Expansion of Piped Water and Sewer Networks: The Effects of Regulation*

Carolina Tojal R. dos Santos[1] and Bruna Morais Guidetti[2]

[1]Inter-American Development Bank

[2]University of Michigan

January 17, 2026

Abstract

This paper investigates strategies to expand piped water and sewer through private providers. Using billing data from a major provider in Brazil and a structural model of consumer sanitation demand and service expansion, we assess the viability of connection targets and the welfare effects of connection subsidies and price incentives. We find that universal connection targets are largely unfeasible due to low sewer take-up. Combining connection subsidies with higher sewer prices boosts expansion and adoption but requires government funding. Charging consumers upon sewer availability is self-sustaining, promotes adoption and expansion but shifts costs to households.

JEL Codes: L95, Q25, L51

*We are grateful to Ying Fan, Catherine Hausman, Zach Brown, and Francine Lafontaine for their guidance throughout this project. We also thank participants at seminars and conferences hosted by the University of Michigan, Berkeley Energy Camp, MSU & UM Energy and Environmental Economics Day, the Seminar on Water Economics (SWELL), the Annual Conference on Taxation (NTA), the University of South Carolina, the Inter-American Development Bank, Miami University, PUC Chile, ITAM, University of Chicago Harris, Michigan State University, Insper, Oberlin College, LACEA-LAMES, the Encontro da Sociedade Brasileira de Econometria (SBE), and the WASH Conference for their feedback and suggestions. We are also thankful to Rene Alejandro Nieto for excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank.

1 Introduction

Access to clean water and sanitation remains a critical global challenge, with 2.2 billion people lacking safely managed drinking water and 3.4 billion without adequate sanitation as of 2022 (UNICEF and WHO, 2023). Expanding piped water and sewer systems is particularly difficult in developing countries, where governments often face financial constraints that limit infrastructure investment. A widely adopted strategy to address this issue is attracting private investment in infrastructure development and service provision, as seen in Brazil, Argentina, Chile, the Philippines, Indonesia, and South Africa (Marin, 2009). However, this approach presents a trade-off between financial viability and social inclusion, as under-served populations have a low willingness to pay for services and it is costly to reach them (Fay et al., 2021). While extensive research highlights the benefits of improved access to water services (Gamper-Rabindran et al., 2010; Devoto et al., 2012; Alsan and Goldin, 2019; Kresch et al., 2023), there is limited evidence on policy and regulation to promote infrastructure expansion while ensuring that households connect to available services.

This paper addresses this gap by examining the case of Brazil, where a large share of the population is still not connected to piped water and sewer services, particularly in the Northern region.[1] To tackle this issue, the government introduced the New Sanitation Regulatory Framework in 2020[2], which encourages municipalities to contract with private providers and sets connection targets of 99% of households with piped water and 90% with piped sewer collection within each concession by 2033. These connection targets follow the United Nations Sustainable Development Goals (UN, 2022), but it is unclear if they are viable for private providers. Moreover, the regulation overlooks a crucial factor: while providers install the pipes up to the sidewalk, it is ultimately up to consumers to complete the connection to their homes. Thus, the feasibility of these connection targets hinges on both firms having incentives to expand the pipes and consumers taking up the service when it is available.

[1] In the Northern region, only 54% of the population had access to piped water, and just 14% were connected to piped sewer services in 2017, according to the National Sanitation Survey (Pesquisa Nacional de Saneamento Básico - PNSB) from IBGE. Appendix Figure A1 provides connection rates for other regions of the country.

[2] *Novo Marco Regulatório do Saneamento* - Federal Law 14.026, Brazil, July 15, 2020.

In this paper, we investigate the feasibility of achieving universal access to services through private providers under current regulated prices. Specifically, we assess whether connection targets would be met if the firm expanded infrastructure across all concession areas. Our findings suggest that a substantial share of households would choose to remain unconnected to piped sewer services even if the infrastructure was available. We then explore policies to encourage take-up and expansion, evaluating the effectiveness of sewer connection subsidies, price adjustments, and charging consumers based on sewer availability regardless of connection status. Finally, we analyze the implications of these policies for welfare distribution between consumers and the provider.

We answer these questions using novel billing data from a private provider in Brazil. Our data set includes monthly billing records at the address level from consumers in various municipalities across the country under the firm's concessions, covering three years before the new regulation. This data, combined with demographic information from the Census, provides detailed consumption information and shows which zip codes the firm expanded to and which services they installed (water or water and sewer) within its concessions.

We first use the data to document key patterns in the firm's expansion decisions, consumer connections, and water consumption choices. We show that zip codes that receive expansion of both services have, on average, higher incomes than zip codes that receive only water. Moreover, the firm is more likely to expand in zip codes close to the installed network. We also document that part of households do not connect when the services are available. Within zip codes where both water and sewer services are available, on average, approximately 71% of households take up both services, while roughly 20% opt for water-only connections. Furthermore, higher-income areas exhibit, on average, higher rates of service adoption.

In our setting, connected consumers face non-linear pricing structures[3] for their water and sewer and respond to the average price. Two pieces of evidence suggest that consumers respond to the average price of their bill rather than the marginal price. First, we find no evidence of consumption bunching at price schedule kinks where marginal prices increase. Second, we observe consumption changes in response to price adjustments that affect the average price without altering the marginal rate. Accurately identifying which price con-

[3]Increasing block rates with a fixed fee.

sumers respond to is crucial for modeling demand and assessing the impact of price changes on provider revenues.

We use a structural model to predict both demand in areas without services, based on prices and demographics, and to recover expansion costs, which allows us to simulate the effects of various policies. The demand side consists of two components. First, a discrete-choice model where households select which service to connect to among the available options in their zip code. Second, a continuous-choice model where connected households decide on their water consumption. On the supply side, the firm faces a discrete-choice problem when deciding which services to install in each zip code based on profitability. We use cross-sectional variation in demographics and exogenous price changes to identify consumer preferences and their responsiveness to average prices. Using the estimated demand, we predict the potential revenue from expansion and estimate cost parameters from observed expansions.

Using the estimated model, we first assess the feasibility of meeting the connection targets by simulating the firm expanding water and sewer services across all zip codes in its concessions. We find that household connections would not meet the targets due to limited consumer take-up, with only about 51% of households connecting to piped sewer once it is universally available. We then allow the firm to optimally select expansion areas and examine three policies aimed at increasing the share of connected households: (1) offering a connection subsidy to cover the cost of installing the final segment between homes and the main sewer line; (2) combining this subsidy with an increase in monthly sewer bills; and (3) imposing a charge for sewer availability, irrespective of whether households connect.[4]

The connection subsidy increases sewer take-up in areas with existing infrastructure but does not incentivize firm expansion. In some regions, the price charged to consumers is insufficient to cover provision costs, making additional connections financially detrimental to the firm and deterring further expansion. To address this, we pair the subsidy with a 50% increase in the average sewer price paid by households on their monthly bills. This combination encourages firm expansion, resulting in approximately 82% of households con-

[4]Under this policy, in every zip code with sewer pipes, consumers who receive a water bill will also be charged for sewer service, regardless of whether they are connected.

necting to the sewer network, primarily driven by adoptions in zip codes receiving the service. While this scenario benefits both consumers and providers, it requires government funding for subsidies, which could be politically and fiscally challenging.

An alternative that does not require direct government intervention is the Sewer Availability Charge, a policy already permitted under existing regulations but rarely implemented by providers. Under this policy, consumers in zip codes with piped sewer infrastructure must pay for the service even if they do not connect. It effectively works as an externality tax that internalizes part of the social cost of remaining unconnected. This approach encourages firm expansion and consumer adoption, increasing sewer connections to approximately 55% of the households. However, unlike the subsidy-based approach, this policy shifts the financial burden entirely onto consumers, raising affordability concerns despite its effectiveness in boosting overall connection rates.

The contributions of this paper are threefold. First, we endogenize firm expansion decisions and link them to consumer choices in the water and sanitation market, providing a framework to evaluate incentives to increase service coverage. This approach adds to existing work that focuses on the benefits of improved sanitation access (Coury et al., 2024; Kresch et al., 2023; Devoto et al., 2012; Alsan and Goldin, 2019; Gamper-Rabindran et al., 2010; Barreto et al., 2007) by addressing how expansions can be achieved. Considering both demand and supply allows us to identify trade-offs created by different policies. This perspective is essential for avoiding unintended consequences, such as those observed with energy subsidies in India and Colombia, where underpricing resulted in underinvestment in infrastructure (Mahadevan, 2024; Burgess et al., 2020; McRae, 2015).

Second, we contribute to the growing literature on the adoption of sanitation technologies, which has largely focused on toilets and septic tanks (Deutschmann et al., 2024, 2023; Gautam, 2023). We extend this work by adapting empirical strategies from the energy demand literature (Resende et al., 2025; Barreca and Clay, 2016; Davis and Kilian, 2011; Dubin and McFadden, 1984) and leveraging detailed billing data to model household take-up and consumption of piped water and sewer services. Our findings reveal that low sewer take-up is a major barrier to achieving universal connection targets, an issue that was overlooked in the design of the new regulatory framework. In this context of limited willingness to pay and

high expansion costs, we show that universal coverage may not be welfare-enhancing, which aligns with findings from the rural electrification literature (Burlig and Preonas, 2024; Lee et al., 2020).

Third, we contribute to the literature on demand elasticity under non-linear pricing by providing evidence consistent with consumers responding to average prices rather than marginal prices. Although related work assumes marginal price responsiveness (Szabo, 2015; Olmstead, 2009; Hewitt and Hanemann, 1995), our findings align with studies that highlight the sensitivity of average prices in the water and energy markets (Ito and Zhang, 2020; Sears, 2023; Ito, 2014; Wichman, 2014; Ito, 2013; Borenstein, 2009).

The structure of the paper is as follows. Section 2 provides background information about the water sector in Brazil. Section 3 describes the data, while Section 4 presents the descriptive evidence obtained from the data. The model is presented in Section 5, and the estimation strategy is detailed in Section 6. We conduct counterfactual simulations in Section 7. Finally, Section 8 concludes the paper.

2 Institutional Background

In Brazil, water and sanitation services fall under the jurisdiction of municipalities, which have the authority to choose how to provide these services.[5] The Sanitation Regulatory Framework of 2007[6] allowed municipalities to contract with public companies without a competitive process, reserving auctions for cases involving private providers.[7] Winning companies pay a grant to the government for the service provision rights and consumer billing. Once a contract is signed, the chosen provider becomes a monopolist in the market for a specified duration, typically around 30 years. The price schedule is set at the beginning of the contract, and the main choice faced by the firm throughout the contract is whether and where to invest.

[5] Provision can be made through direct public administration, contracts with public state companies, public-private partnerships, or private providers.

[6] *Marco Regulatório do Saneamento* - Federal Law 11.445, Brazil, January 5, 2007.

[7] As of 2017, private providers were responsible for delivering piped water in approximately 8% of municipalities and piped sewer services in 3.7% – data from the national system of information about sanitation (Sistema Nacional de Informações sobre Saneamento - SNIS).

In 2020, a New Sanitation Regulatory Framework[8] was enacted to encourage private investments and address the challenge of expanding piped water and sewer access. Under this regulation, competitive auctions are mandated for all contracts formed after its implementation. Once existing contracts expire, public companies may face competition, allowing private providers to enter the market more extensively. Concurrently, this new framework establishes connection targets for each concession, requiring 99% of the population to be connected to piped water and 90% to piped sewer by 2033, while maintaining price regulation.[9] These targets are based on the United Nations Sustainable Development Goal to ensure the availability and sustainable management of water and sanitation for all by 2030 UN (2022). Although there are other forms of providing safe water services, the new regulation in Brazil interpreted this goal as requiring a piped network expansion.

The current regulations also allow municipalities to charge for sewer availability; in neighborhoods where piped sewer is available, households receiving water bills can also be charged for sewer regardless of connection. Nonetheless, most municipalities do not implement this policy,[10] and the federal law does not specify any other sanction for non-connections.

Our analysis relies on information from a provider that was in the market before 2020 to capture the underlying patterns in the absence of the federal connection targets, but the results of the simulations speak to what we can expect from future concessions. From 2020 to 2023, 28 auctions awarded service provision rights to private providers.[11] However, 93.7% of the municipalities still have public providers, and it is not clear whether it is viable for private providers to comply with the requirements of the new regulation.

[8] *Novo Marco Regulatório do Saneamento* - Federal Law 14.026, Brazil, July 15, 2020

[9] The regulation also encourages municipalities to form groups and collectively auction concessions while granting greater authority to the national regulatory agency at the expense of municipal and state regulators.

[10] In our sample, only one municipality had this policy implemented.

[11] 2020 Annual Outlook ABCON SINDCON (National Association and Union of Private Concessionaires of Public Water and Sewage Services): ¡https://abconsindcon.com.br/panorama¿.

3 Data

We use novel confidential water bill data from a private provider in Brazil[12] with concessions in different regions. This data, combined with demographic information from the 2010 Census, allows us to capture consumption patterns and the firm's expansion.

We have access to all consumers' billing data from January 2017 to December 2019. The bills are delivered at the address level, but due to privacy concerns, their zip code is the finer location information reported in our data. In urban areas, a zip code usually represents a street or a city block.[13] We combine it with census tract-level data on income, household size, number of households, number of owned versus rented houses, the share of the population with piped water and sewer, and whether the census tract is urban or rural from the 2010 Census. More details about the data construction are described in Online Appendix A3.1.1.

The areas covered by the firm's concessions reflect access to piped water and sewer in the rest of the country, as shown in Appendix Figure A2. Access to sanitation is highly correlated with income, and the firm operates in a wide range of locations, including low-income areas where access remains limited. We focus our analysis on municipalities where the firm provides piped water and sewer collection services.

We supplement the billing records with data from the National System for Research on Construction Costs and Indices (Sistema Nacional de Pesquisa de Custos e Índices da Construção Civil - SINAPI) in Brazil, which we use to calculate the costs of connecting residential buildings to street water and sewer pipes. This dataset details material and labor costs for construction projects, which we use to proxy household connection costs. Further details on the data construction can be found in Online Appendix A3.1.2. The average connection cost to water pipes is R$754.78, and the average connection cost to both water and sewer together is R$1811.70, considering the areas where there are consumers connected during our sample period.

[12]The University of Michigan signed a Non-Disclosure Agreement with the company on our behalf, and the provider requested not to have its name disclosed in the project. We also do not mention names of municipalities or specific information that could allow one to identify the company.

[13]In our data, the median number of addresses in a zip code is 20, and the average is 60.

3.1 Water consumption data

Our dataset includes monthly metered water bills from consumers connected to the company's network. While we can track addresses across billing periods, household turnover is unknown. For simplicity, we use "addresses" and "households" interchangeably. We focus on residential consumers, who account for 92% of bills and 89% of water consumption.[14] Addresses in our dataset may have both water and sewer connections or only water. All households with piped sewer also have piped water. Since we observe only connected households' water use, other water sources and wastewater destinations are treated as outside options.[15]

The price schedules for piped water and sewer are set at the start of each concession contract, with subsequent changes considered exogenous since the firm has limited control over timing and magnitude. Contracts include inflation adjustments, and the firm may request increases for unexpected cost shocks, but municipalities must approve them, often delaying responses and imposing gradual adjustments to limit public impact. Pricing follows increasing block tariffs (IBT) with fixed fees, where consumers pay based on water meter readings, and sewer charges are a percentage of water rates. The first block has a zero marginal price, meaning low-use households pay only the fixed fee. As consumption rises, so does the price per m^3. The firm calculates bills by assigning usage to price blocks, summing costs, and adding the fixed fee. Households with sewer connections pay an additional 50–100% of their water bill for sewer services.

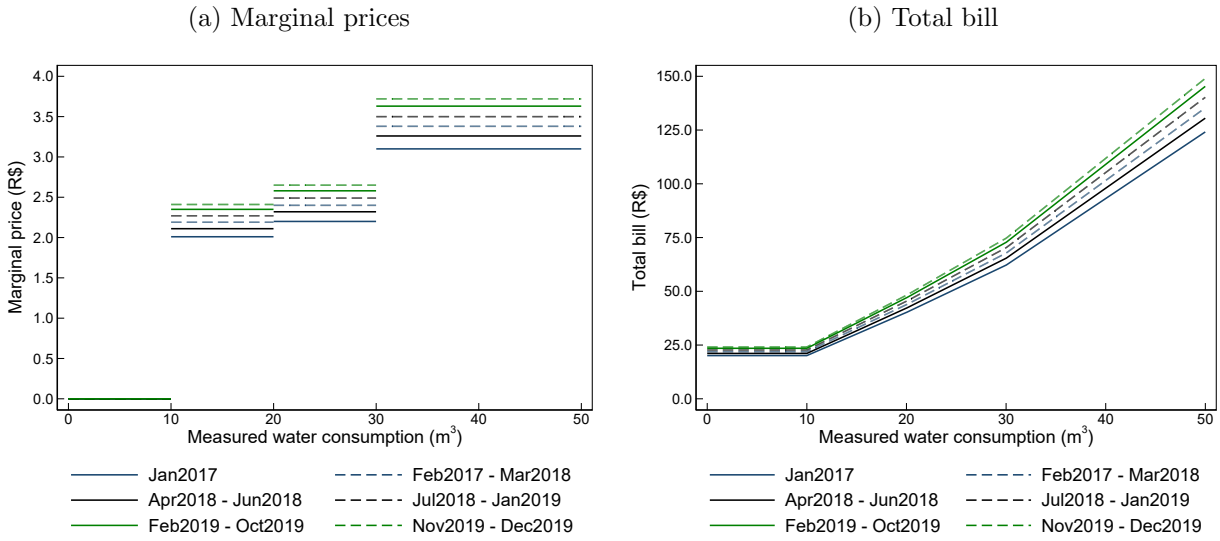
To illustrate the general structure and variation in price schedules, Figure 1 presents the marginal prices (Figure 1a) and the corresponding total bills (Figure 1b) for varying volumes from one of the concessions in our sample. In this case, the firm requested a price increase in 2016, but the municipality determined it would occur in three increments over the next three years. While an increase occurred in 2017, the municipal court blocked further price

[14]To analyze continuous water demand, we exclude water bills with a volume of zero, indicating no occupancy during that month, as well as bills with water consumption exceeding $200m^3$, which are likely due to leaks or other significant issues in the metering process. We also exclude apartment buildings where the consumption of all units is measured jointly.

[15]Non-connected households might obtain water from alternative sources such as cisterns, delivery trucks, wells, or directly from bodies of water, while the wastewater might go to septic tanks and unimproved pits or be thrown directly into the environment.

changes, only allowing adjustments to account for inflation that year. In 2019, the company overturned the previous decision and implemented price increases in two increments.

Figure 1. Price variation from one municipality in the data



Notes: These figures illustrate the variation in water prices observed in the data. Figure 1a displays the marginal prices for different levels of water consumption, while Figure 1b presents the corresponding total water bill for one of the municipalities in our dataset. Other municipalities exhibit similar patterns. The lines represent price levels across different periods in our sample.

Table 1 provides information about the demographic characteristics of households (addresses) connected to either water services only or both water and sewer services. It also includes data on their monthly water consumption and the total amount they are billed. Notably, households with sewer services tend to have higher monthly bills, even with similar water consumption to those without.

Table 1 – Descriptive characteristics of connected households

	Only Water	Water and Sewer
Income (R$)	2523.17	2771.31
	(1567.63)	(2147.27)
Urban	0.76	0.93
	(0.43)	(0.25)
Household size	3.52	3.33
	(0.45)	(0.39)
Water Consumption (m^3)	9.43	11.68
	(55.84)	(26.06)
Total Bill (R$)	55.66	93.83
	(85.28)	(193.53)
Connection Cost (R$)	754.78	1811.70
	(272.94)	(250.40)
Number of households	304194	327089

Notes: This table presents mean demographic and consumption characteristics of households connected to water-only or water and sewer services, with standard deviations reported in parentheses. Demographic data is sourced from the 2010 Census and assigned to households based on the census tract of their zip code. Water consumption and total bill values are drawn from billing records. Connection costs are computed using the SINAPI data. Households that switched service types during the sample period are excluded.

3.2 Pipe Networks

We also use the billing data to infer where the firm expanded its services. Between 2017 and 2019, the firm was not bound by connection targets and could choose where to extend its water and sewer network within its concession areas. However, we lack direct records of expansion locations, and available administrative data on the pipe network is aggregated at the municipality level, making it insufficient to capture key demographics and cost factors influencing expansion decisions.

We analyze service expansion by tracking zip codes that appear in the water and sewer billing data. Since engineering projects are planned at the street level and Brazilian zip codes typically correspond to streets, we use the zip code as the unit of analysis. We define the installed network as those zip codes with water bills in 2017. Expansion is identified in zip codes that first appear in water bills in 2018 or 2019. By examining the bills, we can determine whether the expansion includes both water and sewer services or just water. For example, zip codes that had only water services in 2017 and later show charges for both

water and sewer in subsequent years indicate that sewer services were expanded.

For the cost estimation and the simulations, we focus on concessions in the North and Northeast, where there is still significant room for expansion within the firm's boundaries. In contrast, concessions in the South and Southeast already have near-universal water and sewer coverage in our data.

4 Descriptive evidence

This section provides descriptive evidence of the service expansion, connections, and consumption from our data. First, the firm expands closer to the existing network and to wealthier zip codes, consistent with a profit-maximizing strategy. Second, once the pipes are installed, a significant share of consumers do not connect; demographics, such as income, are good predictors of take-up. Third, for connected households, we also find evidence that the water demand responds to average prices rather than marginal prices which is key to computing consumer's demand price elasticity. These patterns guide the model presented in Section 5.

4.1 Firm expansion

The firm builds pipes of only water, water and sewer in zip codes under its concessions. Table 2 shows that zip codes with both water and sewer pipes tend to have higher average incomes compared to zip codes with only water pipes or no service at all. This pattern holds for zip codes that originally had pipes installed ("old zips") and for zip codes where the firm expanded the pipes ("new zips").

Table 2 – Zip code characteristics by pipe network availability

Pipe network	Number of zips	Avg. income	Distance from Only water network (km)	Distance from Water and sewer network (km)
Old water and sewer	2188	4163		
Old only water	2542	2543		8.41
Old only water/new sewer	67	3148		3.30
New water and sewer	36	4360	0.08	0.15
New only water	219	3252	0.12	20.63
Nothing	539	2827	3.75	22.00

Notes: This table summarizes demographic and infrastructure characteristics (columns) by type of service availability (rows). It includes only zip codes within the firm's concession areas in the North and Northeast regions of the country. Zip codes with at least one bill for the service in 2017 are classified as "old". Zip codes that first appeared in the water billing records in 2018 or 2019 are categorized as "new". The remaining zip codes, with no billing records during the period, are considered to have no service.

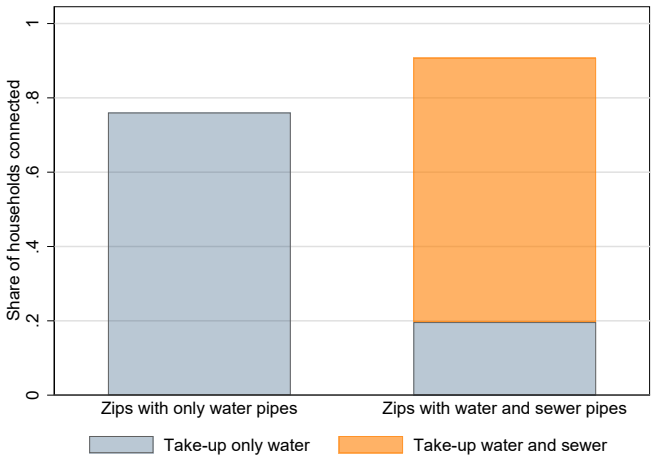
The firm expands closer to the installed network. The last two columns of Table 2 show that the zip codes where the firm expanded are on average closer to the network of the specific service installed. This pattern is unsurprising given the interconnected nature of water and sewer pipelines within a broader network. It is economically advantageous to install pipes near existing infrastructure; the costs related to infrastructure tend to increase as the distance from the installed network grows. Additional evidence that the firm expands service near its existing network is shown in Appendix Figure A3.

4.2 Incomplete service take-up

We show that many households do not take up the services in zip codes with the pipes available. As depicted in Figure 2, on average approximately 24% of households choose not to connect to the water service when it is the only service available in their zip code. In areas where both water and sewer services are available, approximately 71% of households connect to both services, while 20% prefer to connect to water only. Connecting to the sewer system involves connecting the house to the main pipeline, a substantial increase in the bill, and users may not directly perceive benefits. These factors may help explain the incomplete take-up of water and sewer services.

Demographic factors influence the connection to the main water and sewer pipelines. The regression analysis presented in Appendix Table A1 examines the relationship between demographic variables and the share of connected households in each zip code where services are available. Income is positively correlated with complete take-up and negatively correlated

Figure 2. Average service take-up



Notes: This figure illustrates the average service take-up across zip codes. The blue bars represent the average share of households connected to only water, while the orange bars indicate the average share of households connected to both water and sewer. The left bar corresponds to zip codes with only water pipelines, whereas the right bar represents zip codes with both water and sewer infrastructure.

with incomplete take-up.[16] Additionally, larger households are more likely to adopt the services fully, while those with alternative sewer collection methods are, on average, less likely to connect.

4.3 Consumption responds to average price

We investigate whether consumers facing non-linear price schedules respond to marginal or average prices, as this distinction is crucial for accurately estimating their price elasticity. We find suggestive evidence that consumers react to average price, consistent with other work in the water and energy markets (Ito and Zhang, 2020; Sears, 2023; Ito, 2014; Wichman, 2014; Ito, 2013; Borenstein, 2009). However, this result goes against other related papers on water markets that model consumers reacting to marginal prices as Szabo (2015), Olmstead (2009), and Hewitt and Hanemann (1995). The difference might be associated with how the prices are presented to consumers and other particularities of the context where the utility bills are charged.

[16]Complete take-up is defined as connecting to all the services available at the zip code, while connecting to only water when both water and sewer are available is considered incomplete.

1. No bunching at the kinks

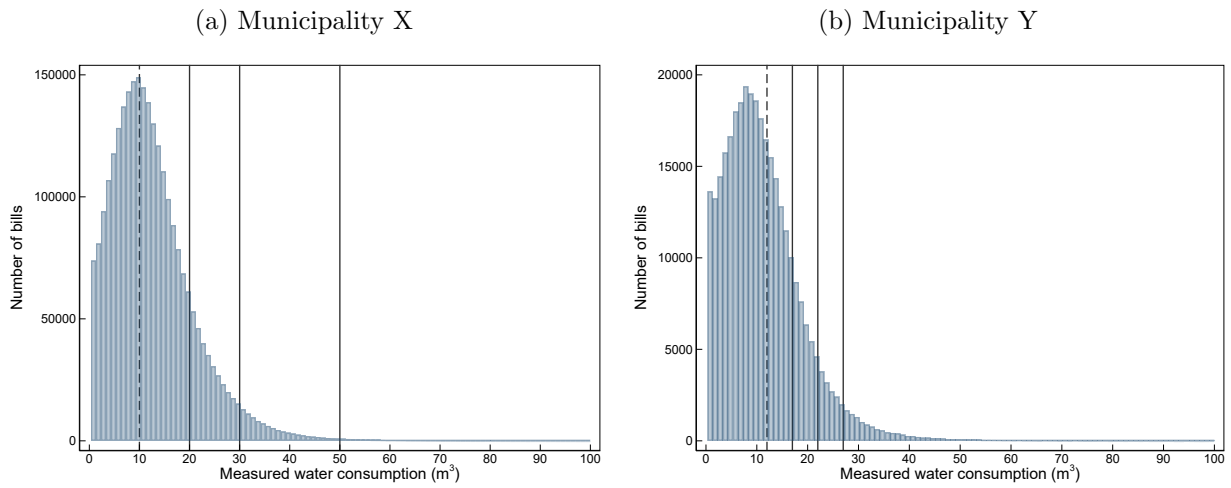
All the concessions included in our sample have non-linear price schedules characterized by increasing block rates. These price schedules result in budget sets that exhibit convex kinks at the points where the marginal price rises. If consumers were responsive to changes in marginal prices, we would expect to observe a bunching of consumption at these kinks, as shown by Hausman (1985) and Moffitt (1990). In our context, the marginal price for the initial consumption bracket is set at zero across all concessions. Consequently, if consumers were responsive to marginal prices, there would be an incentive for them to maximize their consumption without surpassing the threshold of the next bracket, where the marginal prices become strictly positive.

To investigate whether such bunching behavior exists, we plot histograms depicting the residential water consumption patterns for the concessions in our sample. In Figure 3, we present the histograms for two municipalities: Municipality X (Figure 3a) and Municipality Y (Figure 3b), with the price discontinuities represented by the vertical lines. Here, we show the graphs separated by concession because they face different price schedules, although all have the same feature of increasing block rates with zero marginal price in the first block. The histograms reveal a smooth distribution of consumption around the kink points, indicating an absence of bunching.

The absence of bunching can be interpreted in two ways: either consumers exhibit zero elasticity to prices or they respond to an alternative measure of price. To distinguish between these possibilities, we focus on households that consistently consume within the first consumption bracket, where the marginal price is zero, but the average price is strictly positive due to the fixed fee. Analyzing this specific group of households allows us to distinguish if demand responds to average price, given that our setting lacks price variation that would move average prices and marginal prices in different directions, as explored by Ito (2014).

In our setting, when price changes occur, all the marginal prices above the first bracket and the fixed fee change at the same rate, while the marginal price of the first bracket remains constant at zero. Consequently, for consumers who consistently consume below the first threshold, an increase in the fixed fee leads to a variation in the average price but not

Figure 3. No bunching in residential water consumption



Notes: This figure shows the histogram of measured water consumption for water bills from consumers in municipality X (Figure (a)) and municipality Y (Figure (b)). The vertical lines on the graphs represent the end of the brackets, where the marginal prices increase. The marginal price is zero in the first bracket, i.e., for volumes to the left of the dashed vertical line, and increases in the remaining brackets. The graphs include bills for consumers connected since the beginning of our sample and in single-unit residences with individual billing.

in the marginal rate.

To isolate the effect of changes in the fixed fee from weather shocks and other changes that may be happening at the concession level, we leverage the case of one concession that spans across two different states where consumers face different prices depending on which state they are located in. Given the proximity, they are exposed to similar weather events but are charged different prices. Moreover, we include other concessions in the same state to isolate the effect of price changes from other economic changes happening at the state level. Appendix Figure A4 illustrates this scenario.

Using this subset of water bills, we employed a specification similar to the one used by Ito (2014) to test whether consumers react to the fixed fee:

$$\Delta ln(q_{iusjt}) = \alpha \Delta ln(fee_{usjt}) + \Delta ln(I_{ct}) + \delta_{st} + \gamma_{ut} + u_{iusjt} \tag{1}$$

where q_{iusjt} represents the metered water consumption of household (address) i in concession u, state s, and connected to service j during billing month t. fee_{usjt} denotes the minimum

payment required from any address connected to service j in that concession state. I_{ct} represents the income at the census tract where household i is located. δ_{st} denotes state-billing month fixed effects, while γ_{ut} represents concession-billing month fixed effects. We utilized the difference $\Delta ln(q_{iusjt}) = ln(q_{iusjt}) - ln(q_{iusj0})$ between the consumption charged at time t and the same billing month in the previous year t_0, which eliminates household-month of the year fixed effects that account for household characteristics and seasonal components of water demand. $\Delta ln(fee_{usjt}) = ln(fee_{usjt}) - ln(fee_{usj0})$ and $\Delta ln(I_{ct}) = ln(I_{ct}) - ln(I_{c0})$ represent the equivalent difference for the fixed fee and the income, respectively.

If households responded to the marginal price, they would not reduce their consumption in response to increases in the fixed fee, as reducing consumption would not affect the total amount charged in their water bills. In particular, since our sample only includes price increases and no price decreases over the given time frame, we would expect the coefficient α to be zero. However, the results presented in Table 3 indicate that consumers reduce their consumption in response to increases in the fixed fee. The preferred specification described by equation 1 is reported in column (3), but we also report the results for specifications including only time-fixed effects in column (1) and concession-time fixed effects in column (2).

This finding suggests that consumers do not differentiate between fixed and variable costs, which is consistent with evidence found in heating demand in China (Ito and Zhang, 2020). Although consumers do not directly respond to marginal prices, this behavior demonstrates that they react to prices. Considering consumers' misconceptions regarding the non-linear price schedule, we treat them as responding to average prices in the demand model.

5 Model

To explore the feasibility of service mandates and alternative policies, we use a structural model that incorporates the key patterns found in the data. The model encompasses the supply and demand for piped water and sewer collection within the geographic limits under the responsibility of a private firm. Having won the concession, the firm is the sole provider

Table 3 – Consumption response to changes in the fixed fee

$\Delta ln(quantity)$	(1)	(2)	(3)
$\Delta ln(fee)$	−0.193**	−0.211*	−0.250*
	(0.083)	(0.114)	(0.145)
$\Delta ln(income)$	0.024*	0.027*	0.027*
	(0.014)	(0.014)	(0.014)
Time FE	yes	no	no
Concession-time FE	no	yes	yes
State-time FE	no	no	yes
Observations	384,704	384,704	384,704

Notes: This table presents regression estimates for the coefficients in equation 1. The dependent variable is the change in the logarithm of water consumption in a given month relative to the same month in the previous year. The key independent variables are the corresponding change in the minimum water bill payment and the census tract's income. The columns differ in fixed effects specifications: Column (1) includes month-by-year fixed effects, Column (2) adds concession-by-month-by-year fixed effects, and Column (3) further includes state-by-month-by-year fixed effects. The sample is restricted to addresses with continuous water billing throughout the period, no service type changes (water only vs. water and sewer), and consumption consistently within the first bracket. Only single-unit households billed individually are included. Standard errors are clustered at the address level. *** p<0.01, ** p<0.05, * p<0.1.

of these services in the region and operates as a monopolist. The supply side of the model uses a discrete choice approach to represent the firm's decision-making process regarding entry and service offerings at specific zip codes to recover the fixed cost of expansion and variable costs associated with service provision. On the demand side, the model consists on discrete-continuous consumer choice for service take-up and the amount of water consumed after connecting to the network to estimate their preferences regarding the services.

The market outcomes depend on the interplay between the monopolist's expansion decisions and the households' demand decisions. In particular, the availability of services, the share of connected households, and the quantity of water consumed depend on the underlying preferences of households and the fixed costs faced by the monopolist. Overall, this model provides a framework for examining the economic incentives and outcomes of different policies related to the provision of water and sewer services in private monopoly settings with regulated prices.

5.1 Demand

Households (addresses) in each zip code have preferences for piped water and sewer services, which affect their decision to connect to the network and their water demand. We represent the decision using a discrete and continuous model, where each household decides whether to connect to only water or water and sewer when the service network is available in their zip code and, conditional on being connected, households choose their water usage.

1. Take up

Households in a given zip code choose to connect to only water, $j = w$, both water and sewer services, $j = s$, or remain unconnected, $j = o$. More specifically, if the zip code has water and sewer pipes available, households can choose to connect to both services or only water, while if the zip code has only water pipes, the households can only choose to connect to water or remain unconnected.

The indirect utility of household (address) i, located at zip code z (in concession unit u and census tract c), for service $j \in (w, s, o)$ in year y is:

$$U_{ijzy} = V_{jzy} + \varepsilon_{ijzy} \tag{2}$$

$$V_{jzy} = \begin{cases} \alpha_{0ju s} + \alpha_{1j} c_{jcy} + \alpha_{2j} avgp_{juy} + \alpha_{3j} I_{cy} + \alpha'_{4j} D_{cy} + \xi_{jzy} & \text{if } j = w, s \\ 0 & \text{if } j = o \end{cases} \tag{3}$$

where c_{jcy} represents the installation costs for connecting a household to the street pipe network for service j. $avgp_{juy}$ denotes the average price faced by a representative consumer in concession u for service j. I_{cy} indicates average income, while D_{cy} is a vector of demographic characteristics influencing the decision to connect to the network. These characteristics include household size, urban location, the share of households on paved streets, the proportion of rental units in the census tract, the share of households with access to alternative water sources (such as truck delivery, cisterns, or pits), and the share with access to alternative sewage disposal methods (such as septic tanks, chemical toilets, or composting pits). Finally,

ξ_{jzy} captures an unobserved demand shock at the service-year-zip code level.

The parameter α_{0jus} is a product-concession-state fixed effect that incorporates preferences for specific services common for all consumers in a concession-state. The parameter α_{1j} captures consumers' sensitivity to the installation costs of each service. The parameter α_{2j} captures the willingness to trade off the price per unit of water, with or without sewer, against other service features. Parameters α_{3j} and α_{4j} incorporate interactions between demographic and census tract characteristics, respectively, and service alternatives. While ε_{ijzy} is an idiosyncratic preference shock.

We assume that the idiosyncratic utility shocks have a nested structure with one nest (g) that includes the inside options, $\mathcal{J}_g = \{w, s\}$, which in our setting are the services of only water or water and sewer, respectively. The only option outside of the group is not connecting to any service. Specifically, $\varepsilon_{ijzy} = \zeta_{igzy} + (1-\sigma)\mu_{ijzy}$, where μ_{ijzy} is i.i.d. extreme value and ζ_{igzy} is the same for all products in group g and has a distribution that depends on the nesting parameter $\sigma \in [0, 1)$ such that ε_{izjy} is distributed extreme value following Cardell (1997). As σ approaches 1, the utility within-group correlation goes to one, and only groups matter, meaning households care primarily about whether they are connected to any service rather than the specific type of service. As σ approaches 0, the within-group correlation goes to zero, reducing the model to a standard logit where choices are independent. This structure allows for more flexible substitution patterns. In particular, we expect that if one service option is unavailable, households are more likely to choose the remaining service rather than opt out of connection altogether.

Under these assumptions, the probability of selecting service j, conditional on choosing to connect to any service (g), is given by

$$S_{jzy|g} = \frac{exp(V_{jzy}/(1-\sigma))}{\sum_{j \in J_g} exp(V_{jzy}/(1-\sigma))} \tag{4}$$

The probability of choosing to connect to any service is

$$S_{gzy} = \frac{(\sum_{j \in J_g} exp(V_{jzy}/(1-\sigma)))^{(1-\sigma)}}{1 + (\sum_{j \in J_g} exp(V_{jzy}/(1-\sigma)))^{(1-\sigma)}} \tag{5}$$

Finally, the choice probability of product j, which represents the take-up, when service

j is available at zip code z is given by the following multiplication.

$$S_{jzy} = S_{jzy|g} S_{gzy} \tag{6}$$

2. Water consumption

The demand for water is represented by:

$$ln(q_{ijzt}) = \beta_0 + \beta_1 ln(avgp_{ijut}) + \beta_2 ln(I_{cy}) + \beta_3' D_{cy} + \delta_j + \delta_{m\tau} + \eta_{ijzt} \tag{7}$$

such that q_{ijzt} is the amount of water consumed by a household (address) i at zip code z (located in concession unit u and census tract c), connected to service j at a billing month t. $avgp_{ijut}$ denotes the average price faced by household i. I_{cy} is the average income at the census tract level, and D_{cy} is a vector that includes the number of people per household and an indicator of whether the census tract is in an urban area. δ_j is a service fixed effect and $\delta_{m\tau}$ is a municipality-month-of-the-year fixed effect. η_{ijzt} represents an idiosyncratic demand shock for water.

We model households as responding to the average price, $avgp_{ijut}$, which is computed based on the increasing block schedule of each concession. Households that consume only within the first bracket pay a fixed fee. Households with consumption in higher brackets (b) pay the fixed fee plus the cost of the quantities that exceed each bracket's limit, multiplied by the corresponding marginal prices, mp_{bjut}. The average price faced by household is expressed below as a function of the fixed fee, fee_{jut}, the bracket limits, $\bar{q}_{bu}$, and the marginal prices, mp_{bjut}.

$$avgp_{ijut} = \frac{fee_{jut} + \sum_{b=2}^{B} max(min(q_{ijzt} - \bar{q}_{ub-1}, \bar{q}_{ub} - \bar{q}_{ub-1}), 0) mp_{bjut}}{q_{ijzt}} \tag{8}$$

5.2 Supply

There is a single monopolist firm that offers different services in the zip codes within its concessions. For each zip code, the firm decides which service to offer in order to maximize

expected profits. In zip codes with no existing infrastructure, the firm can choose to provide only piped water (w), both water and sewer (s), or no service at all (o). In zip codes that already have piped water, the firm decides whether to add sewer service ($s|w$) or maintain the status quo with water only (w). The monopolist incurs a sunk cost of expansion if it decides to expand the pipes, which must be recovered through service provision to connected consumers. Since prices are regulated with little flexibility for adjustments once the contract is in place, the firm's primary - and essentially only - decision margin is where to expand services.

The expected profits at the zip code z with service j is given by:

$$E_{\xi,\eta}(\Pi_{jz}) = E_{\xi,\eta}(VP_{jz}) - SC_{jz} \tag{9}$$

where $E_{\xi,\eta}(VP_{jz})$ is the expected variable profit,[17] SC_{jz} is the sunk cost of constructing the network of pipes for service j.[18] More precisely the variable profit VP_{jz} is defined as

$$VP_{jzy} = \sum_{y=1}^{5} \frac{\pi_{jzy}}{(1+r)^{y-1}} \tag{10}$$

which represents the present value of the flow of yearly variable profits, π_{jz}, that the firm collects over the next five years after installing the network, discounted by the interest rate r. Although contracts last on average 30 years, the firm typically considers a 5 year period when making expansion decisions.[19] The variable profit, π_{jzy}, is given by the revenues collected from monthly water and sewer bills charged to connected consumers, minus the costs associated with service provision:

$$\pi_{jzy} = N_{zy} S_{jzy} (R_{jzy} - mc_j Q_{jzy}) \tag{11}$$

[17] The firm does not observe the demand shocks ξ and η when deciding on the expansion.

[18] Although the sunk cost differs between offering water and sewer simultaneously (s) versus adding sewer to a zip code that already has water ($s|w$), the expected variable profits are assumed to be the same: $E_{\xi,\eta}(VP_{js}) = E_{\xi,\eta}(VP_{js|w})$. This is because consumers only observe which services are available at the time of their decision, and the marginal cost of delivering a cubic meter of water or sewer service is considered identical whether installed jointly or sequentially.

[19] We discuss alternative time frames for recovering expansion costs in the estimation section.

where N_{zy} is the number of addresses in the zip code z in year y, and S_{jzy} is the take-up of service j. Q_{jzy} represents the total annual water consumption in year y, calculated as the sum of monthly water usage, q_{jzt}, by a representative household in zip code z connected to service j. Similarly, R_{jzy} represents the total annual revenue in year y in zip code z for households connected to service j. mc_j is the marginal cost per unit of water provided with service j.

Consistent with the descriptive evidence, we assume that when a service is available in a zip code, only a fraction of consumers, S_{jzy}, choose to connect to service j. Among those connected, we assume they behave like the average consumer in the zip code, demanding q_{jzt} cubic meters of water per billing month based on the service they use, or Q_{jzy} annually. Their bills are determined by their monthly measured consumption q_{jzt}, the service j to which they are connected, and the regulated increasing block rates with fixed fees, as described earlier. The resulting monthly bills generate the revenue R_{jzy} collected by the firm. Thus, the elements S_{jzy}, Q_{jzy}, and R_{jzy} are determined by the demand side of our model and influence the firm's revenue.

The marginal cost, mc_j, represents the cost of providing service j per cubic meter of water, which may include expenses related to water delivery, as well as sewer collection and treatment. Since most of the costs associated with service expansion stem from building the necessary infrastructure, we model the sunk cost as

$$SC_{jz} = \omega_j dist_{jz} + \nu_{jz} \tag{12}$$

where $dist_{jz}$ captures the distance from zip code z to the network of service j and ω_j represents the cost per kilometer of pipes of service j. ν_{jz} is a sunk cost shock that the firm observes when deciding to build pipes.

The monopolist makes a discrete choice in zip codes without service: whether to install both water and sewer, only water, or nothing. In zip codes with existing water infrastructure, the firm decides whether to expand sewer services. Expansion occurs if it is profitable. For instance, the firm expands only water if the expected profits from providing water alone exceed both the profits from providing no service and the profits from expanding both water

and sewer. Alternatively, the firm will expand both water and sewer if the expected profits exceed those of providing no service and also those of providing only water without sewer. Assuming ν_{jz} is i.i.d. and follows an Extreme Value Type I distribution, we can compute the probability of zip code z receiving service j as:

$$Pr_{jz} = \frac{exp(E_{\xi,\eta}(VP_{jz}) - \omega_j dist_{jz})}{1 + \sum_{k=w,s} exp(E_{\xi,\eta}(VP_{kz}) - \omega_k dist_{kz})} \tag{13}$$

While we assume independence in the firm's decisions across zip codes, we acknowledge that this simplification may not fully capture strategic expansion behavior. In reality, a firm might initially enter less profitable zip codes as a stepping stone toward more lucrative areas further away. Additionally, our calculation of distances to the installed network in 2017 does not account for the possibility of gradual expansion, where the firm first extends services to nearby zip codes before reaching more distant ones. If such a pattern were present, we would expect to see a more incremental outward expansion. However, as shown in Appendix Figure A5, our data does not exhibit this trend. Despite this simplification, the model accurately captures observed expansion patterns, particularly in the North and Northeast regions, where there are significant expansion opportunities.

The decision rule in our model assumes that a firm will expand if the variable profit over the next five years exceeds the sunk costs. However, some firms may also require a minimum return on investment, as discussed in Wollmann (2018). In such cases, our model would overestimate sunk costs. Nevertheless, our estimate would effectively capture a combination of the true sunk cost and the minimum return on investment requirement, ensuring that our counterfactuals remain valid under the same return threshold.

6 Estimation

In this section, we estimate consumers' preference parameters for service take-up decisions and water consumption. Using these estimates, we predict the quantities demanded and the firm's expected revenue. We then estimate the firm's costs of providing water and sewer services and expanding the pipe network.

6.1 Take up

We estimate the Nested Logit discrete service choice problem using the log differences in observed market share following Berry (1994), and adjust for zero market shares using a Bayes estimator following Li (2019). The market is defined as a zip code z, with shares determined by the take-up of each service j. We estimate the following equation:

$$ln(S_{jzy}) - ln(S_{ozy}) = \alpha_{0ju} + \alpha_{1j}c_{jcy} + \alpha_{2j}avgp_{juy} + \alpha_{3j}I_{cy} + \alpha'_{4j}D_{cy} + \sigma ln(S_{jzy|g}) + \xi_{jzy} \tag{14}$$

In addition to the variables previously outlined in the model section, we include $S_{jzy|g}$, which represents the share of households selecting service j, conditional on being connected to the network (group g). The parameter σ denotes the nesting parameter, capturing the substitution between having only water and water and sewer.

We estimate the model using a linear instrumental variable regression. The conditional share $S_{jzy|g}$ is endogenous because it is correlated with the take-up shock, ξ_{jzy}, which is unobserved by the econometrician. To address this, we use the availability of piped sewer in zip code z as an instrument for $S_{jzy|g}$. Since the firm installs the network before consumers make their decisions, the availability of sewer pipes is uncorrelated with unobserved demand shocks while directly influencing consumer choices among services within group g.

Some zip codes have all addresses connected to the service. In such cases, the observed take-up of the specific product is one, and the share of the outside option is zero. However, this does not reflect the true probability of a household connecting, given prices and demographic characteristics. Following Li (2019), we use a parametric empirical Bayes or shrinkage estimator that generates strictly positive posterior estimates of the true take-up probabilities by leveraging information from similar markets.[20] We define similar markets as the 100 zip codes that are closest in terms of income per capita and that offer the same type of service (only water or both water and sewer). More details on this method can be found in Online Appendix A3.2. Therefore, instead of using the observed take-up, we use

[20]Zero market shares are a common challenge in different settings. Li (2019) and Gandhi et al. (2023) discuss alternative methods for addressing this issue.

the posterior estimates as the dependent variable in the nested logit estimation.

We estimate the nested logit model using data at the zip code-year level, focusing on zip codes that had installed pipes in all three years of the available data.[21] The estimated coefficients for each product are reported in Table 4. Column (1) displays the estimates for the "only water" service, while column (2) presents the estimates for the "water and sewer" service. The nesting parameter is repeated in both columns since it remains constant across products.

Table 4 – Take up estimation results

	(1) Only water	(2) Water and sewer
Nesting param	0.907***	0.907***
	(0.022)	(0.022)
Connecting cost (1000 R$)	−2.550***	−2.274***
	(0.208)	(0.420)
Avg. price	−0.078***	−0.192***
	(0.019)	(0.068)
Income (1000 R$)	0.031**	0.085***
	(0.014)	(0.017)
Urban	−0.125	0.222
	(0.389)	(0.440)
Household size	−0.531***	−0.600***
	(0.105)	(0.113)
Share rented	−0.009	0.862**
	(0.306)	(0.385)
Share other water	0.006	−0.095
	(0.234)	(0.381)
Share other sewer	−0.352***	−0.463***
	(0.119)	(0.151)
Share paved	2.014***	1.856***
	(0.162)	(0.285)
State-concession-service FE	yes	yes
F-statistic	6,025	6,025
Observations	48,528	48,528

Notes: This table reports the estimated coefficients from equation 14. Column (1) presents the parameters for households connecting only to water, while Column (2) reports the parameters for households connecting to both water and sewer. The outside option is not connecting. The analysis includes zip codes where pipes were already installed at the beginning of the sample period and no further expansion occurred. The nesting parameter captures the correlation in consumer utilities among service connection options. The conditional take-up rates are instrumented using an indicator for whether the zip code has sewer service. The reported F-statistics correspond to the Kleibergen-Paap rk Wald F-statistic. Standard errors are clustered at the zip code level. *** p<0.01, ** p<0.05, * p<0.1.

[21] Consumers might take some time to connect once new pipes are installed, so we focus on zip codes that already had pipes such that our take-up estimates reflect the demand after this adjustment period.

As expected, our estimated nesting parameter is close to 1, indicating a high within-group correlation. This suggests that households are more likely to substitute between "only water" and "water and sewer", rather than opting for the outside option of not connecting to any services. The other estimated coefficients also align with the expected directions, such as households being less likely to connect if the prices for both water and connection increase.

6.2 Water consumption

Water consumption is modeled as a linear function of the average water price and demographics, as specified in equation 7. Estimating this model presents two endogeneity challenges.

First, the average price consumers face is endogenous due to the structure of increasing block rates. Households pay different marginal prices depending on their consumption level, which directly influences the average price paid. This creates a simultaneity issue, leading to biased price coefficients if estimated using ordinary least squares (OLS).[22]

Second, selection bias may arise as the regression is estimated using only households connected to the network. Households that consume more water may also be more likely to connect, meaning the distribution of unobserved demand differences, η_{ijzt}, could differ between connected and unconnected households. In this case, failing to account for selection would lead to biased coefficients, making them unreliable for predicting consumption among households that may eventually connect.

We use a simulated instrumental variable for the average price to address the simultaneity issue. The instrument is constructed by considering the consumption of similar households in other concessions. We divide households into 16 groups based on income quartiles, service type (only water or water and sewer), and urban or rural areas. Then, we calculate the average consumption of households in the same group but located in different concessions, where they face different price schedules (fixed fees, marginal prices, and bracket limits). The instrumental variable is the average price the household would pay under the price schedule

[22]The coefficient would be downward biased if most households consume within the first price bracket since higher consumption leads to lower average prices. Conversely, it would be upward biased if most households consume above the first bracket, where higher consumption results in higher average prices.

of their own concession and time if they consumed the average consumption of households in the same group but located in other concessions. This instrument is convenient because it captures exogenous variation in the price schedule, but it is not affected by the quantity consumed by the household.

Additionally, we assess the robustness of our average price instrument by comparing it to an alternative approach that uses observed marginal prices for each bracket as instruments, following Olmstead (2009). The advantage of using marginal price-based instruments in our setting is that consumers generally do not switch brackets, as shown in Appendix Figure A7. By holding consumption levels fixed, we can leverage variation in price schedules over time and across concession states to identify price elasticity.

We follow Barreca and Clay (2016), Davis and Kilian (2011), and Dubin and McFadden (1984) to address the potential selection problem, allowing the discrete and continuous components of demand to be correlated. Specifically, the expected value of the continuous water demand shock, η_{ijzt}, is assumed to be a linear function of the demand shock of the service choice, ε_{jzy}, to compute the selection controls based on the estimated take-up of the service, $\hat{S}_{jzy}$. For households in zip codes with only water available, we include a single selection term accounting for the choice between connecting to water and the outside option of remaining unconnected. In zip codes where both water and sewer services are offered, we include two selection terms: one capturing the choice relative to the outside option and another reflecting the inside option not chosen (e.g., choosing water only vs. both water and sewer).[23]

The results are presented in Table 5. Column (1) reports estimates without an instrument or selection controls. Column (2) instruments for the average price using marginal prices for each bracket. Column (3) replaces this with the simulated average price instrument. Finally, Column (4) builds on Column (3) by accounting for potential selection, making it our preferred specification. Appendix Table A2 reports the first-stage results, while Appendix Table A3 presents the reduced-form results.

[23]In our setting, for a household connected to any inside option j, the selection term for the outside option is given by $\hat{S}_{ozy}\frac{ln(\hat{S}_{ozy})}{(1-\hat{S}_{ozy})+ln(\hat{S}_{jzy})}$. If there is another inside option k available, which happens when there is both water and sewer, there is another selection term given by $\hat{S}_{kzy}\frac{ln(\hat{S}_{kzy})}{(1-\hat{S}_{kzy})+ln(\hat{S}_{jzy})}$.

In summary, the results indicate that households reduce their water consumption in response to higher average prices, though the elasticity is small. Higher-income households consume more water, with a stable coefficient across IV specifications. Households with piped sewer connections use more water than those with only water connections. As expected, water consumption increases with household size. Additionally, households in urban areas tend to consume less water on average than those in rural areas. Note that the price elasticity remains similar regardless of the instrument used or the inclusion of the selection control. However, it differs significantly from the OLS estimates, suggesting that our instrument effectively corrects the downward bias.

Our analysis focuses on selection bias arising from consumers' decisions to connect to the pipe network. We assume the firm decides where to expand pipes based on observable demand factors and costs, with demand shocks occurring only after these decisions are made. This implicitly assumes that the firm does not have access to demand shocks that are unobserved by us as econometricians. However, if the firm did anticipate positive demand shocks, it might prioritize expansion in those areas, introducing an additional selection issue that could lead us to overestimate unconditional water demand. We believe this concern is minimal in our context. While the firm may receive additional input from technicians on the ground, we rely on the same administrative data they use. Given the scale of operations, it is unlikely this information is systematically incorporated. Therefore, any bias from unobserved demand shocks known to the firm but not to us is likely minimal.

To predict the consumption of connected addresses, we use the reduced-form estimates presented in Appendix Table A3. Using the reduced form simplifies the problem by allowing us to rely on the simulated average price to determine a single price that households respond to, enabling straightforward consumption predictions. In contrast, using the 2SLS results for prediction would require jointly solving for consumption and the average price. Given the nonlinear shape of the average price function in our setting, this approach would generate two equilibrium consumption quantities—one in the first bracket and another in the higher brackets—forcing us to rely on an ad hoc rule to select between them.

Using the estimated model, we compute $\hat{q}_{jzt}$, which represents the predicted consumption

Table 5 – Continuous demand model estimates

	(1) OLS	(2) Mg. prices IV	(3) Simulated IV	(4) Simulated IV with selection
ln(Avg. price)	−1.030***	−0.249***	−0.214***	−0.213***
	(0.002)	(0.006)	(0.009)	(0.009)
ln(Income)	0.157***	0.123***	0.122***	0.123***
	(0.003)	(0.003)	(0.003)	(0.003)
Piped sewer	0.548***	0.090***	0.070***	0.070***
	(0.003)	(0.004)	(0.006)	(0.006)
Urban	−0.112***	−0.158***	−0.160***	−0.154***
	(0.020)	(0.023)	(0.023)	(0.023)
Household size	0.053***	0.054***	0.054***	0.053***
	(0.004)	(0.005)	(0.005)	(0.005)
Selection inside opt.				−0.013***
				(0.005)
Selection outside opt.				0.014***
				(0.005)
Municipality-month FE	yes	yes	yes	yes
F-statistic		13,545	32,034	32,241
Observations	8,643,951	8,643,951	8,643,951	8,643,951

Notes: This table presents the parameter estimates from the water consumption model. The dependent variable is the logarithm of water consumption, and the key independent variable is the logarithm of the average price. Column (1) reports OLS estimates. Column (2) uses the marginal price of water as an instrument. Column (3) instruments for price using a simulated average price, constructed by grouping households into 16 categories based on income quartile, service type (water only or water and sewer), and urban or rural location. For each group, the average consumption of households in different concessions is computed, then used to calculate the price each household would face under its own concession's pricing schedule if it consumed the group's average usage in other concessions. Column (4) builds on Column (3) by accounting for potential selection bias among households that opted to connect. The sample includes only addresses with water bills for all periods in the dataset that did not switch service type. The reported F-statistics correspond to the Kleibergen-Paap rk Wald F-statistic. Standard errors are clustered at the household level. *** p<0.01, ** p<0.05, * p<0.1.

of a representative consumer in a given billing month, located in zip code z, and connected to service j. This calculation assumes average income and household size within the zip code. We then aggregate this monthly consumption to obtain the yearly consumption, Q_{jzy}, and calculate the corresponding revenue generated by the firm, denoted as R_{jzy}, based on the price schedule.

6.3 Cost

To estimate the costs associated with providing a service in different zip codes, we consider the water consumption and revenue a firm would generate for the next five years if they installed only water or both water and sewer pipes in that zip code. We assume that the firm has perfect foresight of future interest rates, population growth, and income in the areas they have concession over the provision of piped water and sewer. The firm chooses where to install pipes considering their *ex-ante* profit, given by their expected take-up and water demand, the marginal cost to provide the services and the sunk cost of building the pipes.

We consider that the fixed fees and marginal prices are updated annually based on inflation projections from 2017. The population of each zip code grows at the same rate as the municipal population projections. The income per capita also grows at the same rate as the municipal income projections reported by the Brazilian Institute of Geography and Statistics (IBGE).

Using the demand estimates and random draws from the empirical distribution of ξ and η, we compute the expected predicted take-ups ($\hat{S}_{jzy}$), the average monthly household water consumption ($\hat{q}_{jzt}$) and the associated revenue ($R(\hat{q}_{jzt})$).

We estimate the marginal cost associated with each service, mc_j, and the fixed cost parameter, ω_j, via maximum likelihood using the predicted choice probability of connecting zip code z with service j (equation 13) and the observed expansion choices in 2018 and 2019. The key idea is to identify the cost parameter values that maximize the likelihood of observing the firm's actual choices, given the model's assumptions.

The estimation results are available in Table 6. They indicate that the cost of supplying one cubic meter of water is roughly 6.13 Brazilian reais (R$). When including the collection of piped sewer with the same amount of water, the cost increases to about 12.71 Brazilian reais. These costs are based on the metered water consumption at each address and cover expenses associated with water treatment, delivery, and sewer collection, and account for potential water losses during distribution.

The sunk cost for constructing one kilometer of water pipes is approximately R$10869.47, R$12569.58 for a kilometer of combined piped water and sewer and R$2141.60 to extend

sewer pipes to zip codes that had only water. These costs encompass not only the pipes but also all the materials and labor required for excavation and restoring the path after pipe installation. Appendix Tables A4 and A5 present alternative cost estimates, assuming the firm recovers sunk investment costs over 10 or 30 years.

Table 6 – Cost estimates

	Costs
Mg. cost water (m^3)	6.130***
	(0.157)
Mg. cost water and sewer (m^3)	12.706***
	(0.472)
Cost per distance water (km)	10869.465***
	(3982.539)
Cost per distance water and sewer (km)	12569.585**
	(5350.286)
Cost per distance sewer (km)	2141.603***
	(376.737)
Number of zip codes	3,279

Notes: This table presents cost estimate parameters under the assumption that firms consider projected profits over the next 5 years when making decisions. The estimation includes zip codes within the firm's concessions in the North and Northeast regions of the country. From the total 3403 zip codes without service or with only water in 2017, we missed 124 where we could not predict the demand based on our estimated model. These zip codes are dropped either because when matched with the census, they are missing relevant demographics or because there was only one zip code in the municipality, so we could not estimate municipality-moth fixed effects. All estimated costs are reported in Brazilian reais (R$). In 2017, the exchange rate was approximately 3.3R$ per 1U$. *** p<0.01, ** p<0.05, * p<0.1.

Comparing our estimates with existing literature is challenging. Engineering studies, as von Sperling and Salazar (2013), typically consider only accounting costs and focus on a limited number of projects. Additionally, studies based on survey data from different countries, as Brichetti et al. (2021), often do not distinguish between the costs incurred by firms when installing the network and the costs consumers bear to connect their homes to street pipes.

7 Counterfactual simulations

For the simulations, we assume connection targets of 99% of households with piped water and 90% with sewer and use the estimated model to analyze incentives for achieving them. Reaching these targets depends on both the company's expansion decisions and consumers' choices to adopt services. To disentangle these factors, we first simulate a scenario where the firm expands services in all zip codes, such that the gap between connection rates and the targets depends solely on consumer decisions. Next, we allow the firm to endogenously select expansion areas and introduce different policies, focusing on sewer connection subsidies and sewer availability charges. These policies encourage consumer adoption and may incentivize the firm to extend services to uncovered zip codes.[24]

To predict the outcomes, we use demand estimates and the empirical distribution of ξ and η to compute the predicted take-up rates, water consumption, and the resulting firm revenue under the different scenarios. By combining predictions with the cost estimates, we calculate the variable profit the firm would generate and the sunk costs involved in the expansion.

We also measure the changes in consumer surplus and infant deaths that arise from these policy changes. Consumer surplus is computed using the discrete-choice component of the model, where consumers decide which service to connect to when it is available. While consumer surplus is a commonly used welfare measure, the interpretation requires caution in contexts with high inequality. The willingness to pay for the services may not fully capture the benefits consumers would experience upon connecting. Nevertheless, we present this measure to understand its impact on consumers who can afford the service and may not choose to connect under the different simulations.

We compute the number of averted infant deaths in each simulation to capture consumer health benefits. We create a back-of-the-envelope measure using the estimated impact of piped water and sewer in Brazil from Gamper-Rabindran et al. (2010) and the number of live births from DATASUS. This measure incorporates both private benefits from water

[24]In the counterfactual scenarios, we assume that the firm cannot remove services from areas where pipes had already been installed by 2019.

connections and externalities from sewer connections. It is important to note that it does not encompass all dimensions of external benefits, as discussed by Kresch and Schneider (2020), but provides an extra dimension to compare the policies.[25]

7.1 Full expansion of piped water and sewer

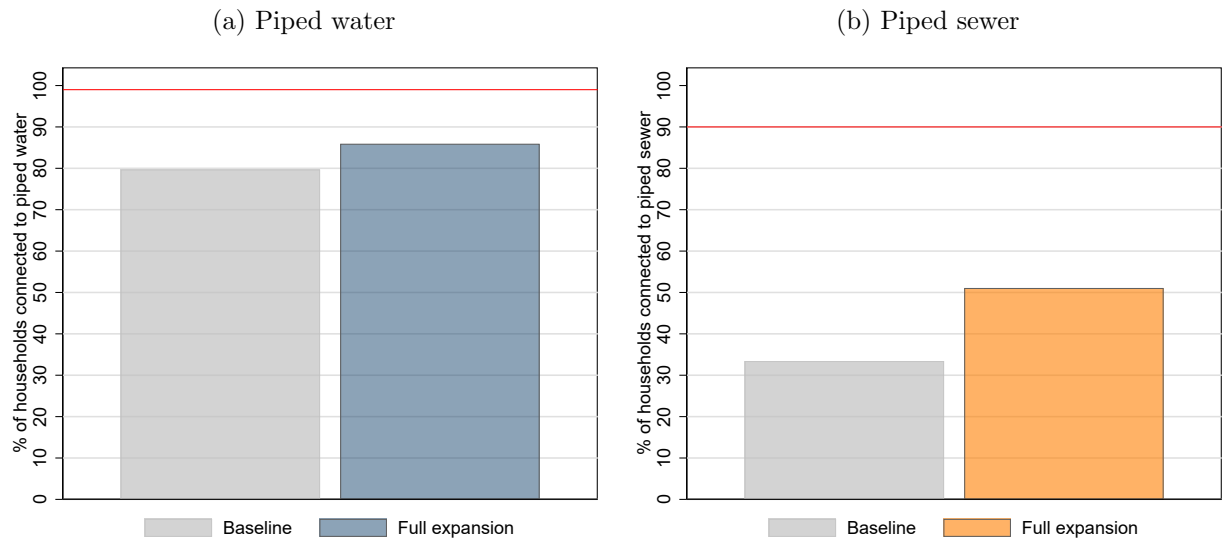
In the first exercise, we simulate a scenario where the company expands water and sewer services to all zip codes within its concessions in the northern region. In this setting, connection rates depend solely on consumer take-up, as all households could potentially connect to the services.

Our results show that connection targets would not be met at current pricing levels even if the firm expanded to all zip codes, due to low take-up, especially for sewer. Despite universal availability in this scenario, 85.95% of the households would connect to water and only 51.09% would connect to the sewer network. Figure 4a illustrates the 99% piped water coverage target (red line), the pre-policy connection rate (gray bar), and the simulated expansion outcome (blue bar). Similarly, the right panel of Figure 4b depicts the 90% sewer coverage target (red line), with the baseline and full-expansion connection rates shown in gray and orange, respectively.

Table 7 provides more details of the simulation results. Panel A displays information on the share of zip codes within each concession where each service is available, while Panel B presents welfare measures of each alternative policy relative to the baseline situation. We present variable profit and consumer surplus over a five-year period in line with our cost estimation framework. Column (1) presents the baseline scenario, while column (2) shows the outcomes under full expansion. By design, the latter ensures all zip codes have access to water and sewer services, as reflected in the first two rows. However, not all households adopt these services, as indicated in the third and fourth rows and the colored bars in Figure 4.

[25]The number of averted infant deaths represents a lower bound of the externalities generated by increasing connections to piped water and sewer. For instance, the services might also reduce the incidence of waterborne diseases, such as diarrhea, which do not always result in child death (Barreto et al., 2007). Social externalities could influence the decisions of neighbors to adopt alternative methods for water sanitation and wastewater disposal (Deutschmann et al., 2024). Additionally, externalities could manifest as increased housing prices in neighborhoods with the service (Coury et al., 2024).

Figure 4. Share of household connections



Notes: This figure compares the baseline percentage of households connected to piped water (Figure (a)) and piped sewer (Figure (b)) to a scenario in which the firm expands service everywhere. The red lines indicate the connection targets set by the 2020 Sanitation Regulatory Framework. The connection share is calculated as the number of households connected to the service divided by the total number of households within the North and Northeast concession areas. Under the "Full Expansion" scenario, all zip codes have piped water and sewer, and the predicted take-up determines the share of connections for each service.

Additionally, we show in column (2) of Table 7 that extending services to all zip codes is not viable for the company, as the substantial sunk costs outweigh the increased variable profit from new connections. However, the expansion increases consumer surplus, as more consumers have the services available and decide to connect. The full expansion also generates a reduction of 10.37% in infant mortality among children below 1-year-old, amounting to roughly 22 fewer deaths when contrasted with the baseline scenario without expansion.

7.2 Endogenous expansion of piped water and sewer

In this set of simulations, we allow the firm to determine which zip codes to expand water and sewer services to while providing incentives for consumers to connect to sewer via subsidies and the sewer availability charge. We focus on sewer adoption and expansion as it presents the largest gap to the connection targets, while for water most people connect to it when it is available.

Table 7 – Simulations

	(1) Baseline	(2) Full expansion	(3) End. expansion Connection subsidy	(4) End. expansion Connection subsidy Price increase	(5) End. expansion Availability charge
Panel A: Service coverage and household connections					
% of zips with water and sewer	41.36	100.00	41.36	78.20	71.75
% of zips with only water	48.77	0.00	48.77	12.00	18.41
% of households connected to water	79.71	85.95	83.38	84.13	78.86
% of households connected to sewer	33.43	51.09	55.17	81.58	54.77
Panel B: Welfare impact relative to the baseline					
Δ Variable profit (mi R$)		−95.77	−127.15	188.62	282.81
Sunk cost (mi R$)		249.69	0.00	4.34	3.83
Δ Consumer surplus (mi R$)		3.37	53.03	65.46	−4.40
Connection subsidy (mi R$)		0.00	127.16	281.94	0.00
Δ% Infant deaths		−10.37	−2.60	−2.66	−0.02
Expansion: firm choice	yes	no	yes	yes	yes
Expansion: all zips with sewer	no	yes	no	no	no
Subsidy sewer connection	no	no	yes	yes	no
Sewer price increase	no	no	no	50%	no
Sewer availability charge	no	no	no	no	yes

Notes: This table presents counterfactual results, with each column corresponding to a different counterfactual simulation. Column (1) shows the baseline. Column (2) considers full expansion. Column (3) allows the firm to expand endogenously while providing consumers with a subsidy to connect. Column (4) adds a price increase to the policy implemented in Column (3). Column (5) combines endogenous firm expansion with charging for sewer service based on availability rather than connection. Panel A reports service coverage outcomes, while Panel B presents welfare outcomes. Differences in variable profit, consumer surplus, and infant mortality are measured relative to the baseline. Variable profit and consumer surplus are calculated over five years, while the change in infant deaths is based on the estimated increase in household connections, following Gamper-Rabindran et al. (2010). The baseline scenario includes 216 infant deaths. All cost estimates are in Brazilian reais (R$). In 2017, the exchange rate was approximately 3.3 R$ per 1 U$.

1. Connection subsidies

The second simulation examines the impact of a one-time sewer connection subsidy, which covers the cost of connecting homes to street pipes. This subsidy applies only to households connecting to both water and sewer, excluding those opting solely for water. Once connected, consumers must continue paying their bills to remain in the system and cannot receive the subsidy again if they disconnect.

The subsidy effectively increases household sewer connections but has no impact on firm expansion. As shown in Table 7, column (3), the share of connected households rises to approximately 55.17%, yet service coverage remains at 41.36%. This indicates that the new connections occur in areas where sewer infrastructure was already available. Furthermore, the subsidy reduces firm profits in locations where sewer pipes were previously installed, suggesting that the revenue from water and sewer bills is insufficient to cover the costs of service provision for newly connected households.

The subsidy significantly increases consumer surplus, valued at 53.03 million Brazilian reais, by enabling more households to access services. However, the policy is costly, amounting to 127.16 million Brazilian reais. We do not take a stand on how the government would finance this subsidy or whether it would compensate firms for incurred losses.

Appendix Figure A8 shows that subsidies covering more than 50% of the cost do not further improve connection rates, which is mainly driven by consumer take-up. Appendix Figure A9 suggests that offering subsidies only to low-income households does not significantly boost connections because the firm does not substantially expand.

2. Connection subsidies with price increases

In our third simulation, we show that combining connection subsidies with an increase in sewer prices can effectively boost connection rates. In some concessions, the new sewer price is high enough to cover the costs of sewage collection, which encourages firms to expand their network. As shown in Table 7, column (4), pairing the subsidy with a 50% increase in the sewer price raises sewer connections to 81.58% of households. Appendix Figure A10 shows that further price increases beyond 50% yield only a marginal increase in connections, as the firm does not expand significantly beyond this threshold.

On the consumer side, the increase in consumer surplus from the subsidy more than offsets the higher sewer price. However, the expansion makes the subsidy more expensive for the government, as more consumers use it to connect to newly available sewer infrastructure. Despite the significant rise in sewer connections, the reduction in infant mortality remains small.

3. Sewer availability charge

In our fourth simulation, we introduce the "Sewer Availability Charge", which requires consumers to pay for sewer in their monthly bills once pipes are available in their zip code, even if they are only connected to water. Although established under the 2007 Sanitation Regulatory Framework, this policy has been rarely implemented, with only a few municipalities adopting it. In our dataset, just one municipality employs this pricing strategy.[26] This charge works as a tax, incentivizing consumers to internalize the externality created when they choose not to connect to the sewer system despite its availability, thereby mitigating the negative effects on their neighbors

[26]The availability charge was introduced in this municipality midway through the study period, so it is excluded from demand and cost estimations.

This policy increases sewer connections, as shown in Table 7, column (5), raising overall sewer connections to approximately 54.77%. The increase results from both firm-driven sewer expansion and higher consumer adoption. Appendix Table A7 shows that without expansion, the connection rate would reach 42.03%, indicating that while firm expansion plays a role, the majority of new connections stem from consumers in already-served areas opting to connect.

7.3 Policy implications

Our analysis highlights the crucial role of consumer take-up in this context, a factor overlooked in the New Sanitation Regulatory Framework. We also show that even with demand policies to stimulate take-up, the firm does not have incentives to expand sewer because, in some regions, the price charged to consumers is insufficient to cover provision costs. Therefore, addressing both demand and supply barriers is essential for designing effective policies that increase connection rates.

If the goal is solely to maximize the share of connected households, providing connection subsidies while increasing the monthly price is the most effective option among the alternatives considered. We do not take a stance on whether taxation and redistribution should be used to compensate firms and consumers for their losses or how the government would finance subsidies. In principle, these mechanisms could help achieve a more equitable outcome, and the policy with the highest overall welfare gains should be prioritized.

The availability charge is the policy that yields the highest net private benefits as the increase in profits to the firm outweighs the loss in consumer surplus. However, it results in the smallest reduction in infant mortality, suggesting that alternative policies may generate larger positive health externalities.

Full expansion could be a superior policy in terms of total welfare if reductions in infant mortality are valued as externalities and the value of saving one infant exceeds 27.58 million Brazilian reais.[27][28] This simplification facilitates policy comparisons, though a comprehen-

[27] Full expansion reduces infant mortality by 22.40 deaths, while the availability charge reduces it by 0.04. With net private benefits of -342.09 million and 274.58 million reais, respectively, full expansion is justified if the value of preventing one infant death exceeds $27.58 = \frac{274.58-(342.09)}{22.40-0.04}$. In 2017, 27.58 million Brazilian reais corresponded to around 8.35 million U.S. dollars.

[28] As a reference, the U.S. Environmental Protection Agency considers the value of a statistical life to be

sive welfare assessment would require accounting for all externalities, which is beyond the scope of the paper.

8 Conclusion

This paper evaluates policies to improve water and sanitation connections through private providers, addressing a critical challenge in many developing countries. We leverage the case of Brazil, where the 2020 New Sanitation Regulatory Framework promotes private-sector involvement and sets ambitious connection targets for 2033. Using novel billing data from a major private provider and a structural model, we examine household decisions on service connection and consumption and firm decisions on expansion.

Our analysis reveals that meeting connection targets requires both firm expansion and consumer adoption. Even with full infrastructure expansion, only approximately 51.09% of households would connect to piped sewer, highlighting the need for demand-side incentives. However, we show that subsidizing household connections alone does not substantially increase connections because the firm, constrained by pricing structures where the price of sewage collection is lower than the provision cost, lacks incentives to expand.

Effective policies must provide incentives for both consumers and firms to make sewer provision financially viable. Among the alternatives analyzed, the most effective strategy to increase connections is consumer sewer connection subsidies with a moderate increase in sewer prices, as this approach encourages both expansion and adoption. However, it requires government funding, which may be politically and fiscally challenging. The Sewer Availability Charge is a viable alternative that incentivizes firm expansion at a lower cost to the government but shifts the financial burden onto consumers, raising affordability concerns. Policymakers must carefully balance these trade-offs to ensure expansion strategies are sustainable.

This study examines the effectiveness of policies currently under consideration and proposes improvements to better achieve connection targets. Further research should explore additional policy alternatives, including reforms to utility auctions and innovative infrastruc-

7.4 million US dollars in 2006 EPA (2010).

ture solutions. Additionally, future studies could investigate the effects of these policies on water quality, which, alongside access, is a crucial factor in improving household well-being.

References

Alsan, M. and C. Goldin (2019). Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy 127*(2), 586–638.

Barreca, A. and K. Clay (2016). Adapting to Climate Change : The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century Olivier Deschenes Michael Greenstone Joseph S . Shapiro. *Journal of Political Economy 124*(1), 105–159.

Barreto, M. L., B. Genser, A. Strina, A. M. O. Assis, R. F. Rego, C. A. Teles, M. S. Prado, S. M. Matos, D. N. Santos, L. A. dos Santos, et al. (2007). Effect of city-wide sanitation programme on reduction in rate of childhood diarrhoea in northeast brazil: assessment by two cohort studies. *The Lancet 370*(9599), 1622–1628.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Borenstein, S. (2009). To what electricity price do consumers respond? residential demand elasticity under increasing-block pricing.

Brichetti, J. P., L. Mastronardi, M. E. Rivas, T. Serebrisky, and B. Solís (2021). The infrastructure gap in latin america and the caribbean: Investment needed through 2030 to meet the sustainable development goals.

Burgess, R., M. Greenstone, N. Ryan, and A. Sudarshan (2020). The consequences of treating electricity as a right. *Journal of Economic Perspectives 34*(1), 145–169.

Burlig, F. and L. Preonas (2024). Out of the darkness and into the light? development effects of rural electrification. *Journal of Political Economy 132*(9), 2937–2971.

Cardell, N. S. (1997). Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory 13*(2), 185–213.

Coury, M., T. Kitagawa, A. Shertzer, and M. A. Turner (2024). The value of piped water and sewers: Evidence from 19th century chicago. *Review of Economics and Statistics*, 1–47.

Davis, L. W. and L. Kilian (2011). The allocative cost of price ceilings in the us residential market for natural gas. *Journal of Political Economy 119*(2), 212–241.

Deutschmann, J. W., J. Gars, J.-F. Houde, M. Lipscomb, and L. Schechter (2023). Privatization of public goods: Evidence from the sanitation sector in senegal. *Journal of Development Economics 160*, 102971.

Deutschmann, J. W., M. Lipscomb, L. Schechter, and J. Zhu (2024). Spillovers without social interactions in urban sanitation. *American Economic Journal: Applied Economics 16*(3), 482–515.

Devoto, F., E. Duflo, P. Dupas, W. Parienté, and V. Pons (2012). Happiness on tap: Piped water adoption in urban morocco. *American Economic Journal: Economic Policy 4*(4), 68–99.

Dubin, J. A. and D. L. McFadden (1984). An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica 52*(2), 345.

EPA (2010). Guidelines for preparing economic analyses.

Fay, M., D. Martimort, and S. Straub (2021). Funding and financing infrastructure: The joint-use of public and private finance. *Journal of Development Economics 150*, 102629.

Gamper-Rabindran, S., S. Khan, and C. Timmins (2010). The impact of piped water provision on infant mortality in brazil: A quantile panel data approach. *Journal of Development Economics 92*(2), 188–200.

Gandhi, A., Z. Lu, and X. Shi (2023). Estimating demand for differentiated products with zeroes in market share data. *Quantitative Economics 14*(2), 381–418.

Gautam, S. (2023). Quantifying welfare effects in the presence of externalities: An ex-ante evaluation of sanitation interventions. *Journal of Development Economics*, 103083.

Hausman, J. A. (1985). The econometrics of nonlinear budget sets. *Econometrica: Journal of the Econometric Society*, 1255–1282.

Hewitt, J. A. and W. M. Hanemann (1995). A discrete/continuous choice approach to residential water demand under block rate pricing. *Land Economics*, 173–192.

Ito, K. (2013). How do consumers respond to nonlinear pricing? evidence from household water demand. *Stanford Institute for Economic Policy Research*.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review 104*(2), 537–63.

Ito, K. and S. Zhang (2020). Do consumers distinguish fixed cost from variable cost?"schmeduling" in two-part tariffs in energy. Technical report, National Bureau of Economic Research.

Kresch, E. P. and R. Schneider (2020). Political determinants of investment in water and sanitation: Evidence from brazilian elections. *Economics Letters 189*, 109041.

Kresch, E. P., M. Walker, M. C. Best, F. Gerard, and J. Naritomi (2023). Sanitation and property tax compliance: Analyzing the social contract in brazil. *Journal of Development Economics 160*, 102954.

Lee, K., E. Miguel, and C. Wolfram (2020). Experimental evidence on the economics of rural electrification. *Journal of Political Economy 128*(4), 1523–1565.

Li, J. (2019). Compatibility and investment in the us electric vehicle market. *Unpublished manuscript, MIT*.

Mahadevan, M. (2024). The price of power: Costs of political corruption in indian electricity. *American Economic Review 114*(10), 3314–3344.

Marin, P. (2009). Public private partnerships for urban water utilities: A review of experiences in developing countries.

McRae, S. (2015). Infrastructure quality and the subsidy trap. *American Economic Review 105*(1), 35–66.

Moffitt, R. (1990). The econometrics of kinked budget constraints. *Journal of Economic Perspectives 4*(2), 119–139.

Olmstead, S. M. (2009). Reduced-form versus structural models of water demand under nonlinear prices. *Journal of Business & Economic Statistics 27*(1), 84–94.

Resende, D., G. Richter, M. Castello Branco Sant'Anna, and A. Trindade (2025). Pricing and informality: Evidence from energy theft in brazil. *Available at SSRN*.

Sears, J. (2023). Fluid demand responses: Long-and short-run urban water price elasticities. Technical report, Working Paper.

Szabo, A. (2015). The value of free water: analyzing south africa's free basic water policy. *Econometrica 83*(5), 1913–1961.

UN (2022). *The Sustainable Development Goals: Report 2022*. United Nations.Department of Economic and Social Affairs.

UNICEF and WHO (2023). Progress on household drinking water, sanitation and hygiene 2000–2022: special focus on gender.

von Sperling, M. and B. L. Salazar (2013). Determination of capital costs for conventional sewerage systems (collection, transportation and treatment) in a developing country. *Journal of water, sanitation and hygiene for development 3*(3), 365–374.

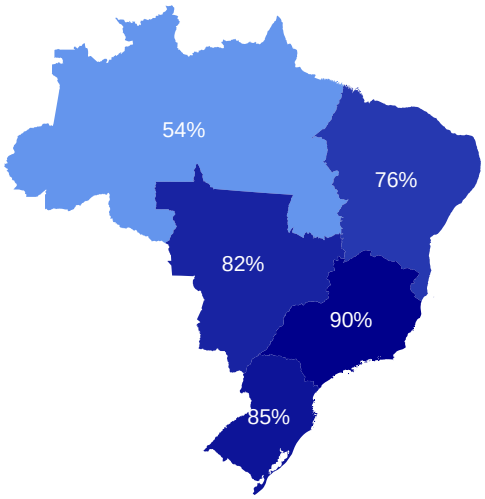
Wichman, C. J. (2014). Perceived price in residential water demand: Evidence from a natural experiment. *Journal of Economic Behavior & Organization 107*, 308–323.

Wollmann, T. G. (2018). Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review 108*(6), 1364–1406.

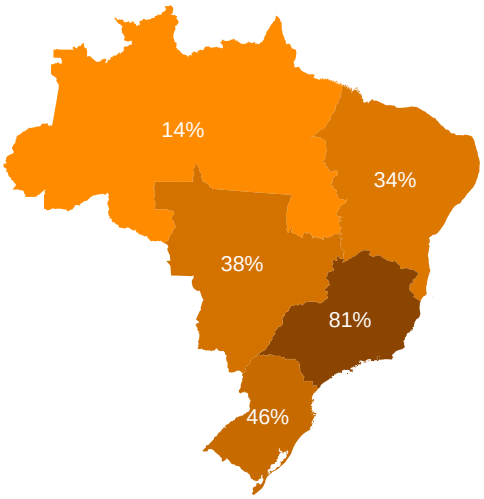
A1 Appendix figures

Figure A1. Share of the population connected by region in Brazil

(a) Piped water

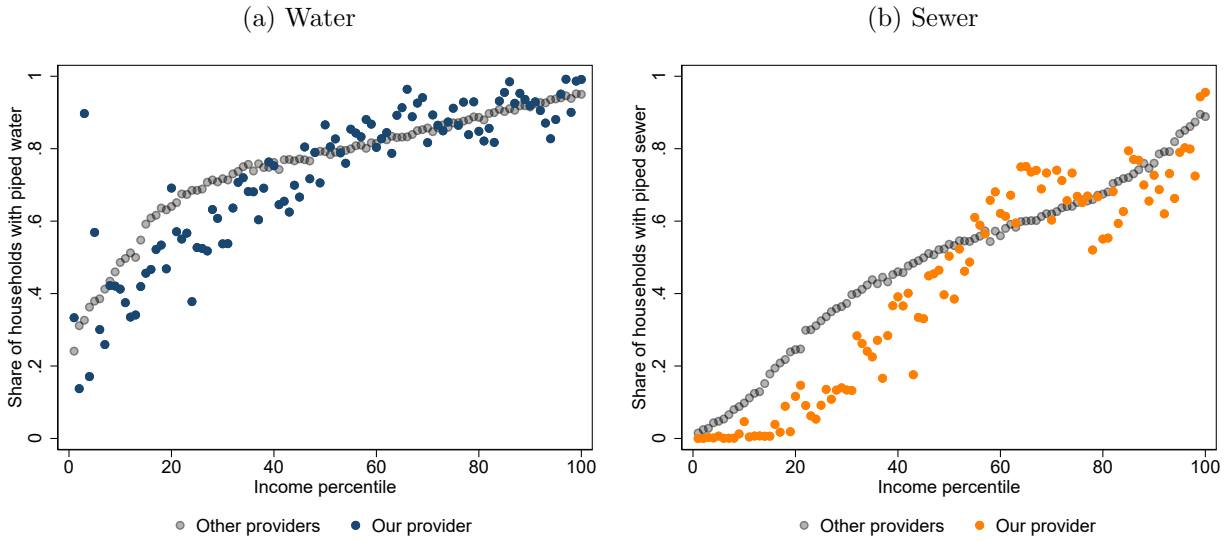


(b) Piped sewer



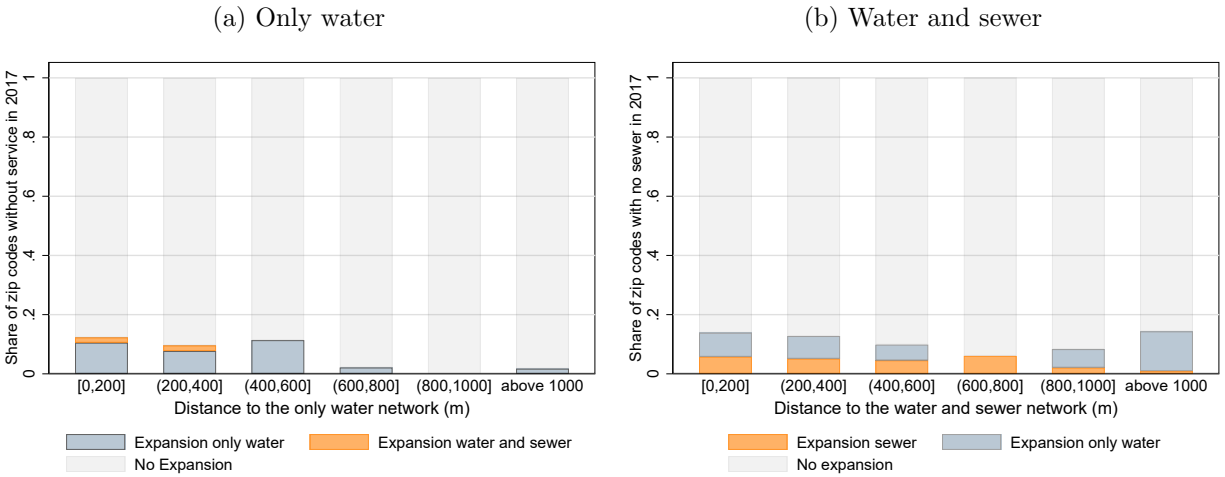
Notes: This figure illustrates the share of people connected to piped water (Figure (a)) and sewer collection (Figure (b)) across the country's regions in 2017. Data source: National Sanitation Survey (Pesquisa Nacional de Saneamento Básico - PNSB) from IBGE.

Figure A2. Share of households with sanitation services by income percentile



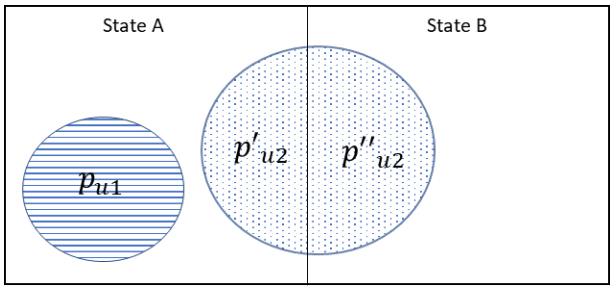
Notes: This figure illustrates connections to piped water (Figure (a)) and sewer services (Figure (b)). Each observation represents the average share of households connected across census tracts within income percentiles. In the left graph, areas within our firm's concession are highlighted in blue, while others are shown in gray. The right-hand graph marks our firm's concession areas in orange, with all other areas in gray. The data comes from 2010 Census.

Figure A3. Distances to the installed network



Notes: This figure depicts the distance from new expansions to the existing network. Figure (a) shows the distance to the installed water network, while Figure (b) includes both water and sewer networks. These graphs cover all zip codes within the firm's concession areas in the North and Northeastern regions of the country that lacked service in 2017.

Figure A4. Differences in the price schedule across concessions and states

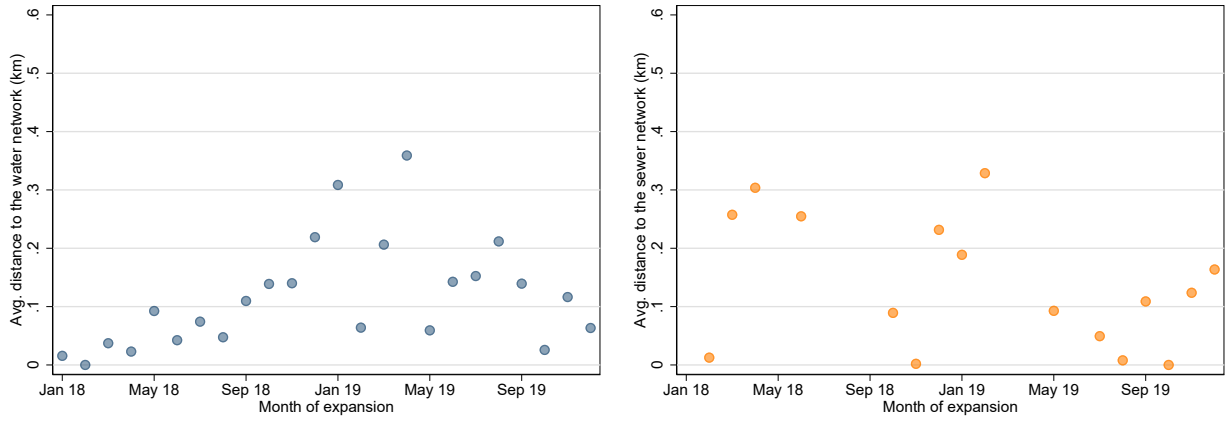


Notes: This figure illustrates the variation in the price schedule across concessions and states. Concession u_1 is entirely within State A, where consumers face price p_{u_1}. Concession u_2 spans State A and State B, leading to different prices (p'_{u_2} and p''_{u_2}) for consumers in the same concession but in different states. This setup generates both within-concession and within-state price variation, which helps identify consumer responses to fixed fee changes.

Figure A5. Distance of new zip codes per month

(a) Only water

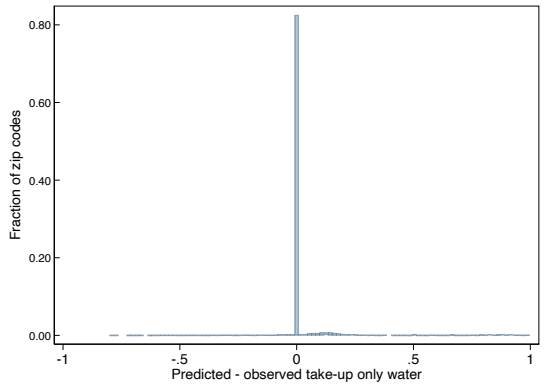
(b) Water and sewer



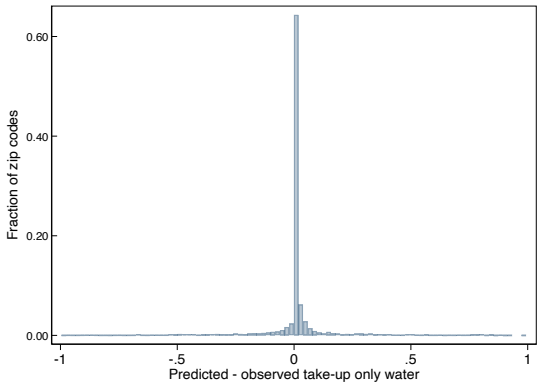
Notes: This figure depicts the distance of network expansions from the existing infrastructure in 2018 and 2019. Figure (a) includes zip codes where only water pipes were expanded, while Figure (b) on the right highlights zipcodes that received expansions of both water and sewer pipes in the North and Northeast regions.

Figure A6. Predicted vs. observed service take-up

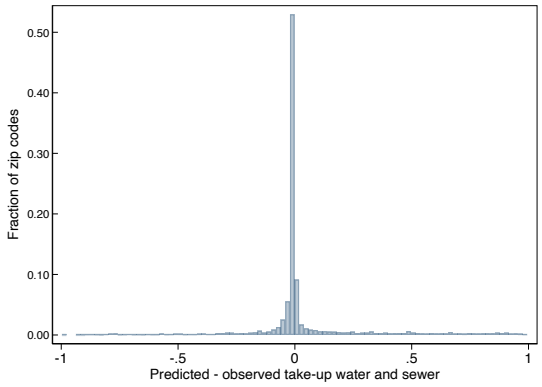
(a) Water take-up in zip codes with only water pipes



(b) Water take-up in zip codes with water and sewer pipes

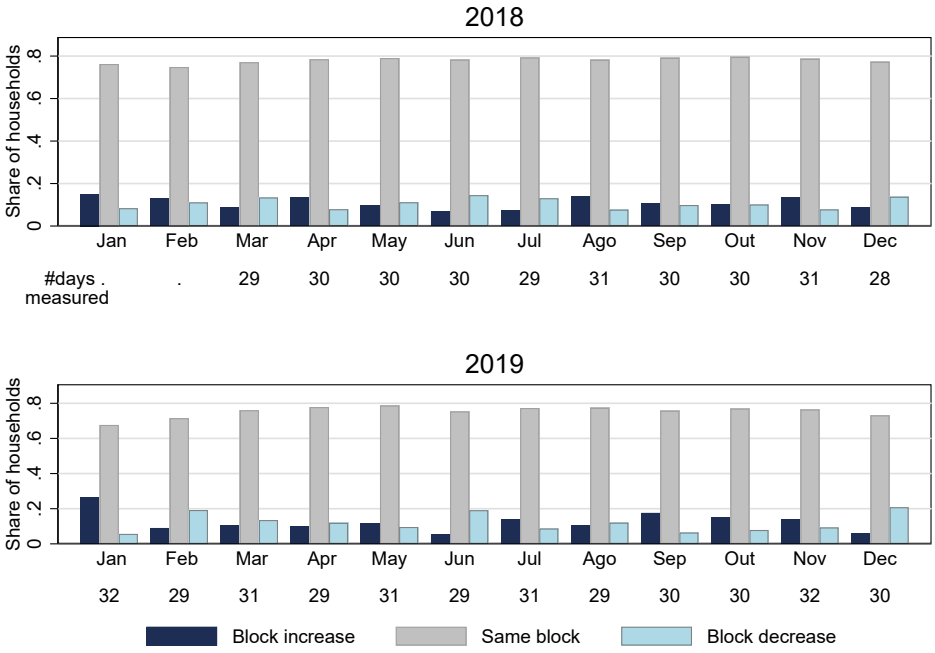


(c) Water and sewer take-up in zip codes with water and sewer pipes



Notes: These figures show the difference between the observed and estimated take-up rates. Figure (a) shows this difference for only water take-up in zip codes where only water pipes were available. Figure (b) presents this difference for only water take-up in zip codes with both water and sewer pipes. Figure (c) presents this difference for water and sewer take-up in zip codes where both water and sewer pipes are available.

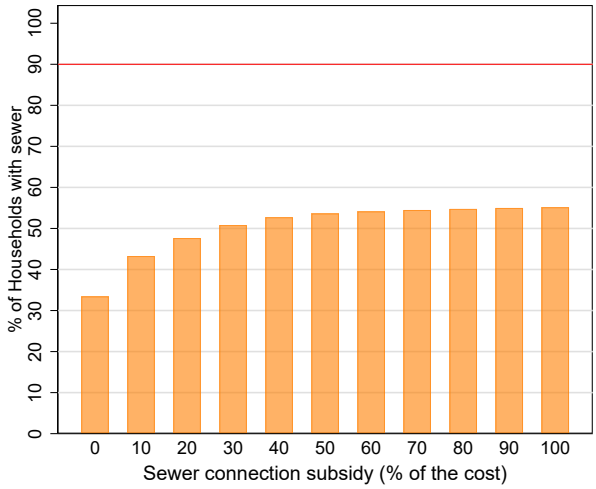
Figure A7. Bracket change



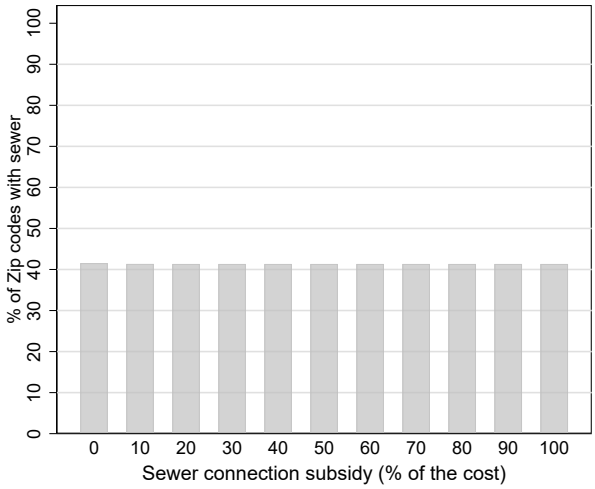
Notes: This figure illustrates the share of households that switch brackets from one month to the next. The data only includes households already connected in 2017 and had water bills for all months of 2017 and 2018. The graphs also include the number of days in the billing cycle from March 2018 to December 2019, which where not available for previous months.

Figure A8. Varying levels of subsidy

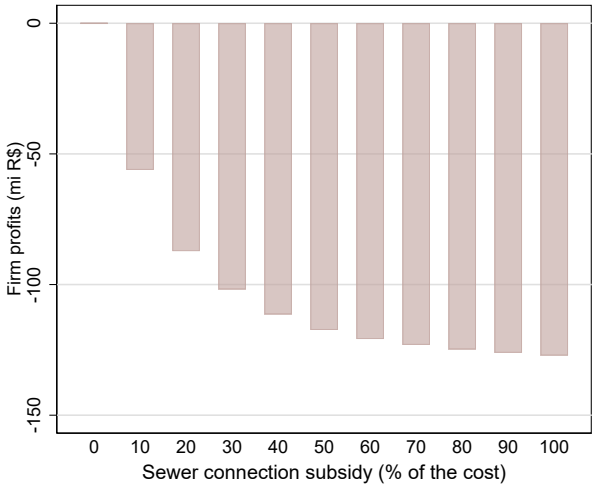
(a) Share of households connected to sewer



(b) Share of zip codes with sewer



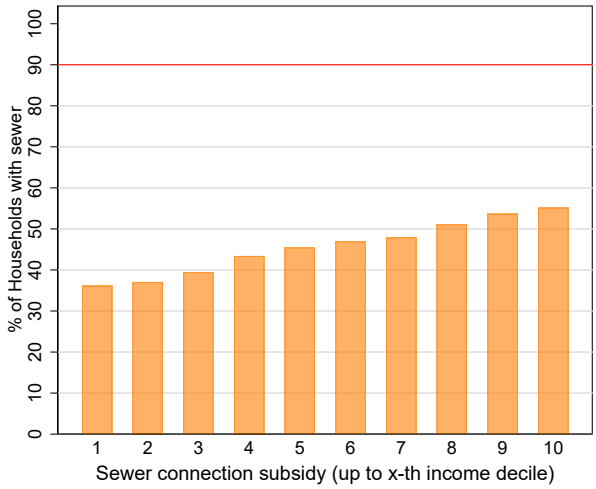
(c) Firm Profits



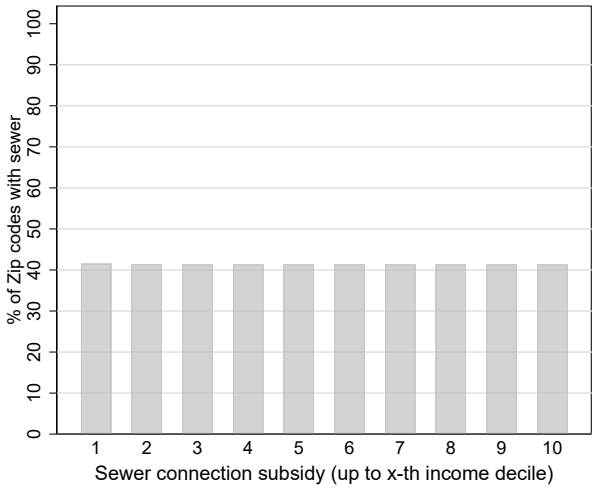
Notes: These figures present outcomes from counterfactual simulations under varying subsidy levels, ranging from 0% to 100% of the connection costs for the sewer network. Figure (a) illustrates the impact on the percentage of households with sewer access. Figure (b) shows the effects on the share of zip codes with sewer pipes. Figure (c) depicts changes in the firm's profit.

Figure A9. Subsidies to different income levels

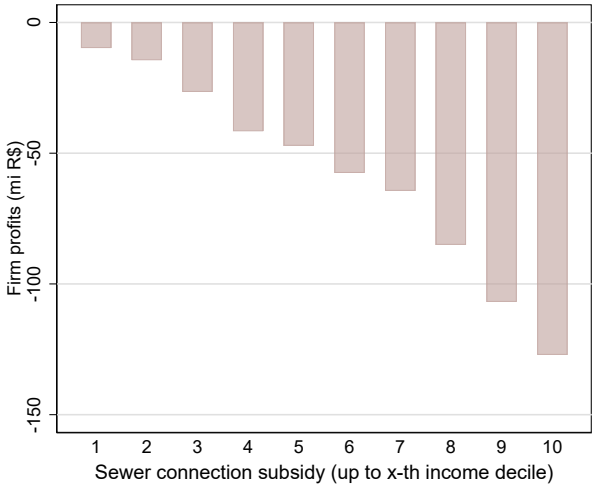
(a) Share of households connected to sewer



(b) Share of zip codes with sewer



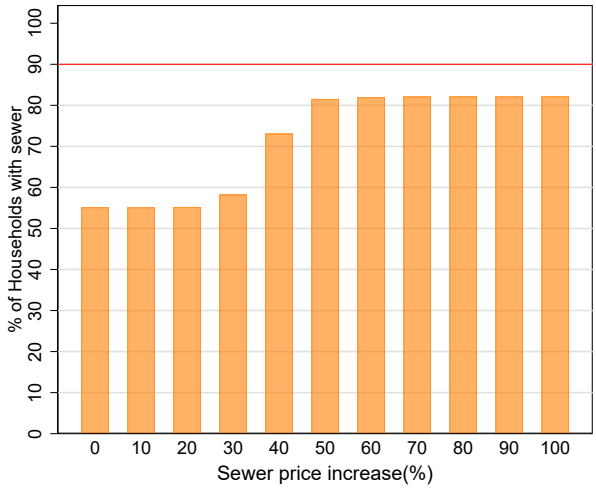
(c) Firm Profits



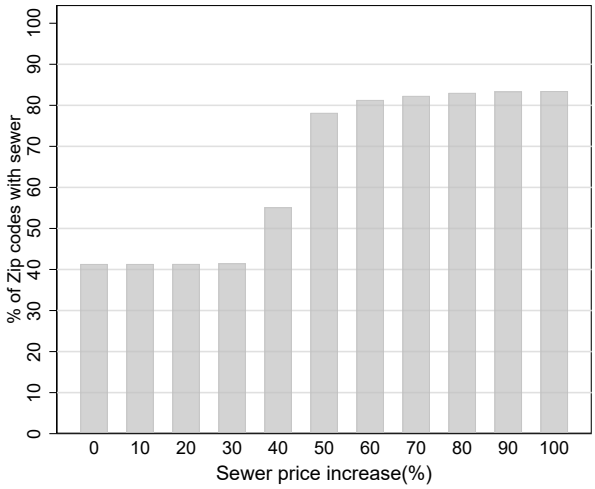
Notes: These figures present outcomes from counterfactual simulations where subsidies are targeted to different income deciles. Figure (a) illustrates the impact on the percentage of households with sewer access, Figure (b) shows the effects on the share of zip codes with sewer pipes, and Figure (c) depicts changes in the firm's profit. The subsidy is allocated to households up to the x-th income decile, meaning that for the 10th decile, all households receive the subsidy.

Figure A10. Subsidies with price increases

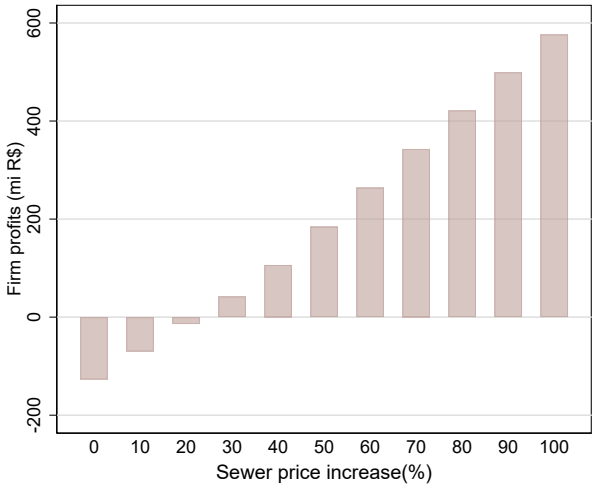
(a) Share of households connected to sewer



(b) Share of zip codes with sewer



(c) Firm Profits



Notes: These figures present outcomes from counterfactual simulations with a 100% subsidy on connection costs, combined with varying increases in sewer prices. Figure (a) illustrates the impact on the percentage of households with sewer access. Figure (b) shows the effects on the share of zip codes with sewer pipes. Figure (c) depicts changes in the firm's profit.

A2 Appendix tables

Table A1 – Take up regression

	(1) Take-up only water (Zips with only water)	(2) Take-up water and sewer (Zips with water and sewer)	(3) Take-up only water (Zips with water and sewer)
ln(Connection cost)	−0.073***	−0.080***	−0.073***
	(0.014)	(0.020)	(0.005)
ln(Income)	0.121***	0.126***	−0.039***
	(0.010)	(0.005)	(0.003)
Urban	−0.215***	0.113***	−0.051***
	(0.049)	(0.034)	(0.019)
Household size	0.199***	0.174***	−0.043***
	(0.018)	(0.008)	(0.004)
Share rented	0.347***	0.443***	−0.052***
	(0.042)	(0.022)	(0.013)
Share other water	−0.024	0.319***	−0.181***
	(0.033)	(0.029)	(0.017)
Share other sewer	0.147***	−0.092***	0.070***
	(0.014)	(0.009)	(0.005)
Municipality-year FE	yes	yes	yes
Observations	7,068	24,335	24,335
R-squared	0.614	0.2	0.371

Notes: This table reports the relationship between demographic variables and the share of connected households in each zip code where services are available. In column (1), the dependent variable is the probability of a household being connected to only water in zip codes that have only water pipes. In column (2), the dependent variable is the probability of a household being connected to both water and sewer in zip codes that have water and sewer pipes. In column (3), the dependent variable is the probability of a household being connected to only water in zip codes that have both water and sewer pipes. *** p<0.01, ** p<0.05, * p<0.1.

Table A2 – Water consumption - first stage

	(1) First stage Simulated IV	(2) First stage Simulated IV with selection
$ln(Avg.price(q_{sim}))$	0.908***	0.909***
	(0.005)	(0.005)
ln(Income)	0.011***	0.014***
	(0.002)	(0.002)
Piped sewer	0.027***	0.024***
	(0.004)	(0.004)
Urban	0.069***	0.085***
	(0.011)	(0.011)
Household size	−0.013***	−0.015***
	(0.002)	(0.002)
Selection inside opt.		−0.037***
		(0.003)
Selection outside opt.		0.042***
		(0.003)
Municipality-month FE	yes	yes
F-statistic	32,034	32,241
Observations	8,643,951	8,643,951

Notes: This table presents the first-stage results from the estimation of water demand. The dependent variable is the logarithm of water consumption, and the endogenous independent variable is the logarithm of the average water bill price. The instrumental variable is a simulated average price, constructed by grouping households into 16 categories based on income quartile, service type (water only or water and sewer), and urban or rural location. For each group, we compute the average consumption of households in different concessions, then calculate the price each household would face under their own concession's pricing schedule if they consumed the group's average usage in other concessions. The second column accounts for potential selection bias among households that opted to connect. Standard errors are clustered at the address level. *** p<0.01, ** p<0.05, * p<0.1.

Table A3 – Water consumption - reduced form

	(1) Reduced form Simulated IV	(2) Reduced form Simulated IV with selection
$ln(Avg.price(q_{sim}))$	−0.194***	−0.194***
	(0.009)	(0.009)
ln(Income)	0.119***	0.120***
	(0.003)	(0.003)
Piped sewer	0.064***	0.065***
	(0.006)	(0.006)
Urban	−0.175***	−0.172***
	(0.024)	(0.025)
Household size	0.057***	0.057***
	(0.005)	(0.005)
Selection inside opt.		−0.005
		(0.005)
Selection outside opt.		0.006
		(0.006)
Municipality-month FE	yes	yes
F-statistic	471	470
Observations	8,643,951	8,643,951

Notes: This table presents the reduced-form results from the estimation of water demand. The dependent variable is the logarithm of water consumption, and the independent variable is the logarithm of the simulated average price, constructed by grouping households into 16 categories based on income quartile, service type (water only or water and sewer), and urban or rural location. For each group, we compute the average consumption of households in different concessions, then calculate the price each household would face under their own concession's pricing schedule if they consumed the group's average usage in other concessions. The second column accounts for potential selection bias among households that opted to connect. Standard errors are clustered at the address level. *** p<0.01, ** p<0.05, * p<0.1.

Table A4 – Cost estimates under a 10-year profit consideration

	Costs
Mg. cost water (m^3)	10.725***
	(3.576)
Mg. cost water and sewer (m^3)	43.805*
	(25.237)
Cost per distance water (km)	241196.978
	(205039.312)
Cost per distance water and sewer (km)	148947.455
	(128611.363)
Cost per distance sewer (km)	16140.828
	(15476.696)
Number of zip codes	3,279

Notes: This table presents cost estimate parameters under the assumption that firms consider projected profits over the next 10 years when making decisions. The estimation includes zip codes within the firm's concessions in the North and Northeast regions of the country that lacked service in 2017. All estimated costs are reported in Brazilian reais (R$). In 2017, the exchange rate was approximately 3.3 R$ per 1 U$. *** p<0.01, ** p<0.05, * p<0.1.

Table A5 – Cost estimates under a 30-year profit consideration

	Costs
Mg. cost water (m^3)	8.668***
	(0.288)
Mg. cost water and sewer (m^3)	17.649***
	(0.785)
Cost per distance water (km)	14582.078***
	(1087.310)
Cost per distance water and sewer (km)	21367.639**
	(9275.147)
Cost per distance sewer (km)	3085.373***
	(686.804)
Number of zip codes	3,279

Notes: This table presents cost estimate parameters under the assumption that firms consider projected profits over the next 30 years when making decisions. The estimation includes zip codes within the firm's concessions in the North and Northeast regions of the country that lacked service in 2017. All estimated costs are reported in Brazilian reais (R$). In 2017, the exchange rate was approximately 3.3 R$ per 1 U$. *** p<0.01, ** p<0.05, * p<0.1.

Table A6 – Simulations with full expansion

	(1) Baseline	(2) Full expansion	(3) Full expansion Connection subsidy	(4) Full expansion Availability charge
Panel A: Service coverage and household connections				
% of zips with water and sewer	41.36	100.00	100.00	100.00
% of zips with only water	48.77	0.00	0.00	0.00
% of households connected to water	79.71	85.95	96.61	85.59
% of households connected to sewer	33.43	51.09	96.58	65.65
Panel B: Welfare impact relative to the baseline				
Δ Variable profit (mi R$)		−95.77	−382.94	207.31
Sunk cost (mi R$)		249.69	249.69	249.69
Δ Consumer surplus (mi R$)		3.37	88.63	−0.88
Connection subsidy (mi R$)		0.00	365.44	0.00
Δ% Infant deaths		−10.37	−24.44	−10.20
Expansion: firm choice	yes	no	no	no
Expansion: all zips with sewer	no	yes	yes	yes
Subsidy sewer connection	no	no	yes	no
Sewer price increase	no	no	no	no
Sewer availability charge	no	no	no	yes

Notes: This table presents counterfactual results assuming the firm expands the water and sewer network to all zip codes. Each column corresponds to a different counterfactual simulation. Column (1) shows the baseline. Column (2) considers full expansion. Column (3) adds a household subsidy for sewer connections. Column (4) assumes full expansion with the firm charging for sewer service based on availability rather than connection. Panel A reports service coverage outcomes, while Panel B presents welfare outcomes. Differences in variable profit, consumer surplus, and infant mortality are measured relative to the baseline. Variable profit and consumer surplus are calculated over five years, while the change in infant deaths is based on the estimated increase in household connections, following Gamper-Rabindran et al. (2010). The baseline scenario includes 216 infant deaths. All cost estimates are in Brazilian reais (R$). In 2017, the exchange rate was approximately 3.3 R$ per 1 U$.

Table A7 – Simulations without expansion

	(1) Baseline	(2) No expansion Connection subsidy	(3) No expansion Availability charge
Panel A: Service coverage and household connections			
% of zips with water and sewer	41.36	41.36	41.36
% of zips with only water	48.77	48.77	48.77
% of households connected to water	79.71	83.38	79.51
% of households connected to sewer	33.43	55.17	42.03
Panel B: Welfare impact relative to the baseline			
Δ Variable profit (mi R$)		−127.15	156.98
Sunk cost (mi R$)		0.00	0.00
Δ Consumer surplus (mi R$)		53.03	−2.47
Connection subsidy (mi R$)		127.16	0.00
Δ% Infant deaths		−2.60	0.00
Expansion: firm choice	yes	no	no
Expansion: all zips with sewer	no	no	no
Subsidy sewer connection	no	yes	no
Sewer price increase	no	no	no
Sewer availability charge	no	no	yes

Notes: This table presents counterfactual results assuming the firm does not expand either the water or sewer network. Each column corresponds to a different counterfactual simulation. Column (1) shows the baseline. Column (2) considers a household subsidy for sewer connections. Column (3) assumes the firm charges for sewer service based on availability rather than connection. Panel A reports service coverage outcomes, while Panel B presents welfare outcomes. Differences in variable profit, consumer surplus, and infant mortality are measured relative to the baseline. Variable profit and consumer surplus are calculated over five years, while the change in infant deaths is based on the estimated increase in household connections, following Gamper-Rabindran et al. (2010). The baseline scenario includes 216 infant deaths. All cost estimates are in Brazilian reais (R$). In 2017, the exchange rate was approximately 3.3 R$ per 1 U$.

A3 Online appendix

A3.1 Data

1. Billing records and census

To combine the water bill data with demographic information from the Census, we define that a zip code is located in the census tract where its centroid falls[29]. The variables at the census tract level are income, household size, number of households, number of owned versus rented houses, the share of the population with piped water and/or sewer, and whether the census tract is urban or rural. To better reflect the economic conditions of 2017-2019, we utilized population and GDP growth projections from IBGE (Brazilian Institute of Geography and Statistics) to update the information from the 2010 census at the municipal level.

We made the following assumptions to estimate demographic characteristics at the zip code level. Firstly, we assumed that the population of each census tract grows at the same rate as the population of its corresponding municipality. Secondly, we assumed that the income at the census tract level grows at the same rate as the municipal GDP per capita. All zip codes within the same census tract were assigned the same demographic characteristics based on these assumptions.

One of the challenges we encountered while working with the data was determining the number of households at the zip code level. Unfortunately, no administrative record is available that specifies the number of addresses in each zip code across the country. To tackle this issue, we implemented the following algorithm. Using the water bill data, we calculated the number of households connected to the piped water network in each zip code. Next, we determined the number of unconnected households at the census tract level by subtracting the number of connected addresses from the total number of households within that census tract. Finally, we distributed the number of unconnected households equally among the zip codes within the respective census tract. Consequently, the total population within a particular zip code is obtained by summing the number of connected households and the proportionate share of unconnected households from the census tract.

[29]Census tracts are the finest geographic area available in the Census-2010.

2. Connection costs

We computed the costs of connecting households to street pipes by developing the budget presented in Table OA.1. This budget outlines the items and quantities typically required for water and sewer connection projects, considering both paved and unpaved streets. The budget is based on comparable projects from the SINAPI dataset.

Table OA.1 – Budget water and sewer connection projects

Product Code (SINAPI)	Description	Quantity	Category	Pavement	Average Unit Price (R$)
89807	Short bend 90 degrees, PVC, normal series, DN 75 mm	3	sewer	any	18.50
89796	Tee, PVC, normal series, DN 100 x 100 mm	2	sewer	any	28.64
89798	PVC pipe, normal series, DN 50 mm	20	sewer	any	8.39
73658	House-to-box sewage connection, 10.0m PVC pipe, DN 100mm	1	sewer	any	526.17
89803	Short bend 90 degrees, PVC, normal series, DN 50 mm	4	sewer	any	9.68
95673	Water meter DN 20 (1/2)", 1.5 m^3/h - installation	1	water	any	111.43
95634	Water tripod kit, main entrance, weldable PVC DN 20 (1/2)"	1	water	any	104.42
89401	Pipe, PVC, weldable, DN 20mm, installed in distribution branch	10	water	any	5.89
95676	Concrete box for water meter, DN 20 (1/2)" - installation	1	water	any	71.47
89404	Elbow 90 degrees, PVC, weldable, DN 20mm	4	water	any	3.80
89405	Elbow 45 degrees, PVC, weldable, DN 20mm	4	water	any	4.02
89402	Pipe, PVC, weldable, DN 25mm, installed in distribution branch	5	water	any	7.17
92970	Demolition of asphalt paving, jackhammer, up to 15 cm thick	10	water	paved	11.86
92391	Execution of interlocking tiles, 35 x 25 cm, 6 cm thickness	10	water	paved	50.51

Notes: In 2017, the exchange rate was about 3.3 R$ to 1 U$.

Using this budget, we calculated the unit price for each item across different states from 2017 to 2019 to determine the total cost of connecting a household to the street pipes—both for water alone and for combined water and sewer connections.

We then computed the average connection cost for each service, weighting the costs for paved and unpaved streets according to the proportion of households with paved streets in each census tract. As a result, our final variable captures cost variations at the census tract and year level.

A3.2 Estimation

1. Empirical Bayes estimator take-up

One challenge in the service take-up estimation is that in some zip codes, all the addresses connect to the available pipes, generating market shares that are equal to 1 for the inside option and 0 for the outside option. In these cases, we would not be able to use the standard demand estimation methods Berry (1994); Berry et al. (1995) because the inversion step requires strictly positive market shares for each good in the market, in our case, for each service in the zip code. One common alternative is to aggregate markets, but in this setting, aggregating zip codes would not capture the relevant take-up faced by the firm when making expansion decisions. Another simple alternative, such as dropping the zeros/ones, would underestimate the service take-up.

We follow Li (2019) and use a parametric empirical Bayes or shrinkage estimator to generate strictly positive posterior take-up probabilities using information from similar zip codes. The number of addresses connected to service j in zip code z, given by K_{jz}, is modeled as a draw from a binomial distribution with N_z trials, representing the total number of addresses in the zip. Here we omit the year subscripts to facilitate the notation. The take-up probabilities S^0_{jz} for each service in each zip are drawn from a Beta prior distribution with parameters λ_{1jz} and λ_{2jz}. Such that $K_{jz} \sim Binomial(N_z, S^0_{jz})$ and $S^0_{jz} \sim Beta(\lambda_{1jz}, \lambda_{2jz})$. The posterior distribution of the take-up is also a Beta distribution

$$S_{jz} \sim Beta(\lambda_{1jz} + K_{jz}, \lambda_{2jz} + N_z - K_{jz})$$

with posterior mean

$$\hat{S}^P_{jz} = \frac{\lambda_{1zj} + K_{jz}}{N_z + \lambda_{1jz} + \lambda_{2jz}}$$

For each zip code z and service j the Beta prior is formed sing the 100 closest in income per capita that also have pipes for j, $l \in \zeta_z$, where l is a zip code from the set of similar zip codes ζ_z. The parameters of the beta prior distribution λ_{1jz} and λ_{2jz} are estimated from

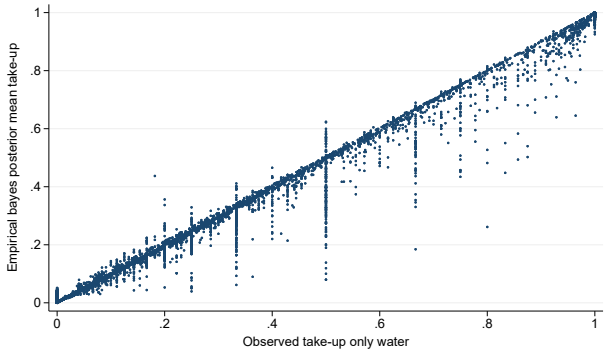
maximizing the log-likelihood over the take-up of similar markets

$$f(K_{jz}, l \in \zeta_z | \lambda_{1jz}, \lambda_{2jz}) = \prod_{l \in \zeta_z} \binom{K_{lj}}{N_l} \frac{\Gamma(\lambda_{1jz} + \lambda_{2jz})\Gamma(\lambda_{1jz} + K_{lj})\Gamma(N_l - K_{lj} + \lambda_{2jz})}{\Gamma(\lambda_{1jz})\Gamma(\lambda_{2jz})\Gamma(\lambda_{1jz} + N_l\lambda_{2jz})}$$

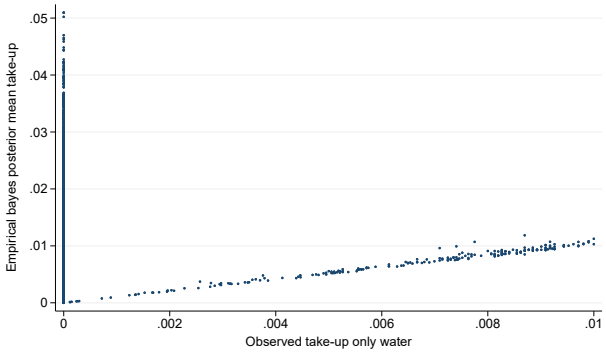
With the estimated parameters, we construct the posterior mean of the take-up probabilities for each zip and service $\hat{S}^P_{jz} = \frac{\hat{\lambda}_{1jz}+K_{jz}}{N_z+\hat{\lambda}_{1jz}+\hat{\lambda}_{2jz}}$, which are strictly between 0 and 1. The figures below show the empirical Bayes posterior mean take-ups and the observed take-ups for only water and water and sewer.

Figure OD.1. Take-up only water: empirical Bayes Posterior vs. observed

(a) All take-ups



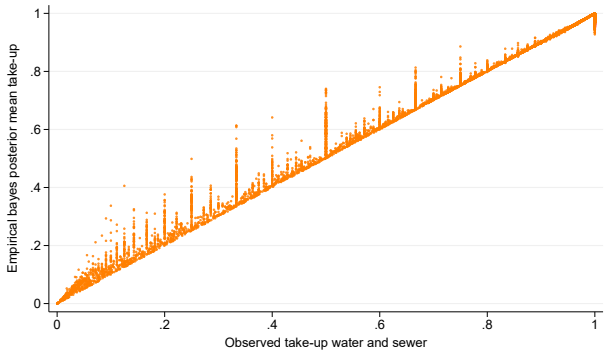
(b) Zooming in on observed take-ups below 0.1



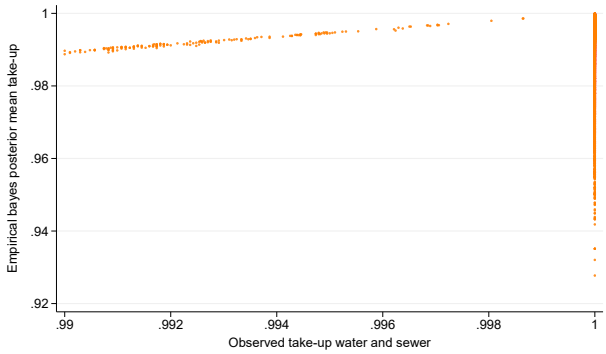
Notes: These graphs show the empirical Bayes Posterior Mean for each observed take-up of only water services.

Figure OD.2. Take-up water and sewer: Empirical Bayes Posterior vs. Observed

(a) All take-ups



(b) Zooming in on observed take-ups above 0.9



Notes: These graphs show the empirical Bayes Posterior Mean for each observed take-up of water and sewer services.