ESSAYS ON MEASURING URBANIZATION AND INFRASTRUCTURE SERVICE LEVELS

Kyle Onda

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of City and Regional Planning in the College of Arts and Sciences.

Chapel Hill
2022

Approved by:

Meenu Tewari

Todd K. BenDor

Nikhil Kaza

Marc Jeuland

Emily Kumpel

ABSTRACT

Kyle Onda: Essays on measuring urbanization and infrastructure service levels
(Under the direction of Meenu Tewari)

Goals, targets, and service level benchmarks play important roles in planning and policy for urban infrastructure. This dissertation includes three studies that examine how such targets are sensitive to underlying conceptual and data problems, and themselves can lead to unanticipated outcomes depending on the institutional context.

In my first study, I implement a method to disaggregate geographically coarse population estimates from the Census of India into a fine population gridded dataset and apply a community detection algorithm and several population density thresholds to identify urbanized areas in India. I find that the Census of India likely undercounted its urban population and the growth rate of that population between 2000 and 2010. My second study examines an experiment by the water utility serving Amravati, a city in Maharashtra, India, to upgrade its service levels from intermittent water supply to continuous water supply in part of its service area. I find that this upgrade resulted in increased water demand for certain subgroups. In my third study, I investigate the spatial distribution of publicly owned, investor-owned, cooperative, and privately owned water systems in California, and the association between ownership type and water rate affordability. I find that publicly owned, and to a lesser extent, investor-owned utilities dominate mid-sized and large cities, but that other systems are prominent in rural and peripheral areas. I find that investor-owned systems do tend to charge higher, more unaffordable water rates, although they cut off fewer of their customers for nonpayment than their publicly owned counterparts when controlling for the affordability of their rates.

ACKNOWLEDGEMENTS

I owe much to my advisor, Meenu Tewari, who took a deep interest in my development as a researcher many years ago, and has demonstrated profound wisdom, kindness, and patience with me over many years. She gave me the opportunity to work on several research projects that allowed me to develop a broad skillset and has been a thoughtful mentor that has challenged and improved my thinking immeasurably.

I would also like to thank my committee members. Todd and Nikhil for teaching me and providing research opportunities, and Marc and Emily for their engagement and advice, despite my difficulties over the past few years. I am grateful to have spent time with all of my professors and classmates at the Department of City & Regional Planning. In addition, Sandra Lazo de la Vega was instrumental in helping me navigate the final phases of graduate school and helping me cross the finish line.

In addition, much of my work, represented in this dissertation as well as other studies, over the last 10 years has been with co-authors and collaborators that I owe my gratitude to for a rich exchange of ideas that have improved my thinking and writing.

This work would have been impossible without my family. My parents, Heidi and Pierre, and siblings Seline and Danielle for their continuous love and support. And of course, Kristine who moved across the country with me so that I could start graduate school and was a source of so much joy and wisdom that made this journey fun even as I struggled.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

United Nations member states have pledged to fulfill the 17 Sustainable Development Goals, which provide a comprehensive framework to improve human well-being for all (United Nations 2015). Many of these goals can be framed as interrelated, depending on the spatial scale of provision (Kroll, Warchold and Pradhan 2019). At the urban scale, especially in dense, rapidly urbanizing cities of the Global South , policies and investments intended to address any one of the SDGs such as 1 (end poverty), 3 (healthy lives), 6 (available and sustainable clean water and sanitation), 7 (affordable, reliable, sustainable energy), 8 (work and economic growth), 9 (industry, innovation, and infrastructure), and 11 (inclusive, safe, and resilient cities) impact all others and are deeply intertwined with each other. For example, the goal of providing affordable water or energy services in a context with high poverty might be addressed through a combination of investments. These investments could include in new infrastructure construction, targeted subsidies to poor people for the purposes of accessing existing and new infrastructure, and investments in workforce development to raise incomes and augment the ability to pay for the costs of services received through taxes or user fees. Real interdependencies exist, and at the same time, as Hirschman has pointed out (1981), realistic investments tend to be incremental and sequential, especially in low-resourced, high service deficit environments.

These conditions of uncertainty, of incremental and piecemeal investments, raise an important overarching question: How do we know that a goal is being realized for a given person, place, or community? That is, what constructs are we using to determine whom an intervention is meant for, measure progress towards the relevant goal, and determine when, where, and for whom certain interventions work better than others (Hak et. al. 2016)? How is performance evaluated, especially when equity in distribution and allocation is considered? Important constructs include the spatial structure and scale over which performance is being evaluated, feasibility, temporality and the criteria for what standards constitutes sufficient performance (Simon 2013). How policy goals are operationalized into targets and progress indicators reveals the limits of what is deemed practical to target and monitor progress over.

For example, the Sustainable Development Goal 11, Target 1 is to "ensure access for all to adequate, safe and affordable housing and basic services and upgrade slums[1]". An associated indicator is the "proportion of urban population living in slums, informal settlements or inadequate housing" (United Nations 2015). Defining and measuring the "urban" population, and its subset in "inadequate" housing, is no simple matter, however. In another example, the Sustainable Development Goal 6 is to "ensure availability and sustainable management of water and sanitation for all" (United Nations 2015). The associated target to "by 2030, achieve universal and equitable access to safe and affordable drinking water for all" makes use of constructs of equity and affordability, while its associated monitoring indicator is the "proportion of population using safely managed drinking water services", which does not provide any mechanism for targeting or measuring equity or affordability. Language about equity in the context of infrastructure planning and investment is often couched in terms of ensuring access for "all", but also often mentions concepts such as the "the worst-off" or "marginalized groups", presumably to maintain focus on groups that are often left out of investments and decision-making processes due to a lack of visibility or historic access to resources in "normal" periods of policy implementation. Information about these groups can be difficult to come by however, as the same processes that have created marginalization can play a role in reducing visibility of groups that would otherwise be targeted by equity initiatives.

Through the lens of providing safe and affordable drinking water to urban households I ask three nested questions in my dissertation to examine the intersections of performance and equity: (1) What is the target population for infrastructure investment and performance monitoring? In order to serve a population, it is often necessary to count the relevant people, but deciding which people to count is not simple. For example, urbanized areas are often not coterminous with the jurisdictional lines over which many policy levers must be based and for which data is collected. Policies may be directed towards people in "inadequate" housing, but the nature of inadequacy needs to be defined. (2) What level of

[1] I use this term throughout this dissertation only when referring to language used in official documentation and data produced by international, government, and water utility organizations.

service is the goal, and what tradeoffs do attempts to achieve this goal generate? For example, in the water sector, continuous water supply is a norm in California. Even one service outage in a three-year period is likely to trigger a vigorous analysis and regulatory activity (California State Water Resources Control Board 2022). By contrast, in India intermittent water supply is the norm and prioritizing continuous water supply competes with expanding coverage to the unserved. Targeting a level of service creates tradeoffs that are often not well understood. (3) At what cost, and under what kinds of subsidy mechanisms, can providers achieve an equity-oriented goal in conjunction with more general performance targets? For example, with respect to water, many organizations have prioritized affordability as an important metric of inclusion. But what constitutes affordability and what kinds of institutional arrangements can help achieve it?

In this dissertation, I address each of these questions in three papers which together also illustrate cross-cutting themes of the importance of data, and of equity. Data of appropriate scale, scope, resolution, and quality is required to credibly define problems and come to meaningful and actionable conclusions. Data specifically designed for infrastructure planning, including for disadvantaged groups, is often not available, and must be cobbled together, integrated from many sources that had disparate primary uses.

My first study addresses the question of who might be served by urban services. If individuals and households are not counted, they are unlikely to be served. In this paper I contend with the problem of urban delineation, a foundational problem in assessing what investments need to be made in urban infrastructure, whether by an international donor agency, a national or subnational government implementing policies or transfers to local governments, or local governments planning for people within (and perhaps outside) of their jurisdiction. The paper examines this problem in the context of urban India, a middle-income country with substantial gaps in access to urban infrastructure services. The government of India estimates that only 31% of the population is urban. We show that this is an artifact of a limited and incomplete definition of urbanity that underestimates the actual level of urbanization in India. A number of people, especially the urban poor in informal settlements and on the peripheries of metros and cities, are therefore left un-served by public service providers because they are not counted as being urban. We use a random forest-based model to create a high-resolution (~ 100 m) population grid from district-level data available from the Indian Census for 2001 and 2011, to create temporally consistent population grids. We then apply a community-detection clustering algorithm to construct urban agglomerations for the entire country. Compared with the 2011 official statistics, we estimate that Indian planners undercount their urban population by 12 percentage points.

My second study addresses the question of what level of service should be provided, given resource constraints, to a diverse urban population to satisfy equity, defined as extension of access to safe and affordable water to those without such access. In this paper we examine the workings of an experiment to upgrade drinking water services to continuous supply in the city of Amravati in Maharashtra from intermittent supplies. New service level goals of providing 24x7 water to cities and regions were instituted at the national level in India in 2005-6 along with nationwide municipal level reforms on the assumption that this upgraded level of service would save water, allowing more of unserved households to be served with piped water supplies (Singh and Roy 2018). These reforms came in the midst of a debate over whether improvements in existing intermittent supplies would better allow more of the under and un-served to be covered with piped water instead of CWS upgrades for some.

Combining consumption, billing and tariff data with an extensive survey conducted by the city's water utility we compared how demand changed under CWS and IWS for the 3 years that the CWS experiment lasted, for whom did it change and by how much and how sensitive demand was to tariff increases in both CWS and IWS neighborhoods. We found that demand actually increased under CWS – in contrast to assumptions that it would decrease due to a combination of reduced water "wasting" and responses to the price signal. Subgroup analysis revealed that households that had previously installed motorized borewells and overhead storage tanks, or households that had done neither, increased their demand for water more than households who had only one or the other. We also found that households that had previously been consuming the lowest and the highest amounts of water increased their consumption under CWS, but not households that were consuming volumes in between. We also found that the way in which CWS supplies were sequenced in the city – via reforms that improved IWS supplies for all – created unanticipated roadblocks to the broader adoption of CWS by households. This mid-level service trap eventually derailed the CWS experiment. Unable to realize anticipated revenue gains from wider CWS adoption (at higher tariffs), rising population pressures in the face of supply constraints led the

CWS experiment to be aborted in six years. Our findings suggest that improved IWS may be a better strategy from a feasibility and equity perspective in a context with resource constraints and supply deficits.

My third study examines affordability of drinking water services, including how it can be defined and measured, and how it plays out spatially and institutionally in the context of California. In particular, this study investigates how different institutional arrangements might be associated with different indicators of affordability (namely, water bills at given levels of consumption as a percentage of income for households in certain income groups, and prevalence of water arrears and shutoffs). Using 2017 data for California I match service area boundaries of water systems with census income data and rate structures to compare the geography, income distribution and affordability of water rates within communities served by systems of different ownership types. I find that for-profit and publicly owned systems serve communities of similar income distributions, while not-for-profit mutual water companies serve higher- income communities. Regulated privately-owned systems charge more for water while providing more low- income assistance and shutting off fewer households than publicly owned systems. Overall, the study suggests that regulation is more important than ownership type when it comes to how households of concern to an equity-minded social planner might experience the costs of water service.

REFERENCES

California State Water Resources Control Board. 2022. “2022 Drinking Water Needs Assessment.” California State Water Resources Control Board. https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/documents/needs/2022needsassessment.pdf.

Hák, Tomáš, Svatava Janoušková, and Bedřich Moldan. 2016. “Sustainable Development Goals: A Need for Relevant Indicators.” *Ecological Indicators* 60 (January): 565–73. https://doi.org/10.1016/j.ecolind.2015.08.003.

Kroll, Christian, Anne Warchold, and Prajal Pradhan. 2019. “Sustainable Development Goals (SDGs): Are We Successful in Turning Trade-Offs into Synergies?” *Palgrave Communications* 5 (1): 140. https://doi.org/10.1057/s41599-019-0335-5.

Simon, Herbert A. 2013. *Administrative Behavior*. Simon and Schuster.

Singh, Ravi S, and Dipankar Roy. 2018. “Service-Level Benchmarking: Some Emerging Global Lessons for Indian Water Governance.” In *Globalisation, Environment and Social Justice*, 319–37. Routledge India.

United Nations. 2015. “Sustainable Development Knowledge Platform: Sustainable Development Goals”. https://sustainabledevelopment.un.org/sdgs

CHAPTER 2: MISSING MILLIONS: UNDERCOUNTING URBANIZATION IN INDIA[2]

2.1 Introduction

The global rate of urban transition has been immense in the past half century, with much of that transition and associated population growth occurring across parts of Asia (Ellis and Roberts 2015; Schneider et al. 2015). In 1960, India and China had similar urban population percentages of 18% and 16%, respectively (World Bank 2018). Yet by 2016, according to the World Bank statistics, while the Chinese urban population was at 54%, Indian urban population was at 33% suggesting very different developmental trajectories. In fact, the World Bank, based on Census of India statistics, estimates that urban India is growing at a declining rate (3.8% in the 1970s to 2.7% in the 1990s and 2000s, to 2.4% in the 2010s) (World Bank 2018). Widely varying estimates of such rates can be found from other sources. United Nations figures rely on national statistics that themselves are generated by a wide diversity of definitions of urban, leading to incomparable estimates of urban population and urbanization rates across countries (Uchida and Nelson 2010). In addition, a long-running debate exists in the literature about the relationship between urbanization of a country's population and its economic growth (Fay and Opal 2000; Henderson 2003; Spence et al. 2009). While higher levels of urbanization are observed in countries with higher per-capita GDP, the rates of urbanization have little correlation to economic growth (Bloom et al. 2008; Chen et al. 2014).

Yet much of this literature presumes that urbanization levels, along with the GDP, are measured consistently and appropriately in different contexts (Satterthwaite 2007). Cross-country consistency in urban definitions is necessary for the design and study of urban policies that may vary by nation, such as

[2] This chapter is published as an article in Population and Environment. The citation is: Onda, K., Sinha, P., Gaughan, A.E. *et al.* Missing millions: undercounting urbanization in India. *Population and Environment* 41, 126–150 (2019). https://doi.org/10.1007/s11111-019-00329-2

the organization of public services or the allocation of development finance towards meeting international development goals (OECD 2012). For example, the Sustainable Development Goal 11, to “Make cities and human settlements inclusive, safe, resilient and sustainable”, is associated with a number of indicators and targets, the measured values of which can change substantially when applying different definitions and delineations of cities (Klopp and Petretta 2017).

Definitional differences are not just a matter for comparative convenience; they have both theoretical and policy implications. Studies of agglomeration economies and the determinants of urban economic growth in India often use districts as units of analysis due to a lack of availability of consistent boundaries for metropolitan areas, which would be a more appropriate unit for such research questions (e.g., Desmet et al. 2015; Duranton and Puga 2004; Ghani et al. 2016). This problem could potentially lead to misleading conclusions in cross-country comparative work. For example, Chauvin et al. (2017) conclude that India does not conform to spatial equilibrium, a central idea in urban economics, in a comparative analysis of India, Brazil, China, and the USA. In this study, districts were the unit of analysis for India, while units more analogous to Metropolitan Statistical Areas in the other three countries were used. In contrast, Hasan et al. (2017) find evidence of relatively low agglomeration economies in India based on town and city-level data, but do not account for how such towns may be part of larger metropolitan regions in their analysis.

From a governance standpoint, the delineation of urban areas has consequences for the spatial distribution of infrastructure provision and related institutional arrangements. Urban areas are seen as engines of economic development and infrastructural and resources are concentrated on them (Indian Planning Commission 2011). Even so, urban infrastructure investment is often assessed to be inadequate in India (Ahluwalia et al. 2014). Underestimating the existence of dense population clusters only exacerbates this problem by limiting the political attention, governance reform, and finance necessary to build and maintain appropriate levels of infrastructures such as intra-city transportation, water, sanitation, and health in dense, yet officially rural areas. Areas with high population density require qualitatively different types of infrastructure and necessitate different institutions to govern them than lower-density areas, regardless of whether they are administered as urban or rural units (Rakodi and Lloyd-Jones 2002).

In India, rapid urbanization that was expected to follow economic liberalization policies starting in the 1990s was predicted to hollow out rural areas in favor of large urban areas such as Bengaluru due to migration based on economic opportunity. In part, these conclusions are drawn from undercounting urban areas and ignoring the large in situ urbanization happening over time. Denis et al. (2012) argue that close to two-fifths of the population live in urban settlements and 35% of the urbanites do live in small towns below 100,000 in population. More importantly, the patterns of urban settlements are different regionally, which also lead to regional developmental imbalances. For example, the less developed states of West Bengal and Bihar have substantially more dense settlements in the Denis et al. (2012) approach than the official estimates. Accordingly, Kundu (2011) argues that when optimistic rural-urban migration predictions were not realized, there were adverse consequences for urban livelihoods in smaller towns, which contribute little to national productivity and command little political attention. Indeed, initiatives such as the Jawaharal Nehru National Urban Renewal Mission (JNNURM), one of the largest infrastructure programs ever undertaken by the Government of India, allocated funds disproportionately to large urban areas and may have caused stagnation in smaller towns and their surrounding rural areas (Khan 2016).

Underbounding metropolitan areas has a related policy consequence when combined with India's federalist governance structure. The 73rd and 74th Constitutional Amendments of 1993 devolved many planning and infrastructure provision responsibilities from state to local governments, including urban local bodies (ULBs) for officially urban areas and gram panchayats for officially rural areas. This devolution in some circumstances allowed local communities to organize appropriate institutions and infrastructure packages (Hutchings 2018). However, it also raises barriers for coordination between communities in the provision of some public goods or the management of shared common-pool resources. For example, the highly administratively fragmented Kochi urban area saw many JNNURM projects delayed or applications rejected due to competing priorities and conflict between the Kochi Municipal Corporation and surrounding ULBs and gram panchayats in the region (Kamath and Zachariah 2015). Such phenomena highlight the potential gains to be had from more regional planning structures that incorporate all neighboring clusters of high-density jurisdictions (whether administered by ULB or panchayat) into related infrastructure needs, as suggested by Mukhopadhyay et al. (2017).

The lack of a georeferenced and consistently delineated dataset also poses a problem for studying urban change over time. Official estimates put change in Indian urban population at 3.3% between 2001 and 2011, with 29.5% of this urban growth due to reclassification of rural areas into Census Towns by the Census of India, rather than expansion or densification of existing urban areas. This is higher than the growth in urban population attributable to migration (Pradhan 2013). However, the significance of these invisible urban villages, classified as urban by the Census but administered as rural areas, is not readily apparent due to the unavailability of appropriate georeferenced datasets. Since no fine-grained geographic and demographic data are readily available, researchers must look for clues in various census tables to locate and measure the extent of such in situ urbanization. In this paper, we aim to make this urbanization visible, so that appropriate political and economic institutions can be fashioned to meet their governance needs.

2.2 Background

There is no consistent definition of what constitutes an urban area around the world (Buettner 2015; Cohen 2006; Satterthwaite 2007). Previous efforts to define consistent, global definitions of urban area relied on daytime satellite images (Angel et al. 2011), nighttime lights (Zhou et al. 2011), functional integration (OECD 2012), and population density combined with travel times to the nearest large city (Uchida and Nelson 2010). Others have followed a more hierarchical definition of classifying the urban areas based on density, the proportion of the population living in different density clusters, population size, and contiguity characteristics (Dijkstra and Poelman 2014).

However, different statistical agencies use different definitions and, thus the measurement of urbanization and rates varies considerably from country to country. Some countries do not have specific criteria to delineate urban regions, instead preferring to list the urban areas with independent local governments. While many countries use a minimum population size (200–50,000), a few use minimum population density (~ 6.3 per ha) (Deuskar and Stewart 2016). India is one of the 16 countries that use criteria of economic activity (dominance of non-agricultural activity). India's definition also has a gender dimension by counting only the type of jobs held by male workers. In particular, the Census of India defines urban areas as follows:

1. All places with a municipality, corporation, containment board, or notified town area committee, etc. (referred to as Statutory Towns)
2. All other places which satisfy all of the following criteria (referred to as Census Towns):
   (a) A minimum population of 5000
   (b) At least 75% of the male working population who work more than 6 months of the year engaged in non-agricultural work
   (c) A population density of at least 400 persons per square kilometer (4 persons per ha).

Despite the detailed definition, there exists considerable debate about the urban character of India and its evolution over time (Denis et al. 2012; Ganapati 2014; Sudhira and Gururaja 2012). For example, Denis et al. (2012) use contiguous built-up areas in India (with some leapfrogging) and assign the population of the Census-defined areas (not spatially demarcated). They then use a 10,000-person population threshold to classify urbanity for Indian cities. This definition and delineation allocate 100 million more people to urban areas compared with the Census of India 2011 estimates.

A variation of these approaches can be found in the works of Balk (2009) and McGranahan et al. (2007). Names and population estimates from the National Statistical Organizations (NSO) are merged with geographic coordinates for given administrative units from gazetteers. To define the urban extents, unlike Denis et al. (2012) which uses the daytime impervious surface, these approaches use nighttime lights, as a proxy for electrification, which is itself a proxy for urban service provision. The population from the NSO/gazetteer points within each urban extent is assigned to the polygon. Using urban extents from Balk (2009) to delineate large cities and including peri- and suburban areas that are within a certain distance from these large cities, Uchida and Nelson (2010) construct an agglomeration index as a characterization of the metropolitan region. Uchida and Nelson estimate the urban population of India to be between 42.9 and 51.9% compared with United Nation's estimate of 27.7% (based on the Census of India 2000 estimates).

Each of these different definitions produces different urbanization estimates as well as extents and locations of urban agglomerations, with its own set of limitations. Using nighttime lights exclusively to define urban extent underestimates dense human settlements that are not yet electrified, or suffer from intermittent electrification provision or from light blooms (Abrahams et al. 2018; Small et al. 2005). In contrast, using exclusively daytime satellite imagery to delineate urban extents is constrained by weather conditions (e.g., cloud cover) and trade-offs between spatial and spectral resolutions. The inability of these methods to incorporate other types of data such as slope, hydrology, climatic zones, and other features such as infrastructure that are associated with human settlement patterns is critiqued by Uchida and Nelson (2010). Furthermore, relying on merging geographic coordinates to population data using place names is susceptible to significant error due to mis/multiple spellings and requires significant expert intervention. For example, about 1.8 million people in India were not assigned to a location in the Denis et al. (2012) approach. The contiguity criterion relied upon by Dijkstra and Poelman (2014) relies on a low spatial resolution (of 1 km) to delineate urban areas, resulting in fragmented and therefore small urban settlements, especially at the fringes of a city. In a different but still spatially compromised way, Uchida and Nelson's agglomeration index merges spatially proximate but non-contiguous urban areas into one metropolitan area, changing the boundaries that can be used. Since their approach is to allow for cross-country comparisons of total urban population, the precise location and boundaries are less important.

We provide a methodology that allows us to define and delineate urban areas consistently across various jurisdictions. We propose a method called Metropolitan Agglomerations from Gridded Population Intensity Estimates (MAGPIE) that draws from the above-mentioned approaches to characterize urban regions and their systems. We explicitly use density thresholds combined with size thresholds in a consistent fashion to distinguish between urban and rural settlements. We ignore the gender and economic activity thresholds that the Indian Census uses, for generalizability purposes. With relatively little human intervention, the proposed method produces an urban/rural delineation with an associated urbanization estimate similar to that of Indiapolis in short order. Because we rely upon gridded datasets, including remote sensing images, our conclusions are not bounded by jurisdictional vagaries. The other methods described in this section are also not limited by jurisdictions and allow for comparisons. However, they are limited by resolution and underlying covariates (Dijkstra and Poelman 2014) and imperfect separation of proximate urban areas (Uchida and Nelson 2010). MAGPIE addresses some of these limitations.

2.3 Method

2.3.1 Study area and data processing

Population counts were sourced from the Office of the Registrar General and Census Commissioner in India and population counts were linked to GIS administrative boundaries for each district (source: https://gadm.org/) creating a spatially explicit representation of population distribution at the census unit level. We do not include parts of Kashmir that do not have census data in our study region. We then modeled gridded population at the district level (n = 594) for the years 2001 and 2011, matching administrative boundaries for boundary and data consistency purposes between years, with 2001 as the base year. Fixed census units between years are important to enable a consistent estimation process across time (Gaughan et al. 2016). In doing so, we reduce the potential of under- or over-fitting the model due to heterogeneity in census unit size and associated average population densities.

We matched all covariate data for both years based on either temporally invariant or temporally explicit datasets. The land cover is based on GlobCover data, which is derived from the ENVISAT satellite mission's MERIS (Medium Resolution Image Spectrometer) imagery. The land cover dataset has thirteen categories: cultivated terrestrial lands, woody/trees, shrubs, herbaceous, other terrestrial vegetation, aquatic vegetation, urban area, bare areas, water bodies, rural settlement, industrial area, built area, and no data. We also used digital elevation data and derived slope estimates from SRTM-based HydroSheds data (Lehner, Verdin, & Jarvis, 2013) and the DMSP-OLS (v.4) lights at nighttime series, obtained from NOAA's National Geophysical Data Center (National Oceanic and Atmospheric Administration, n.d.). In addition, the Global Human Settlement Layer (GDAL/OGR Contributors, 2–19) with a spatial resolution of 38 m was collected from the European Commission Joint Research Centre (2014 beta version) for the years' most coincident with 2001 and 2011. To best use the urban extent information, we created a distance-to-built-edge covariate, where distances inside the built land cover class boundary were negative and distances outside the edge were positive. We also used the WorldClim/BioClim 1950–2000 mean annual precipitation (BIO12) and mean annual temperature (BIO1) estimates (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). In addition to land cover, settlement, and associated raster datasets, we included geospatial data that was correlated with human population presence on the landscape, such as protected area delineations (UNEP-WCMC, 2010), networks of roads, and waterways; large water bodies; and infrastructure-related features and settlement or populated locations from open street map 2017. All these covariate data employed in the modeling process are summarized in Table 2.1. These covariates were all summarized to the district polygon level as the average value within each polygon. All datasets were resampled using nearest neighbor to match the same resolution to a square pixel resolution of 8.33×10^{-4} degrees (approximately 100 m at the equator) and projected into UTM 44N projection prior to analysis. All covariates were prepared in ArcGIS (ESRI 2016) and Python programming language (version 2.7) (Python Software Foundation 2013).

Table 2.1. Covariates used in gridded population modeling process

Variable name(s)	**Source and nominal resolution**
District Census Population, 2001, 2011	Open Government Data (OGD) Platform India, district level
Temporally explicit covariates	
Land Cover, 2000, 2010	GlobCover, 300 m
Global Human Settlement Layer, 2000, 2012	ECJRC, 38 m (Pesaresi et al., 2013)
Lights at night, 2001, 2011	DMSP-OLS-derived (National Oceanic and Atmospheric Administration, n.d.)
Common covariates	
Mean temperature, 1950–2000	WorldClim/BioClim (BIO1) (Hijmans et al., 2005)
Mean precipitation, 1950–2000	WorldClim/BioClim (BIO12) (Hijmans et al., 2005)
Sanctuaries, National parks, Game Reserves, World Heritage Sites	World Database on Protected Areas September 2012, UNEP (IUCN, UNEP-WCMC, 2010)
Elevation	USGS HydroSHEDS (Lehner et al., 2013)
Derived Slope	USGS HydroSHEDS (Lehner et al., 2013)
Distance to infrastructures	Open Street Map, 2017–05

Distance to places	Open Street Map, 2017–05
Distance to road networks	Open Street Map, 2017–05
Distance to waterbodies	Open Street Map, 2017–05

2.3.2 Gridded population intensity estimates

We generated gridded population intensity estimates (GPIE) using the methods described by Stevens et al. (2015) to disaggregate the census population for 2001 and 2011. We used grid cells with a resolution of 3 arc sec (approximately 100 m at the equator). We used a random forest (RF) statistical model (Breiman 2001) to generate a population prediction density layer, in conjunction with a dasymetric redistribution of population counts (Stevens et al. 2015) to produce final gridded population outputs at approximately 100 × 100 m grid cells. For the Indian subcontinent, this represents approximately 395 million pixels of land that population is allocated to. The RF statistical model provides a non-parametric platform coupled with an ensemble machine-learning technique for classification or prediction purposes (Breiman 2001). The RF method relies on the use of bagging and random selection of covariates across numerous classification and regression trees (Lung et al. 2013).

For our purposes, we use census counts at the district level and covariate aggregation values for each census unit to create a RF model to predict log population density (Lung et al. 2013). In this method, the dasymetric redistribution weight is produced as a function of different covariates representing the individual covariates such as lights-at-night, slope, elevation, and proximity to land-use types. The resulting RF is used to predict a country-wide, pixel-level map of log population densities that provides a weighting layer for a dasymetric redistribution scheme (Mennis 2003) to redistribute population counts within each unit to the target cells (Stevens et al. 2015). Figure 2.1 portrays the schematic process involved in creating the dasymetric weighting layer. This dasymetric disaggregation then unevenly allocates the district-level population to underlying raster respecting the protected and uninhabitable areas. The result of this process is a 100 × 100 m grid cell (1 ha) resolution population map for 2001 and 2011. RF model fitting at the administrative unit level and prediction at the grid cell level were both performed in the R statistical environment (R Development Core Team, 2017) using the *randomForest* package (Liaw and Wiener 2002). Predicted range of population datasets from random forest models is sensitive to the scale of the training dataset and using a coarse dataset could lead to a small range in dasymetric weighting surface. As a result, with coarse census data, a less heterogeneous population density will be observed with fewer extremes. We also note a disconnection between the level of support between the model estimated for administrative units and the scale of the predictions from the model used to disaggregate census data. However, while no assumptions are placed on the linearity or interactions present in relating ancillary data to population density (a feature of random forest modeling), we assume that the process resulting in those estimated associations at an aggregate level are, on the whole, representative of the process relating covariates to population density at the finer, gridded scale. In the absence of data on population densities at the finer scale of interest, of which we have none to estimate the model with or validate against across time, output based on this assumption has consistently shown to perform better than less complex or less informed disaggregation techniques (Stevens et al. 2015; Gaughan et al. 2016; Nieves, et al. 2017). Despite the "ecological fallacy" inherent to this change-of-support (Gelfland, et al. 2001; Holt, et al. 1996) and likely biased outcome at the pixel level, the approach still manages to achieve comparable results to bottom-up modeling using fine-scale model estimates (e.g., Engstrom, et al. 2019).

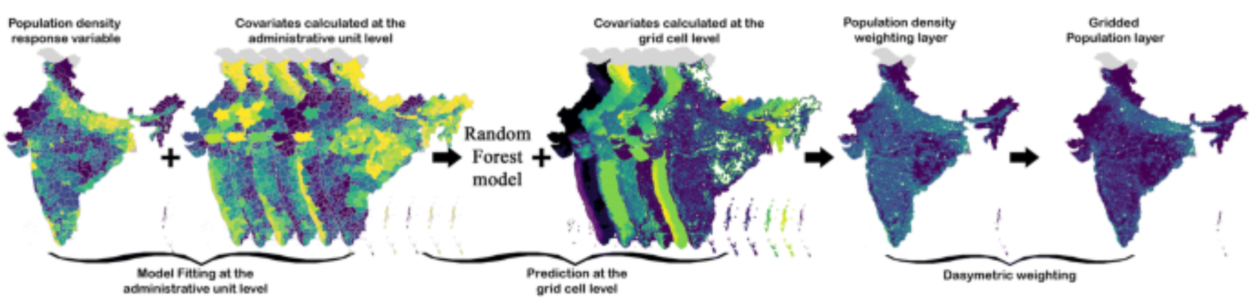


Figure 2.1. Schematic representation of the dasymetric gridded population modeling process

2.3.3 Metropolitan agglomerations

We define urban areas in metropolitan agglomerations (MA) using a three-step process (see Figure 2.2). Based on the GPIE, we first select all cells that are outputs that are above a certain density threshold. We use 7.5 persons per ha as a density threshold and experiment with 5 and 10 persons per ha to test the sensitivity of this threshold. Note that all of these are above the 4 and 3 person per ha thresholds used by Census of India and Dijkstra and Poelman (2014), respectively. We use contiguity of these densely populated cells to construct clusters of urbanized areas using a region grouping algorithm from Geospatial Data Abstraction Library (GDAL/OGR contributors 2019). Holes within each of the polygons are removed. In other words, unpopulated areas that are completely circled by urban areas such as hills, parks, and lakes would be considered to be within the boundary of the urban area. This removal of holes adds 9% more to the urbanized area than otherwise and only has marginal effect on the urban population estimates (~ 2.8%). Because of noise associated with GPIE, we removed areas that are below 2 ha area from consideration. The 2 ha are approximately two contiguous cells that are not adjacent to any other selected cells. We experimented with different thresholds and selected 2 ha as the areal threshold that produces urban population estimates less than 90%.

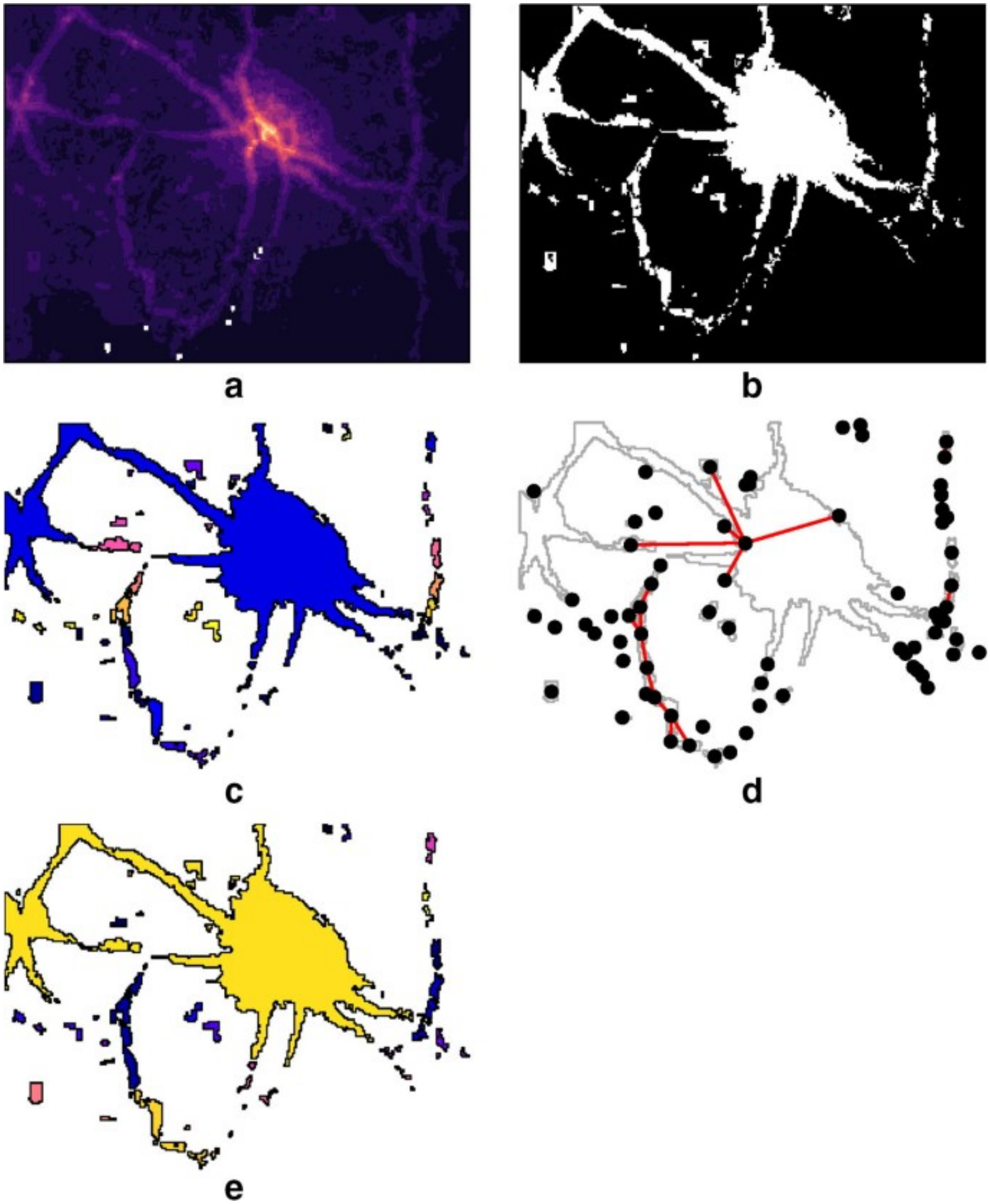


Figure 2.2. Various stages of defining the urban area boundary in MAGPIE. (a) Input population intensity estimates. (b) Urban areas based on density threshold. (c) Removal of holes and polygonization based on contiguity constraint. (d) Construction of graph based on distance threshold to account for non-contiguous polygons. € Construction of clusters based on eigenvector community-detection technique

However, contiguity is an insufficient criterion to delineate metropolitan areas, as they are usually fragmented at the edges. To determine how these constellations of fragments relate to one another and to larger urban areas, we turn to the community-detection algorithms borrowed from network science (Comber et al. 2012; He et al. 2019). We then construct a graph from these polygons with each polygon as a node with the vertex set V(G). Two nodes are connected with an edge if the distance between their boundaries is below a distance threshold of 150 m. This is a distance that is roughly the diagonal of the cell and approximates queen contiguity criterion with one cell skipped over. We then find communities within the components of the graph G using the leading non-negative eigenvector of the modularity matrix of the graph (Newman, 2006). Community-detection techniques allow us to partition the vertices of the graph into groups, where the connections within the groups are denser than the connections between groups. The intuition is that if multiple urban clusters are close to one another, they should be treated a coherent entity. This also allows us to avoid identifying tendril-or dumbbell-like urban patterns unless they are explicitly contiguous. We then combine the polygons represented by the vertices that are part of a community into a single metropolitan agglomeration. This analysis is done using *raster* (Hijmans 2017), *spdep* (Bivand and Piras 2015), and *igraph* (Csardi and Nepusz 2006) packages in R.

2.4 Results

To measure the prediction error of the random forest model, we estimate the out-of-bag (OOB) error from 37% samples with 500 trees. The OOB is an error estimate calculated during the RF model fitting and is based on averaging all mean squared errors. It provides a robust and unbiased measurement of the prediction accuracy of the RF model (Breiman 2001) and informs the accuracy of the final gridded population datasets produced using the RF-based approach (Gaughan et al. 2016; Stevens et al. 2015). The pseudo-r-square value for training model based on mean population density at district scale is 0.88 and 0.87 for 2001 and 2011, respectively. The median values of predicted population counts are 1.4 to 1.6 persons per pixels (see Figure 2.3). To assess the final accuracy of the GPIE estimates, we matched 500 randomly selected village/town boundaries (level 4 administrative units, available from Bhuvan, a high-resolution web mapping service focused on the region of India (National Remote Sensing Center 2019)) with the census population counts using name of the village/town and district as an identifier. We aggregated the GPIE results to the village/town boundaries after correcting for boundary errors and projection issues. The correlation coefficient is 0.86 between census counts and GPIE results providing confidence in the spatial representation of gridded population outputs.

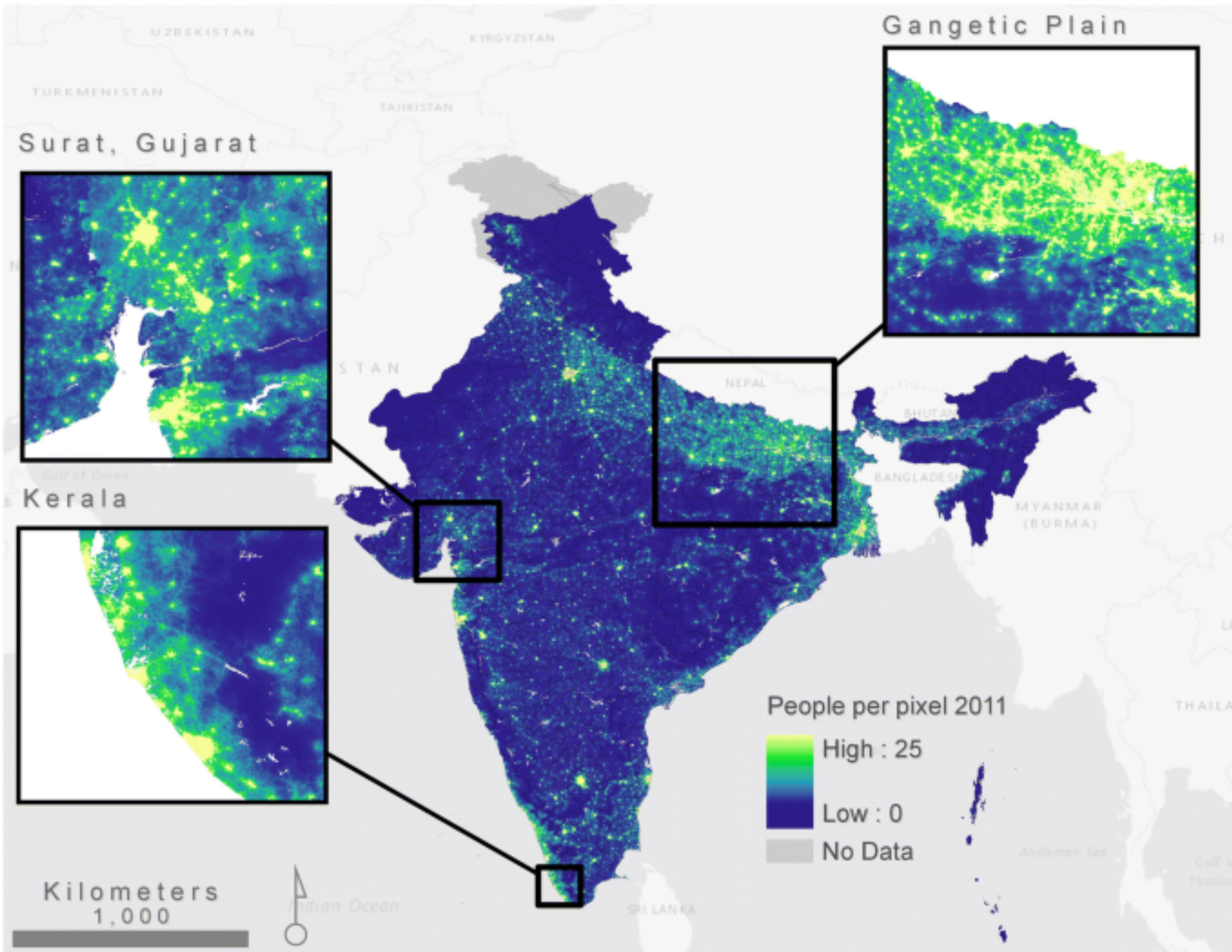


Figure 2.3. Gridded Population Intensity Estimates for India (2011). Maximum value is restricted to 25 for visualization purposes.

We report our estimates for metropolitan agglomerations with each of the three density thresholds (5, 7.5, and 10 persons per hectare) with a 150-m distance threshold. We compare our results with those of three other urbanization estimates; the Census of India, Indiapolis (Denis et al. 2012) and GHS-POP/GHS-SMOD. We produce the GHS-POP/GHS-SMOD estimate by aggregating the 2015 estimate of the 2019 version of the Global Human Settlement population grid (GHS-POP) (Schiavina et al. 2019) with the urban settlements of the 2019 version of the GHS settlement model grid (GHS-SMOD) (Pesaresi et al. 2019). To create the settlement clusters from GHS-SMOD, we combine raster cells classified to be in the “urban domain” that are contiguous at the edges (and not only the corners) into polygons representing discrete contiguous settlements. It should be noted that these are the best available estimates at the time of publication and that they are subject to continual updates.

We divide our results into the following sections: (2.4.1) characterization of the location and type of urbanization for 2011, (2.4.2) comparison of our method with the Census of India, Indiapolis and GHS-POP/GHS-SMOD estimates of urbanization and urban hierarchy for 2011; and (2.4.3) comparison of estimates for urbanization rates between 2001 and 2011 with the Census and Indiapolis.

2.4.1 Patterns of urban settlements in 2011

Figure 2.4 shows the spatial extent of urbanization and the sensitivity to the minimum density thresholds. Lowering the density threshold results in larger solitary cities and the coalescence of cities into larger and more populous agglomerations (see Figure 2.4a, b). Regions where estimated urbanized areas greatly increase in size when the density threshold changes generally indicate large areas of a relatively uniform population density in between denser urban centers. These regions are easily identifiable as the large contiguous megalopolises with greater than 40 million people including the extended agglomeration of Delhi and western Uttar Pradesh, the Gangetic plain through eastern Uttar Pradesh and Bihar (labeled for urban centers Patna, Varanasi, and Gorakhpur in the top row of Figure 2.5), and most of West Bengal (labeled Kolkata in Figure 2.5). Kerala exhibits a similar pattern of contiguous medium-density settlement, but without the numerous dense urban centers in between that would push the total population above 40 million. These regions (labeled urban centers Kozhikode and Thiruvananthapuram in Figure 5) have comparable population sizes to large cities like Mumbai and Bangalore but much less dense, suggesting coalescent urbanization that knits together many villages, towns, and cities.

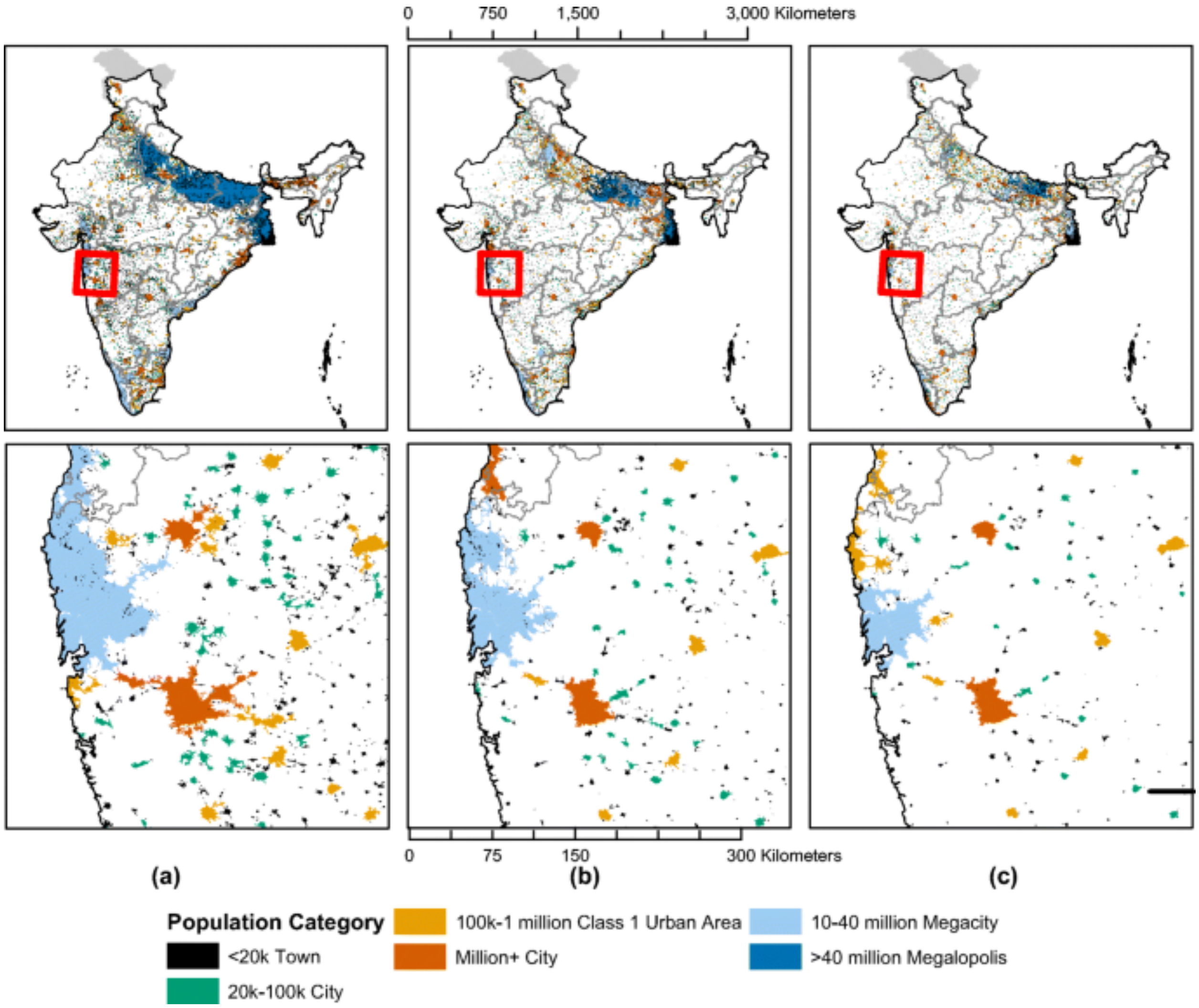


Figure 2.4. Spatial extent of urbanization according to MAGPIE. Top row for the entire map of India, bottom row inset for the detail of smaller areas. Estimated with a minimum density threshold of (a) 5 person/ha., (b) 7.5 person/ha. (c), and 10 person/ha.

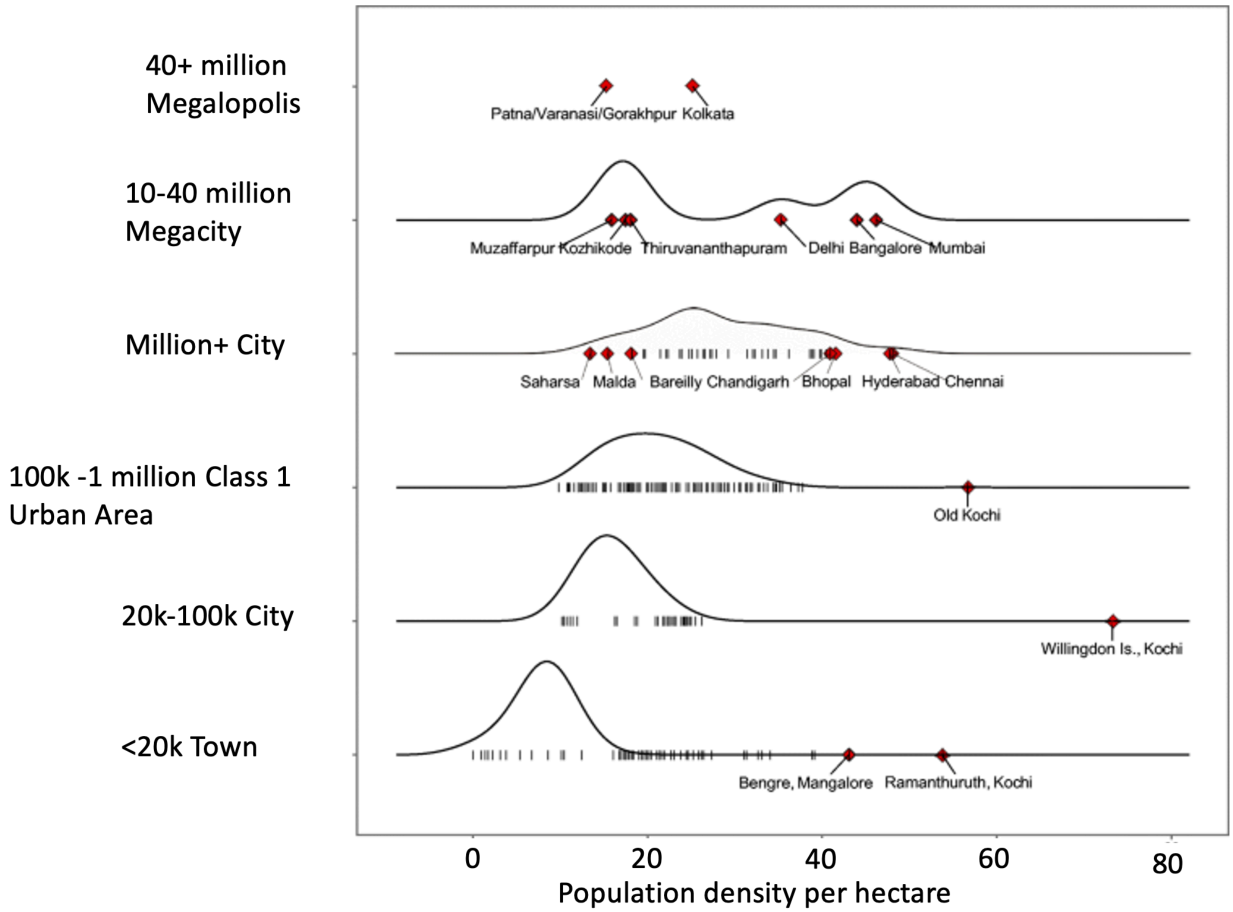


Figure 2.5. Relationships among MAGPIE estimates of population and population density in India in 2011 (7.5 pp./ha threshold). Vertical lines represent individual urban agglomerations, positioned on the x-axis according to population density (population per hectare) and on the y-axis according to population size category. Red diamonds represent particularly high- or low-density urban agglomerations of interest.

Outside of these regions, the pattern of urbanization is different. The relative lack of change in the size of urban areas of at least 100,000 in population when the density thresholds are changed indicates that populations are more highly concentrated in urban areas. Estimates of the urban population in these regions rise as the density threshold falls, but this is due to a combination of two reasons: (1) higher numbers of distinct settlements are counted as urban and (2) the periphery of urban areas is now included within the boundary of an existing urban area. An example of this kind of urbanization pattern is throughout the state of Maharashtra, where areas of high population density are concentrated around Mumbai, Pune, Nagpur, and several cities between 100,000 and 1000,000 in population and new agglomerations are not created by lowering the density threshold (see insets in bottom row in Figure 2.4).

Metropolitan agglomerations such as Chennai, Bangalore, and Hyderabad also follow this pattern, being the densest population centers in India, as they are not connected to other large cities and so are not as populous as the megalopolises (see third row of Figure 2.5). Another way to understand the pattern of urbanization is to see how the density thresholds affect the proportion of people that are considered urban in each state. Uttar Pradesh is dramatically affected by the threshold, suggesting a 45-percentage point difference; changing the density threshold from 5 to 10 decreases the urban population by 45 percentage points (see Fig. 2.6). A similar, but less dramatic, effect is observed in Bihar, Assam, and West Bengal. In contrast, due to high densities, many union territories and the Delhi region are not affected by the threshold.

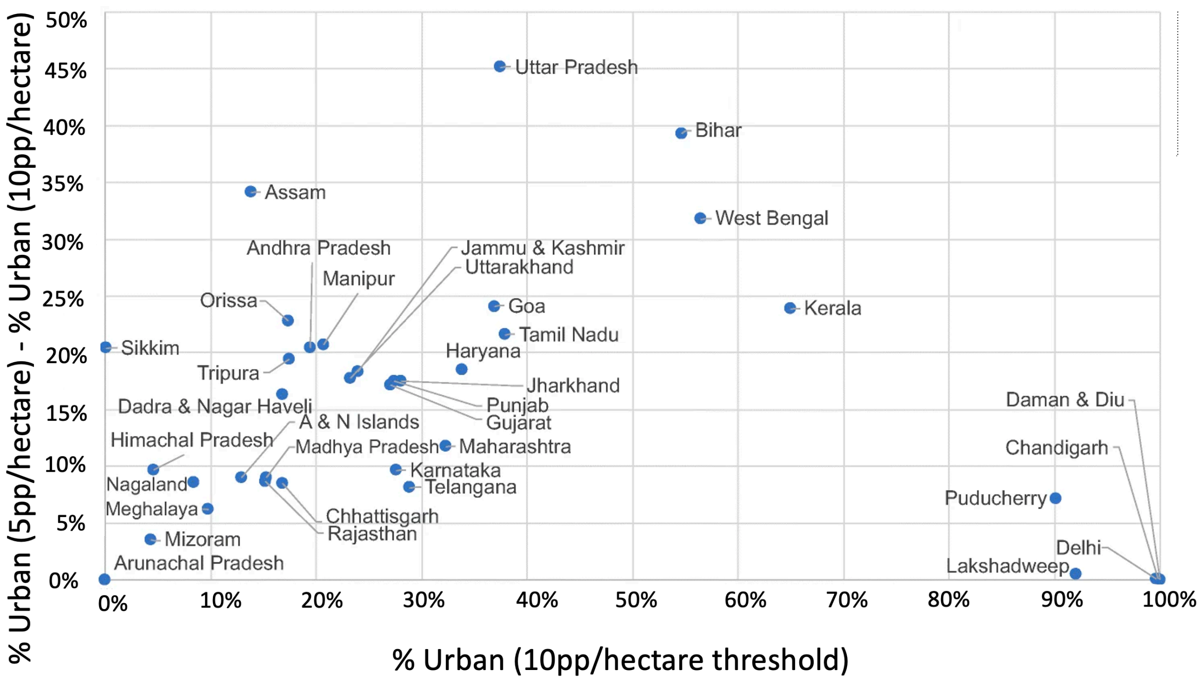


Figure 2.6. Characterizing the type of urbanization in 2011 by State and Union Territory according to MAGPIE. Bottom right corner indicates that regions that are highly urbanized and the density thresholds matter little for urban population counts. Top left corner are regions where density thresholds matter greatly for the counts.

2.4.2 Comparison with other estimates

Different methods produce vastly different estimates of urbanization in India (see Table 2.2). The categories in Table 2.2 are based on various size thresholds used by the Census of India in publications of the populations of urban areas (Registrar General and Census Commissioner of India 2011). Urban areas with less than 5000 people do not meet the qualifications to be counted as a Census Town, so any tabulated are statutory towns in the Indian census. There is no published count of the ones that are not considered part of larger urban areas. The next category is urban areas with between 5000 and 20,000 people but there is similarly no specific tabulation of them. “Urban agglomerations” are composed of combinations of towns, cities, and “out growths” with a combined population of at least 20,000 in 2001. “Class 1 urban agglomerations” have at least 100,000 people, “million plus cities” at least 1 million, and “megacities” at least 10 million people. Thus, based on these thresholds, we tabulate counts and populations of urban areas based on population size breaks of 5000; 20,000; 100,000; 1 million; and 10 million. Our methods in some cases produced large agglomerations with more than twice the population of what are normally considered India’s largest cities of Delhi, Mumbai, and Kolkata. We categorize these agglomerations, with populations greater than 40 million, as “megalopolises.”

Table 2.2. Comparison of counts of urban agglomerations and population estimates in 2011 (2015 for GHSL) for different size categories by the Census of India, Indiapolis, GHS-SMOD, and MAGPIE at 10 pp/ha, 7.5 pp/ha and 5 pp/ha population density thresholds

	Census	Indiapolis	GHS-POP/SMOD	MAGPIE		
Counts				10 pp./ ha	7.5 pp./ ha	5 pp./ha
1. Megalopolis (> 40 million)				1	2	2
2. Megacity (10–40 million)	3	4	5	5	6	6
3. Million+ urban area (1–10 million)	50	47	72	47	46	30
4. Class 1 (100k–1 million)	415	495	1556	330	329	275
5. 20–100k town	1846	2866	5275	782	769	739
6. 5–20k town	NA	4272	15,182	1000	1018	983
7. < 5k town	NA	Not urban	194	10,269	20,804	32,693
Population						
1. Megalopolis (> 40 million)				40,251,032	116,483,274	364,555,198
2. Megacity (10–40 million)	48,802,734	73,894,637	104,736,702	105,197,661	122,513,999	113,912,852
3. Million+ urban area (1–10 million)	111,813,450	116,694,703	179,271,190	127,104,331	136,097,571	95,663,062
4. Class 1 (100k–1 million)	104,293,727	117,489,953	358,730,669	87,419,632	91,539,353	70,548,571
5. 20–100k town	74,112,244	110,112,901	226,556,092	35,992,827	34,445,915	32,843,228
6. 5–20k town	38,098,826	58,464,875	141,688,151	10,634,852	10,799,146	10,115,771
7. < 5k town			462,143	4,919,263	6,886,963	7,548,863
Total	377,120,981	476,657,069	1,011,444,947	411,519,597	518,766,222	695,187,545
% Urban	31%	39%	77%	34%	43%	57%

The 2019 GHS-POP/GHS-SMOD estimates, based on the method of Dijkstra and Poelman (2014), place urbanization (in 2015) in India at 77%, which is much higher than MAGPIE, Indiapolis, and the Census of India. This is likely because the underlying dasymetric population disaggregation only uses one binary covariate, the GHSL built-up area indicator (Corbane et al. 2018; Florczyk et al. 2019). Combined with the relatively coarse 1-km spatial resolution, the effect is to consider a large proportion of square kilometer grid cells in India with any built-up area as urban. We find this to be an implausibly high estimate of urbanization in India.

MAGPIE tends to produce different urban hierarchies than the Census as well as the Indiapolis project in three respects. First, our method produces large numbers of isolated, small urban areas with less than 5000 people, although all of these towns together only amount to 5–8 million people. These settlements are generally not considered urban by the Census and are categorically not considered urban by Indiapolis.

Second, MAGPIE tends to consider large areas of relatively high population density (though not necessarily concentrated around traditional core cities) as urban. This results in the reallocation of the Indian population from small towns as well as areas the Census considers rural into larger urban areas with more than 10 million people. For instance, our most conservative threshold combination of 10 persons/ha with a distance threshold between settlements of 150 m produces two large urban agglomerations of roughly 30 million people each in Bihar and considers almost all of the coast of the state of Kerala as one contiguous agglomeration with over 40 million people. Our method's characterization of urbanization in Kerala is similar to that of Indiapolis.

Third, our method tends to agglomerate populous municipalities with dense networks of smaller settlements in between them into larger agglomerations, reducing the number of mid-size cities and increasing the number of megacities relative to the census. For instance, our method folds many areas that the Census and Indiapolis consider cities with populations between 100,000 and 10 million into larger megacities, while combining almost all of Uttar Pradesh, Bihar, and West Bengal into megalopolises.

MAGPIE also results in different urbanization estimates at the state level. Figure 2.7 summarizes the urbanization rates of the Indian states as calculated by (a) the Census of India, (b) Indiapolis, and (c)

MAGPIE estimates with 7.5 persons/ha threshold. MAGPIE generally estimates lower urbanization rates for each state than the Census. The major exceptions are Kerala, Bihar, Uttar Pradesh, and West Bengal, where we estimate much higher urbanization rates than the Census. We characterize urbanization in Kerala similarly to Indiapolis, although we estimate a much higher degree of urbanization in Bihar, Uttar Pradesh, and West Bengal than does Indiapolis. By contrast, MAGPIE tends to estimate lower urbanization than the Census or Indiapolis in mountainous states such as Mizoram, Nagaland, and Sikkim, as well as in Gujarat and Maharashtra. However, since Bihar, Uttar Pradesh, and West Bengal are very populous states, our estimates of higher urbanization in these states outweigh our lower estimates in the other states to create a higher estimate of national urbanization.

The overall effect is a higher proportion of the Indian population being urban than in the official figures. For 2011, at the 7.5 pp./ha threshold, we estimate overall urbanization at 43% (compared with the Census estimate of 31%). This amounts to a difference of 140 million people from the Census. This also suggests a much different urban hierarchy than the Census implies, with much larger proportions of the Indian population allocated into urban areas with greater than 10 million people and relatively fewer people in cities with less than 100,000 people.

2.4.3 Temporal change: urbanization between 2001 and 2011

We estimate a change of 4.7 percentage points in the proportion of urban population in India between 2001 and 2011 (see Table 2.3). This is not significantly different from other estimates such as Census (3.3 points) and Indiapolis (2.4 points). However, there is significant heterogeneity in estimates of rates of urbanization at the scale of the state. While the Census estimates significant urbanization in the south of India, it undercounts the rate of urbanization in Gujarat relative to Indiapolis (see Figure 2.8). According to our estimates, while southern Indian states have experienced higher urbanization rates, they are dwarfed by the urbanization rates in Uttar Pradesh, Bihar, and West Bengal. While these states have not traditionally been at the forefront of urbanization, they seem to be densifying quite rapidly in a way that is not being captured by the Census estimates. However, unlike Indiapolis, we do not estimate a marginal decline in the urbanization in the heavily urbanized state of Kerala (see Figure 2.8 and Table 2.3).

Table 2.3 Comparison of estimates of change in urbanization between 2001 and 2011 at the state level, as estimated by the Census of India, Indiapolis, and MAGPIE

	Census			**Indiapolis**			**MAGPIE (7.5 persons/ha)**		
State	**2001**	**2011**	**Diff. (%)**	**2001**	**2011**	**Diff. (%)**	**2001**	**2011**	**Diff. (%)**
Andaman & Nicobar Islands	33%	38%	5%	30%	33%	3%	15%	16%	1%
Andhra Pradesh	24%	30%	5%	36%	41%	5%	23%	26%	3%
Arunachal Pradesh	21%	23%	2%	16%	16%	0%	0%	0%	0%
Assam	13%	14%	1%	21%	22%	0%	16%	24%	8%
Bihar	10%	11%	1%	31%	36%	5%	63%	74%	11%
Chandigarh	90%	97%	7%	99%	99%	0%	100%	100%	0%
Chhattisgarh	20%	23%	3%	21%	21%	0%	18%	20%	2%
Dadra and Nagar Haveli	23%	47%	24%	44%	53%	9%	12%	26%	14%
Daman and Diu	36%	75%	39%	87%	95%	8%	121%	115%	− 6%
Delhi	93%	98%	4%	97%	97%	1%	100%	100%	0%
Goa	50%	62%	12%	57%	57%	1%	40%	45%	6%
Gujarat	37%	43%	5%	43%	53%	10%	30%	33%	3%
Haryana	29%	35%	6%	38%	43%	5%	37%	39%	2%
Himachal Pradesh	10%	10%	0%	8%	9%	0%	8%	7%	−1%
Jammu and Kashmir	25%	27%	3%	31%	31%	0%	23%	31%	7%
Jharkhand	22%	24%	2%	25%	25%	0%	28%	34%	5%
Karnataka	34%	39%	5%	38%	43%	4%	26%	31%	4%
Kerala	26%r	48%	22%	97%	96%	− 1%	77%	78%	1%
Lakshadweep	44%	78%	34%	34%	51%	17%	84%	93%	9%
Madhya Pradesh	26%	28%	1%	26%	27%	1%	17%	18%	1%
Maharashtra	42%	45%	3%	48%	51%	3%	35%	36%	1%
Manipur	25%	29%	4%	47%	52%	5%	24%	27%	3%

	Census			Indiapolis			MAGPIE (7.5 persons/ha)		
State	**2001**	**2011**	**Diff. (%)**	**2001**	**2011**	**Diff. (%)**	**2001**	**2011**	**Diff. (%)**
Meghalaya	20%	20%	0%	13%	5%	− 9%	11%	12%	0%
Mizoram	50%	52%	2%	45%	47%	2%	7%	6%	− 1%
Nagaland	17%	29%	12%	25%	29%	5%	10%	11%	1%
Orissa	15%	17%	2%	16%	17%	1%	20%	23%	4%
Puducherry	67%	68%	2%	74%	74%	0%	90%	93%	2%
Punjab	34%	37%	4%	37%	40%	2%	34%	32%	− 2%
Rajasthan	23%	25%	1%	26%	29%	3%	17%	18%	0%
Sikkim	11%	25%	14%	13%	18%	5%	3%	3%	0%
Tamil Nadu	44%	48%	4%	50%	53%	3%	40%	44%	4%
Telangana	32%	39%	7%	45%	52%	7%	30%	31%	1%
Tripura	17%	26%	9%	64%	63%	− 1%	21%	22%	1%
Uttar Pradesh	21%	22%	1%	25%	25%	1%	45%	55%	10%
Uttarakhand	26%	30%	5%	31%	35%	4%	23%	28%	5%
West Bengal	28%	32%	4%	47%	48%	1%	62%	69%	7%
India	28%	31%	3%	37%	39%	2%	38%	43%	5%

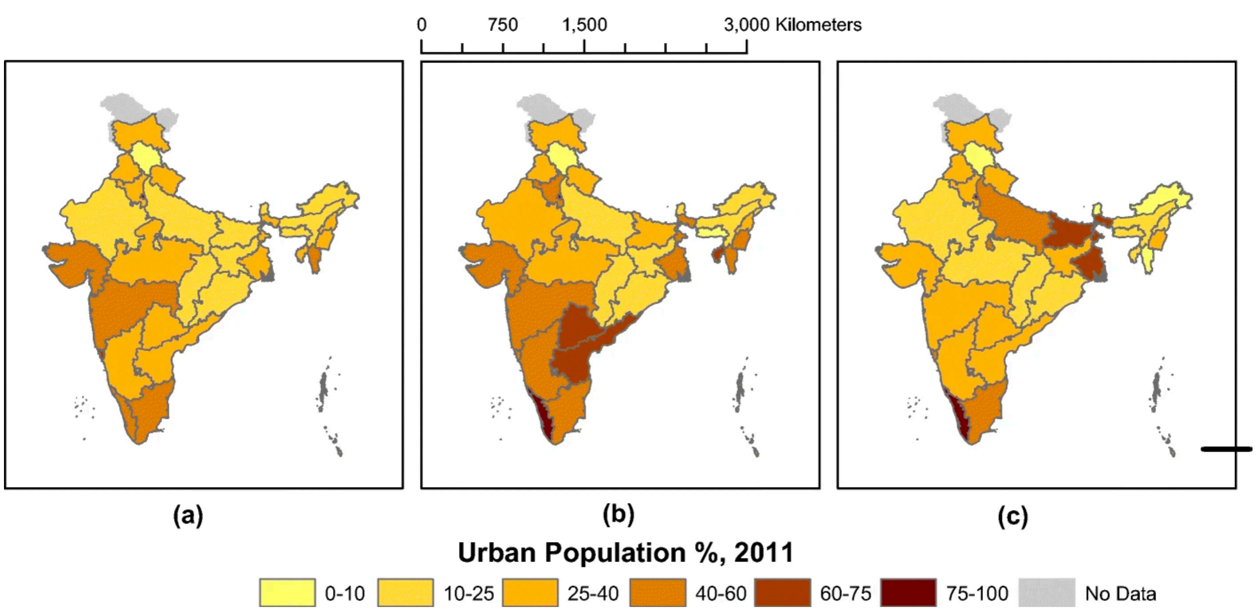


Figure 2.8. Comparison of the 2011 urbanization level by state. a Census of India. b Indiapolis. c MAGPIE (7.5 persons/ha)

The rapid urbanization in Uttar Pradesh, Bihar, and West Bengal is characterized by continuing population growth in small settlements. As these settlements grew in population between 2001 and 2011, they are more likely to pass the population density threshold set by MAGPIE. MAGPIE does not use a threshold on agricultural employment share (as does the Census) or settlement population size (as does Indiapolis). Thus, MAGPIE categorically classifies small, dense settlements as urban. So, larger shares of the population in areas with this development pattern are considered urbanizing as more settlements pass the density threshold and as these relatively small settlements experience population growth. In contrast, the Census measures rapid urbanization in Kerala, because it considered only 26% of the population urban in 2001. Both Indiapolis and MAGPIE considered most of Kerala's land area and population to be urbanized in 2001 already, so there is less potential for further urbanization.

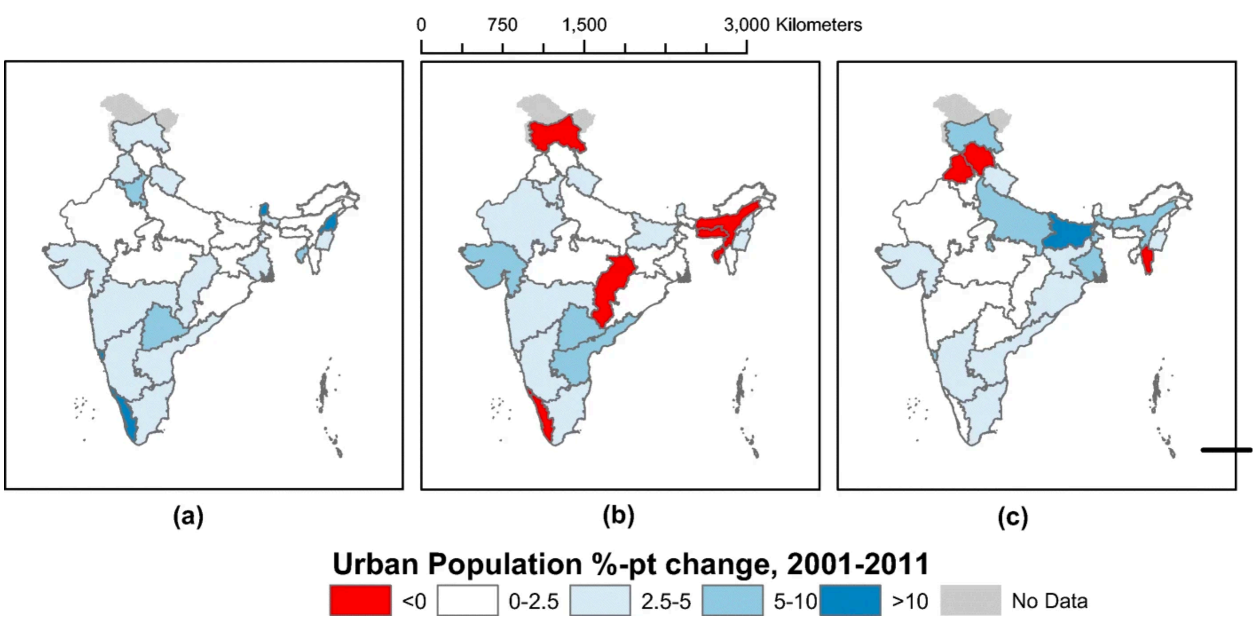


Figure 2.8. Change in urbanization by state, 2001–2011. (a) Indian Census. (b) Indiapolis. (c) MAGPIE estimates

2.5 Discussion

We estimate that in 2011, India's population was 43% urban, or 140 million more urban residents than estimated by the Census of India. MAGPIE places 18% of the total population and 48% of the urban population into very large, often polycentric urban agglomerations of greater than 10 million people while the Census of India considers the bulk of the urban population to be in mid-sized cities with populations between 100,000 and 1,000,000. While the Census' definition may characterize how these urban populations are administered, our method implies a much more spatially interconnected urban system, as well as divergent urbanization processes taking place in different regions of India. Like previous efforts to estimate India's urbanization without reference to gender and employment categories (Denis et al. 2012), the implications are that substantial investment in services to support life in dense settlements will be required, whether the official figures classify populations living within networks of proximate, small, and dense settlements as urban or rural. This problem is only partially addressed in Indian development planning. Perhaps the most visible related policy initiative is the National Rurban Mission begun in 2016, which aims to identify 300 "rurban" clusters of 20 villages each across the country and target each with a variety of local workforce training activities and urban amenities such as water, sanitation, public transport, and street lighting. However, much of the official documentation about the National Rurban Mission implies that the initiative is designed to facilitate the urbanization-in-place of villages, while many of the selected clusters are in fact spatially proximate to large urban agglomerations and could be considered peri-urban or suburban (Singh and Rahman 2018). Similarly, many of the state nodal development authorities in West Bengal, Bihar, and Uttar Pradesh, to the extent that they address development in gram panchayats, are generally targeted towards the fringes of urban agglomerations rather than networks of villages undergoing in situ densification independent of a large city. Some major exceptions include the Gangasagar Bakkhali Development Authority in West Bengal that covers much of the Hooghly River estuary (Gangasagar Bakkhali Development Authority 2019); the Patharchapuri, Barkreswar, Furfura Sharif, and Tarapith Development Authorities in West Bengal covering small rural regions covering 30–100 km2 (Urban Development Branch 2019); and the Kerala Local Government Service Delivery Project, allocating resources to for governance, capacity building, and infrastructure to all local governments outside of the six largest cities in the state (Local Self Government Department 2019).

We also make a methodological contribution to the problem of urban system definition. When provided with population counts at a sufficient spatial granularity, MAGPIE can rapidly and easily delineate urban areas and measure their populations in a much less labor-intensive process than the locality-based methods of e-Geopolis (the global project of which Indiapolis is part). While our implementation here depends on dasymetric population disaggregation, this method could be applied to other population grid products (such as Landscan or GHS-POP) to construct urban hierarchies that account for non-contiguous urban interconnectivity.

MAGPIE has an added advantage of automatically determining the edge of the city without relying solely on the contiguity criterion. The intuition behind this method is that urban areas can be non-contiguous at the edges (Schnieder and Woodcock 2008) and this method allows for them to become part of the urban region. This is similar but not identical to US Census delineation of urban areas using a "hop and jump" criteria to account for discontinuous urbanization (Ratcliffe et al. 2016). One way to account for discontinuous urbanization is to merge urban areas that are within a certain distance from one another. However, if we simply merge urban areas that are within a distance threshold, then there will be situations where urban extents will have tendril/tail or hourglass-like forms due to sparse connections at the edges or between two large urban areas. Tendrils are observed when urban development is caused by linear infrastructure expansions such as highways. MAGPIE allows for these tails to separate urban clusters as there are only a few edges between them. This method also would merge two large urban regions into one, only when there is sufficient number of smaller polygons that are in close proximity to both. Furthermore, if many small urban areas are non-contiguous, by virtue of them being close to one another, they can form an urban cluster and could be treated as a single unit.

Urbanity is a continuum and the standard dichotomy between urban and rural is not adequate to characterize the human settlement patterns and their changes (Hugo et al. 2004; Wratten 1995). However, because the level of urbanization is considered a proxy for development, we argue that consistent characterizations of urban boundaries are useful. Hugo et al. (2004) argue that settlements ought to be measured on different dimensions, including size, concentration, and accessibility within the region. While our method accounts for the first two characteristics explicitly, we do not account for the access characteristics, which should be addressed in the future.

Another limitation of MAGPIE as currently implemented is an inconsistent agglomeration of peninsular or island settlements into surrounding urban areas from which they are separated by water features. This is illustrated by the high-density, low population identified agglomerations of Old Kochi, Willingdon Island, Ramanthuruth, and Bengre, all of which are part of the cities of Kochi or Mangalore (see lower three rows of Figure 2.5).

Furthermore, our method of agglomerating urban settlements depends on the accuracy of the dasymetric disaggregation of census counts, which are only available to us, at the relatively coarse spatial unit. This could contribute to our method's production of large urban agglomerations over areas that are traditionally considered rural, if densely populated. This could also contribute to error in the other direction, as our disaggregation may allocate population growth that actually occurred in concentrated cities throughout the district in which a given city is located. Other data products have become recently available at finer geographic scale that could have improved the results (e.g., Balk et al. 2019; Meiyappan et al. 2018). However, they also suffer from poor spatial precision of spatial units and reconciling them to create temporally consistent units is an arduous task. In any case, all gridded population estimates depend crucially upon the underlying official district-level geographies and counts. Some limitations related to modeling approach could also affect this estimation. As RF is a tree-based estimator, it is restricted by the range of training. The prediction using RF model trained on district-level population density and the zonal mean of covariates will have lesser range and heterogeneity than the RF model trained on actual pixel scale population counts and covariate values. In other words, as the variability among the district-level population density is used to model the variability inside the districts, it will lead to less heterogeneous predictions with smaller variance. In this regard, the sensitivity of MAGPIE with the resolution of training census data needs to be evaluated. Forecasting future urbanization based on current non-linear relationships among the underlying covariates might be problematic. However, scenario-based forecasting that estimates future urbanization based on relationships among subsamples might provide some direction for future research. In addition to the errors associated with other environmental and remote sensing datasets, we ought to be mindful of this limitation.

2.5 Conclusion

In this paper, we show that definitional differences and seemingly innocuous choices of thresholds matter a great deal for the delineation and categorization of urban settlements. Not only the population thresholds matter but also the density cut-offs are important in distinguishing urban from rural. The density thresholds also affect the contiguity and delineation of urban areas, which in turn affect the total population thresholds. Depending on the density cut-off, 35% to 57% of India's population lives in "urban" areas in 2011 (contrast with 31% estimated by the Census of India). Additionally, about 5–7 million people live in about 20,000 distinct small towns (< 5000 population) with relatively high density. Highly dense regions in the Gangetic plain are contiguous enough to form large agglomerations. Furthermore, because we do not rely on the political and jurisdictional boundaries and instead rely on a contiguity criterion, our estimates on the number of medium-sized towns (less than 100k population) are significantly lower than the Census or the Indiapolis estimates by Denis and Zérah (2017). Instead, these small towns are coalesced into much larger urban agglomerations, thus changing the conclusions that can be drawn about the type and extent of urbanization in India. The differing boundaries of urban areas and much larger agglomeration of small towns can be attributed to the accuracy of the gridded surface. Still, the contiguity-based criterion provides a meaningful way to compare the urban agglomerations.

We find that the results from our method also challenge the idea about the declining urbanization rate in India. While there may be strong political and governance reasons for large and dense "rural" areas to be classified as urban, they pose a problem for comparative statistics. Our contribution lies in the methodology to harmonize the differences and provide a consistent characterization of urban across large regions. This work can be extended in a few ways. One extension could be analyzing the sensitivity of this approach with the change in the scale of training data and by using other types of gridded population dataset products. Another possible extension could be modeling multiple contiguous countries together with a different economic status and analyzing the difference in urbanization. This work demonstrates the importance of seemingly benign and arcane definitional matters to the measurement of urbanization. Recognizing them would help us fashion institutions and jurisdictions that are better aligned to manage urban growth.

REFERENCES

Abrahams, Alexei, Christopher Oram, and Nancy Lozano-Gracia. "Deblurring DMSP Nighttime Lights: A New Method Using Gaussian Filters and Frequencies of Illumination." *Remote Sensing of Environment* 210 (June 2018): 242–58. https://doi.org/10.1016/j.rse.2018.03.018.

Ahluwalia, Isher R., and P.K. Mohanty. "Challenges of Urbanisation in India: An Overview." In *Urbanisation in India: Challenges, Opportunities, and the Way Forward*, 1–28. New Delhi: SAGE Publications, 2014.

Angel, Shlomo, Jason Parent, Daniel L. Civco, Alexander Blei, and David Potere. "The Dimensions of Global Urban Expansion: Estimates and Projections for All Countries, 2000–2050." *Progress in Planning* 75, no. 2 (February 2011): 53–107. https://doi.org/10.1016/j.progress.2011.04.001.

"ArcGIS Desktop (Version Release 10.5)." Redlands, CA: ESRI, 2016.

Balk, D. "More than a Name: Why Is Global Urban Population Mapping a GRUMPy Proposition?" In *Global Mapping of Human Settlement: Experiences, Data Sets, and Prospects*, 145–61. Taylor and Francis, 2009.

Balk, Deborah, Mark Montgomery, Hasim Engin, Natalie Lin, Elizabeth Major, and Bryan Jones. "Urbanization in India: Population and Urban Classification Grids for 2011." *Data* 4, no. 1 (February 26, 2019): 35. https://doi.org/10.3390/data4010035.

Bivand, Roger, and Gianfranco Piras. "Comparing Implementations of Estimation Methods for Spatial Econometrics." *Journal of Statistical Software* 63, no. 18 (2015). https://doi.org/10.18637/jss.v063.i18.

Bloom, David E., David Canning, and Günther Fink. "Urbanization and the Wealth of Nations." *Science* 319, no. 5864 (February 8, 2008): 772–75. https://doi.org/10.1126/science.1153057.

Breiman, Leo. "Random Forests." *Machine Learning* 45, no. 1 (2001): 5–32. https://doi.org/10.1023/A:1010933404324.

Buettner, Thomas. "Urban Estimates and Projections at the United Nations: The Strengths, Weaknesses, and Underpinnings of the World Urbanization Prospects." *Spatial Demography* 3, no. 2 (November 2015): 91–108. https://doi.org/10.1007/s40980-015-0004-2.

Chauvin, Juan Pablo, Edward Glaeser, Yueran Ma, and Kristina Tobio. "What Is Different about Urbanization in Rich and Poor Countries? Cities in Brazil, China, India, and the United States." *Journal of Urban Economics* 98 (March 2017): 17–49. https://doi.org/10.1016/j.jue.2016.05.003.

Chen, Mingxing, Hua Zhang, Weidong Liu, and Wenzhong Zhang. "The Global Pattern of Urbanization and Economic Growth: Evidence from the Last Three Decades." Edited by Alejandro Raul Hernandez Montoya. *PLoS ONE* 9, no. 8 (August 6, 2014): e103799. https://doi.org/10.1371/journal.pone.0103799.

Cohen, Barney. "Urbanization in Developing Countries: Current Trends, Future Projections, and Key Challenges for Sustainability." *Technology in Society* 28, no. 1–2 (January 2006): 63–80. https://doi.org/10.1016/j.techsoc.2005.10.005.

Comber, A.J., C.F. Brunsdon, and C.J.Q. Farmer. "Community Detection in Spatial Networks: Inferring Land Use from a Planar Graph of Land Cover Objects." *International Journal of Applied Earth Observation and Geoinformation* 18 (August 2012): 274–82. https://doi.org/10.1016/j.jag.2012.01.020.

Corbane, Christina, Aneta Florczyk, Martino Pesaresi, Panagiotis Politis, and Vasileios Syrris. "GHS-BUILT R2018A - GHS Built-up Grid, Derived from Landsat, Multitemporal (1975-1990-2000-2014) - OBSOLETE RELEASE," July 18, 2018. https://doi.org/10.2905/jrc-ghsl-10007.

Csardi, Gabor, and Tamas Nepusz. "The Igraph Software Package for Complex Network Research." *InterJournal* Complex Systems (2006): 1695.

Denis, Eric, and Marie-Hélène Zérah, eds. *Subaltern Urbanisation in India: An Introduction to the Dynamics of Ordinary Towns*. Exploring Urban Change in South Asia. New Delhi: Springer India, 2017. https://doi.org/10.1007/978-81-322-3616-0.

Denis, Eric, Marie-Hélène Zérah, and Partha Mukhopadhay. "Subaltern Urbanisation in India." *Economic and Political Weekly* 47, no. 30 (2012): 52–62.

Desmet, Klaus, Ejaz Ghani, Stephen O'Connell, and Esteban Rossi-Hansberg. "The Spatial Development of India." *Journal of Regional Science* 55, no. 1 (January 2015): 10–30. https://doi.org/10.1111/jors.12100.

Dijktstra, Lewis, and Hugo Poelman. "A Harmonised Definition of Cities and Rural Areas: The New Degree of Urbanisation." European Commission Directorate-General for Regional and Urban Policy, n.d.

Duranton, Gilles, and Diego Puga. "Chapter 48 Micro-Foundations of Urban Agglomeration Economies." In *Handbook of Regional and Urban Economics*, 4:2063–2117. Elsevier, 2004. https://doi.org/10.1016/S1574-0080(04)80005-1.

Ellis, Peter, and Mark Roberts. *Leveraging Urbanization in South Asia: Managing Spatial Transformation for Prosperity and Livability*. Washington, DC: World Bank, 2016. https://doi.org/10.1596/978-1-4648-0662-9.

Engstrom, Ryan, David Locker Newhouse, and Vidhya Soundararajan. "Estimating Small Area Population Density Using Survey Data and Satellite Imagery: An Application to Sri Lanka." Poverty and Equity Global Practice Working Paper Series 194. Washington, D.C., 2019. http://documents.worldbank.org/curated/en/256241552483977593/Estimating-Small-Area-Population-Density-Using-Survey-Data-and-Satellite-Imagery-An-Application-to-Sri-Lanka.

European Commission. Joint Research Centre. *GHSL Data Package 2019: Public Release GHS P2019.* LU: Publications Office, 2019. https://data.europa.eu/doi/10.2760/062975.

Fay, Marianne, and Charlotte Opan. "Urbanization without Growth: A Not-so-Uncommon Phenomenon (English)." Research working paper. World Bank Group, n.d.

Ganapati, Sukumar. "Shrinking Cities: A Global Perspective." In *Shrinking Cities: A Global Perspective*, edited by Harry W. Richardson and Chang Woon Nam, 169–81. Regions and Cities 71. London; New York: Routledge/Taylor & Francis Group, 2014.

Gaughan, Andrea E., Forrest R. Stevens, Zhuojie Huang, Jeremiah J. Nieves, Alessandro Sorichetta, Shengjie Lai, Xinyue Ye, et al. "Spatiotemporal Patterns of Population in Mainland China, 1990 to 2010." *Scientific Data* 3, no. 1 (December 2016): 160005. https://doi.org/10.1038/sdata.2016.5.

GDAL/OGR Contributors. "GDAL/OGR Geospatial Data Abstraction Software Library." Accessed September 3, 2019. https://gdal.org/index.html.

Gelfand, A. E. "On the Change of Support Problem for Spatio-Temporal Data." *Biostatistics* 2, no. 1 (March 1, 2001): 31–45. https://doi.org/10.1093/biostatistics/2.1.31.

Ghani, Ejaz, Arti Grover Goswami, and William R. Kerr. "Spatial Development and Agglomeration Economies in Services--Lessons from India." Working Paper. Washington, DC: World Bank, June 2016. https://doi.org/10.1596/1813-9450-7741.

Hasan, Rana, Yi Jiang, and Radine Michelle Rafols. "Urban Agglomeration Effects in India: Evidence from Town-Level Data." *Asian Development Review* 34, no. 2 (September 2017): 201–28. https://doi.org/10.1162/adev_a_00100.

He, Mark, Joseph Glasser, Nathaniel Pritchard, Shankar Bhamidi, and Nikhil Kaza. "Demarcating Geographic Regions Using Community Detection in Commuting Networks with Significant Self-Loops." *PLOS ONE* 15, no. 4 (April 29, 2020): e0230941. https://doi.org/10.1371/journal.pone.0230941.

Henderson, Vernon. "The Urbanization Process and Economic Growth: The so-What Question." *Journal of Economic Growth* 8, no. 1 (2003): 47–71. https://doi.org/10.1023/A:1022860800744.

Hijmans, R.J. "Raster: Geographic Data Analysis and Modeling." Accessed September 3, 2019. https://CRAN.R-project.org/package=raster.

Hijmans, R.J., S. Cameron, J. Parra, P.G. Jones, A. Jarvis, and K. Richardson. "WorldClim." University of California, Berkeley, 2005.

Holt, D., D. G. Steel, M. Tranmer, and N. Wrigley. "Aggregation and Ecological Effects in Geographically Based Data." *Geographical Analysis* 28, no. 3 (September 3, 2010): 244–61. https://doi.org/10.1111/j.1538-4632.1996.tb00933.x.

Hugo, Graeme, Anthony Champion, and Alfredo Lattes. "Toward a New Conceptualization of Settlements for Demography." *Population and Development Review* 29, no. 2 (June 2003): 277–97. https://doi.org/10.1111/j.1728-4457.2003.00277.x.

Hutchings, Paul. "Community Management or Coproduction? The Role of State and Citizens in Rural Water Service Delivery in India," June 1, 2018. https://dspace.lib.cranfield.ac.uk/handle/1826/13409.

Indian Planning Commission. *Mid-Term Appraisal of Eleventh Five Year Plan, 2007-2012*. New Delhi: Oxford University Press, 2011.

IUCN and UNEP-WCMC. "The World Database on Protected Areas (WDPA)," 2010.

Kamath, L., and Yacoub Zachariah. "Impact of JNNURM and UIDSSMT/ IHSDP Programmes on Infrastructure and Governance Outcomes in Cities/ Towns in India." TISS Working Paper, 2015.

Khan, Sama. "The Other Jawaharlal Nehru National Urban Renewal Mission: What Does It Mean for Small Town India?" In *Subaltern Urbanisation in India*, edited by Eric Denis and Marie-Hélène Zérah, 337–70. New Delhi: Springer India, 2017. https://doi.org/10.1007/978-81-322-3616-0_13.

Klopp, Jacqueline M, and Danielle L Petretta. "The Urban Sustainable Development Goal: Indicators, Complexity and the Politics of Measuring Cities." *Cities* 63 (March 2017): 92–97. https://doi.org/10.1016/j.cities.2016.12.019.

Kundu, Amithab. "Politics and Economics of Urban Growth." *Economic and Political Weekly* 46, no. 20 (2011).

Lehner, Bernhard, Kristine Verdin, and Andy Jarvis. "New Global Hydrography Derived From Spaceborne Elevation Data." *Eos, Transactions American Geophysical Union* 89, no. 10 (2008): 93. https://doi.org/10.1029/2008EO100001.

Liaw, A., and M. Wiener. "Classification and Regression by RandomForest." *RNews* 2, no. 3 (n.d.).

Lloyd-Jones, Tony, and Carole Rakodi. *Urban Livelihoods: A People-Centred Approach to Reducing Poverty*. Routledge, 2014.

Local Self Government Department. "Kerala Local Government Service Delivery Project," 2019. http://klgsdp.org/.

Lung, Tobias, Tillmann Lübker, James K. Ngochoch, and Gertrud Schaab. "Human Population Distribution Modelling at Regional Level Using Very High Resolution Satellite Imagery." *Applied Geography* 41 (July 2013): 36–45. https://doi.org/10.1016/j.apgeog.2013.03.002.

McGranahan, Gordon, Deborah Balk, and Bridget Anderson. “The Rising Tide: Assessing the Risks of Climate Change and Human Settlements in Low Elevation Coastal Zones.” *Environment and Urbanization* 19, no. 1 (April 2007): 17–37. https://doi.org/10.1177/0956247807076960.

Meiyappan, P., P.S. Roy, A. Soliman, T. Li, P. Mondal, S. Wang, and A.K. Jain. “India Village-Level Geospatial Socio-Economic Data Set: 1991, 2001.” Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC), 2018. https://doi.org/10.7927/H4CN71ZJ.

Mennis, Jeremy. “Generating Surface Models of Population Using Dasymetric Mapping.” *The Professional Geographer* 55, no. 1 (2003): 31–42.

National Oceanic and Atmospheric Administration. “Version 4 DMSP-OLS Nighttime Lights Time Series,” 2019. https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html.

National Remote Sensing Centre. “Bhuvan. Indian Geo-Platform of ISRO,” 2019. https://bhuvanapp1.nrsc.gov.in/bhuvan2d/bhuvan/bhuvan2d.php.

Newman, M. E. J. “Modularity and Community Structure in Networks.” *Proceedings of the National Academy of Sciences* 103, no. 23 (June 6, 2006): 8577–82. https://doi.org/10.1073/pnas.0601602103.

Nieves, Jeremiah J., Forrest R. Stevens, Andrea E. Gaughan, Catherine Linard, Alessandro Sorichetta, Graeme Hornby, Nirav N. Patel, and Andrew J. Tatem. “Examining the Correlates and Drivers of Human Population Distributions across Low- and Middle-Income Countries.” *Journal of The Royal Society Interface* 14, no. 137 (December 2017): 20170401. https://doi.org/10.1098/rsif.2017.0401.

———. “Examining the Correlates and Drivers of Human Population Distributions across Low- and Middle-Income Countries.” *Journal of The Royal Society Interface* 14, no. 137 (December 2017): 20170401. https://doi.org/10.1098/rsif.2017.0401.

OECD. *Redefining “Urban”: A New Way to Measure Metropolitan Areas*. OECD, 2012. https://doi.org/10.1787/9789264174108-en.

Pesaresi, M., A. Florczyk, M. Schiavina, M. Melchiorri, and L. Maffenini. “GHS Settlement Grid, Updated and Refined REGIO Model 2014 in Application to GHS-BUILT R2018A and GHS-POP R2019A, Multitemporal (1975–1990–2000-2015), R2019A.” 2019-09-03, 2019. http://data.europa.eu/89h/42e8be89-54ff-464e-be7b-bf9e64da5218.

Pesaresi, Martino, Guo Huadong, Xavier Blaes, Daniele Ehrlich, Stefano Ferri, Lionel Gueguen, Matina Halkia, et al. “A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6, no. 5 (October 2013): 2102–31. https://doi.org/10.1109/JSTARS.2013.2271445.

Pradhan, K.C. “Unacknowledged Urbanisation: New Census Towns of India.” *Economic and Political Weekly* 48, no. 36 (2013): 43–51.

Python Software Foundation. “Python Language Reference,” 2013.

R Development Core Team. “R: A Language and Environment for Statistical Computing,” 2017.

Ratcliffe, M., C. Burd, K. Holder, and A. Fields. “Defining Rural at the U.S. Census Bureau.” Washington, D.C.: US Census Bureau, 2016.

Registrar General and Census Commissioner of India. “Census of India 2011: Provisional Population Totals-Urban Agglomerations and Cities,” 2012. http://censusindia.gov.in/2011-prov-results/paper2/data_files/India2/1.%20Data%20Highlight.pdf.

Satterthwaite, David. *The Transition to a Predominantly Urban World and Its Underpinnings*. London: Human Settlements Programme, IIED, 2007.

Schiavina, M., S. Freire, and K. MacManus. “GHS Population Grid Multitemporal (1975, 1990, 2000, 2015) R2019A,” 2019. https://data.jrc.ec.europa.eu/dataset/0c6b9751-a71f-4062-830b-43c9f432370f.

Schneider, A, C M Mertes, A J Tatem, B Tan, D Sulla-Menashe, S J Graves, N N Patel, et al. “A New Urban Landscape in East–Southeast Asia, 2000–2010.” *Environmental Research Letters* 10, no. 3 (March 1, 2015): 034002. https://doi.org/10.1088/1748-9326/10/3/034002.

Schneider, Annemarie, and Curtis E. Woodcock. “Compact, Dispersed, Fragmented, Extensive? A Comparison of Urban Growth in Twenty-Five Global Cities Using Remotely Sensed Data, Pattern Metrics and Census Information.” *Urban Studies* 45, no. 3 (March 2008): 659–92. https://doi.org/10.1177/0042098007087340.

Singh, Chandni, and Andaleeb Rahman. “Urbanising the Rural: Reflections on India’s National Rurban Mission: Urbanising the Rural.” *Asia & the Pacific Policy Studies* 5, no. 2 (May 2018): 370–77. https://doi.org/10.1002/app5.234.

Small, C, F Pozzi, and C Elvidge. “Spatial Analysis of Global Urban Extent from DMSP-OLS Night Lights.” *Remote Sensing of Environment* 96, no. 3–4 (June 30, 2005): 277–91. https://doi.org/10.1016/j.rse.2005.02.002.

Spence, Michael, Patricia Clarke Annez, and Robert M. Buckley. “Urbanization and Growth : Commission on Growth and Development.” Washington, DC: World Bank, 2009. https://doi.org/10.1596/978-0-8213-7573-0.

Stevens, Forrest R., Andrea E. Gaughan, Catherine Linard, and Andrew J. Tatem. “Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data.” Edited by Luís A. Nunes Amaral. *PLOS ONE* 10, no. 2 (February 17, 2015): e0107042. https://doi.org/10.1371/journal.pone.0107042.

Stewart, B., and C. Deuskar. “Measuring Global Urbanization Using a Standard Definition of Urban Areas: Analysis of Preliminary Results.” Presented at the World Bank Land & Poverty Conference, Washington, D.C., n.d.

Sudhira, H.S., and Gururaja K.V. “Population Crunch in India: Is It Urban or Still Rural?” *Current Science* 103, no. 1 (2012): 37–40.

Uchida, Hirotsugu, and Andrew Nelson. “Agglomeration Index: Towards a New Measure of Urban Concentration.” Washington, DC: World Bank, 2009. https://openknowledge.worldbank.org/handle/10986/9039.

World Bank. “Urban Population,” 2018. https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS.

Wratten, Ellen. “Conceptualizing Urban Poverty.” *Environment and Urbanization* 7, no. 1 (April 1995): 11–38. https://doi.org/10.1177/095624789500700118.

Zerah, Marie-Helene, Eric Denis, and P. Mukhopadhyay. “Subaltern Urbanization Revisited.” *ICC Quarterly*, no. 43 (n.d.): 28–41.

Zhou, Weiqi, Ganlin Huang, and Mary L. Cadenasso. “Does Spatial Configuration Matter? Understanding the Effects of Land Cover Pattern on Land Surface Temperature in Urban Landscapes.” *Landscape and Urban Planning* 102, no. 1 (July 2011): 54–63. https://doi.org/10.1016/j.landurbplan.2011.03.009.

CHAPTER 3. FROM INTERMITTENT TO CONTINUOUS WATER SUPPLY: WHAT IS THE EFFECT ON WATER DEMAND?

3.1 Introduction

How do service-level standards impact household consumption decisions? How do service providers balance standards that may reflect tradeoffs, such as meeting goals of inclusive coverage and at the same time managing the costs of providing upgraded levels of service? In this chapter, I consider a type of urban infrastructure investment project that has become prominent over the past 10 years in low and middle-income countries and is directly motivated by a model of infrastructure upgrading and commercialization: the conversion of prevailing intermittent water supplies to 24x7, or continuous service. An important premise of the policies that promote this model is that in addition to upgrading service levels, continuous service can reduce water demand, thus making more water available for unserved households. In this chapter, I estimate the impact of such a project on residential water demand. I ask the extent to which the introduction of continuous water supply in an intermittently operating system changed residential consumption of piped water, and to what extent does this change depend on existing customers' previous water consumption levels, and use of storage and alternative water sources.

3.2 Background

There are long-standing and continuing disparities in the quality of water and sanitation services available to citizens in the Global South, despite considerable investments made by governments and multilateral aid organizations since the 1950s to improve service networks (WHO, 2015). These gaps include lack of access altogether to piped water and safe sanitation by the poor in urban and rural areas, as well as inadequate service levels characterized by poor reliability and inadequate water quality for those already connected to piped networks (Lee and Schwab 2005). WHO & UNICEF currently estimate that 578 million urban residents (15 % of the global urban population) do not have access to clean, reliably supplied water in the home. Moreover, 2.38 billion people (60% of the global urban population) lack access to sanitation systems that adequately treat and/or remove excreta from the home. Governments, financial institutions, donors, civil society groups, and academics have advanced various perspectives explaining why these gaps persist, along with similarly varied strategies to address them.

A prominent view in the policy literature is that water is a scarce resource that should be treated as an economic good. This view was encapsulated by Principle 4 of the 1992 Dublin Statement on Water and Sustainable Development r, which was officially adopted by many international organizations and governments. Principle 4 highlights an enduring tension in the theory and practice of closing water and sanitation gaps: "Water has an economic value in all its competing uses and should be recognized as an economic good. Within this principle, it is vital to recognize first the basic right of all human beings to have access to clean water and sanitation at an affordable price…."

This principle of water as a basic right was and remains controversial, with many viewing water as an excludable, rivalrous, private good most efficiently provided when users pay at least the full costs of service (Rogers, De Silva, and Bhatia 2002). Others reject the implied commodification of water when viewed as a private good, on a variety of technical, ethical, and spiritual grounds, and advance alternative visions for types of state, collective, and community management of water (Bakker, 2010). Those with the view of water as a market good tend to consider the main problem to be a lack of efficiency and accountability of water utilities as organizations. Those with the view of water as a collective good also place significant importance on how providers can function sustainably and with responsibility.

This study takes a fresh look at both these views to provide empirical evidence on how, and under what conditions providers may engage with both goals – of achieving equity (in terms of coverage and access) and sustainability of service through improvements in supply efficiencies. I provide a conceptual framework in Figure 3.1 to orient and guide this chapter. Figure 3.1 represents the problem of the "vicious cycle" of water supply (see Spiller & Savedoff, 1999). The cycle is a situation where given poor cost recovery of services provided, water utilities have insufficient revenues to cover the costs of maintenance and service expansion, leading to poor service, generating low willingness to pay for piped water, which in turn leads back to insufficient revenues. Less than full cost-recovery from users undermines effective and efficient service delivery.

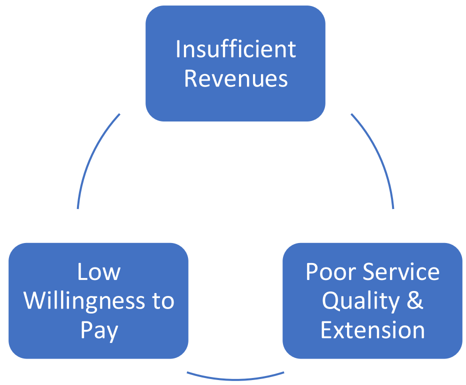


Figure 3.1. The vicious cycle of poor urban water service (from Spiller and Savedoff (1999))

Policy recommendations associated with this framework of supply are oriented around reforming the institutions governing water – primarily utilities -- such that they augment access by behaving more like commercial enterprises -- improving efficiencies, responding to consumer willingness to pay in order to secure the revenues to maintain and expand service (Spiller and Savedoff 1999; World Bank 2003b). Reforms in this line of thinking include, but are not limited to:

- separating utilities from political control through privatization or ring-fencing of finances from local government.
- metering connections.
- charging sufficient volumetric tariffs (prices) for water to at least cover operations costs and ideally debt service for capital expenditures.
- investing in service improvements to those willing to pay the costs of those improvements (Baietti, Kingdom, and van Ginneken 2006). This can include, for example, upgrading intermittent water delivery to continuous supply for higher tariffs.

The reform pathway implied by the “vicious cycle” is that simultaneous investments in service improvements for households willing to pay, volumetric pricing at cost-recovery levels, and political and financial separation of utilities from local governments (who can potentially direct water revenues to other politically pressing uses) will allow utilities to build the resources required to expand service over time.

Starting in 2006, the government of India, like several other rapidly urbanizing countries, adopted the policy of encouraging cities to upgrade drinking water supplies from intermittent to continuous service based on just such a premise. Drawing on a unique dataset from the Indian city of Amravati, Maharashtra, which implemented a continuous water supply pilot in 2010, I provide new empirical evidence on an aspect of this premise: how do service-level upgrades impact household demand for water and their willingness to pay higher tariffs? The rest of the paper is organized as follows. I explore the debates in the literature on intermittent and continuous water supply (CWS) and the benefits and challenges of shifting to CWS (section 3.3). I then present my research methods and describe the study site, the dataset, and the statistical model I use (section 3.4). This is followed by a discussion of the results in conversation with the relevant literature (section 3.5). I conclude with policy implications and avenues for future work (section 3.6).

3.3 Literature Review

Intermittent water supply (IWS) is a service mode where water is available from taps, and is pressurized in all pipes in the network, for less than 24 hours per day. Continuous water supply (CWS), where water is continuously pressurized and available at all times, is considered to be the ideal operating mode for a piped water service and is the norm in most areas in high-and middle-income countries. Indeed, no modern water system is designed to operate intermittently (Galaitsi et al., 2017). However, IWS remains common in many low-income countries. It is the dominant mode of supply in India, affecting at least 300 million people (almost all urban residents with access to piped water (McKenzie and Ray 2009)) and at least an additional 300 million throughout the rest of world (Kumpel and Nelson 2016). Water utilities often operate IWS as a deliberate rationing strategy in systems that lack either the bulk water supply capacity or piped infrastructure capacity to supply pressurized water continuously. Alternatively, utilities can operate intermittently simply due to uncontrolled leakage and interactions with electricity supplies that are also intermittent (Lee and Schwab 2005). There is a general debate in the literature as to the desirability of moving from IWS to CWS, with its effects on water resources sustainability, equity, and affordability often in contention, although empirical work is limited as to the conditions under which CWS affects outcomes along these criteria in either direction.

IWS has several disadvantages. The engineering and public health literatures document problems of water quality associated with IWS. Since no pipe is perfectly sealed from the environment, periods of time where pipes are not full of water represent opportunities for contamination of the interior of pipes, leading to the delivery of unsafe water, even if it has been treated before distribution (Kumpel and Nelson 2014). CWS has been shown to improve water quality at the tap and reduce waterborne disease incidence relative to IWS in India (Ercumen et al. 2015; Kumpel and Nelson 2013), although effects are more ambiguous at the point of use due to the possibility of recontamination in storage containers (Kumpel and Nelson 2013). Many waterborne disease outbreaks in high and middle-income countries have been directly linked to episodes where CWS systems had service interruptions (Ercumen, Gruber, and Colford 2014). Bivins et al. (2017) estimate that over 17 million annual infections of waterborne disease are attributable to IWS.

IWS is also criticized as an unreliable service that reduces the quantity of water available for consumption and is considerably less convenient than CWS. IWS can also be more costly for consumers as it necessitates coping mechanisms such as underground and overhead tanks, pumping, and investment in alternative water supplies such as private wells and tanker-truck delivery services (Zérah 2000). These expenditures may even exceed the cost of providing adequate, continuous, clean, piped water service (Atlaf 1994; Pattanayak et al. 2005). Intermittent service, moreover, often disproportionately burdens the poorest households (if they are even connected to the piped network). They invest the most time supplementing water by waiting at their taps during supply hours, traveling to and waiting to collect it from public standposts, or pay high unit prices for water from street vendors.

For these reasons,"24x7" water supply or CWS is an aspiration for many urban water utilities and is often seen as an *a priori* goal (Galaitsi et al., 2017). In India in particular, which may represent most of the global population with piped connections to IWS (Kumpel and Nelson 2016), and where no urban area provides water on a continuous basis, "24x7" has become a policy priority (Ahluwalia et al., 2011). This priority is embodied in the adoption of "Service Level Benchmarks" for water supply including CWS as a goal, and the allocation of central and state government transfers for this purpose (Ministry of Urban Development (Government of India), 2014; The World Bank, 2012). In government and donor initiatives that track and prioritize funding, “average service hours” or “service continuity” has become a key indicator, with numbers closer to 24 being considered better (Jaladhi et. al. 2016).

As of 2012, 43 cities in India were actively pursuing 24x7 water supply in at least some areas of the city (Mitra 2012). The following benefits of converting to CWS or 24x7 are repeatedly stated in the policy literature both within India and internationally (see Charalambous & Laspidou, 2017; Rana, 2013; World Bank, 2010; World Bank, 2003a):

1. *24x7 supply delivers better quality water for public health:* High levels of bacterial contamination are experienced in the first 10 minutes of re-pressurization of an intermittent system. The contaminated water ends up in the consumers’ storage tanks, which become also highly contaminated. Maintaining full pressure removes that risk. 24 x7 supply gives significantly better service to all customers.
2. *24x7 supply revolutionizes service to the poor:* Consumers can access more water for improved health and hygiene while saving time in queuing and carrying, and gainfully using the time thus saved for employment opportunities
3. *24x7 supply converts household coping costs into resources for the service provider*: Coping costs that consumers need to incur are reduced; they pay for a better service
4. *24x7 supply reduces the burden on water resources*: Continuous supply reduces water wastage arising from overflowing storage systems and open taps. It saves on household stored water that is discarded when new supply comes in. Because the network is renewed where needed, it also reduces losses arising from leaks.
5. *24x7 delivers effective supply management and demand management*: CWS makes possible the effective management of leakage through pressure management and flow measurement. Water conservation is also encouraged through metering and price signals via volumetric tariff to consumers
6. *24x7 supply enables improved efficiency of service provision*: Operational efficiencies are achieved because of a reduced need for valve-men, and a conversion of these jobs into more efficient ones of meter reading and customer care. It also makes possible the management of illegal connections.

The theory of change underlying these claims is that improvements in water supply hours will reduce water contamination, leakage and operations and maintenance costs while increasing willingness-to-pay, which can be used to justify increases in water tariffs. The proceeds of these higher tariffs, as well as the increased water availability from reduced leakage, can then be used to increase investment in service expansion over time. Thus, the IWS to CWS conversion has become an aspirational pathway to transform water utilities with poor service and low levels of coverage into utilities that eventually provide universal services comparable to those in high-income countries, thus reversing the ‘vicious cycle’ of supply. The “24x7” aspect of water service is conceived of by proponents as part of a broader suite of management reforms including metering, aggressive leak detection, volumetric pricing, and the removal of standposts and illegal connections and replacing them with household connections. Thus, equity (in that CWS will benefit the poor the most, as IWS water supply is often inadequate, if even accessible to the poor at all) is achieved through the greater efficiency that operating a CWS system results in.

However, setting aside the positive effect of CWS on the quality of water delivered at the tap, many doubt the universal suitability of CWS for cities in low-income countries. Critics of CWS water supply interventions in India focus on both the purported behavioral changes and the institutional changes associated with 24x7 supply along with the more technical changes that are needed to deliver continuous service. For example, the success of CWS conversion could depend on how consumers change how much water they demand from the system, how much revenue is required to maintain that level of service given other organizational revenue requirements, and the extent to which expanding service is required by laws or contractual arrangements.

These criticisms address both issues of equity and sustainability. In terms of sustainability, a common criticism is that CWS will result in higher water consumption by those that are already connected in the short and long term, leading to exhaustion of limited water resources. This not only makes this strategy unsustainable in the context of supply constraints; it also leaves less water available for service extension to unconnected, and usually comparatively poor residents (Sangameswaran, Madhav, and D’Rozario 2008).

In terms of equity, a concern is that the upgrading to CWS which requires large expenditures to improve service, may end up benefiting only a minority of residents, most of whom are already connected, and are often among the richest of those in the urban area already, exhausting the revenues needed for extension of service to others (Mitra 2008). Criticism is also leveled at "24/7" projects for their role in the commodification and privatization of urban water through metering, tariff increases, and elimination of free public water sources, as led by the priorities of transnational institutions in partnership with national governments, in opposition to local priorities and conceptions of water as a non-market good or a human right (Mitra, 2008, 2012). Implicit in this criticism is the prediction that the components of 24/7 projects involving metering, volumetric pricing, and removing public standposts would render not only piped water, but any clean water, unattainable for the bulk of urban poor (Sangameswaran, Madhav, and D'Rozario 2008; Bommier and Renouard 2014). For instance, in Nagpur, Maharashtra, a pilot CWS project removed public standposts from areas receiving CWS, while providing subsidized water connections to only "notified" or officially recognized settlements, leaving non-notified informal settlements with no legal recourse to access network water (Bommier & Renouard, 2014). In this view, CWS is only a technical fix, and with commercialization as the attendant management solution, is incapable of leading to equitable or efficient outcomes. Burt, Billava, & Ray (2018) integrate the equity and sustainability criticisms in their analysis of the CWS pilot in Hubli-Dharwad, Karnataka. They argue that without alternatives to private connections such as standposts, unconnected households embody a risk for illegal connections and threaten the sustainability of the upgrade.

Proponents of "24x7" water in India point to the success of commercial water reforms in Phnom Penh, Cambodia, a city that is poorer than many Indian cities. Over 20 years, the water utility there successfully transitioned from an unreliable, unsafe system delivering water to a small portion of the city to a near-universal CWS service generating enough profits to subsidize connections for its poorest households and even finance loans to other government entities (Biswas & Tortajada, 2010; Chan, 2009)

There is thus a debate in the policy literature, wherein two broad bodies of theory present competing narratives about how 24x7 investments can be integrated within water systems, and what the results might be for consumers and utilities. These two approaches are simplified and visualized in Figure 3.2:

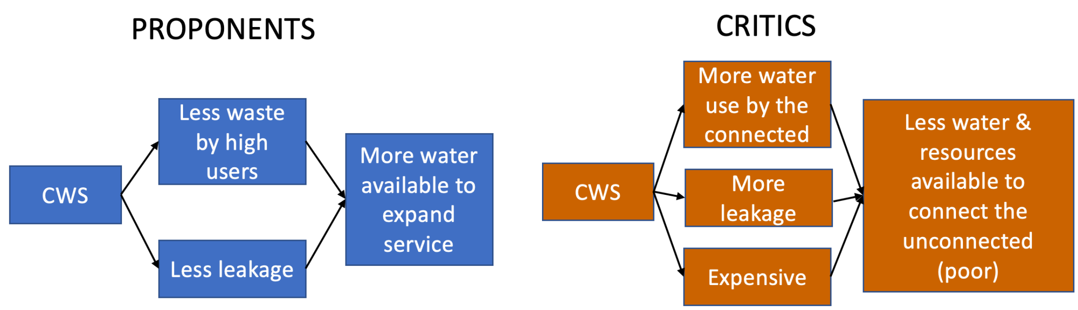


Figure 3.2. Competing theories of the results of converting from IWS to CWS

In the face of these competing logics, sound empirical evidence about the conditions under which each scenario is likely to proceed from CWS investment is scarce. There are very few studies that explicitly compare an IWS scenario to CWS scenarios in the same setting. I discuss prominent findings from some key studies below.

Christodoulou & Agathokleous (2012) found that when drought led water utilities in Cyprus to ration water and reduce service from 24 hours per day to 12 hours every two days, water consumption fell by 9%, while incidences of pipe breaks and operating costs rose dramatically. This suggests that IWS does indeed constrain water consumption from a piped network, while increasing maintenance and operating costs. In the opposite direction, Hastak, Labhasetwar, Kundley, & Gupta (2016) compared aggregate billing records in a neighborhood of Nagpur, India before and after its conversion from IWS to CWS, finding an average 3.6% rise in water billed per customer, even while total water consumption by the utility per household fell by 26%, suggesting large leakage reductions accompanying CWS. By contrast, Andey & Kelkar (2009) found mixed results in a comparison of water consumption in four neighborhoods across India that converted from IWS to CWS. They found that water consumption rose dramatically in cities where intermittency was severe enough to constrain water availability, while households that could collect large amounts of water from the system under IWS already, did not see their consumption rise after CWS was implemented. However, this study relied on small sample sizes given the variability in water demand between households and was thus less conclusive.

On issues of water quality and willingness to pay, many studies suggest a positive link between service quality and willingness to pay, although most empirical work focuses on contingent valuation surveys of households with intermittent or otherwise poor-quality water supplies, and few have observed actual outcomes of a transition to CWS (e.g. Altaf & Hughes, 1994; Dutta, 2005; Grazhdani, 2015; Whittington, Pattanayak, Yang, & Kumar, 2002). These studies tend to find greater willingness to pay for greater water services reliability than water quality or continuity *per se*. Other studies use cross-sectional surveys between customers of varying service quality and measure their bill payment rates or accumulated arrears, and generally find that customers more frequently miss payments and outright refuse to pay bills in areas where service is less reliable. These studies also find that willingness to pay is related to perceived cognitive and time costs of bill payment and short-term financial constraints (e.g. Mugabi, Kayaga, Smout, & Njiru, 2010; Vasquez, 2015).

The most-studied CWS conversion project is that of Hubli-Dharwad, Belgaum, and Gulbarga, three cities in Karnataka, India. This project remains the largest 24x7 project studied to date and is notable for being implemented by a private operating concessionaire. The project converted 10% of households from an unreliable intermittent supply to CWS in 2008 but has not been significantly scaled up (Buurman and Santhanakrishnan 2017). Here the billed water volumes of those with CWS were three times higher than those with IWS, although water bills were only available from small portions of studied households (Burt, Billava, and Ray 2018). There was also little evidence for wasteful water use by IWS customers in the form of draining unused water stored in water tanks before the next round of intermittent water was delivered, or leaving taps open during supply hours (Kumpel et al. 2017). Overall water demand from the areas served with CWS is on track to outstrip the bulk water resources currently available (Jayaramu, Burt, and Kumar 2015).

Satisfaction among those with CWS was high, and there were significant benefits to the utility and their customers (Burt, Billava, and Ray 2018), although households continued to store water due to habit or convenience, and nonpayment of bills was as high as 60% (Burt and Ray 2014; Buurman and Santhanakrishnan 2017). Moreover, these benefits accrued disproportionately to high-income households, with the time savings of low-income households being outweighed by their increased water bills (Burt, Billava, and Ray 2018). The net present value of the pilot project hinges crucially on whether the costs of an associated supply augmentation are attributed solely to the CWS pilot areas or the entire urban area (including those that saw no benefits), and it seems that bulk water savings from the project were quite modest after accounting for increased demand in CWS areas (Burt, Billava, and Ray 2018). The costs were high, with complete replacement of the piped network in CWS areas. Meanwhile, services for the 90% of the population that are still on IWS have not improved in years in spite of tariff increases (Commission, 2011).

In terms of the planning and implementation process of CWS project, the Hubli-Dharwad initiative has been criticized as being undemocratic (Mitra 2008; Sangameswaran, Madhav, and D'Rozario 2008; Walters 2013). However, it has also been praised for certain pro-poor policies including tariff and connection subsidies, as well as allowing connections to households with insecure land-tenure (Jayaramu, Devadiga, and Kumar 2015)

Overall from the few cases of CWS conversion that have been studied, it seems clear that CWS projects should be evaluated on a case-by-case basis, as initial conditions, costs, water resource constraints, customer reactions, and policies that may or may not be within utility control can interact in ways that produce variable outcomes. In addition, all previous studies of CWS have either been cross-sectional designs or pre/post designs without control groups. There has been no study measuring the impact of changing from IWS to CWS on customer behavior that uses both spatial and temporal variation in the intervention. In the study that follows below, I present empirical evidence to the question of CWS with a thorough analysis of several factors. This study is the first of which I am aware to leverage data following the same water customers before and after they were converted from IWS to CWS, and after CWS was reverted to IWS, as well as customers in the same setting that were never converted to CWS. Thus, in a single research design, I compare trends in consumption between households moving from IWS to CWS and back again and those that remained on IWS throughout the study period.

3.4 Methods

3.4.1 Research Questions

The purpose of this study is to estimate the effect that moving from intermittent to continuous water supply in Amravati, and back again, had on the demand for piped water of residential customers, and how this effect interacted with (a) initial water demand before any improvements were made to the system, and (b) previous use of coping mechanisms (storage tanks) and access to alternative water supplies (private wells). This study is of interest to scholars interested in urban water supplies in a number of disciplines, including civil engineering, public health, and urban planning. It contributes to our understanding of the conditions under which improvements in service hours can result in changes in water consumption at metered connections, which has important implications for the feasibility of such projects that face high capital costs for augmenting water supply. By including a matched control group of IWS households that were never converted to CWS but experienced improvements in IWS (improved reliability and water pressure) that were made for all residents as part of the CWS rollout, I can reflect on the perceived benefits to customers of improved IWS versus continuous water.

3.4.2 Study Site and Research Design

This study leverages a quasi-experimental design, using a difference-in-differences identification strategy involving the temporary upgrading of a portion of a mid-sized Indian city from intermittent to continuous water supply that began in 2009. Four of 16 operating zones were converted to CWS between 2009 and 2011. In 2012, one zone was allowed to lapse back to IWS for several reasons. This allows me to compare the consumption behavior of the *same* households before and after conversion to CWS and then after they returned to IWS; allowing for the comparison of trends in behavior between those that received CWS and those that did not. This section explains the research design and lays out the context of water services in my study site, using information from government documents and personal communications with water utility staff and their regulators.

Amravati is a city with a population of 646,000 as of the 2011 Census of India, located in the eastern part of the state of Maharashtra. Its water system is owned and operated by the Maharashtra Jeevan Pradhikaran (MJP), the state public health engineering department. MJP used to operate all (251) urban water utilities in the state other than those in Greater Mumbai, until 1996 when all but 25 were devolved to local government management. Amravati is the largest of the remaining cities whose water supply is controlled by MJP. MJP sets common water tariffs for all water systems that it operates. In addition, MJP uses a cross-subsidy scheme, whereby all revenue from its systems is centralized and redistributed to cover local operating cost shortfalls and capital works across the projects it manages. Amravati is one of MJP's profitable systems.

Beginning in 2008, the State government of Maharashtra began a comprehensive reform of its water and sanitation sector. The reform, called the MSNA[3], aimed to improve service quality and coverage levels by making the water utilities in the state more efficient and effective, primarily through a more commercial orientation. The MSNA provided grant funding for infrastructure improvement and capacity building conditioned on management reforms including comprehensive metering and cost-recovery based on volumetric tariffs. Previously, metering in Maharashra's utilities (and throughout India) was sparse, and tariffs were generally collected as flat monthly fees that did not recover operating costs.

As part of the MSNA, MJP entered into a Water Operator Partnership with a Malaysian water utility, through which CWS was implemented in Badlapur, a town of 40,000 near Mumbai. Following this learning experience, MJP implemented CWS in another town, Malkapur, on its own. After this success, MJP received funds for a CWS pilot project in its largest utility, Amravati, where water had traditionally been supplied for two hours per day on average.

In 2009, MJP conducted a census of every building in Amravati, collecting basic demographic information and data on water infrastructure and usage, in order to complete comprehensive meter installation, the regularization of existing illegal connections, and to project water demand. MJP also fully digitized its water network maps (the network had been laid in 1994) and used computer simulation models to plan and implement several infrastructure improvements. These included the construction of elevated storage reservoirs (ESRs) throughout Amravati, dividing the network into 16 operating zones around these reservoirs, and further subdividing these zones into smaller District Metering Areas (DMAs). Following this subdivision, MJP could identify the pipes, pumps, and valves in most immediate need of repair. After doing so, Amravati's water supply throughout the city improved, delivering water 2 hours

[3] Maharashtra Sujal Nirmal Abhiyan (Maharashtra Water Improvement Mission)

each morning and 2 hours each evening with sufficient pressure to reach elevated storage tanks up to two stories.

Following six months of further modeling and preparation, in June 2010 MJP, converted the two least leakage-prone operating zones in the city to CWS using a least-cost first strategy. While many criteria were weighted, the four zones corresponded to those with the highest proportion of metallic (as opposed to asbestos-cement) water mains, lowest number of standposts, and lowest non-revenue water rates. All standposts in these zones were removed, and connections were offered to (only) slum households willing to pay connection fees. At this time, the flat-rate tariff that was in force city-wide was raised to an increasing block tariff. Two more zones were converted to CWS in June 2011. In July of 2012, one of these new CWS zones developed a large leak in its ESR. Without sufficient funding to repair this problem, this zone was reverted to the two-hour morning/ two-hour evening IWS rotation. A second tariff increase occurred in July of 2012. By mid-2014, there had been a change in leadership at MJP, and rising overall leakage rates and demand for new connections resulted in the decision to revert all of the zones back to IWS.

3.4.3 Data

This study primarily uses the aforementioned census of Amravati water services and infrastructure carried out by MJP in 2009, as well as MJP's water billing database. MJP has provided me with their census (which is continually updated as new customers are connected), billing records, and records of their tariff structures from October 2009 through March 2013. They generously provided these data under a confidentiality agreement in July of 2013.

Census

The census includes the following information on all potential customers in Amravati, whether or not they are MJP customers:

- Building type (house, slum tenement, Apartment complex, commercial building, industrial facility, school, hospital, etc.)

- Operating Zone (the storage reservoir from which the building would ultimately receive water if connected)
- Number of residents (where applicable)
- Monthly family income (3 categories (Rs.): <100,000; 100,000-300,000; >300,000)
- Whether the male head of household is employed, and in what sector (government, private, self, informal)
- Whether the female head of household is employed and in what sector
- Education level of the male and female head of household (8 categories from non-literate through PhD)
- Building footprint
- Garden (yard) footprint
- Whether the structure is considered "slum"
- Whether the structure had an illegal water connection
- Whether the structure accesses water from, and the monthly expenditures on water from:
    - A handpump
    - A motorized borewell
    - An open well
    - Tanker trucks
    - Private vendors
    - Public standposts
    - An "other water source" remarks field describing the any other water sources used (most commonly including neighboring households, businesses, and institutions)
- Approximate daily water consumption for:
    - Cooking
    - Bathing
    - Washing Clothes
    - Cleaning

    - Outdoor uses
- Whether a diarrheal disease was suffered by a household member in the last year

The census also includes this information for water customers.

- Number of connections to the piped network
- Number of indoor and outdoor water taps on the property
- Type of connection (retail, bulk, nondomestic, institutional, other)
- Meter size
- Distance of the structure to the nearest water main
- Opinion of the water pressure (Low, Medium, High)
- Opinion of the water quality (Satisfactory, Non-satisfactory)
- Opinion of water quantity (Excess, Adequate, Insufficient)

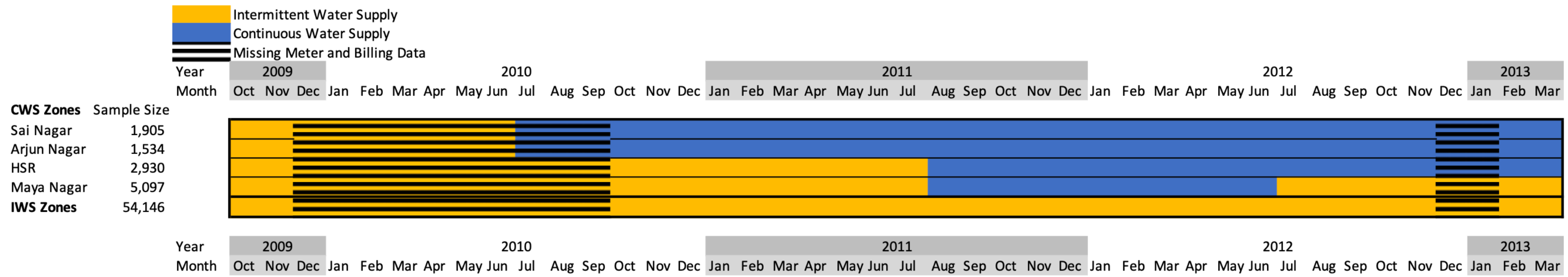
Billing records

MJP's billing records consist of meter readings and associated water bills and payment records for all customers. 99% of customers are metered. Meters are read on a bimonthly schedule, with about 70% of meters being read for a given billing period. The remaining 30% are instances of broken or inaccessible meters. The billing records are matched with a common unique identifier to the census. Each billing record includes the following information

- Previous and current meter read dates
- Measured volume
- Whether the meter was read, and if not, the reason why (meter inaccessible, meter broken, etc.)
- The total bill
- The date of bill delivery
- The amount of the bill that was paid
- The date the bill was paid

Unfortunately, billing records for certain bimonths are missing throughout the study period due to data management problems at MJP's Amravati office. Figure 3.3 illustrates the overall research design and data availability with an annotated timeline indicating the water service level over time for each of the four operating zones that ever received CWS, and the number of customer records in each zone, separated by residential and non-residential customers. In addition, the overlapping periods of different price structures, missing data, and periods of data to be collected are shown. Figure 3.4 provides a map of the service areas.

In addition to all the above quantitative information, we conducted a series of unstructured interviews with MJP officials at the Amravati utility operations office and the agency headquarters in Mumbai, as well as 50 customers throughout the city. In this study, we focus on the quantitative data and associated models, but the interviews do provide additional context valuable in interpreting our results.



DOMESTIC WATER TARIFF

2 month use(kL)	Charge (Rs.)
0-20	204 (flat)
20+	10.2 (Rs/L)

2 month use(kL)	Charge (Rs.)
0-20	224 (flat)
20-30	10.2/kL
30-40	12.3/kL
40-50	16.8/kL
50+	22.4/kL

2 month use(kL)	Charge (Rs.)
0-20	260 (flat)
20-30	13/kL
30-40	20/kL
40-50	26/kL
50+	52/kL

Figure 3.3 Summary of IWS and CWS conditions, water tariffs, and meter data availability in Amravati from October 2009-March 2013

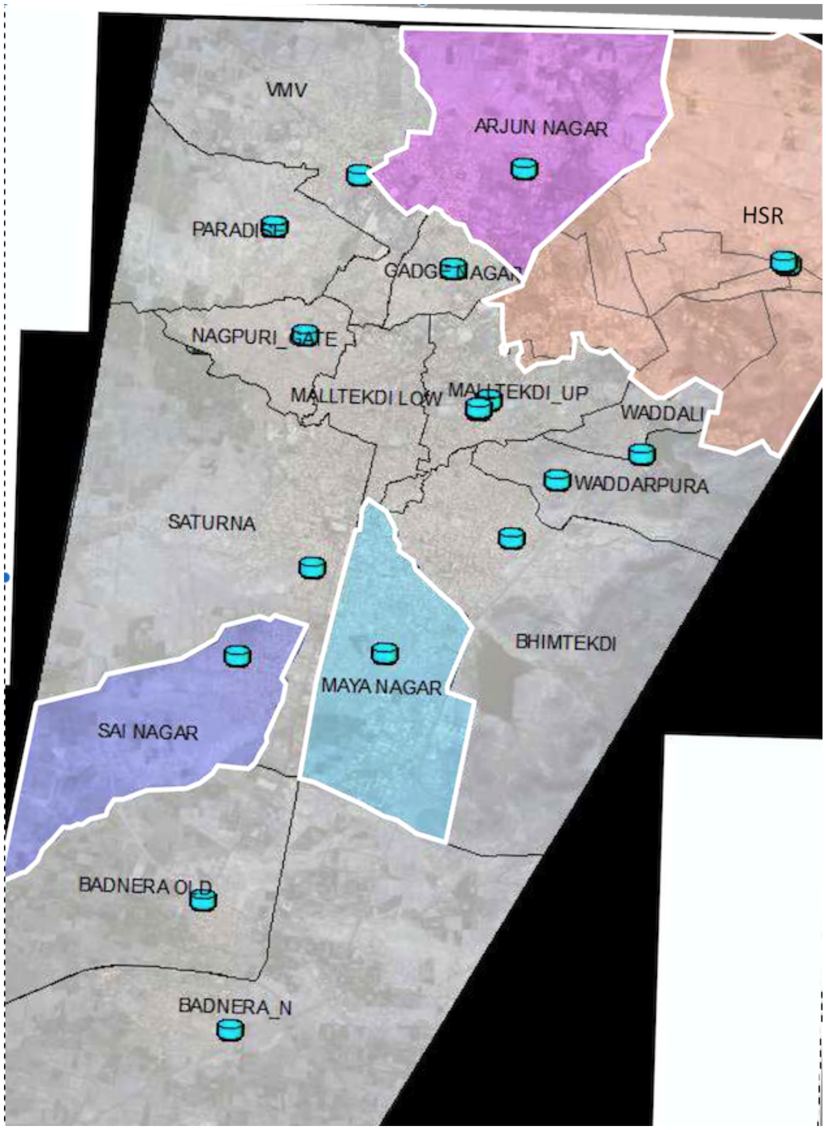


Figure 3.4. Map of Amravati service zones, with CWS zones highlighted.

3.4.4 Statistical Model

To estimate the effects of CWS I estimate a two-way fixed effects regression model, represented by the equation below:

$$ln(q_{it}) = \alpha_i + \beta\ CWS_{it} + \gamma MP_{IV,it-1} +\ \theta_t +\ \epsilon_{it}$$

Where

ln(q_{it}) is the natural logarithm of water consumption in liters per capita per day (LPCD) in a given household *i* in billing period *t*

CWS_{it} is an indicator for whether household *i* had CWS in billing period *t*

α_i represents household fixed effects

θ_t represents bi-month fixed effects

$MP_{IV,it}$ is the natural logarithm of the marginal price of water lagged by one billing period

It is desirable to control for the marginal price due to the price changes that took place over the study period, which involve increasing block tariffs that changed to a greater degree for large consumed volumes than smaller ones, and thus could affect consumers who received CWS differently than those that didn't if CWS enabled higher water consumption. However, there is considerable debate in the water and energy demand estimation literature as to how to specify price under complex pricing structures such as increasing block tariffs, due to the simultaneity involved when prices change with consumption volume, and whether to use marginal or average price (Wichman 2014; Ito 2012; Auffhammer et al. 2014). While there is substantial evidence that when faced with the cognitive demands of complex price schedules, consumers tend to respond to the average rather than marginal price of water, there is evidence in this case that consumers tend to respond to marginal price. Figure 3.5 below is a histogram of water consumption volumes in liters in Amravati in December 2012-January 2013. Note the peaks in density around consumption levels around the kink points in the price schedule of 20kL, 30kL, 40kL, and 50kL, suggesting customers targeted consumption points to avoid sharp marginal price increases. Regressions were performed using the *felm* function from the R package *lfe*.

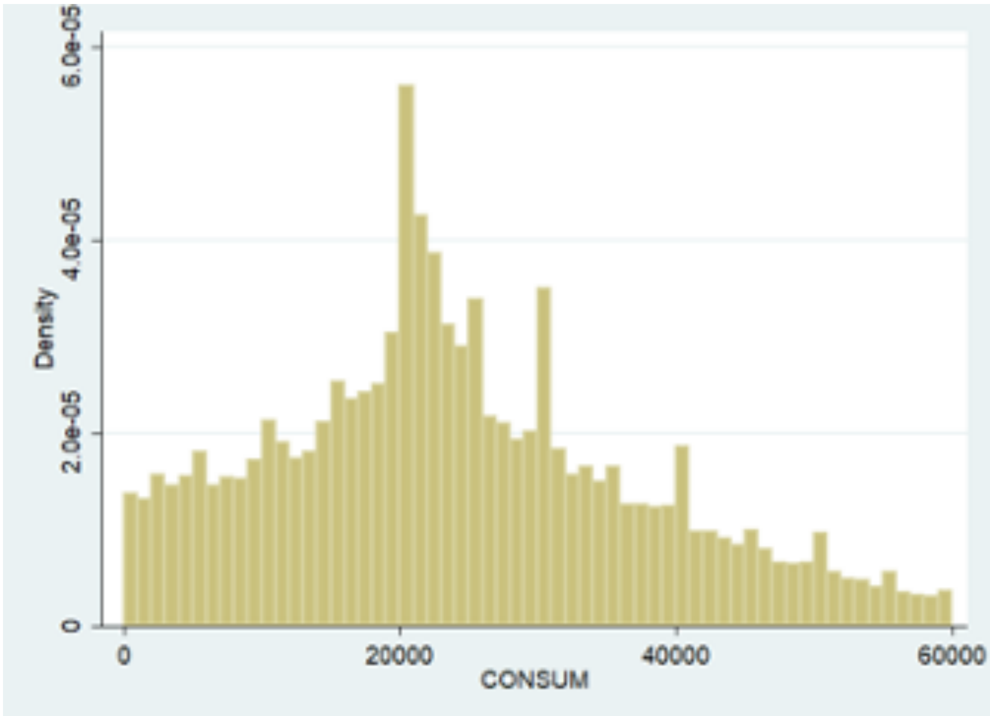


Figure 3.5: Histogram of water consumption (liters) in December 2012-January 2013

I follow the standard approach in the literature since Nieswiadomy & Molina (1989) and use the lagged block prices and consumption blocks as instruments for the marginal price variable in order to control for this endogeneity in the marginal price. I lag the prices by one billing period since consumers can only use the information from their most recent bill in their consumption decision making for the current period. Since there is only data from one billing period before the first two zones Arjun Nagar and Sai Nagar were converted from IWS to CWS, these zones cannot contribute to the identification of the effect of CWS using this approach. Thus, the basic specification includes only IWS zone households and households in the HSR and Maya Nagar zones where CWS continued till 2014 and 2012 respectively

However, I also leverage the Arjun Nagar and Sai Nagar data to estimate the effect of reverting to IWS from a CWS condition, by estimating the models on a subset of data including only the CWS zones, and only data from billing periods on or after June 2011, when Maya Nagar and HSR began receiving CWS. In this specification, CWS is recoded so that IWS is the "treatment" condition, meaning Arjun Nagar, Sai Nagar, and HSR serve as control groups being compared to Maya Nagar, which was "treated" with IWS after May 2012.

To control for potential selection bias in the households that received CWS, I use coarsened exact matching (Iacus, King, and Porro 2012) on observable household characteristics from the census data to ensure there is common support between IWS and CWS households. I also investigate heterogeneity in the effect of CWS by estimating models that include interaction terms for the CWS indicator for the following time-invariant variables:

- indicators for deciles of initial water consumption as measured in October-November 2009. This specification attempts to reveal whether CWS affected demand differently for households that were consuming relatively small or large quantities of water before they began receiving CWS. It is possible that any measured effects could amount to reversion to the mean. However, both IWS and CWS households are well-represented in each decile, I use only Oct-Nov 2009 consumption to specifically calculate the initial deciles, and use only meter data from Oct-Nov 2010 onward to estimate the model.
- 4 household infrastructure configurations composed of combinations of the presence of private wells and overhead storage tanks

That is:

$$ln(q_{it}) = \alpha_i + \sum_{k}^{10} \beta_k \; CWS_{it}Decile_{ik} + \; \theta_t + \; \epsilon_{it}\,.$$

$$ln(q_{it}) = \alpha_i + \sum_{k}^{4} \beta_k \; CWS_{it}WellTankCombination_{ik} + \; \theta_t + \; \epsilon_{it}\,.$$

Finally, I estimate versions of each model that include and do not include the peak season billing periods of April-May and June-July, when temperatures and water consumption are generally highest, to account for the possibility that demand effects of CWS might be more pronounced during these periods.

3.5. Results

In this section I first show descriptive statistics about the households included in the study, and then the demand patterns over the study period for each treatment group. Then, I show and explain results for each model specification.

3.5.1 Summary Statistics

The households in the IWS zones and in the zones that were ever converted to CWS do have different distributions of social status and household infrastructure (Table 3.1). The coarsened exact matching procedure did mitigate the difference in these distributions substantially (Table 3.2).

Table 3.1. Mean and standard deviations (in parentheses) of sociodemographic and house infrastructure variables between IWS and CWS zones.

Variable	IWS Zones	All CWS Zones
N	**23,643**	**8,434**
Total water taps in house	2.14	2.65
	(1.98)	(2.15)
Garden size (sq. m)	0.99	0.99
	(11.89)	(8.90)
No. People in HH	5.00	4.50
	(2.14)	(1.67)
Income category (<100k Rs.)	0.75	0.61
Income category (100k-300k Rs.)	0.22	0.34
Income category (>300k Rs.)	0.03	0.05

Plot area (sq. m)	74.14	79.81
	(82.34)	(56.86)
HH has borewell (dummy)	0.10	0.11
HH has dug well (dummy)	0.11	0.14
HH has handpump (dummy)	0.02	0.01
Overhead tank (dummy)	0.45	0.59
Underground tank (dummy)	0.46	0.42
HH in slum area (dummy)	0.06	0.03
Head of household education (years)	11.74	13.05
	(4.13)	(3.35)
Log of distance from meter to main (m)	1.71	1.42
	(1.12)	(1.07)
Female head of HH works outside the home (dummy)	0.06	0.07

Table 3.2. Mean difference in matching variables between IWS and CWS Zones before and after matching

Variable	Full Sample	Matched
N	23,643 IWS | 8,434 CWS	13,172 IWS | 6,265 CWS
No. People in HH	-0.495	-0.0804
Income Category (<100k Rs)	0.136	0.0000
Income Category (100k-300k Rs)	0.020	0.0000
Income Category (>300k Rs)	0.114	0.0000

Head of Household Education (years)	1.318	0.0361
Total water taps in house	0.514	0.0111
log of distance from meter to water main (meters)	-0.293	-0.0071
Garden size (sq. meters)	-0.0439	-0.006
Plot area (sq. meters)	5.673	1.6112
HH in slum area (dummy)	-0.028	0.0000
Overhead tank (dummy)	0.141	0.0000
Underground tank (dummy)	-0.041	0.0000
Female head of HH works outside home (dummy)	0.010	0.0000
HH has motorized borewell	0.010	0.0000
HH has handpump (dummy)	-0.012	0.0000
HH has dug well (dummy)	0.023	0.0000

Figures 3.6 and 3.7 below show the trend in the mean water consumption (in liters per capita per day) by treatment group (IWS group and each CWS zone), using the full sample (Figure 3.6) and the sample pruned to only include matches (Figure 3.7) using the coarsened exact matching procedure. By visual inspection of the full sample, it is evident that Arjun Nagar and Sai Nagar Zones, either by virtue of selection or because of having been previously converted to IWS, were consuming much higher quantities of water per connected household than either of the later CWS zones, and more still than the IWS zones. The matched sample results in a closer alignment of consumption trends between HSR/Maya Nagar and the CWS zones before they were converted to CWS in May 2011 (billing period 5), although Arjun Nagar and Sai Nagar were still well above the other groups.

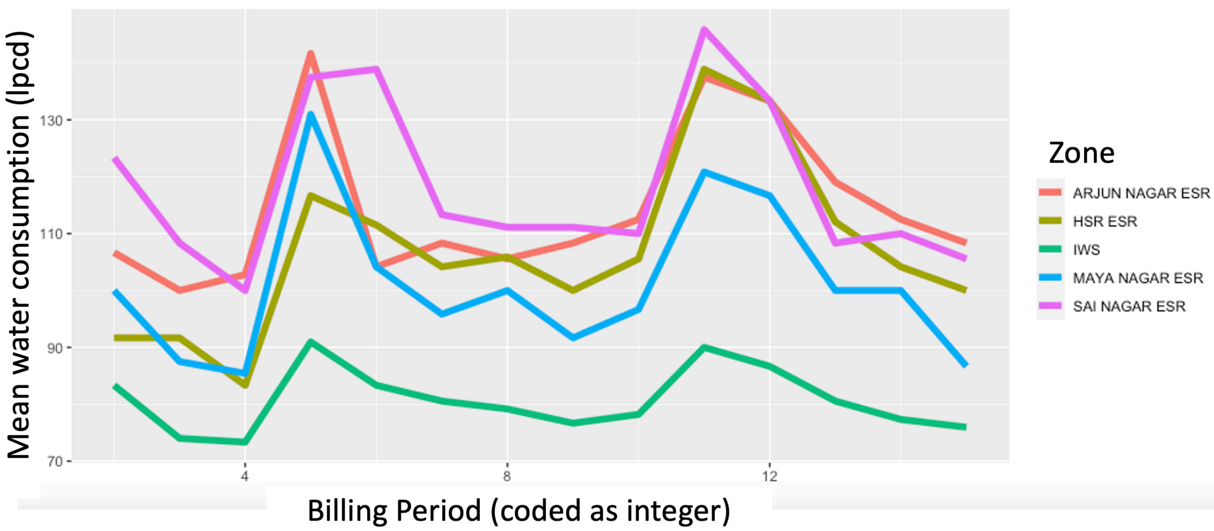


Figure 3.6. Average water consumption amongst connected, metered households in liters per capita per day by treatment group, full sample.

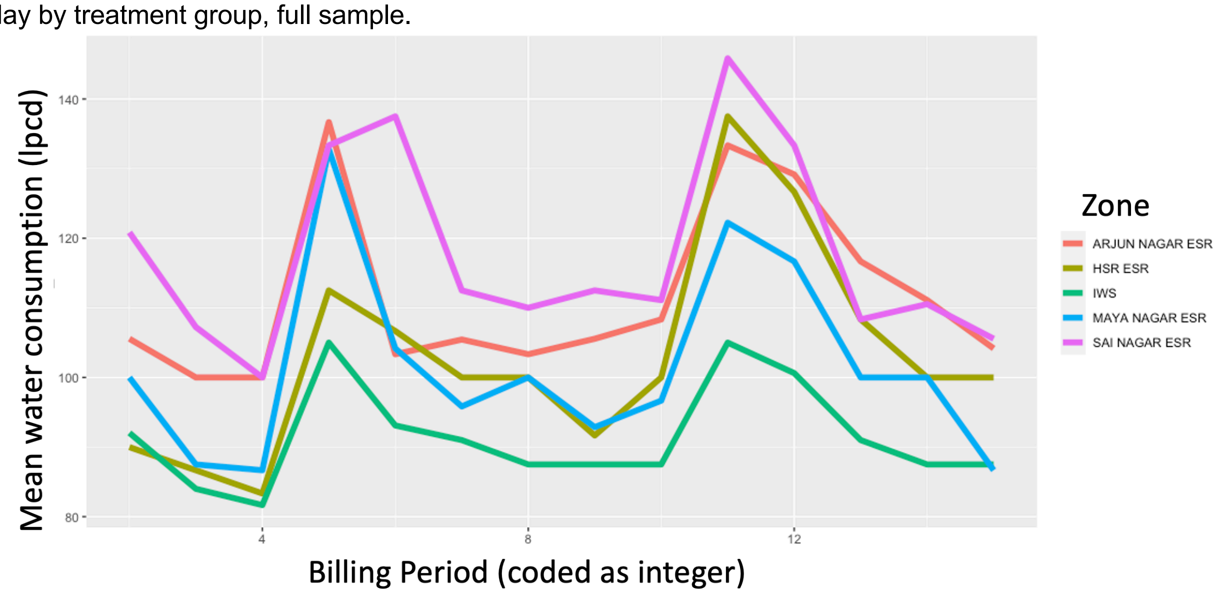


Figure 3.7. Average water consumption amongst connected, metered households in liters per capita per day by treatment group, coarsened exact matching sample.

3.5.2 Basic specification results

Overall, transitioning from IWS to CWS in Amravati was associated with a 13-16% increase in water demand, whether or not the sample was matched or included peak demand periods, when controlling for the contemporaneous price changes (Figure 3.8). Price changes were estimated to have negative (-0.11 to -0.13), but statistically insignificant effects, indicating very inelastic demand. Reverting from IWS to CWS in Amravati was associated with a 2-6% decrease in water demand (Figure 3.9).

3.5.3 Effect of CWS by initial consumption volume decile

The lower bounds of the deciles of consumption in liters per capita per day are shown in Table 3.3.

Table 3.3 Lower bounds of water consumption for deciles of water consumption in October-November 2009

Decile Consumption Oct-Nov 2009	Lower bound (lpcd)
1	6
2	50
3	66
4	70
5	83
6	96
7	111
8	133
9	167
10	217

When parsing the effect by initial decile of consumption (Figure 3.10), the first three deciles showed substantial increases in demand associated with the change to CWS – of 15-55%. This effect is not likely to be regression to the mean alone, as the 10th decile (and no others) also showed an increase in demand of 7-12%, depending on whether the matched sample and peak periods alone were used. As in the overall regression, the coefficient on the lagged price instrument was negative and non-significant, indicating very inelastic price responses over the study period. Reverting from IWS to CWS (Figure 3.11) was associated with a 5-25% decrease in water demand for the decile that consumed the most water in October-November 2009, prior to the conversion.

3.5.4 Effect of CWS by household water infrastructure

When parsing the effect by infrastructure (Figure 3.12), among households that had neither storage tanks nor private wells, CWS was associated with a 13-16% increase in demand. Among households that had both storage tanks and private wells, CWS was surprisingly associated with a far higher – 15-30% increase in demand, with the 30% increase in demand being estimated when omitting peak periods. Having a well and no tank, or a tank and no well, was not associated with a statistically significant change in demand with a change from IWS to CWS. Reverting from IWS to CWS (Figure 3.13) generally was not associated with statistically significant changes in demand for any household infrastructure type. Elasticities were small (-0.11 to -0.13) and not statistically significant

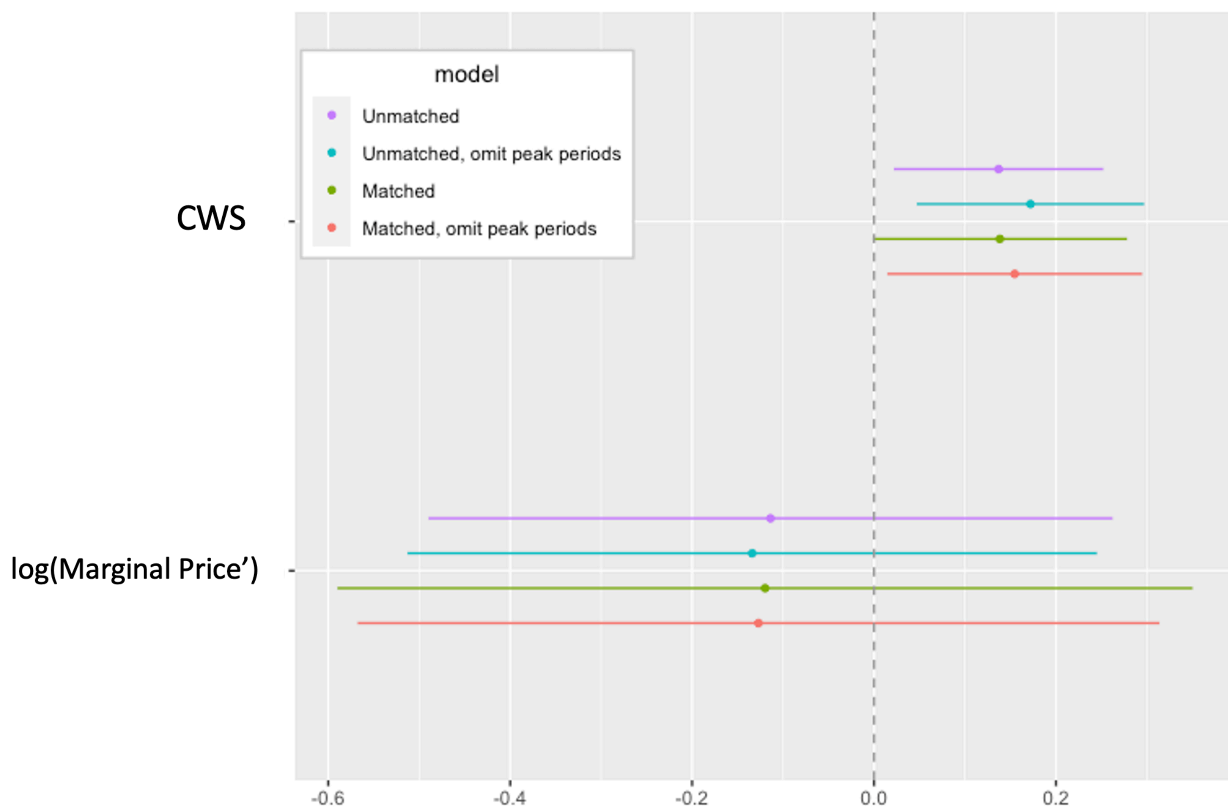


Figure 3.8. Two-way fixed effects regression comparing CWS (HSR and Maya Nagar) with IWS households. Full regression results reported in Table A1 in Appendix.

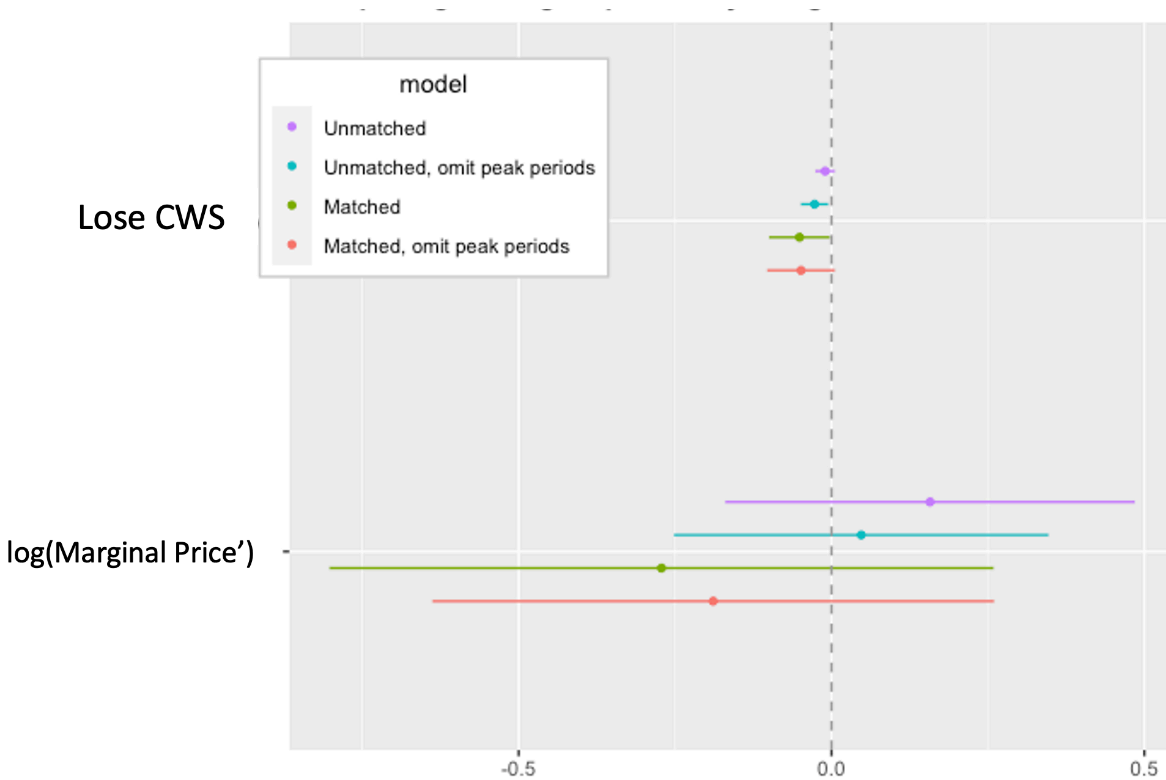


Figure 3.9. Basic regression comparing households that reverted to IWS (Maya Nagar) to CWS households between June-July 2011 and February-March 2013. Full regression results reported in Table A2 in the Appendix.

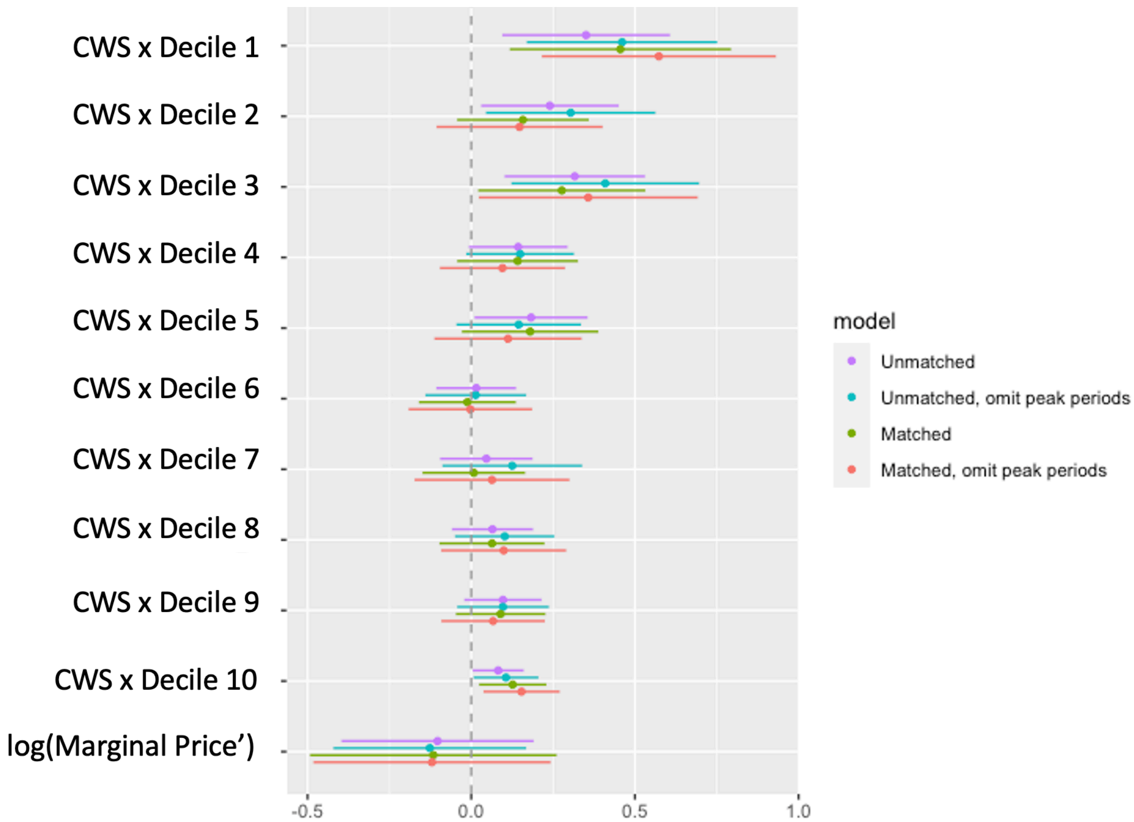


Figure 3.10. Effect of transition from IWS to CWS by decile of consumption in October-November 2009, comparison of HSR to IWS groups. Full regression results reported in Table A3 in Appendix. Note: Decile 3 corresponds to consumption below 66 liters per capita per day. Decile 10 corresponds to consumption above 170 liters per capita per day.

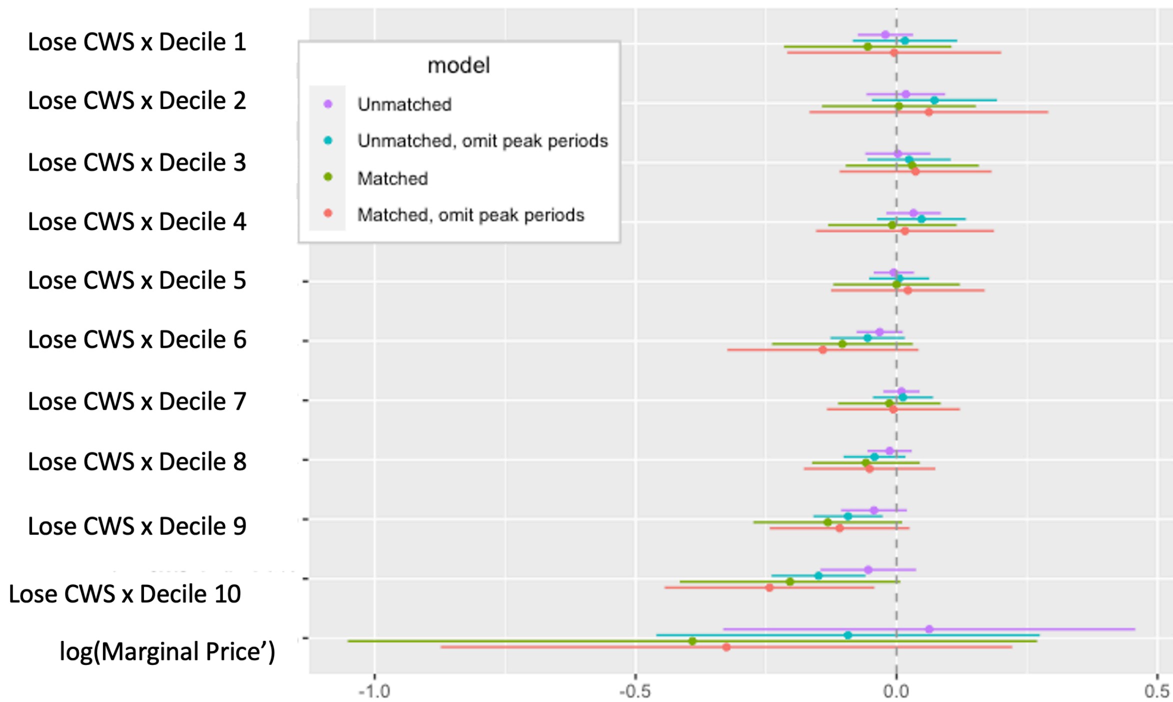


Figure 3.11. Effect of transition from CWS to IWS by decile of consumption in October-November 2009, comparison of Maya Nagar to other CWS groups. Full regression results reported in Table A4 in Appendix.

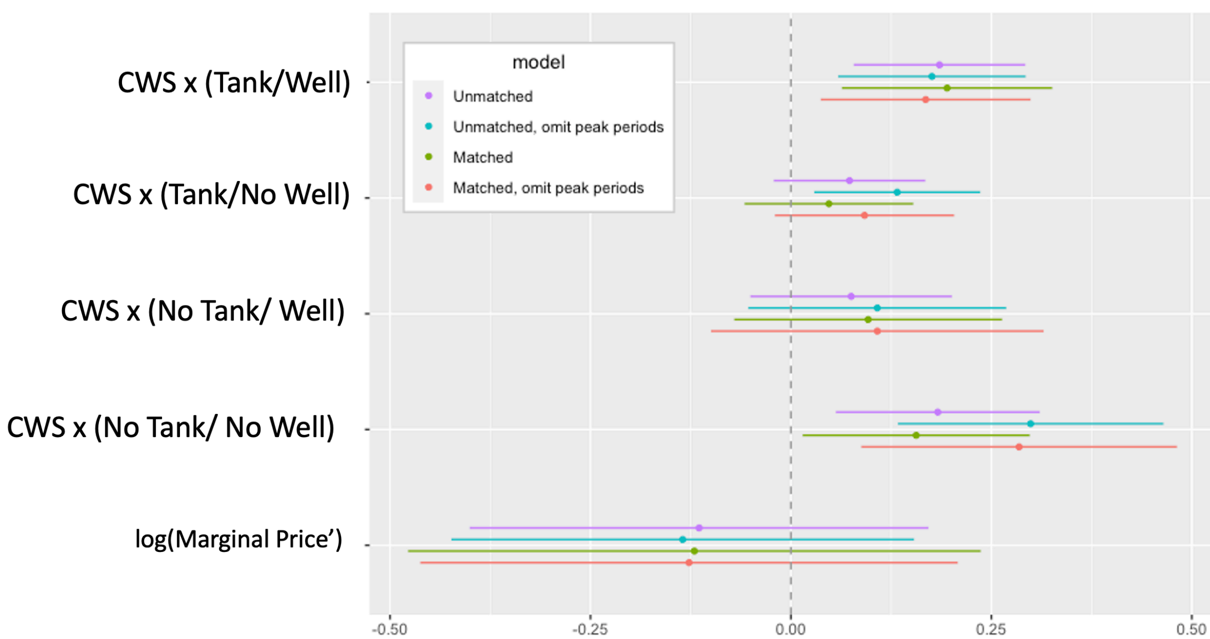


Figure 3.12. Effect of transition from IWS to CWS by storage tank and well, comparison of HSR to IWS groups. Full regression results reported in Table A5 in Appendix.

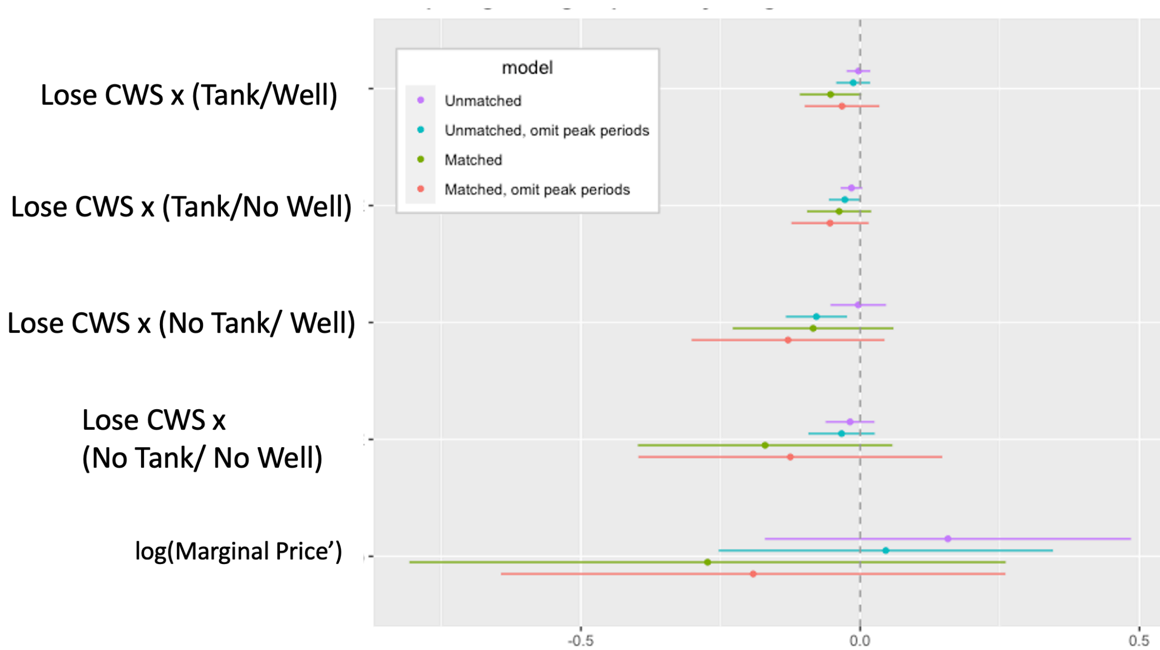


Figure 3.13. Effect of reversion from CWS to IWS by storage tank and well, comparison of Maya Nagar to CWS groups. Full regression results reported in Table A6 in Appendix.

3.6 Discussion

The state of IWS at the time of the CWS experiment in Amravati was a relatively high level of service, such that 30% of households with private connections were consuming at least 133 lpcd from the piped water network (with some exceeding the 135 lpcd Indian service level benchmark (CPHEEO 1999). For comparison, over much of the CWS period in Hubli-Dharwad, water demand overall, including leakage, was averaging 110-120 lpcd (Jayaramu et. al. 2015), while in IWS areas households with private connections were consuming 97 lpcd (Kumpel et. al 2017).

In general, we estimate that water demand in Amravati was price inelastic. The estimated elasticities between -0.11 and -0.13 are on the lower end, but consistent with other estimates of price elasticity of demand for piped water in low and middle-income country urban settings (World Health Organization et al. 2020). However, part of this may be due to the lack of consistent billing enforcement measures. Households in Amravati received their water bills at least two months after the relevant water was consumed, and disconnection proceedings would generally not begin until at least one year after nonpayment of a bill. Under such conditions, it is possible many households would not react to price changes at all, or that the effective price they made consumption decisions based on was zero.

We found that CWS increased water demand substantially (~15%) for households that had neither an overhead storage tank nor a private borewell, as *well as* for households that had both. Households that had one or the other did not have statistically significant effects. It would seem straightforward that households with neither tanks nor wells would have been limited in the supply of water that they could consume from their service connections under IWS due to a lack of (automatic) storage. The lack of a well would also indicate that the service connection would be the only source of piped water to taps inside the home or to a storage tank, so a conversion to CWS could be highly salient to such households in how they consume water for domestic purposes.

Our estimate that households that had both tanks and wells also increased their consumption under CWS while the other types of households did not, presents a puzzle. One possibility is that such households simply have higher demands for water in general, hence their investment in both types of household water infrastructure, while maintaining a preference for piped water over well water. Such households would be limited in their supply of piped water under IWS, and supplement with well water, which they would substitute away from under CWS. In contrast, a household with a well and no tank may be reserving piped water under IWS for select purposes like drinking and cooking, while using well water for bathing, cleaning, gardening, and productive or commercial uses. In this scenario, household plumbing may be set up to use well water for more purposes, giving structure to a water consumption pattern that does not make much use of piped water. On the other hand, a household with a tank and no well under IWS would probably be having its piped water needs met under that condition. This setup often involves the piped water connection going straight to the overhead tank, activated by a float valve that automatically closes when the tank is full. Most water is abstracted for use through household plumbing that is served by this tank. Thus, the change from IWS to CWS would not be expected to change behavior if the quantity of water demanded daily under IWS was less than the storage capacity of the tank.

Our additional subgroup analysis by decile of water demand more than one year before the HSR and Maya Nagar zones began receiving CWS revealed that the bottom three deciles and the top decile increased their water demand by 20-40% and 5-10%, respectively. Note that the bottom three deciles correspond to consumption levels in October-November 2009 under IWS of less than 70 lpcd, and the top decile corresponds to consumption levels above 217 lpcd under IWS, much higher than the Indian Service Level Benchmark. One possible explanation for this pattern is that it is both relatively low and high-volume water users under IWS that are the most likely to have been physically or temporally limited in their piped IWS water consumption, and thus for both groups, transitioning to CWS increased the volume of piped water they could consume. For example, it is possible that relatively high-volume users were more likely to have persistent indirect usages such as high leakage rates in their plumbing or selling or sharing their tap water with neighbors that were limited only by IWS service hours, as documented by Kumpel et. al. (2017). Initial water consumption was not correlated with the coarse 3-income categorical variable, or years of education of the head of household.

This same analysis on the effect of reverting from CWS to IWS reveals a 10-25% reduction in consumption by the top decile, but no reduction in the lower deciles. This is an interesting finding and could suggest that part of the nature of the demand shift due to CWS came from alleviating physical limits. Another part may have come from changed preferences in favor of consuming more piped water, even though this higher consumption level may have been physically possible under Amravati's reliable and predictable IWS regime–improvements that were put in place for all of Amravati before any of the four CWS zones were converted.

A middle-service-level trap?

Amravati's style of reliable, predictable, high-pressure IWS—which Taylor et. al. (2019) might refer to as "satisfied IWS" -- at first glance, provides a model to balance service quality and extension of coverage in the context of urban India. This service level allows most connected people to consume reasonable quantities of water that they are willing to pay for. This should result in revenues that can be used to improve and extend the network.

During the study period, the median connected household was consuming 96 lpcd, near a quantity corresponding to what the World Health Organization considers “optimal” access (World Health Organization et. al. 2019). Only 25% of Amravati’s connected population, and 16% of its total population between 2008 and 2014 ever had CWS. However, between 2008 and 2021, the overall service population coverage under IWS expanded from 50% of the population to 68%, coverage of the population in informal settlements expanded from 21% to 30%, and average per capita water deliveries rose from 95 lpcd to 117 lpcd (CEPT C-WAS 2022). This is doubly impressive considering that the city’s population grew by 18% over the same period. Throughout this period, Amravati recovered 150-200% of its operating costs from water bills.

However, “satisfied IWS” may also represent a trap. Taylor et. al. (2019) elaborate theoretical reasons and a mathematical model for why, and Amravati may be a real-world case of such a trap. Under satisfied IWS, most connections are able to consume as much water as desired, but high leakage rates are likely still present throughout the distribution network and will be chronically exacerbated over time due to the physical properties of the IWS service mode (Galaitsi et. al. 2016). In Amravati, between 10-30% of households increased their consumption under IWS, and only the highest decile of water consumers reduced their consumption when they reverted from CWS to IWS. Under such conditions, any zone that was converted from IWS to CWS without comprehensive leakage reduction would thus result in only marginal increases in billed water (and revenue), while leakage could increase dramatically (volumes leaked under a 4 hour/day service schedule would be multiplied by 6 under CWS). Depending on the increase in leaked volume, the utility may face strong financial incentives to reduce service hours, allocating saved water to new connections. Our interviews with utility management in 2015 corroborated this, indicating that the organization did not have enough revenues to augment water supply or maintain the system, and could not justify the 40% non-revenue water rate they were experiencing in the CWS zones in the face of demand for new connections. They instead chose to endeavor “to improve the current system while giving access to all” (personal communication with an MJP official). Future work will take advantage of bulk meter data and financial records made available by MJP to potentially validate these statements and the model of Taylor et. al. (2019) by tracing the development of revenue and water leakages in the IWS and each of the CWS zones in Amravati over time.

An implication of this trap (or less judgmentally, equilibrium) is that the sequence of investment in an unreliable, unpredictable, “unsatisfied” IWS system should be carefully considered in light of existing conditions, available capital, and desired end state. Satisfied IWS is not a clear stepping-stone on the way to CWS in a context where current available water supply is not sufficient for anticipated population growth. While it may be more immediately financially feasible to upgrade a system from unsatisfied to satisfied IWS rather than CWS, it may then become unlikely for CWS to be achievable for quite some time. If CWS is a desired end state, it may make more sense to work to assemble the capital necessary to invest in water supply augmentation and rigorous water leak detection and control programs that could support CWS and a reasonable leakage rate under realistic population growth projections. This choice should depend on a number of factors, not least of which is the relative importance placed on prioritizing bringing sustainable water services to the unconnected, especially the unconnected poor, at a rapid pace (namely, equity). If this is a priority, whatever choice is taken should comport with reasonable financial and water budgets for effective mechanisms to be inclusive of the poor, such as means-tested connection subsidies and maintaining standpipes (Cook et al 2020).

There are additional institutional complications in Amravati’s case. MJP centralizes the revenue for all of the systems it operates and reallocates it. Thus, a fraction of Amravati’s water bill revenues was being used to cross-subsidize other water systems and centralized MJP expenses such as pensions. Whether or not full allocation of Amravati’s profits to its own system would have been sufficient to maintain and expand CWS is uncertain and should be investigated in future work. However, this arrangement certainly did not improve the prospects of the long-term success of Amravati’s CWS experiment. Even so, it is not clear that the outcome Amravati experienced, of steadily improving access to “satisfied IWS” over time is even a poor one, despite MJP seeming to combine many of the worst aspects of private and public ownership often discussed in the literature. Amravati’s water system is publicly owned, yet is operated by a state-level agency, rather divorced from local government politics. It had one of the highest water rates in India during the study period. However, its revenues were not ring-fenced at the level of the water system, and were directed to both other water systems and non-water-related purposes. This reveals that debates in the literature regarding the merits of public or private ownership of water utilities (e.g. Bakker 2010), and administrative and financial separation of water utilities from local government (Pierce et al 2022), need to be more nuanced.

Note that this study only focuses on metered connections. During the study period, 51% of Amravati's households were connected to the water network, 75% of which were metered. The remaining households procured water from public standpipes (which were all removed in the CWS areas to reduce non-revenue water from standpipes and incentivize metered connections, for which no connection subsidies were implemented), wells, or neighbors' connections. Moreover, the intervention targeted areas with the smallest population living in informal housing. Overall, the Amravati intervention seems to have routed more water to connected households, and much of that to households with already higher consumption levels within this group. The equity implications of this change are unclear without further investigation of the affected households. For example, it could be considered highly inequitable to allow much more water to flow to those with the highest water demands while roughly half of the city's residents don't have access to the network at all. On the other hand, if these higher end water users represent those sharing or selling water to their neighbors, this could represent more of an informal network extension that improves access to unconnected households fortunate enough to be near willing tap sharers (Zuin et. al. 2011).

3.7 Conclusion

We found evidence of statistically significant shifts in household water demand for certain subgroups of Amravati's connected households that we cautiously attribute to the transition from IWS to CWS. We showed that this shift varies greatly with initial levels of water demand and with the type of household water infrastructure that has been previously installed under an IWS regime. We could not identify the exact mechanisms for why particular groups of households were affected in this way due to a lack of detailed information such as volumes used from wells for households that had them, the combination of interior and exterior taps served by service connections and wells, and the extent of water sharing or reselling to neighbors.

We showed that CWS certainly does not reduce demand, as might be expected if water wasting via storage tank and vessel emptying and constant tap opening were a consistent, pervasive behavior under IWS. This confirms cross-sectional analyses of households (Burt et. al. 2018; Kumpel et. al .2017) and bulk water delivery data (Jayaramu et. al. 2015) from Hubli-Dharwad. This casts doubt on the suitability of CWS as a short-term performance target for Indian cities that don't have access to debt finance or donor funds for water supply augmentation. This is because these cities also must contend with demands for service connections within their nominal or planned service areas, whether due to normative commitments to equity or due to political pressures that accompany growing unserved populations. This is particularly concerning for Amravati, where under IWS water arrived on a predictable schedule for four hours every day – a high level of service in the Indian context. Indeed, eventually Amravati's utility management decided to return the entire system to IWS to reallocate available water supplies towards new neighborhoods to be served. Continuous water supply efforts should make realistic water demand projections and fully account for the costs of augmenting water supply and/or maintaining a vigorous leak detection and control program under that service level standard.

Overall, this study complicates the notion of a service level benchmark for water service continuity. In this case, we show that a seemingly unambiguous improvement in a benchmark represents different benefits to different types of households. Reliability and predictability may be more appropriate concepts to operationalize and measure than an average of service hours to reflect the usefulness of a piped water supply to connected households. The suitability of a service level can be moderated by household water infrastructure as well, and this should perhaps be accounted for in conjunction with measures of service continuity and reliability. Moreover, it is not clear how the improvement should be weighed against its tradeoffs relative to other benchmarks, such as the proportion of households with access to service, or the affordability of that service for households of different income levels when accounting for the increase in tariffs (both actual and what might be required to maintain the service level).

REFERENCES

Abu-Madi, Maher, and Nemanja Trifunovic. “Impacts of Supply Duration on the Design and Performance of Intermittent Water Distribution Systems in the West Bank.” *Water International* 38, no. 3 (May 2013): 263–82. https://doi.org/10.1080/02508060.2013.794404.

Altaf, Anjum. “The Economics of Household Response to Inadequate Water Supplies.” *Third World Planning Review* 16, no. 1 (1994).

Andey, Subhash P., and Prakash S. Kelkar. “Performance of Water Distribution Systems during Intermittent versus Continuous Water Supply.” *Journal American Water Works Association* 99, no. 8 (2007): 99–106.

Bakker, Karen. 2010. *Privatizing Water: Governance Failure and the World’s Urban Water Crisis*. Cornell University Press.

Batish, Rajiv. “A New Approach to the Design of Intermittent Water Supply Networks.” In *World Water & Environmental Resources Congress 2003*, 1–11. Philadelphia, Pennsylvania, United States: American Society of Civil Engineers, 2003. https://doi.org/10.1061/40685(2003)123.

Berg, Caroline van den, and Alexander Danilenko. *The IBNET Water Supply and Sanitation Performance Blue Book: The International Benchmarking Network for Water and Sanitation Utilities Databook*. The World Bank, 2010. https://doi.org/10.1596/978-0-8213-8582-1.

Bivins, Aaron W, Trent Sumner, Emily Kumpel, Guy Howard, Oliver Cumming, Ian Ross, Kara Nelson, and Joe Brown. “Estimating Infection Risks and the Global Burden of Diarrheal Disease Attributable to Intermittent Water Supply Using QMRA.” *Environmental Science & Technology* 51, no. 13 (2017): 7542–51.

Burt, Zachary, Narayana Billava, and Isha Ray. “From Intermittent to Continuous Service: Costs , Benefits , Equity and Sustainability of Water System Reforms in Hubli-Dharwad , India” 109 (2018): 121–33. https://doi.org/10.1016/j.worlddev.2018.04.011.

Burt, Zachary, and Isha Ray. “Storage and Non-Payment: Persistent Informalities within the Formal Water Supply of Hubli-Dharwad , India.” *Water Alternatives* 7, no. 1 (2014): 106–20.

Buurman, Joost, and Deepak Santhanakrishnan. “Uncorrected Proof Opportunities and Barriers in Scaling up of 24 / 7 Urban Water Supply: The Case of Karnataka , India Uncorrected Proof” 2003 (2017): 1–17. https://doi.org/10.2166/wp.2017.191.

Charalambous, Bambos, and Chrysi Laspidou. *Dealing with the Complex Interrelation of Intermittent Supply and Water Losses*. Scientific and Technical Report / International Water Association 25. London: IWA Publishing, 2017.

CEPT C-WAS. 2022. “PAS - Know Your City.” Performance Assessment System. 2022. https://www.pas.org.in/web/ceptpas/knowyourcity?p_p_id=Knowyourcity_WAR_Portal&p_p_lifecycle=1&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=1&actionVal=Retrieve&SkipAccessChecking=false.

Cook, Joseph, David Fuente, and Dale Whittington. 2020. “Choosing Among Pro-Poor Policy Options in the Delivery of Municipal Water Services.” *Water Economics and Policy* 06 (03): 1950013. https://doi.org/10.1142/S2382624X19500139.

Colombo, Andrew F., and Bryan W. Karney. “Energy and Costs of Leaky Pipes: Toward Comprehensive Picture.” *Journal of Water Resources Planning and Management* 128, no. 6 (November 2002): 441–50. https://doi.org/10.1061/(ASCE)0733-9496(2002)128:6(441).

CPHEEO. *Central Public Health and Environmental Engineering Organization Manual of Water Supply and Treatment*. 3rd ed. New Delhi: Ministry of Urban Development and Poverty Alleviation, Govt. of India, 1999.

Dasgupta, Purnamita, and Rajib Dasgupta. “Economic Value of Safe Water for the Infrastructurally Disadvantaged Urban Household: A Case Study in Delhi, India: VALUING SAFE WATER IN URBAN HOUSEHOLD.” *Water Resources Research* 40, no. 11 (November 2004). https://doi.org/10.1029/2003WR002461.

De Marchis, M., C. M. Fontanazza, G. Freni, G. La Loggia, E. Napoli, and V. Notaro. “A Model of the Filling Process of an Intermittent Distribution Network.” *Urban Water Journal* 7, no. 6 (December 2010): 321–33. https://doi.org/10.1080/1573062X.2010.519776.

De Marchis, Mauro, Gabriele Freni, and Barbara Milici. “Experimental Evidence of the Discharge Law in Private Tanks Connected to Water Distribution Networks.” *Procedia Engineering* 154 (2016): 115–22. https://doi.org/10.1016/j.proeng.2016.07.428.

Ercumen, Ayse, Benjamin F. Arnold, Emily Kumpel, Zachary Burt, Isha Ray, Kara Nelson, and John M. Colford. “Upgrading a Piped Water Supply from Intermittent to Continuous Delivery and Association with Waterborne Illness: A Matched Cohort Study in Urban India.” Edited by Clarissa Brocklehurst. *PLOS Medicine* 12, no. 10 (October 2015): e1001892. https://doi.org/10.1371/journal.pmed.1001892.

Ercumen, Ayse, Joshua S. Gruber, and John M. Colford. *Water Distribution System Deficiencies and Gastrointestinal Illness: A Systematic Review and Meta-Analysis*. Vol. 122. 7, 2014. https://doi.org/10.1289/ehp.1306912.

Erickson, John J., Charlotte D. Smith, Amador Goodridge, and Kara L. Nelson. “Water Quality Effects of Intermittent Water Supply in Arraiján, Panama.” *Water Research* 114 (May 2017): 338–50. https://doi.org/10.1016/j.watres.2017.02.009.

Fan, Liangxin, Guobin Liu, Fei Wang, Coen Ritsema, and Violette Geissen. “Domestic Water Consumption under Intermittent and Continuous Modes of Water Supply.” *Water Resources Management* 28 (2014): 853–65. https://doi.org/10.1007/s11269-014-0520-7.

Fontanazza, C. M., G. Freni, and G. La Loggia. “Analysis of Intermittent Supply Systems in Water Scarcity Conditions and Evaluation of the Resource Distribution Equity Indices.” In *Water Resources Management IV*, I:635–44. Kos, Greece: WIT Press, 2007. https://doi.org/10.2495/WRM070591.

Galaitsi, S., Robert Russell, Amahl Bishara, John Durant, Jennifer Bogle, and Annette Huber-Lee. “Intermittent Domestic Water Supply: A Critical Review and Analysis of Causal-Consequential Pathways.” *Water* 8, no. 7 (June 30, 2016): 274. https://doi.org/10.3390/w8070274.

Guragai, B., S. Takizawa, T. Hashimoto, and K. Oguma. “Effects of Inequality of Supply Hours on Consumers’ Coping Strategies and Perceptions of Intermittent Water Supply in Kathmandu Valley, Nepal.” *Science of The Total Environment* 599–600 (December 2017): 431–41. https://doi.org/10.1016/j.scitotenv.2017.04.182.

Hamilton, Stuart, and Bambos Charalambous, eds. *Leak Detection: Technology and Implementation*. Second edition. London: IWA Publishing, 2019.

Hastak, Shashikant, Pawan Labhasetwar, Priya Kundley, and Rajesh Gupta. “Changing from Intermittent to Continuous Water Supply and Its Influence on Service Level Benchmarks: A Case Study in the Demonstration Zone of Nagpur , India.” *Urban Water Journal* 9006, no. October (2016): 1–5. https://doi.org/10.1080/1573062X.2016.1240808.

Ilaya-Ayza, Amilkar, Enrique Campbell, Rafael Pérez-García, and Joaquín Izquierdo. “Network Capacity Assessment and Increase in Systems with Intermittent Water Supply.” *Water* 8, no. 4 (March 31, 2016): 126. https://doi.org/10.3390/w8040126.

Jayaramu, K P, Zachary Burt, and B Manoj Kumar. “A Study of the Consumption Pattern in a Continuous Water Service Demonstration Zone and Bulk Water Demand Forecasting for Hubli-

Dharwad, India." *JOURNAL OF WATER SANITATION AND HYGIENE FOR DEVELOPMENT* 5, no. 2 (2015): 201–12. https://doi.org/10.2166/washdev.2015.096.

Klingel, Philipp. "Technical Causes and Impacts of Intermittent Water Distribution." *Water Science and Technology: Water Supply* 12, no. 4 (July 2012): 504–12. https://doi.org/10.2166/ws.2012.023.

Kumpel, Emily, and Kara L Nelson. "Comparing Microbial Water Quality in an Intermittent and Continuous Piped Water Supply." *Water Research* 47, no. 14 (September 2013): 5176–88. https://doi.org/10.1016/j.watres.2013.05.058.

Kumpel, Emily, and Kara L. Nelson. "Intermittent Water Supply: Prevalence, Practice, and Microbial Water Quality." *Environmental Science & Technology*, 2016, acs.est.5b03973. https://doi.org/10.1021/acs.est.5b03973.

———. "Mechanisms Affecting Water Quality in an Intermittent Piped Water Supply." *Environmental Science and Technology* 48, no. 5 (2014): 2766–75. https://doi.org/10.1021/es405054u.

Kumpel, Emily, Isha Ray, and Kara L Nelson. "Measuring Household Water Consumption and Waste in Unmetered, Intermittent Piped Water Systems," no. October (2017). https://doi.org/10.1002/2016WR019702.CITATIONS.

McIntosh, Arthur C. *Asian Water Supplies: Reaching the Urban Poor: A Guide and Sourcebook on Urban Water Supplies in Asia for Governments, Utilities, Consultants, Development Agencies, and Nongovernment Organizations*. Manila, Philippines: London: Asian Development Bank ; International Water Association, 2003.

McIntosh, Arthur C, Ian Makin, Tan Gee Paw, Chuanpit Dhamasiri, Arjun Thapan, Perry Rivera, Roland Liemberger, and Michael White. *Urban Water Supply and Sanitation in Southeast Asia: A Guide to Good Practice*. Manila: Asian Development Bank, 2014.

Mitra, Susanna Ghosh. "Global Governance and Water Policy-Making: Shifting Analytical Lens for 24/7 Urban Water Supply Programmer in India." In *Water Governance*, 24-, 2012.

Pattanayak, Subhrendu K., Jui-Chen Yang, Dale Whittington, and K. C. Bal Kumar. "Coping with Unreliable Public Water Supplies: Averting Expenditures by Households in Kathmandu, Nepal." *Water Resources Research* 41, no. 2 (February 2005): n/a-n/a. https://doi.org/10.1029/2003WR002443.

Pierce, Gregory, Justin McBride, and Jacqueline Adams. 2022. "Subsidized or Subsidizing? Municipal Drinking Water Service Funds in California." *Utilities Policy* 79 (December): 101434. https://doi.org/10.1016/j.jup.2022.101434.

Rana, M.H. "Progression from Intermittent to 24x7 Mode of Water Supply." *Engineering and Technology Journal of S.O.C.E.T* 1, no. 2 (2013): 24–28.

Rosenberg, David E., Tarek Tarawneh, Rania Abdel-Khaleq, and Jay R. Lund. "Modeling Integrated Water User Decisions in Intermittent Supply Systems: INTEGRATED WATER USER DECISIONS." *Water Resources Research* 43, no. 7 (July 2007). https://doi.org/10.1029/2006WR005340.

Sangameswaran, P, R Madhav, and C D'Rozario. "24/7,'Privatisation'and Water Reform: Insights from Hubli-Dharwad." *Economic and Political Weekly*, 2008, 60–67.

Simukonda, Kondwani, Raziyeh Farmani, and David Butler. "Intermittent Water Supply Systems: Causal Factors, Problems and Solution Options." *Urban Water Journal* 15, no. 5 (May 28, 2018): 488–500. https://doi.org/10.1080/1573062X.2018.1483522.

Taylor, David D. J., Alexander H. Slocum, and Andrew J. Whittle. 2019. “Demand Satisfaction as a Framework for Understanding Intermittent Water Supply Systems.” *Water Resources Research* 55 (7): 5217–37. https://doi.org/10.1029/2018WR024124.

The Karnataka Urban Water Sector Improvement Project. “24x7 Water Supply Is Achievable,” no. September (2010): 24.

Totsuka, Natsuko, Nemanja Trifunovic, and Kalanithy Vairavamoorthy. “Intermittent Urban Water Supply under Water Starving Situations,” 2004.

Vairavamoorthy, Kala, Sunil D. Gorantiwar, and S. Mohan. “Intermittent Water Supply under Water Scarcity Situations.” *Water International* 32, no. 1 (March 2007): 121–32. https://doi.org/10.1080/02508060708691969.

Virjee, Kameel, and Susan Gaskin. “Coping with Poor Water Services and the Demand for Change in Trinidad and Tobago.” *Water International* 35, no. 3 (June 15, 2010): 285–97. https://doi.org/10.1080/02508060.2010.487299.

Whittington, Dale, Subhrendu K Pattanayak, Jui-Chen Yang, and K C Bal Kumar. “Household Demand for Improved Piped Water Services: Evidence from Kathmandu, Nepal.” *Water Policy* 4, no. 6 (2002): 531–56.

World Bank. *Making Services Work for Poor People: World Development Report 2004: Overview.* Washington, D.C.: World Bank, 2003.

World Health Organization, Guy Howard, Jamie Bartram, Ashley Williams, Alycia Overbo, David Fuente, and Jo-Anne Geere. 2020. *Domestic Water Quantity, Service Level and Health*. 2nd ed. Geneva: World Health Organization. https://apps.who.int/iris/handle/10665/338044.

Zuin, Valentina, Leonard Ortolano, Manuel Alvarinho, Kory Russel, Anne Thebo, Odete Muximpua, and Jennifer Davis. “Water Supply Services for Africa’s Urban Poor: The Role of Resale.” *Journal of Water and Health* 9, no. 4 (2011): 773–84. https://doi.org/10.2166/wh.2011.031.

CHAPTER 4: GEOGRAPHY, OWNERSHIP, AND AFFORDABILITY: AN ANALYSIS OF THE LANDSCAPE OF WATER SYSTEMS IN CALIFORNIA[4]

4.1 Introduction

Ever since infrastructure privatization policies gained attention in the 1990s and 2000s, there has been an active debate over the political, operational, efficiency, affordability, and environmental implications of private participation in water services. There is a widely expressed collective aspiration for all to have access to safe, adequate, and affordable drinking water (as now enshrined in the 2030 Agenda for Sustainable Development Goal #6 adopted in 2015). In practice, the goal of universal service has come up against the high capital and recurring fixed costs of water infrastructure and the high transactions costs involved in contracting out (Bel et al. 2010; Bakker 2010; Kirkpatrick and Parker 2005). This is complicated by gaps in access due to legal and regulatory frameworks. The objective of universal public provision has kept alive the debate over private participation in the water sector and the conditions under which privatization may or may not further goals of affordability, sustainability, and inclusion. More recently, the growing problems of aging infrastructure, climate change, and global public health crises such as the COVID-19 pandemic have reignited the debate on how to ensure universal coverage of clean and affordable water for all, and the role of water privatization in it (Teodoro et al. 2020).

After nearly thirty years of debate many questions remain unsettled. On the one hand proponents of privatization claim that privately owned utilities bring superior operational performance and efficiency by more flexibly adopting new technologies while being more responsive to customers (World Bank 2003). Some argue that privately owned entities are able to finance improved infrastructure and service delivery by their ability to charge more flexibly for water than public agencies (which might be politically constrained from charging higher rates), allowing tariffs to better reflect what they argue is the reality of

[4] This chapter is published as an article in Utilities Policy. The citation is: Onda, K. and Tewari M. Water systems in California: Ownership, geography, and affordability. *Utilities Policy* 72, 101279 (2021).

water as a scarce resource that should be treated as an economic good (Segerfeldt 2005; Robbins 2003; Renzetti and Dupont 2003; Levin et al., 2002; Winpenny 1994; Rogers 2002). On the other hand, critics of privatization argue that the commodification of water elevates economic efficiency over goals of affordability and social equity in access to this essential resource (Barlow and Clarke 2017; Bakker 2007, 2010). Privately owned utilities, critics point out, have a profit motive in an economic context characterized by natural monopoly cost structures. This makes them more likely to charge higher prices to cover costs and contribute to profits than would utilities under public management. Monopoly control may provide them with few incentives to focus on consistently improving performance and inclusion of the poorest (Subramaniam and Williford 2012; Bakker 2008). Complicating these debates are the varieties of institutional and regulatory frameworks that can govern private sector participation in different contexts (Davis 2005). Public utility commissions exert control over water tariffs and access policies, such as setting thresholds that trigger cutoffs for nonpayment. How these regulatory structures apply to privately owned and publicly owned water systems varies by jurisdiction, making it difficult to generalize (J. A. Beecher and Kalmbach 2013; Homsy and Warner 2020). Most states have public utilities commissions that regulate privately (investor) owned for-profit utility and some regulate both public and privately owned utilities. Some states regulate publicly owned utilities under certain conditions, such as whether or not they serve residents outside of municipal borders. Consequently, how institutional and regulatory variations actually influence affordability outcomes in publicly owned and privately owned water systems across jurisdictions remains an understudied question (Beecher and Kalmbach 2013).

Most studies on water sector privatization find that privately owned water systems charge systematically higher rates than publicly owned providers, although the extent to which this difference can be attributed to private ownership itself remains unclear due to the variety of other factors that contribute to water prices, such as local infrastructure age and costs, scale economies, politics, and regulatory regimes (Wait and Adam, 2017; Teodoro 2018, 2019; Beecher and Kalmbach 2013; Su’arez-Varela and Martínez-Espineira, 2018). Moreover, translating these rate differences to the concept of affordability of those rates for the affected service populations remains elusive due to challenges with operationalizing and measuring the concept of affordability. Wait and Adam (2017) determined that on average, customers of privately owned water systems pay $19.56 more per month than customers of publicly owned utilities; however, it is not clear how this difference translates to affordability given variations in household incomes of populations served by different water systems. In a study examining the drivers of water system privatization, Greiner (2016) focused on economic geography and pricing strategy to find that "communities with the greatest probability of having a privatized water system within them are those whose median household income is roughly $70,000 a year," suggesting that privately owned companies may seek communities that can afford rate increases.

Comparing prices relative to household incomes is an aspect of affordability assessment. Teodoro (2019) is the only study we are aware of that uses a statistically representative sample of privately owned and publicly owned utilities to measure water bills relative to household income and compares public and privately owned utilities explicitly. In contrast to Greiner (2016), they calculate water system-specific affordability metrics that normalize water prices by the 20th percentile of income within each system. While they found that privately owned utilities tend to charge more for water relative to relevant minimum wages in the area, this difference did not translate into water bills constituting higher shares of the incomes of the lowest-income quintile of households in their service areas. However, Teodoro does not account for the effects of differences in regulatory jurisdiction. Thus, even with significant research on water prices, empirical research on affordability remains relatively scarce. Part of the problem may be methodological and related to data availability. Almost all of the prior work on quantifying water affordability in the United States relies on matching water rates data of some sort with income information intended to be representative of the service area of the water system, and computing an "affordability ratio," with a measure of household income or expenditure in the denominator and an estimate of household water-specific expenditures in the numerator (Wallsten and Kosec, 2005; Teodoro et al., 2020; Teodoro, 2018, 2019; Schaider et al., 2019; Goddard, 2021). Usually, income or expenditure data characterizing water system service populations are taken from the U.S. Census Places cartographic boundaries corresponding to the city that the water system is nominally tied to in the U.S. Environmental Protection Agency (EPA) Safe Drinking Water Information System (SDWIS) (Wallsten and Kosec, 2005; Teodoro et al., 2020; Teodoro, 2019; Wait and Adam, 2017). This can be a problematic assumption, given inaccuracies in SDWIS (J. Beecher et al., 2020), the degree of socioeconomic and spatial segregation within U.S. cities (Reardon and Bischoff, 2011), and the fact that water system service area boundaries can comprise only parts of cities or span multiple cities (in whole or in part) (Post and Ray, 2020). A recent study by Goddard, (2021) is among the first to our knowledge that uses actual service area boundary polygon data to compute household incomes, pairing these data with a survey question asking utilities to self-report the water bill for a fixed volume of water consumed.

Our study builds on these recent advances and makes two important contributions to the literatures on water affordability and privatization. First, we measure affordability using service area boundaries and actual rate structures as source data. Thus, we create a more accurate combination of income and water bill estimates with which to quantify affordability and trace differences across publicly and privately owned water systems than has been hitherto created within the United States. Second, we disaggregate the typical public/private distinction into more nuanced categories of privately owned water systems that are regulated differently and provide an analysis of the geography of these different types of systems.

This approach allows us to ask two questions. First, who is being served by different types of water systems? We explore whether local government-owned and privately owned systems of various institutional types tend to serve communities of different population size, household income, or location within metropolitan areas of various sizes, or rural regions. We do this by combining US Census data with actual service area boundaries. This builds on work on the concepts of splintering urbanism (Graham and Marvin 2001) and utility sprawl (Pierce et al. 2020), whereby communities receive differentiated access to urban services depending on their location irrespective of administrative or technical rationales. We attempt to discern if there is any pattern to this phenomenon with respect to system ownership.

Secondly, we ask whether water system ownership types differ with respect to affordability, controlling for community characteristics. We perform this analysis by combining service-area matched census data with water bill estimates computed from actual rate structures. We center our study on systems in California due to the unique availability of a large dataset of water system service area boundaries that can be matched with detailed rate structure information and water system ownership type information, as well as to focus the analysis on one common, statewide regulatory regime. Many factors drive water privatization decisions and water pricing decisions, and our analysis is not meant to be causal. Our aim is to provide an overview of the landscape of privatization and affordability in California given the best data available.

2. Methods

In this section we clarify how we characterize the geography and service area populations of water systems, including how we define water systems of different ownership types (Section 4.2.1). We then discuss how we operationalize affordability and compare it across utility ownership types (Section 4.2.2).

4.2.1 The geography of privately owned systems

Our unit of analysis, and delineation between publicly owned and privately owned water systems is central to our approach. As J. Beecher et al. (2020) point out, the most common unit of analysis in utility studies in the U.S. is the "public water system" (PWS), which is the primary unit of record in the EPA SDWIS database. A PWS is a regulatory unit subject to U.S. Safe Drinking Water Act regulations and is defined by the US EPA as a system that "provides water for human consumption to at least 15 service connections or serves an average of at least 35 people for at least 60 days a year." Physically, a PWS generally corresponds to a single, independent water distribution network, which can range from a rest stop bathroom facility sourced from a single well to a metropolitan water system serving several municipalities. Most research on PWS focuses on a subset called "Community Water Systems" (CWS) that serves water to a non-transient population year-round. It should be noted that a CWS is still primarily a physical-regulatory construct focused on water deliveries to end users and does not necessarily correspond to water utilities as administrative or business entities. It may or may not include water supply development and wholesale functions. A single water utility can operate many CWSs. Beecher et al. (2020) propose a typology of CWS that divides them into independent combinations of governmental and non-governmental ownership, and primary (utilities providing water as a primary business function with finances derived from charges for water services) or ancillary structural-functional systems (those providing water to their facilities as part of some other operational function such as schools, hospitals, prisons, mobile home parks, and commercial businesses). They also propose combining CWS as enumerated in SDWIS into the single utility where such arrangements are relevant.

In our study, which in part endeavors to characterize the kinds of communities that rely on different types of water systems for domestic purposes, our unit of analysis is the CWS of California serving residential areas. We attempted conceptually to restrict our analysis to what Beecher et al. (2020) describe as primary CWS, except for also including mobile home parks, which Beecher et al. describe as the "exception" to the rule that ancillary systems are close to "non transient noncommunity systems for regulatory policy and research purposes." We also elected not to combine water systems that are operated by the same utility, because multi-system water utilities often have different rate structures for different (often spatially disparate) systems depending on the local infrastructure cost structures, local contractual agreements, and state regulations.

Our study's universe was thus the set of 2,773 CWS in California identified in SDWIS whose primary service area (as encoded in the SDWIS Service Area Type Code field of the Service Area Table, see (USEPA 2020)) was one of the following types: "Homeowners Association", "Residential Area", "Other Residential Area", "Municipality" (all generally considered to be "primary systems" in the Beecher et al. (2020) typology), as well as "Mobile Home Parks". This is not strictly equivalent to the "primary" definition of Beecher et al. (2020), since we did not assess the extent to which each CWS had water as its "primary" business purpose. We rather proxy this characteristic by what its primary service area is. This means we excluded systems whose primary service areas were non-residential, such as "Industrial/Agricultural", "Institution", "Highway Rest Area", "Medical Facility", "Service Station", and "Water Bottler", that SDWIS nevertheless codes as "community Water Systems."

Characterizing water system ownership types

Table 4.1 Types of ownership of water systems and price regulation in California

Ownership Type	Examples	Nature of price regulation
Local Government	Local water department; Special utility district; Water company owned by multiple local governments.	CA Proposition 218: water fees must not exceed the cost of service attributable to a parcel without a vote of the electorate.
State or Federal	University; Military installation; Parks with campgrounds	None.

Private, for-profit, PUC-regulated	Investor-owned water utility; Homeowners association that sells water to customers outside the HOA.	Regulated by Public Utilities Commission (PUC) on a cost-of-service basis with an allowed profit margin. Low-Income Rate Assistance (LIRA) program required for utilities serving >10,000 population.
Private, non-PUC regulated	Mobile home parks; Self-supplied condominium complex; Private universities	None; must recover costs only, or complaints may result in PUC regulation.
Private Mutual Water Company	Homeowners association with no external customers; Rural water district owned by landowners.	None; must recover costs only, or complaints may result in PUC regulation.
Private Business (non-community)	Amusement park; Self-supplied motel; RV parks; Private recreational facilities.	None.

Privatization can take a variety of institutional forms. We endeavor to describe the kinds of communities (in terms of geography and income distribution) that are served by water systems of different ownership types, as well as the affordability of water rates within these communities. While institutional forms can also vary within publicly owned systems (Piece and Gmoser-Daskalakis, 2021), we cannot account for these variations with the data available.

Most studies dichotomize a “public” or “private” distinction between systems, whether by design or due to data limitations. SDWIS does not differentiate between investor-owned for-profit systems and not-for-profit (NFP) mutual companies, or nominal “water companies” that are owned by local governments (Beecher et al., 2020). We use data from the 2017 California Water Resources Control Board (CSWRCB) electronic Annual Report (eAR), which explicitly asked water systems to identify their “Water System Ownership” type, which corresponds to specific types of water systems that are regulated differently according to the California Water Code. These regulations have implications for water affordability, in terms of the kinds of rate structures that are allowed, and the degree to which cross-subsidy programs for low-income households are possible or even encouraged (California State Water Resources Control Board and UCLA Luskin Center for Innovation, 2020). The ownership types we use throughout the analysis and their regulatory implications are summarized in Table 4.1. Note that the

CSWRCB considers some privately owned businesses that issue water bills to be non-community systems while SDWIS considers them to be community systems.

Characterizing communities served by water systems

Water system service area boundaries (and their ownership) often do not correspond very closely at all with municipal boundaries, resulting in different distributions of household income within a given service area than the city nominally served (See Fig. 4.1). We measure five variables related to water system service areas: (i) median household income, (ii) the 20th percentile of household income, (iii) two system level variables: the percent of households with incomes <200 % of the Federal Poverty Level, and (iv) the Gini coefficient, our inequality measure, as well as (v) population of the urbanized area that the water system is spatially a part of. To compute these demographic measures, we spatially intersected service area boundaries with block group data from the American Community Survey (2019 ACS 5-year). The service area boundaries dataset is curated at the PWS level by the CSWRCB (this same data was most recently used for similar purposes by Goddard (2021)).

Since many census block groups only partially overlapped with service area boundaries, we calculated the measures using an area-based weighting method similar to that employed by Goddard (2021). We computed a weighted sum of the census variable of interest by household count within the block group, multiplied by the proportion of the area of the block group that is overlapped by the service area. This is equivalent to assuming that households are uniformly spatially distributed throughout census block groups, and that household census variables are randomly distributed among these households. This is undoubtedly a limitation but is the best available estimate using publicly available data. To compute the 20th percentile of income and median income, we used census block group count estimates of total households and the number of households within 16 income brackets to create a distribution of household income. We then carried out a linear interpolation between the backet midpoints to estimate the 20th percentile or 50th percentile of this distribution. If a water system's total area was smaller than 50 % of the area of any of the block groups it intersected, the demographic variable of interest was coded as NA. We coded a water system as part of an urbanized area if any part of its service area boundary spatially intersected a 2010 Census Urbanized Area or Urban Cluster, essentially adopting the U.S. Census Bureau's definition of urban and rural (U.S. Census Bureau, 2019). We also recorded the urban area or cluster intersecting each water system and the population of that urban area. We then computed descriptive statistics and performed simple correlational ordinary least squares (OLS) regressions to explore the association between each of our demographic measures (denoted Yk) and water system ownership (denoted $Ownership_i$).

$$Y_k = \Sigma\beta_i Ownership_i + \epsilon$$

All spatial data manipulation was performed using the sf package (Pebesma 2018) for R version 4.0.2 (R Core Team, 2020), with which all statistical analysis was also conducted. Of the 2773 systems in the universe, 2506 systems had Water System Ownership type data in the eAR and were included in a tabulation of population distribution by water system ownership type. Our sample of systems with water system ownership and ACS data included 1310 systems covering 36.1 million people or 91.3 % of the population. We dropped 1463 systems averaging 179 people in size. Table A1 in the appendix compares the percentage of the population within each urbanized area category covered by each water system ownership type, in the total sample and in the sub sample we used for income analysis. While the subsample drops many small systems across all combinations, the largest differences in coverage of the population within each category are limited to rural areas. We cautiously considered this sample representative for urban systems generally with at least 1,000 people, and more confidently for those with at least 10,000 people.

4.2.2 Comparing affordability of water systems across ownership types

We measure three aspects of affordability: 1) Affordability of water bills, 2) shutoff rates for nonpayment, and 3) the availability of subsidy programs to assist low-income households with water bill payment.

Water bill affordability

We use affordability ratios to operationalize and measure "water bill affordability" across water systems. Affordability ratios have a measure of water expenditures in the numerator and a measure of income or total expenditures in the denominator. Generally, affordability ratios assume water expenditures for a given volume of water designated to support basic needs (like drinking, cooking, and basic hygiene). For example, Teodoro (2019) assumes this volume is 50 gallons per capita per day. Goddard (2021) give an overview of possibilities for this volume ranging from 26 to 55 gallons per capita per day, while the World Health Organization standard translates to about 25 gallons per capita per day. There are thus a wide variety of affordability ratios that can be calculated, depending on choices as to what volume of water corresponds to “basic needs” and for what household size, whether to use total income or “discretionary” income, and what level of income is to be considered representative for affordability purposes (Goddard, 2021; Teodoro 2018; Gawel et al., 2013).

We elect to use two affordability ratios, AR_{20} and AR_{50}, evaluated at a variety of household sizes to capture a range of potential affordability scenarios. For AR_{20}, the denominator is the 20th percentile of monthly household income, in order to focus attention on the relatively lowest income quintile of each water system. For AR50, the denominator is the monthly median household income, to reflect affordability challenges that might be faced in communities where large portions of the population have absolutely low incomes. For both ratios, the numerator is the water bill evaluated assuming 50 gallons per capita per day (gpcd) at household sizes from 2 to 7.

$$AR_{20,i} = 100\ x\ \frac{Water\ bill\ (50gpcd\ x\ household\ size_i\ x\ 30\ days)}{20th\ percentile\ of\ area\ annual\ HH\ income\ /\ 12\ months}$$

$$AR_{50,i} = 100\ x\ \frac{Water\ bill\ (50gpcd\ x\ household\ size_i\ x\ 30\ days)}{50th\ percential\ of\ area\ annual\ HH\ income /\ 12\ months}$$

To calculate the water bill, we use a dataset of water rate structures from the California Data Collaborative called the Open Water Rate Specification (OWRS) which was used for a rate survey in 2017 for the American Water Works Association California and Nevada Section (California Data Collaborative, Raftelis, and AWWA California-Nevada Section, 2017). This dataset covers 426 of the largest water systems in California. As such, our water bill affordability analysis can only be considered representative for the largest systems in the state. This contrasts with the eAR, used in a previous analysis (Goddard, 2021), which has greater coverage among small systems (~1600), although only for “typical bills” for set volumes of 6 HCF (hundred cubic feet) and 12 HCF and 24 HCF. We used OWRS because it includes a standardized encoding of the full elaboration of flat rates, fixed fees, surcharges, and possible complex multi-tiered volumetric rates, allowing water bills to be calculated for arbitrary volumes. We use these data to calculate water bills assuming 50 gpcd and a range of household sizes for 5/8″ residential connections. To test the association between the various affordability ratios (denoted AR_k), we estimate OLS regressions of the form:

$$AR_k = \Sigma\beta_i OwnershipType_i + \gamma log(Pop) + \lambda \log(Pop. Urbanized\ Area) + \epsilon$$

This tests for whether ownership types differ in their affordability, controlling for two aspects of population size. We control for each system's own size to capture economies of scale, which have been shown to be associated with water bills in a variety of contexts (Teodoro 2019; Wait and Adam, 2017; Goddard, 2021). We also control for the size of the urbanized area systems are located in to capture the effects of regional agglomeration economies (Eberts and McMillen 1999). Our rationale is that water systems located in larger urban areas may have access to larger pools of labor and expertise, supplies and parts at lower costs due to closer integration with regional, national, and international supply chains; and reduced costs for utilities such as energy, telecommunications, and in some cases, wholesale water supplies due to these infrastructure services themselves being characterized by scale economies. Of the 1310 water systems with ownership type and ACS data, 376 had water rate data and were used in the regressions described above. These systems, almost all of which serve at least 1000 people, collectively serve 31.9 million Californians (81%).

Water shutoffs

One key concern for high water rates is the risk of service shutoffs if households are unable to pay bills. Thus, in addition to measuring water bill affordability, we examine the prevalence of shutoff rates and delinquency rates which are also self-reported in the eAR. In order to quantify whether different ownership types exhibit different propensities to shut off customers, we could ideally measure the shutoff rate as a proportion of accounts missing bill payments:

$$\frac{Total\ accounts\ with\ at\ least\ one\ shutoff\ for\ nonpayment\ over\ the\ year}{Total\ residential\ connections\ missing\ payments\ over\ the\ year}$$

The eAR asks water systems about the number of residential connections they have, as well as the number of water service shutoffs for nonpayment they conducted over the year of the survey. For the 2019 survey year only, the eAR also includes self-reported figures for the number of residential accounts that were in arrears at the end of the year. This makes calculating the shutoff rate impossible even for 2019, since many customers that were shut off might have resolved outstanding bills by the end of the survey year. As an alternative, we calculate two measures using the 2019 version of these data. The shutoff prevalence measures the overall rate of shutoff events in the system:

$$\frac{Total\ shutoffs}{Total\ residential\ connections}$$

The end-year arrears prevalence measures the proportion of residential accounts in arrears at the end of the survey year, as a proxy for the general propensity for accounts to not pay bills. Bill nonpayment could be for many reasons, including inability to pay, unwillingness to pay, forgetfulness, or administrative errors. To test the association between ownership and shutoff prevalence, we estimate OLS regressions of the form:

$$\begin{aligned} ShutoffPrevalence \\ &= \Sigma\beta_i OwnershipType_i + \gamma log(Pop) + \tau AR20 \\ &+ \delta PropAccountsArrearsAtYearEnd + \lambda PercPopBelow200\%FPL + \epsilon \end{aligned}$$

This models the shutoff rate as a function of system ownership, system size, the relative affordability of its water rates, and the prevalence of accounts in arrears, and the relative prevalence of people who might have trouble paying water bills as measured by the Federal Poverty Level. We also estimate a similar model with *PropAccountsArrearsAtYearEnd* as the dependent variable. Our sample for this analysis included 277 systems for which we had ownership type and this shutoff data from the eAR. We interpret these regressions with caution, given the mix of 2017 rate data with 2019 shutoff data. The results are presented in Table 4.4.

Water rate subsidies and payment assistance

Our final measure of affordability relates to the prevalence of water rate subsidies. One strategy recommended by water policy experts to mitigate the risks of water shutoffs is for utilities to offer water rate subsidies or water bill payment assistance programs for low-income households (Pierce et al. 2020; Pierce et al., 2021). The eAR asks water systems whether they offer such programs. The offering of such programs depends both on funding availability as well as a water system's relevant regulatory structure for its rate setting. We use the 2017 version of these data to determine whether a water system had a subsidy program and compare the prevalence of such programs across water system ownership types. Unfortunately, information about the participation rates of eligible households, and the amount of subsidy actually taken advantage of by these households is currently unavailable. However, we find it important to ascertain how many water systems of each type at least offer such programs. A detailed discussion of the different types of subsidies offered is available in Table 4.1. Our sample for this analysis included 373 systems for which we had ownership type and water rate subsidy data from the eAR.

3. Results and Discussion

4.3.1 Water system ownership, geography, and community characteristics

Where do water systems with different ownership types tend to be located? In Table 4.2, we tabulate the count and total population of water systems, broken into categories of overall urbanized area (U.S. Census 2010 delineation) population sizes in CA including "very large urban areas" (>1 million), "large urban areas" (100,000 to 1 million), "small urban areas" (10,000 to 100,000), "small towns" (<10,000 urbanized areas), and "rural areas". For example, 317 water systems serve residents of very large urban areas (i.e. the Los Angeles – Anaheim, San Francisco – Oakland, San Jose, San Diego, Sacramento, and Riverside – San Bernardino urbanized areas).

Table 4.2. Count and population of water systems by ownership and urban area size. Population subtotals may be inaccurate due to rounding.

Ownership Type by Urbanized Area Population Category	N (total = 2506)	% of Systems	Pop. (1000s)	% Pop.	Average Population per System
Rural	1181	47 %	508	1.4 %	430
Local Government	319	27 %	287	56.4 %	899
Privately owned business	52	4 %	11	2.1 %	210
Privately owned Mutual Water Company or Association	359	30 %	80	15.7 %	222
Privately owned, non-PUC-regulated (Community Water System)	359	30 %	65	12.9 %	182
Privately owned, PUC-regulated, for profit water company	92	8 %	66	12.9 %	715
Small Towns (Urban,<10k)	160	6 %	457	1.3 %	2854
Local Government	83	52 %	396	86.6 %	4766
Privately owned business	5	3 %	0.3	0.1 %	65
Privately owned Mutual Water Company or Association	18	11 %	12	2.6 %	671
Privately owned, non-PUC-regulated (Community Water System)	39	24 %	16	3.4 %	398
Privately owned, PUC-regulated, for profit water company	15	9 %	33	7.3 %	2208
Small Urban Areas (10k - 100k)	344	14 %	2910	8.0 %	8473
Local Government	164	48 %	2570	88.3 %	15,695
Privately owned business	10	3 %	1	0.0 %	112
Privately owned Mutual Water Company or Association	64	19 %	91	3.1 %	1424
Privately owned, non-PUC-regulated (Community Water System)	75	22 %	114	3.9 %	1525
Privately owned, PUC-regulated, for profit water company	31	9 %	134	4.6 %	4328
Large Urban Areas (100k - 1 million)	504	20 %	8410	23.1 %	16,689
Local Government	165	33 %	7331	87.2 %	44,429
Privately owned business	14	3 %	3	0.0 %	200
Privately owned Mutual Water Company or Association	127	25 %	45	0.5 %	352
Privately owned, non-PUC-regulated (Community Water System)	151	30 %	59	0.7 %	388
Privately owned, PUC-regulated, for profit water company	47	9 %	97	11.6 %	20,730
Very Large Urban Areas (>1 million)	317	13 %	24,086	66.2 %	75,984
Local Government	168	53 %	19,502	81.0 %	116,082
Privately owned business	3	1 %	1	0.0 %	479
Privately owned Mutual Water Company or Association	58	18 %	316	1.3 %	5450
Privately owned, non-PUC-regulated (Community Water System)	36	11 %	207	0.9 %	5758
Privately owned, PUC-regulated, for profit water company	52	16 %	4060	16.9 %	78,084

Across all urbanized area population categories, local government systems are the largest on average in terms of population per system, with PUC-regulated privately owned utilities being the second largest, and the other privately owned system types being much smaller on average. For example, in very large urban areas the average size of a local government system is 116,000 people, for PUC-regulated systems the average size is 78,000, while other types are on average less than 6000 people. Mutual Water Companies and Non-PUC privately owned systems are smaller and similarly sized on average, but generally larger than Privately Owned Businesses (e.g., private facilities and recreational resorts, etc.) other than in rural areas.

Local government owned systems are also the most numerous and serve the greatest proportion of the population (always >80 %) across all urbanized area size categories except for rural areas where they serve 56 % of the population and where Mutual Water Companies and nonPUC privately owned systems (generally Mobile Home Parks and other small-scale providers) outnumber them. Privately owned businesses remain a small minority across all urbanized area size categories and are generally among the smallest in terms of average size. On the face of it, given the assumption that privately owned systems would exist where the potential for profits is greatest, it was surprising to find that privately owned systems outnumber publicly owned utilities in scattered rural communities where scale economies are difficult. This is where unpacking the types of private institutional arrangements is helpful. We find that among privately owned systems in rural geographies all institutional types are present. Of these however, NFP mutual water companies and non-PUC regulated systems tend to dominate. This is also consistent with the idea that small privately owned systems exist in rural areas by necessity due to the lack of incorporated local governments.

Interestingly, while the average size (in terms of system population) for Local Government and PUC-regulated systems increases intuitively with the size category of the urbanized area, non-PUC systems and Mutual Water Companies are roughly four times larger on average within small urban areas than in small towns. The trend reverses with the next size category, with such systems serving ~1500 residents on average in small urban areas and only ~360 in large urban areas. This may suggest a niche for larger mutual water companies and non-PUC regulated privately owned systems in smaller cities that may indicate a combination of a lack of preexisting public infrastructure with high enough incomes to enable the finance of relatively large water systems in the form of local club goods. For example, Chester, near Lake Almanor, is served by a local government system with 2,300 people, while neighboring resort communities and country clubs around other parts of the lake are served by mutual water companies, one of which serves 6000 people. In larger cities such arrangements may be more suitable to much smaller developments outside the main service areas of existing large water systems that do not expand fully in concert with urban growth. An example of this situation is the constellation of mutual water companies serving subdivisions on the outskirts of the Bakersfield area (Fig. 4.1) most of which serve between 100 and 600 people.

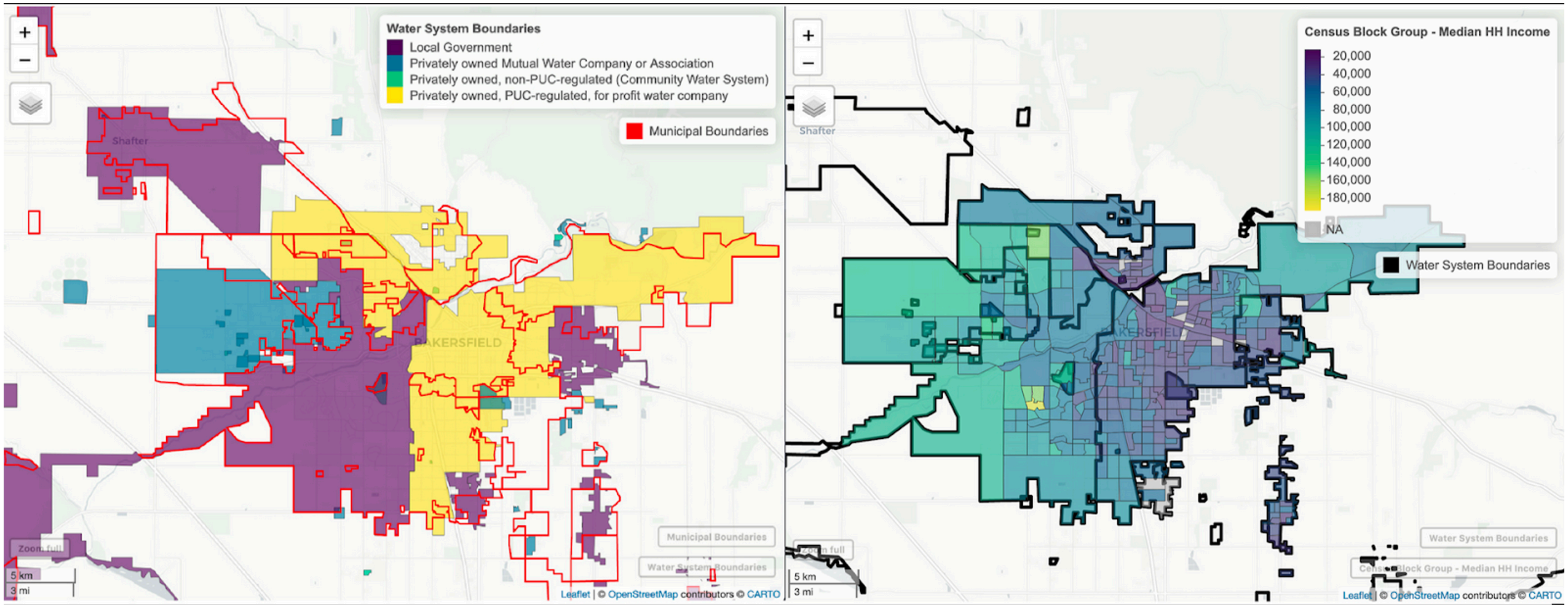


Figure 4.1. Water service areas and municipalities around Bakersfield, CA.

We find that for-profit PUC-regulated utilities exhibit a U-shaped relationship with urbanized area population, serving 12–16% of the population of large and very large urban areas as well as the rural population, but only 5% of the population of small urban areas. The reasons for this merit further research. A possibility is that the scale of small cities is not particularly advantageous for investor-owned utilities due to the lack of potential customers and their preference for the economies of scale to be found in larger cities. It could also reflect the relative lack of presence of local government institutions with the capacity the construct and operate water systems in rural areas.

Thus, with respect to the urban geography of different types of utilities, we show that it is important to unpack the different types of privatization forms. We found that publicly owned and for-profit systems together dominate large urban areas. However, NFP systems are prominent in terms of population covered in small cities and rural areas. In addition, there appears to be an interesting niche for relatively large mutual water companies and non-PUC privately owned systems serving more than 1,000 people in small cities, while they tend to be much smaller in both rural areas and larger urban areas.

In trying to understand differences between publicly owned and PUC-regulated privately owned systems we find that in general, across all urban area population categories, local government and PUC-regulated systems serve households that exhibit similar median incomes as well as similar 20th percentile household incomes. This indicates that local government and PUC-regulated systems tend to serve economically similar communities on average across the metropolitan areas and the smaller cities and towns of California, in contrast to our assumption going on that for-profit water systems might serve communities with higher incomes or include fewer low-income households. In the 10k-100k size category, however, the Mutual Water Companies serve households with higher median and 20th percentile incomes. They also have lower percentages of the population living below 200 % of the Federal Poverty Level than do local government and PUC systems. Interestingly, the relationship is reversed in the 100k-1 million and >1 million urban population size categories. This is suggestive of the Mutual Water Company and non-PUC privately owned systems serving different socioeconomic communities depending on the context of urbanization. We speculate that this could be related to a greater degree of under-bounding of disadvantaged communities in larger cities, who are then served by these less formalized arrangements. In smaller urban areas with less extensive systems, relatively wealthy communities are the ones organizing their own supplies. There is no pattern in the distribution of Gini coefficients between system types.

Regression analyses (Table 4.3) revealed that Mutual Water Companies had higher median and 20th percentile household incomes than Local Government systems ($p < 0.1$), and that Privately Owned Business-type systems had lower median incomes ($p < 0.05$) than Local Government systems, which was a surprise. We found no other statistically significant relationships of note.

Table 4.3. Regression analysis of community income characteristics on water system ownership

	Dependent variable:							
	Median HH Income		20th Perc. HH Income		% Pop. < 200% FPL		Gini HH Income	
County Fixed Effects and Clustered Standard Errors?	No	Yes	No	Yes	No	Yes	No	Yes
Reference Category: Local Government								
Privately owned business (non-community)	-10,852.53** (5,514.54)	-13,591.76*** (4,995.43)	-3,089.66 (3,264.62)	-5,236.90* (2,699.61)	1.32 (3.44)	3.16 (3.35)	0.02 (0.01)	0.02** (0.01)
Mutual Water Company or Association	9,643.10** (4,814.78)	5,250.69* (3,091.93)	5,167.49** (2,407.12)	2,965.11* (1,510.25)	-3.89 (2.53)	-1.13 (1.68)	-0.01* (0.01	-0.01* (0.005)
Privately owned, non-PUC-regulated	2,487.94 (4,371.64)	1,654.79 (2,960.60)	407.32 (2,277.30)	-412.17 (1,517.28)	0.18 (2.6)	0.35 (2.21)	-0.01 (0.01	-0.01 (0.01)
PUC-regulated, for-profit water company	4,778.24 (3,797.96)	482.5 (2,537.37)	2,326.58 (1,764.57)	438.56 (1,149.94)	-2.08 (1.83)	-0.42 (0.38)	-0.01* (0.005)	-0.004 (0.003)
Constant	72,219.51*** (3,211.94)		31,633.49*** (1,417.13)		32.73*** (1.87)		0.39*** -0.004	
Observations	1,310	1,310	1,310	1,310	1,310	1,310	1,310	1,310
R^2	0.01	0.42	0.02	0.35	0.01	0.38	0.01	0.23
Adjusted R^2	0.01	0.4	0.01	0.32	0.01	0.35	0.01	0.19
Residual Std. Error	33,243.07 (df = 1304)	26,005.51 (df = 1247)	16,356.88 (df = 1305)	13,558.49 (df = 1248)	17.05 (df = 1305)	13.83 (df = 1248)	0.05 (df = 1305)	0.05 (df = 1248)
Note:								$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.0$

4.3.2 Affordability

We report our analysis using figures for the affordability ratios calculated assuming a household size of 4 and a monthly water consumption of 6,000 gallons. The analysis was also completed for household sizes of 2–7, but the results were not qualitatively different. Both affordability measures were significantly associated ($p < 0.05$) with PUC-regulated for-profit water companies, indicating that such water systems charge more for water than their local government counterparts, and even other privately owned systems (Table 4.4). This makes sense, given that the Public Utilities Commission regulates their rates including an allowed profit margin, while other types of water systems in the state are nominally not supposed to make a profit. A PUC regulated system on average charges 0.4 percentage points more as a percentage of the income for a family of four at the 20th percentile of household income in its service area than a comparatively sized and located local government system (our AR20 measure of affordability). The difference is 0.2 percentage points for AR50 (where we evaluate affordability for a household at the median income). However, we also find that both system size and the size of the urbanized area in which systems are located are associated with lower affordability ratios. On average, a 10% increase in the population of a system (i.e., system size), or in the size of the urbanized area, is associated with a 0.1 percentage point decrease in AR_{20}, and about half that decrease in AR_{50}. This is consistent with larger systems being able to take advantage of economies of scale to charge lower rates, and systems in larger urban areas being able to take advantage of more ready access to resources, parts, expertise, or even the presence of wholesale water suppliers.

Our results here should be interpreted with caution given that this information is self-reported, and often not reported (with 277 systems out of 376 with OWRS rate data reporting shutoff and arrears data in 2019, and only 289 only shutoff data in 2017). There is no evidence of a strong linkage between the prevalence of accounts with arrears and the prevalence of shutoffs. Neither is there an association between system ownership or AR_{20} and the prevalence of accounts with arrears, although the proportion of the population living under 200% of the Federal Poverty Level is associated ($p < 0.05$) (Table 4.4, column 5).

When it comes to the proportion of households experiencing shutoffs for nonpayment, a counterintuitive pattern emerges. While AR_{20} is strongly positively associated ($p < 0.01$) with shutoff prevalence, as might be expected (a 1 percentage point increase in AR_{20} is associated with a 0.85 percentage point increase in the percentage of households being shut off at least once), PUC-regulated privately owned systems are negatively associated ($p < 0.01$), controlling for AR_{20} and the prevalence of accounts with arrears (Table 4.4, column 4).1 This means that on average, a PUC-regulated privately owned system shuts off fewer accounts than a local government system of equivalent size, water rates and prevalence of arrears. There could be many reasons for this, including the greater ability of for-profit systems to absorb arrearages through de-facto cross-subsidization via higher rates on paying customers and more aggressive collections processes, or greater scrutiny by local water rate advocates, than other systems face. In addition, California CPUC regulations require a number of exemptions, appeal processes, and alternative payment arrangements of for-profit systems that do not universally apply to other system types, although individual local governments may adopt similar measures.

Another important possibility suggested by the eAR data (Table 4.5) is that PUC-regulated systems offer low-income rate assistance programs while most local government systems do not. In fact, almost all people served by PUC-regulated systems are served by systems with such programs, while only 57 % of those served by local government systems are. This is unsurprising – the California PUC requires the systems it regulates with more than 10,000 people to offer rate assistance programs to low-income households. Meanwhile, Proposition 218 limits the ability of local government systems to finance rate subsidy programs via cross-subsidy. The 28% of the (generally very large) local government systems that do have such programs either finance with external sources such as voluntary donations and property tax revenues or use cross-subsidies that are enabled by local electoral measures (UNC Environmental Finance Center, 2017).

Even if enrollment is low relative to the eligible population of customers, the availability of such subsidies may help offset some of the more expensive rates that customers of PUC-regulated systems experience, helping them avert shutoffs. That said, measuring the extent to which these programs define eligibility and benefit amounts appropriately given their contexts and successfully enroll the eligible population requires more data than is available from the eAR and should be a topic for further mixed methods research. Moreover, the persistent gap in availability of these programs to residents of local government systems may motivate measures such as modifications to Proposition 218 or a statewide water rate assistance program (Pierce et al., 2020).

Regarding affordability, our combination of service area boundaries, census data, and detailed rate information, allows us to confirm, as well as complicate, several ideas from the broader affordability literature. We find evidence consistent with economies of scale in this sector, with larger systems in larger urban areas having more affordable water rates on average. We also find that for-profit systems are somewhat less affordable than their local government and privately owned, NFP counterparts. At the same time, we find that PUC regulated privately owned systems offer more low-income subsidies than their publicly owned counterparts as is mandated by California's regulatory system. There are also fewer shutoffs among privately owned systems than publicly owned ones, suggesting tentative evidence for affordability being related to shutoff rates, and for subsidy programs mitigating financial stress that may come with higher prices and/or associated shutoffs for nonpayment.

Table 4.4. Affordability Regressions

	Dependent variable:				
	AR_{20}	AR_{50}	Proportion Connections Shutoff (2017)	Proportion Connections Shutoff (2019)	Proportion Connections Arrears (2019)
	(1)	(2)	(3)	(4)	(5)
Reference Category: Local Government					
Mutual Water Company	0.00	0.00	0.00	-0.01	-0.07
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Privately owned, non-PUC-regulated	0.00	0.00	-0.02*	-0.02*	-0.04
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)

PUC-regulated, for-profit water company	0.004**	0.002***	-0.01***	-0.02***	-0.02
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
log(Pop of water system)	-0.001**	-0.0005**	-0.01*	-0.005**	-0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
log(Pop of urbanized area where system is located)	-0.001***	-0.0005***			
	(0.00)	(0.00)			
AR_{20}			0.66***	0.85***	0.28
			(0.17)	(0.22)	(0.92)
Percent Pop < 200% FPL			0.00	0.00	0.00
			(0.00)	(0.00)	(0.00)
Percent Connections in Arrears				0.01	
				(0.01)	
Constant	0.05***	0.02***	0.06*	0.06*	0.06*
	(0.01)	(0.00)	(0.03)	(0.03)	(0.03)
Observations	376	376	289	277	277
R^2	0.16	0.16	0.09	0.12	0.03
Adjusted R^2	0.15	0.15	0.07	0.1	0.01
Residual Std. Error	0.01 (df = 370)	0.005 (df = 370)	0.04 (df = 282)	0.04 (df=268)	0,.04 (df=269)
Note:	*p<0.1; **p<0.05; ***p<0.01				

Table 4.5. Water rate subsidy offering by water system ownership types

Ownership	Total Systems in Sample	Total Systems answering eAR regarding subsidies	Total Systems with Subsidies	% of aggregate population in systems with subsidies
Local Government	314	296	83	57%
Mutual Water Company	14	9	0	0%
Privately owned, non-PUC-regulated system	9	6	3	27%
Privately owned, PUC-regulated system	69	62	57	99%

4.4 Conclusions

Our study makes several substantive and methodological contributions to the literature on the affordability and geography of water system privatization, including insights into different institutional types of privately owned water systems. This institutional unpacking allows us to clarify where different kinds of water systems are most prevalent and who is served by them across urban and rural geographies as well as within exact service area boundaries (in terms of income distribution and population size). Overall, our results suggest that state level regulatory structures (such as requiring for-profit systems to offer low-income assistance programs) and the variety of institutional forms of water systems might be more important in terms of how households experience affordability than who owns the systems.

4.4.1 Affordability

For-profit systems in California, while serving similar communities in terms of income distribution to NFP and local government systems, tend to have more expensive water charges, including for poorer households (roughly 0.4 percentage points more than other water system types as a percentage the 20th percentile of income within a given service area). This is perhaps unsurprising, given that they must return profits, pay taxes, and may incur higher capital and financing costs (Beecher 2021). In spite of their higher water charges, for-profit systems exhibit the lowest shutoff prevalence among their service accounts in California. In addition to direct regulations of shutoffs in for-profit systems by the PUC, 99 % of residents served by for-profit systems at least nominally have access to a low-income rate subsidy program, while only 57 % of residents served by local government, and less than 30 % of residents served by privately owned NFP systems do. All of the above points to why the nature of regulation of both publicly owned and privately owned water systems affects how rates translate to affordability as experienced by households more so than whether or not a water system is owned by a public or private entity. In the literature, some say privately owned systems represent a grave affordability threat to low-income customers (Barlow and Clarke 2017; Bakker 2010), while others contend private ownership is an important antidote to inefficient local government ownership (World Bank 2003; Galiani et al., 2005).

Our findings instead suggest that the institutional arrangements surrounding these systems might be at least as important to how consumers receive water services as who in particular owns the systems, or what nominal rates may be, while highlighting the potential importance of low-income rate subsidies. In California, the particular set of institutions that have been built around for-profit privately owned systems may serve to protect their customers from shutoffs more than customers of local government systems are theoretically protected via local political processes. Indeed, the California Public Utilities Commission (CPUC) has adopted policies around disconnections that are more protective than the statutory requirement (Feinstein and Warner 2018). These strict and deliberative shutoff rules for investor-owned utilities are more broadly noteworthy. With for-profit utilities there may be concerns of rent-seeking: if left to their own devices, some investor-owned utilities might be more inclined to disconnect based on their business orientation or corporatized models. More stringent regulation of privately owned systems may then, ironically, end up providing greater protection for low-income consumers in for-profit system service areas than under local government run systems. The evidence presented here is enough only to raise this possibility; these relationships should be investigated further.

4.4.2 Geography

There does appear to be a spatial pattern in terms of which institutional types of water systems are most common in which types of urban or rural contexts. We find some evidence for “splintering urbanism” (Graham and Marvin 2001), with echoes of patterns found in low- and middle-income countries, where communities rich and poor alike are served by different combinations of providers rather than by universal water services (Kooy and Bakker 2008; Botton and Gouvello 2008; Z´erah 2008). Small, NFP water systems are important for providing water services to large proportions of the population of California residing in unincorporated communities, including in peripheries and potentially underbounded gaps within large urban areas, as well as core areas of small cities and in rural areas. Such systems seem to serve different roles by context, with some providing service to relatively disadvantaged communities where local governments do not exist (as in rural or underbounded communities) or where for-profit providers are not interested, and others being organized by relatively wealthy communities who self-incorporate or for one reason or another are not served by a local government system. In smaller cities, these NFP providers appear to have a niche serving unincorporated, yet wealthy communities organizing water systems for themselves. In larger urban areas, this niche may be occupied instead by systems owned by fragmented local governments. We note that this pattern likely depends on the particular regulatory frameworks in force in California and may not be present elsewhere.

4.4.3 Methodological Implications

Our study demonstrates both the importance and feasibility of conducting affordability analysis by using high-fidelity data. Water system boundaries can vary considerably from municipal boundaries. To our knowledge no study in the United States has addressed the relationship between privatization and affordability while making use of appropriate demographic information (i.e., service area boundaries) as well as detailed water rate information together in the same analysis dataset. This makes a difference in understanding whom utilities actually serve and how affordability can be calculated. We encourage scholars that use SDWIS to enumerate samples of water systems to seek out water system boundaries if they wish to calculate system-level socioeconomic covariates. In addition, water tariffs and rate structures can make it challenging to calculate typical bills. Thus, we encourage scholars to both use and contribute to open, fully specified water rate datasets such as the OWRS.

4.4.4 Future Work

A key interesting finding in the paper was that among the large water systems in California for which we had price data, for-profit water systems exhibited both relatively high water rates and lower shutoff rates, with data indicative of low-income rate subsidy programs as a mediating factor. This suggests that any water affordability analysis is incomplete without being able to account for such programs. Further research should investigate the extent to which low-income households are eligible for such programs and successfully avail themselves of them across water systems. Ideally, using mixed methods, such research should focus on household-level data rather than system-level aggregates to most accurately measure the relationship between household incomes, size, water consumption levels, water expenditures net of subsidies, and risks of water shutoffs or financial hardships imposed by water bill late payment penalties.

REFERENCES

Bakker, Karen. "Commons versus Commodities: Political Ecologies of Water Privatization." In *Global Political Ecology*, 347–70. Taylor and Francis, 2010. https://www.taylorfrancis.com/books/edit/10.4324/9780203842249/global-political-ecology-richard-peet-paul-robbins-michael-watts.

Bakker, Karen J. "Privatizing Water, Producing Scarcity: The Yorkshire Drought of 1995*." *Economic Geography* 76, no. 1 (February 16, 2009): 4–27. https://doi.org/10.1111/j.1944-8287.2000.tb00131.x.

Barlow, Maude, and Tony Clarke. *Blue Gold: The Battle against Corporate Theft of the World's Water*, 2017. https://www.taylorfrancis.com/books/9781315096216.

Beecher, Janice A., and Jason A. Kalmbach. "Structure, Regulation, and Pricing of Water in the United States: A Study of the Great Lakes Region." *Utilities Policy* 24 (March 2013): 32–47. https://doi.org/10.1016/j.jup.2012.08.002.

Beecher, Janice, Kyle Redican, and Megan Kolioupoulos. "(Mis)Classification of Water Systems in the United States." *SSRN Electronic Journal*, 2020. https://doi.org/10.2139/ssrn.3627915.

Bel, Germà, Xavier Fageda, and Mildred E. Warner. "Is Private Production of Public Services Cheaper than Public Production? A Meta-Regression Analysis of Solid Waste and Water Services." *Journal of Policy Analysis and Management* 29, no. 3 (June 3, 2010): 553–77. https://doi.org/10.1002/pam.20509.

Bel, Germà, and Mildred Warner. "Does Privatization of Solid Waste and Water Services Reduce Costs? A Review of Empirical Studies." *Resources, Conservation and Recycling* 52, no. 12 (October 2008): 1337–48. https://doi.org/10.1016/j.resconrec.2008.07.014.

Budds, Jessica, and Gordon McGranahan. "Are the Debates on Water Privatization Missing the Point? Experiences from Africa, Asia and Latin America." *Environment and Urbanization* 15, no. 2 (October 1, 2003): 87–114. https://doi.org/10.1177/095624780301500222.

Byatt, Ian. "25 Years of Regulation of Water Services" *Utilities Policy* 48 (October 2017): 103–8. https://doi.org/10.1016/j.jup.2017.09.006.

California Data Collaborative, Raftelis, and AWWA California-Nevada Section. "2017 California-Nevada Water Rate Survey." CA-NV AWWA, 2017. https://ca-nv-awwa.org/canv/downloads/2018/CA-NV_RateSurvey-2017_final.pdf.

California State Water Resources Control Board and UCLA Luskin Center for Innovation. "Recommendations for Implementation of a Statewide Low-Income Water Rate Assistance Program," 2020. https://www.waterboards.ca.gov/water_issues/programs/conservation_portal/assistance/docs/ab401_report.pdf.

Camdessus, Michael. "Financing Water for All," 2003.

Cesar, Sylvia. "Privatization of Water: Evaluating Its Performance in the Developing World." *Annals of Public and Cooperative Economics* 90, no. 1 (2019): 5–23. https://doi.org/10.1111/apce.12221.

Cotta, Sabrina A. "Privatization and Water Service Provision in the United States: A Recommendation for Expanded Oversight and the Development and Adoption of Best Practices." *Water International* 37, no. 7 (November 1, 2012): 818–30. https://doi.org/10.1080/02508060.2012.731719.

Davis, Jennifer. "PRIVATE-SECTOR PARTICIPATION IN THE WATER AND SANITATION SECTOR." *Annual Review of Environment and Resources* 30, no. 1 (November 21, 2005): 145–83. https://doi.org/10.1146/annurev.energy.30.050504.144635.

Association of Clean Water Administrators. "Developing a New Framework for Household Affordability and Financial Capability Assessment in the Water Sector." Accessed April 29, 2021. https://www.acwa-us.org/documents/developing-a-new-framework-for-household-affordability-and-financial-capability-assessment-in-the-water-sector/.

Dieter, Cheryl A., Molly A. Maupin, Rodney R. Caldwell, Melissa A. Harris, Tamara I. Ivahnenko, John K. Lovelace, Nancy L. Barber, and Kristin S. Linsey. "Estimated Use of Water in the United States in 2015." Report. Circular. Reston, VA, 2018. USGS Publications Warehouse. https://doi.org/10.3133/cir1441.

"Drinking Water Partnerships and Consolidation | California State Water Resources Control Board." Accessed April 29, 2021. https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/waterpartnership.html.

Eberts, Randall W., and Daniel P. McMillen. "Chapter 38 Agglomeration Economies and Urban Public Infrastructure." In *Handbook of Regional and Urban Economics*, 3:1455–95. Elsevier, 1999. https://doi.org/10.1016/S1574-0080(99)80007-8.

"Estimated Use of Water in the United States." Circular. Circular, 2018.

Goddard, Jessica J., Isha Ray, and Carolina Balazs. "Water Affordability and Human Right to Water Implications in California." *PLOS ONE* 16, no. 1 (January 20, 2021): e0245237. https://doi.org/10.1371/journal.pone.0245237.

Greiner, Patrick Trent. "Social Drivers of Water Utility Privatization in the United States: An Examination of the Presence of Variegated Neoliberal Strategies in the Water Utility Sector: Drivers of Water Utility Privatization." *Rural Sociology* 81, no. 3 (September 2016): 387–406. https://doi.org/10.1111/ruso.12099.

Helm, Dieter. "Thirty Years after Water Privatization—Is the English Model the Envy of the World?" *Oxford Review of Economic Policy* 36, no. 1 (2020): 69–85.

Homsy, George C., and Mildred E. Warner. "Does Public Ownership of Utilities Matter for Local Government Water Policies?" *Utilities Policy* 64 (June 2020): 101057. https://doi.org/10.1016/j.jup.2020.101057.

Kirkpatrick, Colin, David Parker, Yin-Fang Zhang, Colin Kirkpatrick, David Parker, and Yin-Fang Zhang. "State versus Private Sector Provision of Water Services in Africa: A Statistical, DEA and Stochastic Cost Frontier Analysis," 2004. https://doi.org/10.22004/AG.ECON.30604.

Levin, Ronnie B, Paul R Epstein, Tim E Ford, Winston Harrington, Erik Olson, and Eric G Reichard. "U.S. Drinking Water Challenges in the Twenty-First Century." *Environmental Health Perspectives* 110, no. suppl 1 (February 2002): 43–52. https://doi.org/10.1289/ehp.02110s143.

Mack, Elizabeth A., and Sarah Wrase. "A Burgeoning Crisis? A Nationwide Assessment of the Geography of Water Affordability in the United States." Edited by Jaymie Meliker. *PLOS ONE* 12, no. 1 (January 11, 2017): e0169488. https://doi.org/10.1371/journal.pone.0169488.

Meehan, Katie, Wendy Jepson, Leila M. Harris, Amber Wutich, Melissa Beresford, Amanda Fencl, Jonathan London, et al. "Exposing the Myths of Household Water Insecurity in the Global North: A Critical Review." *WIREs Water* 7, no. 6 (November 2020). https://doi.org/10.1002/wat2.1486.

Megdal, Sharon B. "The Role of the Public and Private Sectors in Water Provision in Arizona, USA." *Water International* 37, no. 2 (March 2012): 156–68. https://doi.org/10.1080/02508060.2012.662732.

Nauges, Céline, and Caroline van den Berg. "Economies of Density, Scale and Scope in the Water Supply and Sewerage Sector: A Study of Four Developing and Transition Economies." *Journal of Regulatory Economics* 34, no. 2 (October 2008): 144–63. https://doi.org/10.1007/s11149-008-9063-2.

Pebesma, Edzer. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10, no. 1 (2018): 439. https://doi.org/10.32614/RJ-2018-009.

Pierce, Gregory, Nicholas Chow, and J.R. DeShazo. "The Case for State-Level Drinking Water Affordability Programs: Conceptual and Empirical Evidence from California." *Utilities Policy* 63 (April 2020): 101006. https://doi.org/10.1016/j.jup.2020.101006.

———. "The Case for State-Level Drinking Water Affordability Programs: Conceptual and Empirical Evidence from California." *Utilities Policy* 63 (April 2020): 101006. https://doi.org/10.1016/j.jup.2020.101006.

Pierce, Gregory, Ahmed Rachid El-Khattabi, Kyra Gmoser-Daskalakis, and Nicholas Chow. "Solutions to the Problem of Drinking Water Service Affordability: A Review of the Evidence." *WIREs Water*, March 25, 2021. https://doi.org/10.1002/wat2.1522.

Pierce, Gregory, and Kyra Gmoser-Daskalakis. "Multifaceted Intra-City Water System Arrangements in California: Influences and Implications for Residents." *Utilities Policy* 71 (August 2021): 101231. https://doi.org/10.1016/j.jup.2021.101231.

Post, Alison, and Isha Ray. "Hybrid Modes of Urban Water Delivery in Low- and Middle-Income Countries." In *Oxford Research Encyclopedia of Environmental Science*, by Alison Post and Isha Ray. Oxford University Press, 2020. https://doi.org/10.1093/acrefore/9780199389414.013.679.

R Core Team. "R: A Language and Environment for Statistical Computing." Vienna: R Foundation for Statistical Computing, 2020.

Reardon, Sean F., and Kendra Bischoff. "Income Inequality and Income Segregation." *American Journal of Sociology* 116, no. 4 (January 2011): 1092–1153. https://doi.org/10.1086/657114.

Renzetti, Steven, and Diane Dupont. "Ownership and Performance of Water Utilities." *Greener Management International*, no. 42 (2003): 9–19.

Risse, Mathias. "The Human Right to Water and Common Ownership of the Earth." In *Political Theory Without Borders*, 46–73. John Wiley & Sons, Ltd, 2015. https://doi.org/10.1002/9781119110132.ch4.

Robbins, Peter T. "Transnational Corporations and the Discourse of Water Privatization." *Journal of International Development* 15, no. 8 (November 2003): 1073–82. https://doi.org/10.1002/jid.1054.

Rogers, P. "Water Is an Economic Good: How to Use Prices to Promote Equity, Efficiency, and Sustainability." *Water Policy* 4, no. 1 (2002): 1–17. https://doi.org/10.1016/S1366-7017(02)00004-1.

Schaider, Laurel A., Lucien Swetschinski, Christopher Campbell, and Ruthann A. Rudel. "Environmental Justice and Drinking Water Quality: Are There Socioeconomic Disparities in Nitrate Levels in U.S. Drinking Water?" *Environmental Health* 18, no. 1 (January 17, 2019): 3. https://doi.org/10.1186/s12940-018-0442-6.

Segerfeldt, Fredrik. *Water for Sale: How Business and the Market Can Resolve the World's Water Crisis*. Washington, D.C.: Lanham, MD: Cato Institute ; Distributed to the trade by National Book Network, 2005.

Suárez-Varela, Marta, and Roberto Martínez-Espiñeira. "A Proposal for the Analysis of Price Escalation within Water Tariffs: The Impact of the Water Framework Directive in Spain." *Environment and Planning C: Politics and Space* 36, no. 4 (June 2018): 726–49. https://doi.org/10.1177/2399654417719558.

———. "A Proposal for the Analysis of Price Escalation within Water Tariffs: The Impact of the Water Framework Directive in Spain." *Environment and Planning C: Politics and Space* 36, no. 4 (June 1, 2018): 726–49. https://doi.org/10.1177/2399654417719558.

Subramaniam, Mangala, and Beth Williford. "Contesting Water Rights: Collective Ownership and Struggles against Privatization." *Sociology Compass* 6, no. 5 (2012): 413–24. https://doi.org/10.1111/j.1751-9020.2012.00461.x.

Teodoro, Manuel P. "Measuring Household Affordability for Water and Sewer Utilities." *Journal AWWA* 110, no. 1 (2018): 13–24. https://doi.org/10.5942/jawwa.2018.110.0002.

———. "Water and Sewer Affordability in the United States." *AWWA Water Science* 1, no. 2 (March 2019): e1129. https://doi.org/10.1002/aws2.1129.

Teodoro, Manuel P., Youlang Zhang, and David Switzer. "Political Decoupling: Private Implementation of Public Policy." *Policy Studies Journal* 48, no. 2 (May 2020): 401–24. https://doi.org/10.1111/psj.12287.

UNC Environmental Finance Center. "Navigating Legal Pathways to Rate-Funded Customer Assistance Programs: A Guide for Water and Wastewater Utilities," 2017. https://efc.sog.unc.edu/wp-content/uploads/sites/1172/2021/06/Nagivating-Pathways-to-Rate-Funded-CAPs.pdf.

U.S. Census Bureau. "2010 Census Urban and Rural Classification and Urban Area Criteria." 2010 Census Urban and Rural Classification and Urban Area Criteria, 2019. https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/2010-urban-rural.html.

USEPA. "Search Results Help - Drinking Water." ECHO, 2020. https://echo.epa.gov/help/facility-search/drinking-water-search-results-help#address.

Wait, Isaac W., and William Adam Petrie. "Comparison of Water Pricing for Publicly and Privately Owned Water Utilities in the United States." *Water International* 42, no. 8 (November 17, 2017): 967–80. https://doi.org/10.1080/02508060.2017.1406782.

Wallsten, Scott, and Katrina Kosec. "Public or Private Drinking Water? The Effects of Ownership and Benchmark Competition on U.S. Water System Regulatory Compliance and Household Water Expenditures." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, March 1, 2005. https://papers.ssrn.com/abstract=707131.

Winpenny, J. T. *Managing Water as an Economic Resource*. Development Policy Studies. London; New York: Routledge, 1994.

World Bank. "World Development Report 2004: Making Services Work for Poor People," 2003. http://hdl.handle.net/10986/5986

CHAPTER 5: CONCLUSION

Indicators, metrics, and service level benchmarks are important tools in planning and policy generally, and for services such as water and sanitation. They are important for identifying problems, setting goals, implementing, evaluating, and legitimizing investments, programs and projects that meet shared goals. I summarize below the major findings and limitations of each of the three studies presented here, and then discuss implications for future work.

Whom should policy serve? The first paper is motivated by the premise that those that cannot be counted are often not adequately served. Focusing on urban services, in many low resourced, data scarce countries, there are significant gaps in accurately representing local urban populations due to data collection mechanisms that focus on arbitrary jurisdictional categories and boundaries. In the first paper I showed that different criteria for deciding what is 'urban' can have profound impacts on estimations of the urban population, which in turn, has implications for infrastructure finance, service provision and urban governance. The paper also highlighted the potential of using diverse sources of available data in creative ways to address the problem of accurate delineation and counting.

In my second paper, I showed that only certain subgroups of households increase their demand for water upon transitioning from a reliable, high-quality IWS to CWS. In so doing, I raised the possibility that service level standards and upgrades can become traps that result in Janus-faced outcomes in terms of sustainability and access, and ultimately have equity consequences. Converting from IWS to CWS represents a large expenditure, but the nature of the starting point for IWS matters, as does the sequence in which the transition to CWS occurs. CWS does not appear to represent a service level improvement over predictable IWS for many households. Such households are unlikely to be substantial sources of additional revenue for utilities that invest in these marginal improvements. Without additional revenues, sustainability suffers as upgrades cannot be maintained in the face of growing population pressures. This is the middle-service-level trap. The existence of this trap suggests that CWS conversion efforts should budget for several key investments so as not to slide to Satisfied IWS in short order: bulk water supply sufficient to serve projected population growth, an aggressive and continuous leak detection and repair program, subsidies necessary to serve poor people by standpipe or private connection.

My third paper asked how the ownership of water systems in California is associated with metrics of affordability. This paper illustrates that measurement of affordability is itself a combined conceptual and data problem like that of urban determination. It engages with questions such as: How should affordability be defined? Assuming it is linked to a relationship between an expenditure and relative income, what income thresholds should be used for planning purposes and for considering tariff and subsidy structures? I show that while private, for-profit water utilities charge more for water on average relative to the incomes of their customers, they also tend to cut off their customers at lower rates than public utilities with similarly "unaffordable" rates. I speculate that this may be due to regulations regarding customer assistance programs and closer scrutiny on disconnections faced by investor-owned utilities in California. My findings suggest that governance and regulations are probably more important to outcomes of affordability, like service cutoffs, than the nominal ownership type of a water system.

Considering my findings throughout these papers, I describe three future streams of work. First, concepts around environmental justice are being operationalized into initiatives like the Biden Administration's Justice40, which is supposed to direct federal funding to historically marginalized communities in the United States, but there are a number of thorny questions about how to create and monitor such a project (Baptista et al 2021). There is a focus on nation-scale metrics that can be used in "screening tools" for identifying and targeting communities that are marginalized and thus particularly eligible for funding under Justice40 (White House 2022). There is currently no coherent suite of indicators regarding water and sanitation being considered in the processes surrounding Justice40 that could be used to direct funding from federal programs such as Drinking Water and Clean Water Revolving Funds. One possible reason is a lack of nation-scale data regarding the service area boundaries of drinking water systems that would be necessary to calculate metrics about socioeconomic status or other indicators relevant to historic marginalization. Many U.S. states have agency programs that collect these data, but most do not. I am currently involved in efforts to collate existing data from relevant states, engage with state agencies to help them create programs to collect drinking water boundaries, and create web-based tools that enable the crowdsourcing, validation, and improvement of these boundary datasets over time. I hope the resulting datasets become widely available to practitioners in the environmental justice space as well as academics who study water utilities in the United States.

In a related future stream of work, I want to engage with questions regarding the appropriate spatial scale of water and sanitation governance in an urban area, and pathways forward given current states of fragmentation. Drinking water governance is quite fragmented in the United States (and elsewhere). There are over 50,000 water systems in the United States operated by diverse entities that include local government departments, multi-government umbrella authorities, investor-owned utilities, and small-scale sole proprietor systems like mobile home parks. Moreover, all of these might have complex relationships with each other, such as wholesale/resale relationships, concessions, and operating contracts. These systems also often spatially cross multiple jurisdictional boundaries. There is evidence that this fragmentation of water infrastructure management can lead to highly inequitable outcomes in terms of water quantity and quality. For example, small systems do not enjoy the economies of scale that larger systems do, and often have trouble meeting water quality regulations or securing water supply mixes resilient to drought. Conversely, historical processes such as white flight have led to the hollowing out of the finances and technical and managerial capacity of many urban water systems that lost customers to wealthier suburban municipalities (and their fragmented water systems). I would like to study processes of water system consolidation. I want to investigate the conditions under which efforts to consolidate disadvantaged water systems together or with better-resourced systems succeed, and the outcomes of these consolidations for their service populations.

Finally, I would like to continue my work on Amravati and continuous water supply in low- and middle-income countries more broadly. I have access to bulk meter volume and pressure readings for the operating zones and district metering areas from MJP and would like to trace the trends in non-revenue water in the IWS and CWS zones over time as the CWS zones were converted. This will allow me to validate the model of Taylor et. al. (2019). I would also be interested in investigating other cases around the world of IWS to CWS conversions using similar methods as this dissertation's Chapter 3 where data is available.

REFERENCES

Baptista, Ana Isabel, Nicky Sheats, and Cathleen Kelly. 2021. “Call for Papers: Special Environmental Justice Issue on the Biden Administration’s Justice40 Initiative: Deadline for Manuscript Submission: December 31, 2021.” *Environmental Justice* 14 (6): 437–38.

White House. “CEQ Publishes Draft Climate and Economic Justice Screening Tool, Key Component in the Implementation of President Biden’s Justice40 Initiative.” White House press release, February 18, 2022. https://www.whitehouse.gov/ceq/news-updates/2022/02/18/ceq-publishes-draft-climate-and-economic-justice-screening-tool-key-component-in-the-implementation-of-president-bidens-justice40-initiative/, accessed 10 October 2022

.

APPENDIX

Table A1. Two-way fixed effects regression comparing CWS (HSR and Maya Nagar) with IWS households (full results)

	Dependent variable:			
	log(bi-monthly consumption)			
	Unmatched	Unmatched, omit peak periods	Matched	Matched, omit peak periods
	(1)	(2)	(3)	(4)
CWS	0.137***	0.172***	0.138**	0.155***
	(0.045)	(0.048)	(0.054)	(0.054)
\`lag_MP(fit)\`	-0.114	-0.134	-0.120	-0.127
	(0.146)	(0.147)	(0.182)	(0.171)
Observations	301,963	207,938	142,096	97,759
R^2	0.566	0.558	0.566	0.572
Adjusted R^2	0.527	0.499	0.528	0.515
Residual Std. Error	0.582 (df = 277302)	0.597 (df = 183281)	0.569 (df = 130551)	0.575 (df = 86218)
Note:				*p<0.1; **p<0.05; ***p<0.01
				robust standard errors, clustered by household

Table A2. Basic regression comparing households that reverted to IWS (Maya Nagar) to CWS households between June-July 2011 and February-March 2013 (full results)

	Dependent variable:			
	log(bi-monthly consumption)			
	Unmatched	Unmatched, omit peak periods	Matched	Matched, omit peak periods
	(1)	(2)	(3)	(4)
loseCWS	-0.010	-0.027**	-0.051**	-0.049*
	(0.008)	(0.011)	(0.025)	(0.028)
\`lag_MP(fit)\`	0.158	0.048	-0.272	-0.189
	(0.167)	(0.153)	(0.271)	(0.229)
Observations	91,368	63,931	23,925	16,783
R^2	0.632	0.675	0.441	0.564
Adjusted R^2	0.586	0.614	0.371	0.482
Residual Std. Error	0.521 (df = 81165)	0.500 (df = 53782)	0.634 (df = 21236)	0.571 (df = 14110)
Note:				*p<0.1; **p<0.05; ***p<0.01
				robust standard errors, clustered by household

Table A3. Effect of transition from IWS to CWS by decile of consumption in October-November 2009, comparison of HSR to IWS groups (full results)

	Dependent variable:			
	log(bi-monthly consumption)			
	Unmatched	Unmatched, omit peak periods	Matched	Matched, omit peak periods
	(1)	(2)	(3)	(4)
CWS:decile_init1	0.325***	0.388***	0.412***	0.488***
	(0.092)	(0.098)	(0.113)	(0.119)
CWS:decile_init2	0.222***	0.265***	0.155*	0.168
	(0.077)	(0.095)	(0.080)	(0.106)
CWS:decile_init3	0.318***	0.386***	0.279***	0.348***
	(0.064)	(0.083)	(0.075)	(0.099)
CWS:decile_init4	0.136***	0.156***	0.123**	0.111*
	(0.045)	(0.054)	(0.053)	(0.062)
CWS:decile_init5	0.142**	0.098	0.123*	0.053
	(0.059)	(0.061)	(0.069)	(0.068)
CWS:decile_init6	0.028	0.030	0.008	0.012
	(0.050)	(0.062)	(0.058)	(0.076)
CWS:decile_init7	0.057	0.119*	0.019	0.073
	(0.046)	(0.064)	(0.054)	(0.078)
CWS:decile_init8	0.081	0.121**	0.083	0.123
	(0.052)	(0.061)	(0.066)	(0.078)
CWS:decile_init9	0.082*	0.079	0.074	0.062
	(0.048)	(0.054)	(0.053)	(0.060)
CWS:decile_init10	0.071**	0.083*	0.107***	0.124**
	(0.034)	(0.044)	(0.039)	(0.049)
Observations	293,642	209,288	139,872	99,512
R^2	0.600	0.604	0.609	0.616
Adjusted R^2	0.568	0.557	0.577	0.571
Residual Std. Error	0.554 (df = 271438)	0.558 (df = 187088)	0.536 (df = 129331)	0.537 (df = 88975)
Note:				*p<0.1; **p<0.05; ***p<0.01
				robust standard errors, clustered by household

Table A4. Effect of transition from CWS to IWS by decile of consumption in October-November 2009, comparison of Maya Nagar to other CWS groups (full results)

	Dependent variable:			
	log(bi-monthly consumption)			
	Unmatched	Unmatched, omit peak periods	Matched	Matched, omit peak periods
	(1)	(2)	(3)	(4)
loseCWS:decile_init1	-0.021	0.016	-0.055	-0.005
	(0.027)	(0.051)	(0.082)	(0.104)
loseCWS:decile_init2	0.018	0.073	0.005	0.062
	(0.038)	(0.061)	(0.075)	(0.117)
loseCWS:decile_init3	0.002	0.024	0.030	0.036
	(0.032)	(0.041)	(0.065)	(0.074)
loseCWS:decile_init4	0.033	0.048	-0.008	0.016
	(0.027)	(0.043)	(0.063)	(0.087)
loseCWS:decile_init5	-0.005	0.005	-0.00001	0.022
	(0.020)	(0.029)	(0.062)	(0.075)
loseCWS:decile_init6	-0.032	-0.055	-0.104	-0.141
	(0.022)	(0.036)	(0.069)	(0.093)
loseCWS:decile_init7	0.009	0.012	-0.014	-0.006
	(0.018)	(0.029)	(0.050)	(0.065)
loseCWS:decile_init8	-0.013	-0.042	-0.059	-0.052
	(0.022)	(0.030)	(0.053)	(0.064)
loseCWS:decile_init9	-0.043	-0.093***	-0.132*	-0.109
	(0.032)	(0.034)	(0.073)	(0.068)
loseCWS:decile_init10	-0.054	-0.150***	-0.204*	-0.243**
	(0.047)	(0.046)	(0.108)	(0.102)
\`lag_MP(fit)\`	0.063	-0.093	-0.391	-0.326
	(0.201)	(0.187)	(0.337)	(0.279)
Observations	81,126	56,735	21,697	15,213
R^2	0.647	0.639	0.267	0.414
Adjusted R^2	0.603	0.571	0.173	0.303
Residual Std. Error	0.506 (df = 72036)	0.522 (df = 47692)	0.722 (df = 19243)	0.658 (df = 12773)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

robust standard errors, clustered by household

Table A5. Effect of transition from IWS to CWS by storage tank and well, comparison of HSR to IWS groups (full results)

	Dependent variable:			
	log(bi-monthly consumption)			
	Unmatched	Unmatched, omit peak periods	Matched	Matched, omit peak periods
	(1)	(2)	(3)	(4)
CWS:otankFALSE:wellFALSE	0.185***	0.176***	0.195***	0.168**
	(0.055)	(0.060)	(0.067)	(0.067)
CWS:otankTRUE:wellFALSE	0.073	0.133**	0.047	0.092
	(0.048)	(0.053)	(0.054)	(0.057)
CWS:otankFALSE:well	0.075	0.108	0.096	0.108
	(0.064)	(0.082)	(0.085)	(0.106)
CWS:otankTRUE:well	0.183***	0.299***	0.156**	0.285***
	(0.065)	(0.084)	(0.072)	(0.100)
`lag_MP(fit)`	-0.114	-0.135	-0.120	-0.127
	(0.146)	(0.147)	(0.182)	(0.171)
Observations	301,963	207,938	142,096	97,759
R^2	0.566	0.559	0.567	0.573
Adjusted R^2	0.528	0.499	0.529	0.516
Residual Std. Error	0.581 (df = 277299)	0.596 (df = 183278)	0.568 (df = 130548)	0.573 (df = 86215)
Note:				*p<0.1; **p<0.05; ***p<0.01
				robust standard errors, clustered by household

Table A6. Effect of reversion from CWS to IWS by storage tank and well, comparison of Maya Nagar to CWS groups (full results)

	Dependent variable:			
	log(bi-monthly consumption)			
	Unmatched	Unmatched, omit peak periods	Matched	Matched, omit peak periods
	(1)	(2)	(3)	(4)
loseCWS:otankFALSE:wellFALSE	-0.005	-0.012	-0.048**	-0.040
	(0.012)	(0.016)	(0.020)	(0.027)
loseCWS:otankTRUE:wellFALSE	-0.014	-0.026*	-0.030	-0.054*
	(0.011)	(0.014)	(0.021)	(0.028)
loseCWS:otankFALSE:well	-0.020	-0.081***	-0.052	-0.100*
	(0.019)	(0.028)	(0.046)	(0.060)
loseCWS:otankTRUE:well	-0.026	-0.033	-0.131	-0.098
	(0.023)	(0.032)	(0.085)	(0.110)
Observations	91,389	63,931	23,929	16,783
R^2	0.647	0.672	0.630	0.654
Adjusted R^2	0.603	0.610	0.583	0.588
Residual Std. Error	0.510 (df = 81180)	0.503 (df = 53780)	0.516 (df = 21237)	0.509 (df = 14108)
Note:				*p<0.1; **p<0.05; ***p<0.01
				robust standard errors, clustered by household