundwater economy of South Asia: an assessment of size, significance and socio-ecological impacts," *The agricultural groundwater revolution: Opportunities and threats to development*, 2007, pp. 7–36.

— , "Sustainable Groundwater Management," in Meredith Giordano, Frank R Rijsberman, and R Maria Saleth, eds., *More Crop Per Drop: : Revisiting a Research Paradigm: Results and Synthesis of IWMI Research, 1996-2005*, IWA Publishing, 2007, pp. 84–117.

— , *Taming the anarchy: Groundwater governance in South Asia*, Routledge, 2010.

Singh, Sakshi, Shamil Khedgikar, and Arun CM Murali, "Agricultural crop substitution can bridge rural and urban water needs," June 2020.

Sivapalan, Murugesu, Hubert HG Savenije, Günter Blöschl et al., "Socio-hydrology: A new science of people and water," *Hydrol. Process*, 2012, *26* (8), 1270–1276.

— , M Konar, V Srinivasan, A Chhatre, A Wutich, CA Scott, JL Wescoat, and I Rodríguez-Iturbe, "Socio-hydrology: Use-inspired water sustainability science for the Anthropocene," *Earth's Future*, 2014, *2* (4), 225–230.

Subhadra, Bobban, "Water: Halt India's groundwater loss," *Nature*, 2015, *521* (7552), 289–289. Publisher: Nature Publishing Group.

Tapley, Byron D, S Bettadpur, Mo Watkins, and Ch Reigber, "The gravity recovery and climate experiment: Mission overview and early results," *Geophysical research letters*, 2004, *31* (9). Publisher: Wiley Online Library.

Taylor, Richard and Mohammad Shamsudduha, "Groundwater: depleting reserves must be protected around the world," March 2022.

Tiboris, Michael, "Against the Human Right to Water?," *Human Rights Quarterly*, 2019, *41* (4), 916–938.

Varghese, Shalet Korattukudy, Prakashan Chellattan Veettil, Stijn Speelman, Jeroen Buysse, and Guido Van Huylenbroeck, "Estimating the causal effect of water scarcity on the groundwater use efficiency of rice farming in South India," *Ecological Economics*, 2013, *86*, 55–64.

Verma, Shilp and Sanjiv J Phansalkar, "India's Water Future 2050: Potential Deviations from 'Business-as-Usual'," *International Journal of Rural Management*, 2007, *3* (1), 149–179.

Wada, Yoshihide, Min-Hui Lo, Pat J-F Yeh, John T Reager, James S Famiglietti, Ren-Jie Wu, and Yu-Heng Tseng, "Fate of water pumped from underground and contributions to sea-level rise," *Nature Climate Change*, 2016, *6* (8), 777–780.