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

Carolina Tojal R. dos Santos† Bruna Morais Guidetti‡

University of Michigan

January 15, 2024

Click here for latest version

Abstract

Expansion of piped water and sewage collection is a critical issue in developing countries, where public investment is limited. In Brazil, the New Sanitation Regulatory Framework from 2020 intends to stimulate the entry of private providers into the market while setting connection targets: 99% of households with piped water and 90% with piped sewer. We assess the viability of these connection goals under the current regulated prices and compare the distributional impacts of supplemental policies. We use billing data from a major private provider and a structural model encompassing the expansion decisions of the service provider and the connection and consumption choices of consumers. We find that even if the firm expanded water and sewer everywhere, the targets would not be met, with only 43% connecting to piped sewage. Moreover, the expansion is not profitable for the firm. We then simulate charging consumers when piped sewage is available, regardless of connection, and price discounts on the sewer bills. Combinations of these two policies stimulate the firm to expand and consumers to connect, achieving 76% connected to sewer but reducing consumer surplus.

*We are extremely grateful to Ying Fan, Catherine Hausman, Zach Brown, and Francine Lafontaine for their guidance and advice. We also want to thank Hanna Onyshchenko, Agostina Brinatti, Iris Vrioni, Cristiano Carvalho, Fernando Crepaldi, Brock Rowberry, Lea Bart, Bruno Souza, Hiroshi Toma, Emilio Colombi, and Daniel Velasquez for their helpful comments. We are grateful to seminar participants at the University of Michigan, the Berkeley Energy Camp, and the MSU & UM Energy & Environmental Economics Day for their insights.

†Job Market Paper, e-mail: ctsantos@umich.edu , website: www.carolinatojal.com

‡e-mail: guidetti@umich.edu

1 Introduction

Access to clean water and sanitation remains a pressing issue worldwide, with 2.2 billion people lacking safely managed drinking water and 3.4 billion lacking sanitation as of 2022 (WHO, 2023). Extending piped water and sewage collection to households poses a significant challenge for developing countries, where governments often have limited resources to invest in infrastructure. Potential solutions to increase investment in the sector include engaging private firms in infrastructure construction and service provision, as was done in countries such as Brazil, Argentina, Chile, the Philippines, Indonesia, and South Africa (Marin, 2009). In this context, regulation can be used to balance the financial viability of the projects with social inclusion, especially to reach areas further away from the installed network and with lower demand. While numerous studies have documented the advantages of improved water services (Gamper-Rabindran et al., 2010; Devoto et al., 2012; Alsan and Goldin, 2019; Coury et al., 2022; Kresch et al., 2023), there has been little attention given to designing incentives that promote infrastructure expansion and connections in underdeveloped areas.

Brazil faces a significant challenge in expanding piped water and sewer services, particularly in its Northern region.[1] To address this issue, the government introduced the New Regulatory Framework in 2020.[2] This framework encourages municipalities to contract with private providers and sets ambitious connection targets: 99% of households with piped water and 90% with piped sewage collection within each concession by 2033. The new regulation has sparked debates regarding the feasibility of private providers expanding services under the current regulated prices. However, the debates overlook that achieving the connection targets depends on both the firm and consumers: within a concession, the firm is responsible for installing pipes up to the sidewalk, and consumers are responsible for connecting to these pipes. Considering this interaction is crucial when designing further incentives for the firm and consumers. For example, while higher prices may encourage network expansion, they may simultaneously discourage consumer connections.

[1] In the Northern region of the country 54% of the population was connected to piped water and only 14% connected to piped sewer in 2017, according to the National Sanitation Survey (Pesquisa Nacional de Saneamento Básico - PNSB) from IBGE. The share of connections in the other regions of the country is shown in Appendix A.

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

This paper aims to answer the following research questions. Would the connection targets be met if providers expanded water and sewer pipes everywhere? Is it profitable for the private providers to expand across all areas under their concession? Our findings suggest a negative answer to both of these questions under the current price schedules, which leads to further inquiries: What alternative price incentives are viable to attain the desired connection goals? How do these policies impact the welfare distribution among consumers and the provider?

We address these questions using novel billing data from a private provider in Brazil.[3] 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 (only water or water and sewer) within its concessions.

We first use the data to document key patterns in the firm expansion decisions, consumer connections, and consumption choices. We show that zip codes that receive both services have, on average, higher income than zip codes with only water. Moreover, the firm is more likely to expand in zip codes close to the installed network. We also document that some 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 this setting, connected consumers face non-linear price structures[4] for their water and sewage bills. It is essential to determine whether they respond to average or marginal prices in order to estimate the demand price elasticity and assess the impact of price changes on provider profits. We find suggestive evidence that water consumption responds to average prices rather than marginal prices. This conclusion is drawn from the absence of consumption bunching at price schedule kinks where the marginal prices rise. Also, consumption responds

[3]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. This firm was already in the market before the regulation

[4]Increasing block rates with fixed fees.

to price changes that do not change the marginal price but increase the average price.

We then use a structural model to predict the demand where the services are not yet available and recover the costs of expansion, which allows us to simulate the effects of different policies. On the supply side, the firm operates as a monopolist, providing piped water and sewer services in the region. The firm's decision is represented as a discrete-choice model where they choose which services to install based on the profitability of expanding in each potential zip code. The demand side comprises two components. First, we consider a discrete-choice model where households select the specific service they wish to connect to among the available options in their zip code. Second, we consider a continuous choice where connected households determine their water consumption level. We use cross-sectional variation in demographics and exogenous price changes to identify consumer preferences and their responsiveness to average prices. Leveraging the estimated demand, we predict the revenue the firm would generate if it expanded into each zip and estimate its cost parameters using the observed expansions.

Using the estimated model, we first use counterfactual simulations to investigate the viability of the connection targets. We simulate a scenario where the firm expands water and sewer in all zip codes within its concessions. We find that household connections would not achieve the targets due to limited consumer take-up. Specifically, only 43% of households would connect to piped sewers if available everywhere. Moreover, we show that a full expansion of water and sewer is not profitable for the firm without further incentives.

We next investigate three supplemental policies aiming to increase the share of connected households: (1) imposing charges for sewer availability, regardless of consumer connections;[5] (2) offering discounts on sewer prices, and (3) combining the aforementioned policies. When considering full expansion, we would successfully approach the 90% sewer connection target with the sewer availability charges, a 40% discount on sewer price, or a combination of sewer availability with a discount of 10%. However, these approaches diverge in terms of the distribution of welfare when compared to a situation with no expansion: the first results in a reduction in consumer surplus, the second generates negative profits for the firm, and in the

[5] Under this policy in every zip code that has sewage pipes installed, every consumer that receives a water bill will also be charged for sewer regardless of being connected to this service.

third, the expansion is profitable for the firm while there is a smaller reduction in consumer surplus.

When we allow endogenous expansion by the firm, the availability charge stimulates the expansion of the pipes and consumers to connect. Nevertheless, it is infeasible to achieve the targets because the firm would not have incentives to expand to all zip codes. In this case, implementing the sewer availability charge yields the highest sewer coverage, with 81% of households connected. However, this policy places the entire financial burden on consumers, leading to a reduction in their overall consumer surplus. Pairing this policy with a 10% discount on sewer bills would result in 76% of the households connected, reducing the burden on consumers while still promoting significant sewer coverage.

This paper contributes to the Environmental and Industrial Organization literatures by linking infrastructure expansion decisions and consumer demand in the water market. While the benefits of these services have been widely documented – including reductions in child mortality (Alsan and Goldin, 2019; Gamper-Rabindran et al., 2010; Barreto et al., 2007), increases in overall well-being (Devoto et al., 2012), increases in property values (Coury et al., 2022) and greater tax compliance (Kresch et al., 2023) – we are the first to investigate trade-offs associated with promoting network expansion and ensuring consumer connections. Our rich billing data enables us to capture consumer preferences related to take up and usage, akin to studies on electricity demand (Barreca and Clay, 2016; Davis and Kilian, 2011; Dubin and McFadden, 1984) and to incorporate the firm expansion decisions that have been explored in the context of energy, telecommunications, and retail (Granja, 2021; Li, 2019; Holmes, 2011; Jia, 2008).

We show that consumers may not be willing to connect to piped sewage even when it is available, which adds to the results of the adoption of other sanitary technologies (Gautam, 2023; Deutschmann et al., 2022) and of rural electrification (Lee et al., 2020). Incorporating endogenous firm expansion decisions, we show that price discounts to consumers disincentivize infrastructure investment, adding to the literature on the unintended consequences of utility subsidies (Mahadevan, 2021; Burgess et al., 2020; McRae, 2015). More broadly, we contribute to discussions surrounding the consequences of utility privatization (Deutschmann et al., 2023; Galiani et al., 2005), which generally consider provider expansion decisions as exogenous.

Our analysis also contributes to the literature on demand elasticity under non-linear pricing, which is crucial for understanding how consumers respond to price policies in the water market (Wichman et al., 2016; Szabo, 2015; Wichman, 2014; Olmstead, 2009; Hewitt and Hanemann, 1995), but also in the electricity (Ito, 2014; Borenstein, 2009) and gas markets (Ito and Zhang, 2020).

This paper has direct policy implications for Brazil but also offers insights into the comprehensive policy landscape of countries striving to achieve universal access to piped water, sewer services, and other utilities.

The structure of the paper is as follows. Section I provides background information about the water sector in Brazil. Section II describes the data, while Section III presents the descriptive evidence obtained from the data. The model is presented in Section IV, and the estimation strategy is detailed in Section V. We conduct counterfactual simulations in Section VI. Finally, Section VIII concludes the paper.

2 Institutional Background

In Brazil, water and sanitation services fall under the jurisdiction of municipalities, and they can select how to provide these services[6]. The Regulatory Framework of 2007[7] allowed municipalities to contract with public companies without a competitive process, reserving auctions for cases involving private providers.[8] 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.

In 2020, a New Regulatory Framework[9] was enacted to encourage private investments

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

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

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

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

and address the challenge of expanding piped water and sewer access. Under this regulation, competitive auctions are mandated for all contracts. 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.[10].

The targets in the recent regulation are based on connections, which hinge on service providers and consumers. Providers are responsible for constructing pipes until the sidewalk, while users connect their residences to these pipes. In particular, both the 2007 and 2020 laws say that all permanent urban buildings must connect to available public water and sewage networks, and users are subject to tariffs and fees associated with these services.

The regulations also allow municipalities to charge for sewer availability; in neighborhoods where piped sewage is available, households receiving water bills can also be charged for sewer regardless of connection. Nonetheless, most municipalities do not implement this policy[11], and the federal law does not specify any other sanction for non-connections. A survey of the 100 largest municipalities in Brazil (TrataBrasil, 2015) identifies the top three reasons for consumers' under-utilization of the network are the reluctance to pay tariffs, lack of information, and absence of penalties for noncompliance.

From 2020 to 2023, 28 auctions awarded service provision rights to private providers[12]. 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. Our analysis is made with 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.

[10]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.

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

[12]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

In this paper, we use novel confidential water bill data from a private provider in Brazil[13] with concessions in different regions. This data combined with demographic information from the 2010 Census, allows us to capture consumption patterns but also the expansion of the firm.

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 [14].

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[15]. 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 use the municipality GDP and population growth to update the income and number of households, respectively. More details about the data construction at the zip code level are described in appendix B.

The areas under the firm's concessions are representative of access to piped water and sewage compared to the rest of the country. The graphs in figure 1 show the share of the population by census tract with access to piped water and sewage, respectively. In the left graph, the census tracts that are within the firm concession regions are marked in blue, and the other census tracts in the country are in gray. In the graph to the right, the census tracts under the firm's concessions are represented in orange. Access to these services is highly correlated with income, and the firm provides services to a wide range of locations, including in low-income areas where access is still very low.

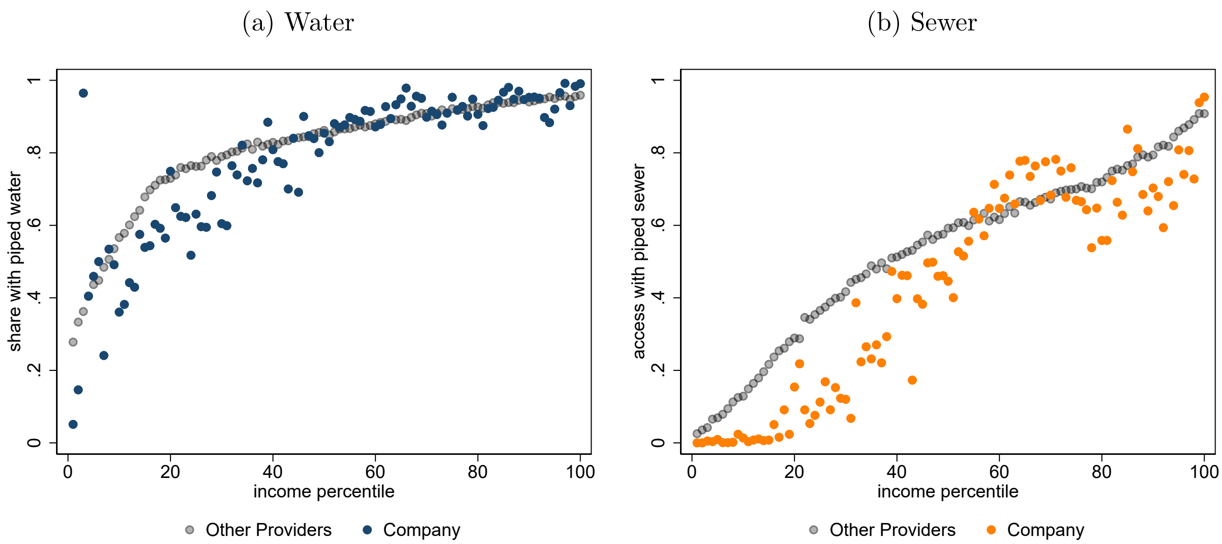
We focus our analysis on municipalities where the firm provides piped water and sewage

[13]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.

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

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

Figure 1. Access to piped water and sewer by census tract



Notes: includes all the census tracts of the country. Each observation is a census tract. The colored ones within the firm's concessions.

collection services. So, we can explore the firm decision to expand these services jointly or separately and understand consumers' preferences with respect to the services since they are charged in the same bill and are managed by the same provider.

3.1 Water Consumption Data

Our data set comprises monthly metered water bills from consumers connected to the network of pipes in areas under the company's concessions. The data set includes a total of 29.2 million water bills. We can track addresses over multiple billing periods, although we cannot determine if there have been changes in households residing at these addresses over time. However, for simplicity, we will refer to addresses and households interchangeably throughout the paper. We restrict our analysis to residential consumers, who account for 92% of the bills in the data and 89% of the volume of water consumed[16].

Addresses in our data set can be connected to water and sewer pipes or solely to the water network. In our sample, any household with piped sewage also has piped water. We only

[16]For the purpose of analyzing 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 200 m^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.

observe the water consumption of connected households, so all the other possible sources of water and wastewater destinations will be treated as outside options. 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.

The price schedules for piped water and sewage are fixed at the start of each concession contract. Subsequent price changes during the contract period are considered exogenous, as the firm has limited control over the timing and magnitude of these adjustments. Contracts typically include provisions for inflation correction over time, and in case of unexpected cost shocks, the firm may request price increases. However, the authorization of such changes is at the municipality's discretion, which often takes several months to respond to requests. If approved, municipalities tend to impose rules requiring delayed or incremental price increases spread across multiple months, aiming to mitigate the impact on the population and reduce the political cost associated with the increase.

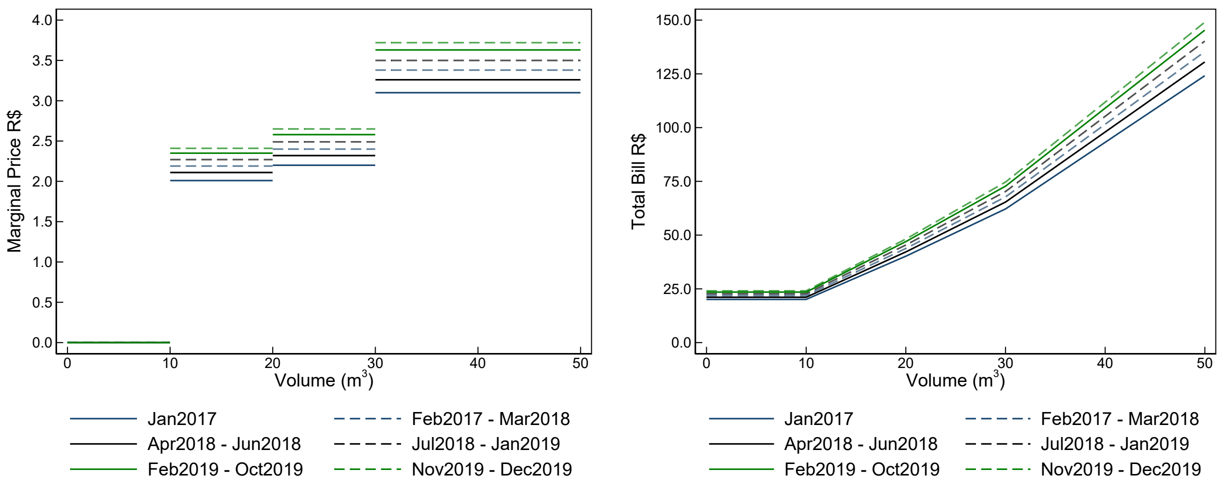
The pricing structure comprises increasing block tariffs (IBT) with fixed fees. Consumers are charged based on their water meter readings, with sewer charges calculated as a proportion of the water rates. Notably, the first block of water usage incurs a zero marginal price, meaning that addresses consuming within this range only pay the fixed fee. As water usage moves into higher blocks, the price per m^3 increases. To calculate the water bill, the firm determines the amount of water consumed by the user during the billing period, assigns the appropriate block for each unit of water consumed, and applies the corresponding price per m^3 to calculate the cost for that block. The costs for all blocks are then summed with the fixed fee to determine the total water bill. Households connected to the sewage network are charged a percentage (ranging from 50% to 100%) of their total water bill for sewer services.

To illustrate the general structure and variation in price schedules, we present figure 2 showcasing the marginal prices (figure 2a) and the corresponding total bill (figure 2b) for different volumes of 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 increases, only permitting adjustments to account for inflation that year. In

2019, the company overturned the previous decision and implemented price increases in two increments.

Figure 2. Price Variation from one municipality in the data

(a) Marginal prices (b) Total bill



Notes: The figures represent the marginal prices and the corresponding total bill for water in one of the municipalities in our data. The other municipalities present similar patterns. The different dates represent periods of time when the regulated prices were at the displayed levels.

The fixed fee and marginal prices exhibit variation among different concessions and over time. When there are inflation corrections or changes authorized by the municipalities, both the fixed fee and the marginal prices are adjusted simultaneously and at the same rate. Concessions may have slightly different limits for each consumption block; however, these limits remained constant throughout the time period covered by our data. Additionally, the percentage of the total water bill charged for sewer services is determined at the concession level and remains unchanged over time.

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. It is important to note that those with sewer services tend to have higher monthly bills even if their water consumption is similar to those without sewer services.

Table 1 – Descriptive Characteristics: Connected Households

	Means (standard deviation)	
	Only Water	Water and Sewer
Income (R$)	2523.13	2770.00
	(1566.93)	(2146.90)
Urban	0.76	0.93
	(0.43)	(0.25)
Household size	3.52	3.33
	(0.45)	(0.39)
Water Consumption (m^3)	9.43	11.68
	(55.84)	(26.06)
Total Bill (R$)	55.66	93.83
	(85.28)	(193.54)
Number of households	304194	327089

Notes: Here we show the average demographic characteristics and the corresponding standard deviations of households connected to only water or water and sewer services. Excludes households that switched services over our sample period.

3.2 Pipe Networks

We also use the billing data to infer which zip codes the firm expanded the services. During the analyzed period from 2017 to 2019, the firm was not obligated by the connection targets, being able to choose the locations where to expand their water and sewage pipe network within their concession boundaries. However, we lack access to the firm's records detailing the expansion locations, and the available administrative data on the pipe network is aggregated at the municipality level, making it insufficient for capturing the relevant demographics and cost factors that influence the firm's decisions. To address this, we examine the addresses appearing in the water and sewage bills to infer the expansion areas.

Our approach is based on the assumption that if there is at least one water bill in a specific zip code during a billing month, it indicates the presence of water pipes in that zip code starting from that period on. Similarly, if there is at least one bill containing sewage charges, it implies the existence of sewage pipes in that location. We focus our analysis on the zip code level because, according to the firm, engineering projects are defined at the street level, and in Brazil, each zip code generally corresponds to a street. Additionally, this is the finest geographical level observable in our data.

We define that the installed network of pipes is in zip codes that appear in the billing records with addresses having water bills in 2017. Regarding the expansion, we consider the construction of pipes in zip codes that did not have an installed network in 2017 but started appearing in the water bills in 2018 and 2019. Again, we may observe zip codes appearing in the water bills with addresses connected to water and sewer services or only water, allowing us to determine the types of services expanded. We also account for the possibility that zip codes with an installed network for water only in 2017 may receive the expansion of sewer services in subsequent years if bills charging for both services appear there.

For the cost estimation, we restrict the data to concessions located in the North and Northeast of the country where there is still considerable space for expansion. The concessions in the South and Southeast have close to universal water and sewer coverage.

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 or 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 – Pipe Network situation by zip code

Situation	Number of Zips	Avg. Income	Distance from Only Water Network (km)	Distance from Water and sewer Network (km)
Old Water and Sewer	2042	4166		
Old Only Water	2246	2530		9.44
Old Only Water/New Sewer	63	2992		3.49
New Water and Sewer	18	4001	0.23	0.18
New Only Water	131	2924	0.18	34.32
Nothing	723	3121	3.41	17.56

Notes: Includes only zip codes located in concession areas of the firm in the North and North East of the country. Zips with at least one bill corresponding to the service in 2017 are considered "old". Zips that only appeared in the water bills in 2018 or 2019 are considered "new". While the remaining zips with no billing record over the period are considered as having no service.

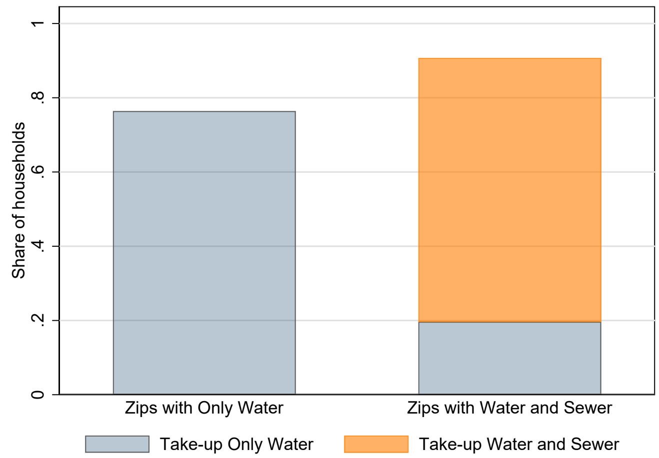
The firm expands closer to the installed network. The last two columns of table 2 show that the zip codes where the firm expanded are on average closer to the network of the specific service installed. Further evidence is displayed in appendix C.1. This pattern is unsurprising given the interconnected 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.

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 3, on average approximately 20% 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. In appendix C.2 we show the averages by income group. Connecting to the sewer system involves connecting the house to the main pipeline, which leads to 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 Table 3 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 with incom-

Figure 3. Average Service Take-up



Notes: In blue, we have the average share of households connected to only wanted, and in orange the average share of households connected to water and sewer. The left bar includes zip codes with pipes of only water. The right bar includes zip codes that have pipes for both water and sewage.

plete take-up[17]. Additionally, larger households positively correlate with complete take-up, while households with alternative sewer collection methods on average connect less. Notably, demographics account for a significant portion of the variation observed in adoption rates.

Table 3 – Take up regression

	(1) Take-up Only Water (Zips with Only Water)	(2) Take-up Water and Sewage (Zips with Water and Sewer)	(3) Take-up Only Water (Zips with Water and Sewer)
ln(Income)	0.110***	0.124***	−0.043***
	(0.010)	(0.005)	(0.003)
Urban	−0.212***	0.115***	−0.047**
	(0.049)	(0.034)	(0.019)
Household size	0.202***	0.176***	−0.039***
	(0.018)	(0.008)	(0.004)
Share rented	0.237***	0.419***	−0.102***
	(0.036)	(0.022)	(0.012)
Share Other Water	0.020	0.351***	−0.110***
	(0.032)	(0.028)	(0.016)
Share Other Sewer	0.142***	−0.086***	0.081***
	(0.014)	(0.009)	(0.005)
Mun-year FE	yes	yes	yes
Observations	7,068	24,335	24,335
R-squared	0.613	0.2	0.365

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.Column (1) includes zip codes with only water pipes. Columns (2) and (3) include zip codes that have pipes for both water and sewer. Includes only zip codes that had pipes installed since the beginning of our sample. The regressions show correlations; these results should not be interpreted as causal.

[17]Here 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.

4.3 Consumption responds to average price

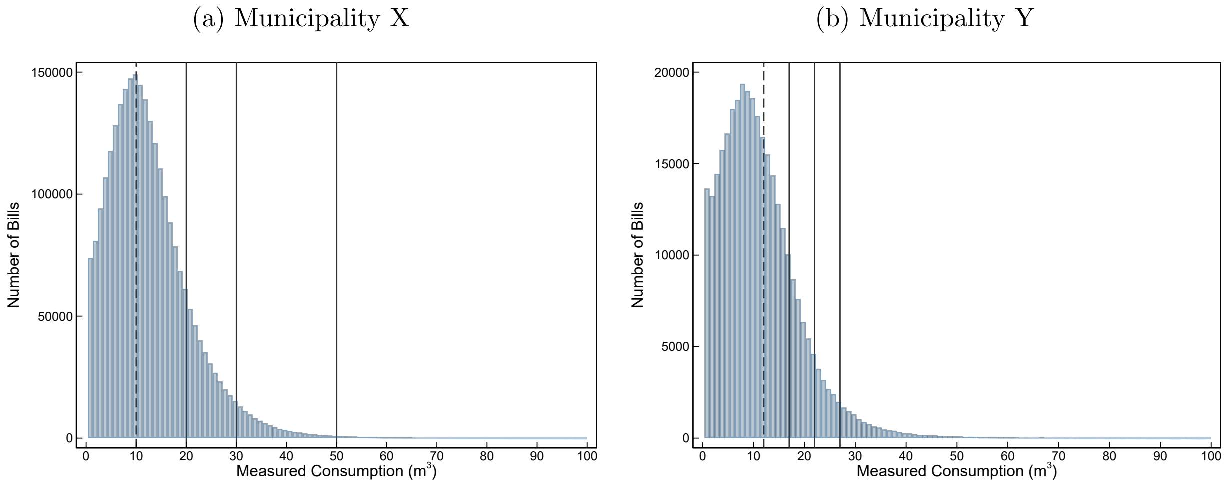
Here we investigate if consumers facing the non-linear price schedules respond to marginal or average prices, which will be important to compute their price elasticities to the perceived prices. In our context, we find suggestive evidence that consumers react to average price, consistent with other work in the water and energy markets (Sears, 2023; Wichman et al., 2016; Ito, 2014). However, this result goes against other related papers on water markets that model consumers reacting to marginal prices as (Szabo, 2015; Olmstead, 2009; 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.

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 4, we present the histograms for two specific municipalities: Municipality X (Figure 4a) and Municipality Y (Figure 4b), 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. Notably, no bunching is observed at the first price discontinuity, where we would expect consumers to bunch when facing a zero marginal price.

Figure 4. Residential Water Consumption



Notes: The vertical lines 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. Here we show the measured consumption for consumers connected since the beginning of our sample and in single units with individual billing.

In figure 4a, there is a concentration of water bills at $10m^3$. Despite the smoothness of the distribution, this could be mistaken as a bunching point. We address this by plotting the consumption graph for another concession where the end of the first bracket is at $12m^3$ instead of $10m^3$. As shown in figure 4b, no bunching is observed at the price discontinuity, and the consumption concentration is still at $10m^3$. Therefore, this concentration is unrelated to the discontinuity in marginal price.

The absence of bunching can be interpreted in two ways: either consumers exhibit zero elasticity with respect to prices, or they respond to an alternative measure of price. To distinguish between these two possibilities, we examine households that consistently fall within the first consumption bracket, where the marginal price is zero but the average price is positive. Analyzing this specific group of households allows us to distinguish if demand indeed responds to average price, given that our setting lacks price variation that would move average prices and marginal prices in different directions, as observed in prior studies (Ito, 2014, 2013).

2. Consumers react to changes in the fixed fee

To further investigate whether consumers respond to price we test if they distinguish the fixed fee from the marginal prices. In our scenario, 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 in the marginal rate.

We specifically focus on consumers who consistently consumed quantities of water within the first bracket of the concession they belonged to throughout the 36-month period covered in our sample. These households faced zero marginal prices but had a positive average price. Additionally, we narrow down our data to two specific concessions (u_1 and u_2) that allow us to isolate the effects of changes in the fixed fees from other confounding factors. Figure 9 in appendix C.3 illustrates these concessions, with municipalities in concession u_1 located in two different states (States A and B), each facing different price schedules. On the other hand, all municipalities in concession u_2 are situated in State B. By leveraging state-time and concession-time fixed effects, we are able to control for economic and weather-related shocks that could impact consumption.

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:

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

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. y_{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 $\Delta ln(y_{ct}) = ln(y_{ct}) - ln(y_{c0})$ 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 4 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).

Table 4 – Fixed fee

	(1)	(2)	(3)
$\Delta ln(Fee_t)$	−0.193**	−0.211*	−0.250*
	(0.083)	(0.114)	(0.145)
$\Delta ln(Income_t)$	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: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.Standard errors clustered at the address level. Includes only addresses with water bills for all periods in the sample, did not switch the type of service (just water/ water and sewer), and consumption was always in the first bracket. Also, it includes only single units that are billed individually.

This finding suggests that consumers do not differentiate between fixed and variable costs, which is consistent with evidence found in the gas market 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.

Our focus is on understanding how consumers on average respond to price in order to forecast consumption and to predict the consumption in unconnected zip codes if they were to receive these services. So even if there are different types of consumers, ranging from those unaware of the pricing structure to others who closely monitor their consumption and respond to marginal prices, the presented results are consistent with consumers on average responding to average prices.

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, this firm is the sole provider 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 incorporates a discrete-continuous approach to analyze consumer preferences for service take-up and the amount of water consumed after connecting to the network.

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 number of connected households, and the quantity of water 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 Supply

There is a single monopolist firm that can offer different services in different zip codes within its concessions. The monopolist decides which service j to offer among the options water and sewer together (s), only water (w), or nothing (o) in each zip code, with the goal of maximizing its profits. The monopolist faces a sunk cost of expansion that must be recovered through the supply of the chosen service and the collection of water and sewage bills from the connected consumers. Since the prices are regulated, the firm can only decide where to expand but cannot choose prices once the contract is in place.

We model the decision rule of the managers responsible for the expansion decisions as follows. They will compare the profits from installing a particular service in a zip code to the profits of installing an alternative service or not building any pipes in that zip code and choose the option that yields higher profits. Each zip code (z) is located in a census tract (c) within a concession unit (u). To simplify the notation, we will only use the subscript z for variables that are at the zip code level. The subscripts c and u will be used only for variables with variation in these more aggregate levels. The profits at the zip code z with service j is given by

$$\Pi_{jz} = VP_{jz} - SC_{jz} \tag{2}$$

where VP_{jz} is the variable profit and SC_{jz} is the sunk cost of constructing the network of pipes for service j.

More precisely the variable profit VP_{jz} is defined by the equation below

$$VP_{jzy} = \sum_{y=1}^{5} \frac{\pi_{jzy}}{(1+r)^{y-1}} \tag{3}$$

which represents the present value of the flow of yearly variable profits, π_{jz}, that the firm collects from the next five years after they install the network, discounted by the interest rate r. Although the contracts last, on average 30 years, the expansion decision of the managers typically accounts for a time frame, on the estimation we discuss alternative decision periods. The variable profits π_{jzy} will be given by the revenues collected from charging monthly water and sewer bills from connected consumers minus the costs associated with the service provision.

$$\pi_{jzy} = N_{zy} S_{jzy} (R_{jzy} - mc_j Q_{jzy}) \tag{4}$$

where N_{zy} is the number of addresses in the zip code in year y, S_{jzy} is the take-up of service j. $Q_{jzy} = \sum_t^y q_{jzt}$ is the sum of monthly water consumption q_{jzt} by the average household in zip z connected to service j. $R_{jzy} = \sum_t^y R(q_{jzt})$ is the revenue associated with the monthly consumption and price schedule in zip code z in each billing month. mc_j is the marginal cost per unit of water consumed with service j.

Consistently with the descriptive evidence, we consider that when service is available at the zip code only a share of consumers S_{jzy} will actually connect to the service j. For the connected consumers, we assume that they behave as an average consumer in the zip code that demands q_{jzt} cubic meters of water per billing month associated with the service they have. They will be charged water and sewage bills based on their measured consumption q_{jzt} and the regulated increasing block rates with fixed fees, as described earlier, which generates the monthly revenue R_{jzy} collected by the firm. Therefore, the elements S_{jzy}, Q_{jzy}, and R_{jzy} will be determined by the demand side of our model.

The marginal cost mc_j incorporates the costs of the water and sewer treatment and the costs of delivering each m^3 of water to the household tap and collecting the sewage that comes out. And it will be recovered together with the fixed costs.

To incorporate the key factors of the fixed cost of installing the network for service j in a zip code z, we parameterize it as

$$SC_{jz} = \omega_{1j} dist_{jz} + \omega_{2j} \bar{N}_z \tag{5}$$

where $dist_{jz}$ is the distance from zip code z to the network of service j and ω_{1j} is the cost per kilometer of pipes of service j. $\bar{N}_z$ is the average number of households in the zip code over the periods of the next 5 years and ω_{2j} is the fixed cost of installing the service j per household. Although some households may not connect, the firm builds pipes that cover all the addresses in the zip code and that will have enough capacity to provide to the neighborhood.

The manager's decision is then a discrete choice between the services in each zip code where the firm expands if it is profitable to do so. For instance, the firm will expand only water if the profits from providing only water are greater than if they do not provide any service and are also greater than the profits they would get by expanding the network of water and sewer and providing both services at the same time. Assuming that there is also a profit shock ε_{jz} that the firm observes but the econometrician does not the inequalities are as follows:

$$VP_{wz} - \omega_{1w} dist_{wz} - \omega_{2w} \bar{N}_z + \varepsilon_{wz} > \varepsilon_{0z} \tag{6}$$

$$VP_{wz} - \omega_{1w} dist_{wz} - \omega_{2w} \bar{N}_z + \varepsilon_{wz} > VP_{sz} - \omega_{1s} dist_{sz} - \omega_{2s} \bar{N}_z + \varepsilon_{sz} \tag{7}$$

Alternatively, the firm will expand water and sewer if the profits are greater than not providing any service and greater than the profits they would get if they provided only water without sewer.

Assuming ε_{jz} is i.i.d and distributed Extreme Value Type I we can compute the probability of a zip code z getting service j as

$$Pr_{jz} = \frac{exp(VP_{jz} - \omega_{1j} dist_{jz} - \omega_{2j} \bar{N}_z)}{1 + \sum_{k=w,s} exp(VP_{kz} - \omega_{1k} dist_{kz} - \omega_{2k} \bar{N}_z)} \tag{8}$$

Having been able to predict the take-up S_{jzt}, water consumption Q_{jzt}, and revenue R_{jzt} from all the zip codes under the two possible service alternatives (only water and water and sewer), we can use the information about the zips to where the firm expanded and from where she did not expand to recover the cost parameters.

We recognize that assuming independence in the firm's decisions across different zip codes is a potentially oversimplified assumption. This assumption overlooks the possibility that the firm might venture into unprofitable zip codes initially, strategically expanding towards more profitable ones located at a greater distance. Additionally, our computation of distances to the installed network in 2017 has limitations, as it does not consider the scenario where the firm initially expands to a nearby zip code before gradually moving on to more distant ones. In such a case, we would anticipate the firm's expansion to occur incrementally, starting in close proximity to the existing network and progressing outward as neighboring areas are covered. However, our data, illustrated in Figure 10 in the appendix D, does not exhibit this pattern

Moreover, despite this simplification, our model accurately captures the observed patterns in the data, particularly in the areas of the North and Northeast regions of the country where significant opportunities for expansion exist. While the assumption may not fully capture the complexities of the firm's decision-making process, it provides a reasonable representation of the expansion in these regions.

The decision rule we model considers managers who would expand if the variable profit over the next five years exceeds the sunk costs. However, some managers might also consider a minimum return on investment, as discussed in Wollmann (2018). In this case, our model would overestimate the sunk costs. This is not a problem when simulating the firm's decision under various scenarios, as our estimated costs will already factor into this minimum return requirement. Therefore, the simulation will correctly predict that the firm will only expand if this requirement is met. Therefore, in our simulations, we will assume that the firm would be able to maintain the same return they obtained in the analyzed period.

The static nature of the model abstracts away from strategic interactions, which is not a problem here given that the firm is a monopolist in the areas where they have the concession to provide the services, so if they do not expand, there is no risk the demand being captured by another firm. It also implies that there is no value in waiting to expand in the future. If this is a factor that managers take into account, we would again overestimate the true sunk cost of the expansion.

5.2 Demand

Households (addresses) in each zip code have preferences for piped water and sewage services, which determine their decision to connect to the network and their demand for water. We use 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 either only water $j = w$, both water and sewage services $j = s$, or remain unconnected $j = o$. More specifically, if the zip code has water and sewage pipes available households can choose to connect to both services or to only water, while if the zip code has only water pipes the households can only choose to connect to water or remain unconnected.

The decision to connect is modeled using the following indirect utility function U_{ijzy} 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.

$$U_{ijzy} = V_{jzy} + \varepsilon_{ijcy} \tag{9}$$

$$V_{jzy} = \begin{cases} \alpha_{0jus} + \alpha_{1j} avgp_{juy} + \alpha_{2j} I_{cy} + \alpha'_{3j} D_{cy} & \text{if } j = w, s \\ 0 & \text{if } j = o \end{cases} \tag{10}$$

The utility function includes the average price $avgp_{juy}$ charged for the average consumption of households connected to service j and I_{cy} is the average income. D_{cy} is a vector of demographic characteristics that affect the decision to connect to the pipes, including the number of people per household, whether they are in an urban area, the share of addresses that are rental units in the census tract, and the share of households with access to other sources of water such as truck delivery, cistern or pits and the share with access to alternative sewage destinations such as septic tanks, chemical toilets, composting pits.

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

We assume that the idiosyncratic 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, $\varepsilon_{ijzy} = \zeta_{izgy} + (1-\sigma)\mu_{ijzy}$, where μ_{ijzy} is i.i.d. extreme value and ζ_{izgy} has a distribution such that ε_{izjy} is distributed extreme value. $\sigma \in [0,1)$ is the nesting parameter. As σ approaches 1, the within-group correlation of the utility goes to one, and only groups matter. So, the households care only about getting connected or not. As σ approaches 0, the within-group correlation goes to zero. The nesting structure allows for more flexible substitution patterns, in particular, we expect that when one type of service is not available households are more likely to connect to the other service option rather than choosing the outside option of not connecting to any network of pipes.

Under these assumptions, the probability of choosing product j is conditional on choosing to connect and the conditional probability of connecting to any service, i.e. choosing a product within group g are given by the following equations, respectively.

$$S_{jzy|g} = \frac{exp(V_{jzy}/(1-\sigma))}{\sum_{j \in J_g} exp(V_{jzy}/(1-\sigma))} \tag{11}$$

$$S_{zgy} = \frac{(\sum_{j \in J_g} exp(V_{jzy}/(1-\sigma)))^{(1-\sigma)}}{1 + (\sum_{j \in J_g} exp(V_{jzy}/(1-\sigma)))^{(1-\sigma)}} \tag{12}$$

Finally, the choice probability of product j which here represents the take-up when it is available at zip code z is given by the following multiplication.

$$S_{jzy} = S_{jzy|g} S_{zgy} \tag{13}$$

From here we obtain the share of consumers that connect S_{jzy} to service j when it is available which enters the monopolist decision on whether to expand to that zip.

2. Water consumption

We specify that the demand for water takes the form

$$ln(q_{ijzt}) = \beta_0 + \beta_1 ln(avgp_{ijut}) + \beta_2 ln(I_{cy}) + \beta_3' D_{cy} + \delta_j + \delta_{m\tau} + \eta_{ijzt} \tag{14}$$

In the equation, q_{ijzt} is the quantity 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. 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 the unobserved difference in the demand for water across households.

As discussed, we model households as responding to the average price $avgp_{ijut}$, which is computed based on the increasing block schedule of each concession. For the households that consume in the first bracket, they pay only a fixed fee. For the households with consumption in the upper brackets (b), they will pay the fixed feed plus the quantities that lie in the above brackets limits times their corresponding marginal prices mp_{bjut}. The price schedule is determined at the concession-state level and includes the fixed fee fee_{jut}, the limits of the brackets $\bar{q}_{bu}$, and the marginal prices mp_{bjut}.

$$avgp_{ijut} = \frac{fee_{jut} + \sum_{b=2}^{B} max(min(q_{ijzt} - \bar{q}_{ub-1}, \bar{q}_{ub} - \bar{q}_{ub-1}), 0)mp_{bjut}}{q_{ijzt}} \tag{15}$$

Here we should note that we only observe consumption for households located in zip codes with a network of pipes for at least water and that decided to connect.

We are modeling the discrete choice of service connection and the continuous choice of how much water to consume as two separate decisions. However, we also allow the choices to be correlated following Dubin and McFadden (1984); Davis and Kilian (2011); Barreca and Clay (2016) where they model the choice of appliances or energy sources and the energy usage as a joint discrete-continuous model. In our case, unobserved characteristics might affect the decision to connect and the water consumption. For example, a family that prefers to take more showers daily is likely to connect to piped water and consume more water relative to other families when connected. Since we only observe the consumption of households that decided to connect, there is a selection concern.

Following this joint discrete-continuous approach, we handle this endogeneity issue by assuming that the expected value of the idiosyncratic error η_{ijzt} in the continuous part of the demand is a linear function of ε_{ijzt} and compute selection correction terms for being connected to only water $\hat{S}_{wzy}\frac{ln(\hat{S}_{wzy})}{(1-\hat{S}_{wzy})+ln(\hat{S}_{szy})}$ or to water and sewer $\hat{S}_{szy}\frac{ln(\hat{S}_{szy})}{(1-\hat{S}_{szy})+ln(\hat{S}_{wzy})}$ that are accounted for in the estimation of the continuous demand for water.

6 Estimation

In this section, we present the estimation of consumers' preference parameters related to service take-up decisions and water demand. We then use these estimates to compute the predicted quantities and revenues the firm faces and estimate its costs of providing water and sewer and expanding the pipe network of these services.

6.1 Take up estimation

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).

In our case, the market is a zip code z and the shares are given by the take-up of each service j, so we estimate the following equation

$$ln(S_{jzy}) - ln(S_{ozy}) = \alpha_{0ju} + \alpha_{1j}avgp_{juy} + \alpha_{2j}I_{cy} + \alpha'_{3j}D_{cy} + \sigma ln(S_{jzy|g}) + \varepsilon_{jzy} \tag{16}$$

where S_{jzy} is the share of households connected to service j in zip code z and year y, S_{ozy} is the share of consumers that are not connected to the services, and $S_{jzy|g}$ is the share of households that chose service j conditioned on being connected to the network (group g). $avgp_{juy}$ is the average price charged for the service j for all the zip codes in concession-state u, I_{cy} is the average income at the census tract level, D_{cy} is a vector of demographic characteristics. σ is the nesting parameter representing the substitution of services that involve connecting to the network vs. not connecting. The parameter α_{1jus} represents product-concession-state fixed effects that capture different service installation costs across regions and services. While the other coefficients interacted with the service-specific characteristics.

Using this transformation, the model is estimated using a linear instrumental variable regression. Here, the conditional share $S_{jzy|g}$ is endogenous, as it may be correlated with unobserved demand shocks in ε_{jzy}. To address this, we use the share of zip codes with sewer in the same census tract as the target zip code z as an instrument for $S_{jzy|g}$. This instrument captures the fact that if there is an installed network of sewage pipes nearby, it is less costly for the firm to also have pipes in the target zip code, allowing households to choose the service. This instrument affects the choice among services within group g but is not correlated with demand shocks at the zip code level.

In our sample, we find that in 60% of zip codes with only water pipes, all addresses are connected to the service, and in 72% of the zips with water and sewer all households are connected. In such cases, the observed take-up of the specific product is one, and the share of the outside option is zero, which prevents us from taking logarithms as required by the model. We address this challenge following Li (2019) with 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.[18]. We define similar markets as the 100 zip codes that are closest in terms of income per capita and offer the same type of service (only water or water and sewer). More details about the method can be found in appendix E.1. Therefore, instead of using the observed take-up, we use the posterior estimates of take-up as the dependent variable in the nested logit estimation.

We estimate the nested logit model using data at the zip code-year level, specifically focusing on zip codes that had installed pipes in all three years of the available data. The estimated coefficients for each product are presented in Table 5. 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 the different products.

As expected, our estimated nesting parameter is close to 1, indicating a high within-group correlation. So households are more likely to substitute between only water and water and sewer when both are available than to the outside option of not connecting to the services.

[18]Having zero market shares is a common challenge in different markets. Li (2019) and Gandhi et al. (2023) discuss alternative methods of dealing with zero market shares

Table 5 – Connection Type Choice

	(1) Only Water	(2) Water and Sewage
Nesting Param	0.927***	0.927***
	(0.017)	(0.017)
Avg. Price	–0.164***	–0.492***
	(0.035)	(0.068)
Income (1000 R$)	0.030***	0.088***
	(0.009)	(0.010)
Urban	–0.032	0.320
	(0.206)	(0.278)
Household size	–0.581***	–0.633***
	(0.059)	(0.064)
Share rented	0.351**	1.173***
	(0.165)	(0.202)
Share Other Water	–0.234	–0.419*
	(0.165)	(0.244)
Share Other Sewer	–0.400***	–0.494***
	(0.074)	(0.084)
State-Unit-Choice FE	yes	yes
F-statistic	1,789	1,789
Observations	49,214	49,214

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.The coefficients on both columns were estimated from equation 16; they are presented separately to facilitate reading. Includes zip codes that had pipes installed since the beginning of our sample period and did not receive any expansion. The reported F-statistics corresponds to the first stage.

The other estimated coefficients also align with the expected directions, such as households are less likely to connect if the prices increase.

After obtaining the estimated coefficients, we compute the predicted take-up rates for the "only water" service, denoted as $\hat{S}_{wzy}$, and the take-up rates for the "water and sewer" service, denoted as $\hat{S}_{szy}$, for all zip codes in the sample. These predicted take-up rates will be used in the firm's decision-making process. Furthermore, the predicted take-up rates for zip codes with an installed network are utilized to compute the selection parameters that will be used to estimate the continuous demand for water.

6.2 Demand Estimation

Here, we estimate the demand for water from connected households using the following model

$$ln(q_{ijzt}) = \beta_0 + \beta_1 ln(avgp_{ijut}) + \beta_2 ln(I_{cy}) + \beta_3' D_{cy} + \delta_j + \delta_{m\tau} + \eta_{ijzt} \quad (17)$$

Where q_{ijzt} is the quantity of water consumed by a household (address) i at zip code z, connected to service j at month-year t. I_{cy} is the average income and D_{cy} is a vector of demographics at the census tract c and year y. δ_j is a fixed effect of being connected to service j. $\delta_{m\tau}$ is a municipality month of the year fixed effect, for example, municipality A in January, that accounts for weather and precipitation patterns that might influence water consumption across municipalities. η_{ijzt} represents the unobserved difference in the demand for water across households.

The average price faced by consumers is endogenous due to the nature of the increasing block rates, rather than correlated shocks between demand and supply as is usual in most demand estimation settings. Depending on the quantity consumed, households are charged different marginal prices, directly affecting the average price paid. This creates a simultaneity issue, leading to biased price coefficients if the model is estimated using ordinary least squares (OLS). Specifically, the coefficient would be downward biased if most households consume in the first price bracket because higher consumption leads to lower average prices. Conversely, the coefficient would be upward biased if most households consume above the first bracket because higher consumption leads to higher average prices.

We use a simulated instrumental variable for the average price to address this 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 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.

We report the results from the estimation of equation on table 6. The 2SLS (two-stage least squares) results using the simulated price IV are displayed in column (3). In column (1), we include as a benchmark the results when estimating the equation via OLS. Since most consumers consume in the first bracket, our proposed instrument corrects for the downward bias in the price elasticity coefficient compared to the OLS estimate.

Additionally, we explore the robustness of our average price instrument by comparing it to using observed marginal prices for each bracket as instruments, as discussed in Olmstead (2009). The results using this specification are presented in Column (2) of the table 6 and they yield similar estimates to our preferred instrument. The advantage of using instruments based on marginal prices in our setting is that consumers generally do not switch brackets, as we can see in figure 14 on appendix E, so by fixing some level of consumption, we can use the price schedule variation across time and concession-states to identify the price elasticity.

Table 6 – Continuous Demand Estimates

	(1)	(2)	(3)	(4)
	OLS	Mg Prices IV	Simulated IV	Simulated IV with selection
ln(AvgP)	–0.838***	–0.217***	–0.210***	–0.221***
	(0.003)	(0.007)	(0.012)	(0.013)
ln(Income)	0.171***	0.134***	0.133***	0.138***
	(0.003)	(0.004)	(0.004)	(0.004)
Piped Sewer	0.445***	0.083***	0.079***	0.091***
	(0.004)	(0.006)	(0.008)	(0.009)
Household size	0.045***	0.052***	0.052***	0.049***
	(0.005)	(0.005)	(0.005)	(0.005)
Urban	–0.128***	–0.167***	–0.167***	–0.156***
	(0.021)	(0.024)	(0.024)	(0.024)
Municipality-Month FE	yes	yes	yes	yes
Selection Correction	no	no	no	yes
F-statistic		9,879	7,038	6,500
Observations	6,685,016	6,685,016	6,685,016	6,685,016

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.Standard errors clustered at the address level. Includes only addresses that had water bills for all periods in the sample and did not switch the type of service (just water/ water and sewer). Column (4) displays our preferred specification.

Furthermore, we account for selection when estimating the IV regression. Since we only observe the consumption of households paying water bills and connected to pipes, there is potential selection bias. Households that consume more water may, for example, also be more likely to connect to the network. The distribution of the unobserved difference in demand η_{ijzt} may differ between connected and unconnected households. In this case, the coefficients estimated without accounting for selection will be biased and cannot be used to predict the consumption for households that may eventually get connected.

Following Barreca and Clay (2016); Davis and Kilian (2011); Dubin and McFadden (1984), we allow the discrete and continuous components of the demand to be correlated. More specifically, the expected value of η_{ijzt} is a linear function of ε_{jzy} and so we use the density of the extreme value distribution to evaluate the conditional expectation of η_{ijzt} analytically. We then compute selection controls using the predicted service take-up from the discrete connection choice and include it in the regression as a new variable associated with the type of service the household is connected to. The results of this specification are displayed in Column (4) of Table 6. The first stage results (table 10) and reduced form (table 11) are included in the appendix E. The price elasticity, accounting for selection, is comparable to the previous IV specifications, and the other coefficients are generally stable.

Our analysis focuses on the selection arising from consumers' decision to connect to the pipe network. We assume that the firm determines which zip codes to build pipes in based on observable factors of the demand and its costs. Implicitly, we assume that there are no demand shocks that the firm observes but are unknown to us as econometricians. However, if demand shocks increase the quantity consumed and make the firm more inclined to build pipes in specific areas, it would introduce another selection issue. In such a case, our estimates would overestimate the unconditional demand for water. Nevertheless, we believe this concern is not significant in our context. While the firm may have access to additional information from their technicians on the ground in each area, we have access to precisely the same administrative data they do. Although the firm's technicians may provide valuable insights, it is unlikely that they can consistently utilize this information across their extensive operations. Therefore, any potential bias introduced by unobserved demand shocks that are known to the firm but not to us is likely to be minimal.

In summary, the results indicate that households respond to increases in the average price by decreasing their water consumption, although the elasticity is small. Higher-income households consume more water, with a stable coefficient across IV specifications. Households connected to piped sewers consume more water than those connected to only water. Water consumption increases with household size, as expected. Additionally, households located in urban areas, on average, consume a lower volume of water compared to rural households.

To incorporate these patterns and predict the consumption of connected addresses we use the reduced form estimates, displayed in table 11 at appendix E.3. In particular, using the reduced form simplifies the problem because we can use the simulated average price to pin down one average price that the households respond to and then predict the consumption volume based on it. If we used the 2SLS results for the prediction, we would need to simultaneously pin down consumption and the average price, which, given the shape of the average price function in our setting, would generate two equilibrium consumption quantities - one in the first bracket and another in the above brackets - and, we would need to rely on had hoc rule to choose one.

Using the estimated model, we can compute q_{jzt}, which represents the consumption of a representative consumer in a given billing month, located in zip code z, and connected to service j. This computation assumes the mean income and number of people in the household at the zip code. We then aggregate this monthly consumption to obtain Q_{jzy} yearly and calculate the firm's corresponding revenue generated, denoted as R_{jzy}, based on the price schedule.

6.3 Cost Estimation

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. In order to calculate these values, we assume 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 these assumptions, we calculate the predicted take-up ($\hat{S}_{jzy}$), representing the expected number of households connecting to service j in zip code z. We also calculate the average household water consumption when connected to service j ($\hat{Q}_{jzy}$) and the average revenue obtained by the firm from a household connected ($\hat{R}_{jzy}$). These values are key components of the variable profits generated.

We then estimate the marginal cost associated with each service mc_j and the fixed cost parameters ω_{1j}, ω_{2j} via maximum likelihood using the predicted choice probability of connecting zip code z with service j (equation 8) and the observed expansion choices in 2018 and 2019. The key idea of the estimation is to find the cost parameter values that maximize the likelihood of observing the actual choices made by the firm, given the model's assumptions.

The estimation results are available in table 7. They indicate that the cost of supplying one cubic meter of water is roughly 6.75 Brazilian reais (R$). When including the collection of piped sewer with the same amount of water, the cost increases to about 9.98 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$9949, and for a kilometer of combined piped water and sewer, it is R$33073. These costs encompass not only the actual pipes but also all the materials and labor required for excavation and restoring the path after pipe installation.

The sunk cost per household in a zip code with only water pipes is around R$686, and in areas with both water and sewer, it's approximately R$1345. It is important to note that while the firm's responsibility extends only to the sidewalk level, the installation costs in a zip code are influenced by the number of households. This is because the pipe's capacity needs to accommodate consumers who choose to connect, and the number of households also serves as a proxy for the extent of the street covered by the provider's pipes.

We are the first to estimate the expansion costs of water and sewer pipes based on firm

Table 7 – Costs Estimates

	MLE
MC W (m^3)	-6.753***
	(0.377)
MC WS (m^3)	-9.975***
	(2.235)
Cost per distance W (km)	-9949.531*
	(5750.140)
Cost per distance WS (km)	-33073.097*
	(19420.501)
Cost per household W	-686.059**
	(330.084)
Cost per household WS	-1345.425**
	(589.206)
Number of Zips	708

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.All estimated costs are in Brazilian Reais (R$). In 2017, the exchange rate was about 3.3R$ to 1 U$. The estimation considers zip codes within the firm's concessions in the north and northeast regions of the country that had no service in 2017. Here we consider that the managers consider the profits they will make over the following 5 years when making decisions.

expansion decisions in Brazil. In contrast, engineering studies rely on cost measures from specific project accounting data. For example, von Sperling and Salazar (2013) examined 47 sewer collection projects in the state of Minas Gerais and found average costs ranging from R$121,691 to R$216,710 per kilometer of pipes[19]. Our cost estimates per unit of pipe length differ from these findings because we distinguish the costs associated with the length of pipes from those linked to the number of households served instead of dividing total costs by the expanded kilometer and we consider the expansion of both water and sewage together.

If we aggregate our cost estimates per km of pipes and household, making some rough

[19]The costs in the paper are originally reported in dollars per meter, so we converted them to reais using the exchange rate from April 10 to ensure comparability with our estimates.

assumptions[20], we find that the average cost of expanding water and sewer is approximately R$107,795 which is comparable in magnitude to the results in von Sperling and Salazar (2013). However, we might still be underestimating the costs considering that they look at only sewer and we are considering both water and sewer. In these case, our counterfactual estimates would represent a lower bound of the expansion costs.

In appendix F, we show alternative cost estimations considering that managers consider 10 or 30 years to recover the sunk cost of the investment; however, the results with 5 years are closer to the results of other studies and the order of magnitude reported by the firm.

7 Counterfactual Simulations

For the simulations, we take the connection targets of having 99% of the households connected to piped water and 90% with sewage as given and use the estimated model to analyze incentives to reach these targets. Meeting the percentage of households connected to piped water and sewer depends on the company's decisions regarding service expansion and consumer decisions to adopt these services. To disentangle these factors, we first simulate the firm expanding the services in all zip codes such that the share of connections will depend only on consumers' choices. We then let the firm endogenously choose where to expand and introduce price incentives for consumers to connect. In particular, we focus on sewer availability charges and discounts without extra government funding, so they are either paid by the provider or by consumers, as the main objective of the regulation was to promote expansion without relying on public investment.

To predict the outcomes, we make use of demand estimates to determine consumer take-up, water consumption, and the resulting revenue for the firm under the current prices and the different price incentives. By combining these estimates with the cost estimates, we calculate the variable profit the firm would generate and the fixed costs involved in the expansion. Additionally, we measure the changes in consumer surplus and infant deaths that arise from these policy changes.

[20]We aggregate our cost estimates per kilometers of pipes and per households considering that the zip codes with water and sewer expansion received on average 0.18 km of pipes and that on average each zip code that received expansion of water and sewer in our sample has 10 households to make them comparable to engineering estimates.

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, its 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 believe it is valuable to present this measure to understand its impact on consumers who can afford the service and may choose to connect when it becomes available.

We also compute the number of averted infant deaths in each simulation to capture consumer health benefits. We create this 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 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). For instance, the services might also reduce the incidence of water-borne diseases, such as diarrhea, which do not always result in child death (Barreto et al., 2007). Moreover, social externalities could influence the decisions of neighbors to adopt alternative methods for water sanitation and wastewater disposal (Deutschmann et al., 2022). Additionally, externalities could manifest as increased housing prices in neighborhoods that have the service (Coury et al., 2022).

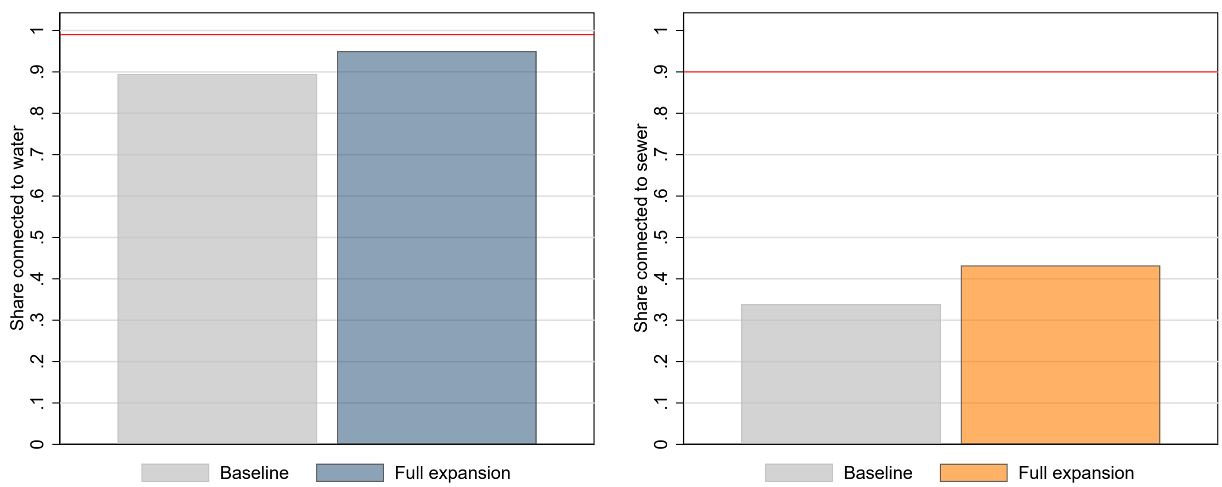
7.1 Simulations with full expansion

The first set of simulations focuses on scenarios where the company expands water and sewer services to all zip codes within its concessions in the country's northern region. By construction, in all these simulations, the hypothetical share of zips with the installed network water and sewer is one, and the share of zips with only water is zero. Therefore, the share of households connected to water and sewer will depend only on the consumer's take-up of these services, given that all households could connect.

In our first counterfactual, we show that the share of households connected would not achieve the mandated targets at the current pricing levels, even with an expansion of water and sewer services to all zip codes. The left panel of Figure 5a illustrates the mandated piped water coverage of 99% as the red line, the share of households connected to sewer pre-policy as the blue bar, and the share of households connected when we simulate the firm expanding to all zip codes in gray. Similarly, the right panel of Figure 5b presents the mandated 90% piped sewer coverage as the red line, with the orange and gray bars depicting households connected to piped sewer. Notably, only 43% of households would connect to the sewer network even though it is available to everyone under this simulation. Therefore, achieving the sewer mandate requires additional incentives for consumers.

Figure 5. Share of household connections

(a) Piped Water (b) Piped Sewer



Notes: The red lines correspond to the connection targets of the 2020 regulatory framework. The shares of connections is the ratio of the number of households connected to that service over the total number of households within the concession contracts. In the baseline, not all have water and sewer available in their zip code. In Simulation 1 all households have water and sewer available in their zip code, and the predicted take-up gives the share of connections.

Furthermore, extending services to all zip codes does not prove financially viable for the company, as the substantial sunk costs outweigh the increased variable profit from new connections. Column (2) of Table 8 presents the outcomes of this first simulation. By construction, this simulation ensures all zip codes possess water and sewer services, as shown in the first and second rows of the table. However, not all households would adopt these services, as indicated by the third and fourth rows of Table 8 and in the gray bars of figure 5. Differences in variable profit, consumer surplus, and infant mortality are calculated relative to the baseline, presented in column (1). Here, we present the variable profit and the consumer surplus over the course of five years to be consistent with our cost estimation.

As anticipated, the company's expansion encompasses zip codes with lower incomes than those initially possessing sewer services. The final row of Table 8 indicates that the mean income of covered zip codes declines by approximately 20% with full expansion. The full expansion generates a reduction of 6.54% in infant mortality among children below 1-year-old, amounting to roughly 14 fewer deaths when contrasted with the baseline scenario without expansion. This measure is computed using estimates from Gamper-Rabindran et al. (2010). The study emphasizes that increasing the proportion of the population with access to piped water is more efficient in reducing infant mortality than expanding sewer services. This distinction arises because households with treated water are less vulnerable to the effects of inadequate sewer systems.

Table 8 – Simulations with full expansion

	(1) Baseline	(2) Full Expansion	(3) Full Expansion Availability Charge	(4) Full Expansion Price Discount	(5) Full Expansion Availability Charge Price Discount
Share of Zips with Water and Sewer	0.44	1.00	1.00	1.00	1.00
Share of Zips with only Water	0.44	0.00	0.00	0.00	0.00
Share Connected to Water	0.90	0.95	0.95	0.96	0.95
Share Connected to Sewer	0.34	0.43	0.89	0.90	0.91
Δ Variable Profit (mi R$)		−25.31	295.22	−93.96	196.35
Sunk Cost (mi R$)		−148.79	−148.79	−148.79	−148.79
Δ Consumer Surplus (mi R$)		1.22	−2.40	3.75	−1.08
Δ% Infant Deaths		−6.54	−6.60	−8.46	−7.03
Mean Income Zips Sewer (R$)	4447.46	3531.33	3531.33	3531.33	3531.33
Expansion: firm choice	Yes	No	No	No	No
Expansion: all zips with sewer	No	Yes	Yes	Yes	Yes
Sewer Availability Charge	No	No	Yes	No	Yes
Water Price	Current	Current	Current	Current	Current
Sewer Price	Current	Current	Current	40% discount	10% discount

Notes: Full expansion means the firm installs water and sewer in all zips. The share of households connected in the simulations is given by the predicted take-up estimates. The differences in variable profit, consumer surplus, and infant deaths were computed with respect to the baseline. The sunk cost appears only when there is expansion. The variable profit and consumer surplus are computed over 5 years, consistently with our cost estimates. The difference in infant deaths is computed using the change in the share of households connected, and estimates from Gamper-Rabindran et al. (2010). There are 216 infant deaths in the baseline. In 2017, the exchange rate was about 3.3 R$ to 1 U$.

In our second simulation, we introduce the "Sewer Availability Charge", which entails charging consumers for sewer availability once the pipes are available in their zip codes, even if they are exclusively connected to water. The 2007 Regulatory Framework established this policy but has seen limited implementation, with only a handful of municipalities adopting this approach. Within our data set, only one municipality employs this pricing strategy, introduced midway through our time frame and thus excluded from preference and cost estimations. In the context of full expansion, this charge applies to all connected households, regardless of whether they choose to connect solely to water.

This policy substantially increases the proportion of households connected to sewer, as shown in column (3) of the table 8, reaching 89% of sewer coverage. However, the policy's costs are entirely borne by consumers, resulting in a reduction in consumer surplus of 2.4 million reais compared to the baseline scenario without expansion or availability charges. Under these circumstances, full expansion becomes profitable for the firm, as the notable increase in variable profits effectively offsets the fixed costs associated with expansion. While expansion may not yield profits in every zip code, the profitable ones outweigh any losses. As we are still considering a full expansion, the mean income of the zip codes with the service remains the same as in the previous simulations.

In the third simulation, we introduce a 40% discount on sewer prices in the context of full expansion, with the outcomes displayed in column (4) of table 8. For example, in a municipality where the sewer charge is initially 80% of the water price, this discount would reduce the sewer price to only 32% of the water price. The discount applies to all marginal prices. In Figure 15 in appendix G.1, we illustrate how different discount levels impact coverage. We focus on a 40% discount on sewer prices since it achieves the 90% coverage mandate. In this case, the firm would bear the entire cost of the policy. Although the discount would encourage more consumers to connect to the network, the variable profit decreases because the firm collects less revenue from all households already connected and the newly connected do not make up for this reduction. As expected, there's an increase in consumer surplus compared to previous simulations. Given that this policy achieves the highest household connection rates for water and sewer, it leads to the greatest reduction in infant deaths among the simulations.

Finally, we simulate a policy that combines charging consumers for sewer availability and giving a discount for sewer. The intuition of this policy is to reduce the burden faced by consumers by reducing the firm's profit while trying to maintain the share of households covered by the services. Column (5) of table 8 shows the outcomes when consumers are forced to pay for sewer, but the price has a discount of 10%. The price discount level is chosen so that the firm could still make profits, as shown in Figure 16 in the appendix G.1. With this combination, 95% of the households connect to water, and 91% connect to sewer, however, consumers still have a decrease in consumer surplus when compared to the situation with no expansion.

We explore alternative pricing strategies involving exclusive discounts for low-income groups coupled with an availability charge. In appendix G.1.1, We show that we can extend discounts of up to 70% to households in the bottom quartile of the income distribution while guaranteeing profits for the firm and the share of households connected to sewer above the target of 90%. We also discuss an alternative policy of providing subsidies to connect to the sewer. Our model is limited to addressing this hypothetical situation. Since we do not have data on the construction costs for households to connect themselves to the street-level pipes, these costs are incorporated by the state-concession-service fixed effect of the take-up decision. In appendix G.1.2, the subsidy is simulated as an income transfer when consumers connect to the sewer network. This policy would be inefficient, incurring transfers that cost more than the sunk cost of expanding the service while generating only a 10 percentage point increase in the share of households connected to sewer when compared to the baseline.

Although the simulations with full expansion are useful to disentangle the incentives, requiring the firm to expand to all zip codes might not be feasible. So, in our subsequent simulations, we delve into scenarios where the firm has the discretion to choose which zip codes to expand into, analyzing the outcomes of different policies within the context of consumers and the firm's decisions.

7.2 Simulations with endogenous expansion

In the second set of simulations, we provide incentives for consumers to connect and let the firm choose which zips to expand water and sewer. The firm's decision to expand is based on whether the expected profits from providing these services over the next 5 years exceed the sunk cost of expanding into those areas. We assume that the firm cannot stop offering services in areas where they had already installed the pipes by 2019. The baseline presented in column (1) of table 9 is the same as the one presented in table 8. The deltas computed in the simulations are relative to this baseline scenario.

The first two rows of table 9 illustrate the percentage of zip codes with access to both water and sewer services under the different scenarios, and they depend solely on the firm's expansion choices. Conversely, the percentages of households connected, as seen in the third and fourth rows of the table, result from a combination of the firm's expansion choices and consumer take-up. In this context, a household can only be connected if the service is available in their zip code and they opt to connect. It's important to clarify that the percentage of connected households is calculated as the ratio of households connected over the total number of households within the concession area, regardless of whether the service is available in their specific zip code.

We implement the "Sewer Availability Charge" policy in the first simulation and introduce endogenous firm expansion. The outcomes are presented in column (2) of table 9. Under this policy, the firm expands water and sewer services, achieving coverage in 74% of the zip codes with both services, while 14% have only water service. Concerning household connections, the share of households with piped water increases to 89%, and the share of households with sewage reaches 81%. The firm only expands into profitable zip codes, resulting in increased variable profits that more than offset the sunk cost of expansion. However, the costs of the policy are borne entirely by consumers, leading to a reduction in consumer surplus of R$3.48 million compared to the baseline. In this scenario, the share of households connected to water decreases as the policy essentially raises the monthly bills of households connected to only water in zip codes where sewer is available, leading some to disconnect. The change in infant mortality is negligible as the positive effects of sewer connections are offset by the negative effects of water disconnections.

In the second simulation, we provide a price discount of 40% for the sewer prices in the same way as the previous set of simulations but allow the firm to choose where to expand[21]. The outcomes are displayed in column (3) of table 9. There is no incentive for further expansions here, so the share of zip codes with water and sewer remains at 44%, and there is no sunk cost. However, the share of households connected to sewer increased from 34%

[21] A price discount could stimulate the expansion if the increased take-up more than compensated the decrease in revenue from the price drop.

Table 9 – Simulations with endogenous expansion

	(1) Baseline	(2) Endogenous Expansion Availability Charge	(3) Endogenous Expansion Price Discount	(4) Endogenous Expansion Availability Charge Price Discount
Share of Zips with Water and Sewer	0.44	0.74	0.44	0.56
Share of Zips with only Water	0.44	0.14	0.44	0.31
Share of Households with Water	0.90	0.89	0.90	0.90
Share of Households with Sewer	0.34	0.81	0.66	0.76
Δ Variable Profit (mi R$)		312.18	−44.35	202.62
Sunk Cost (mi R$)		−42.59	0.00	−22.08
Δ Consumer Surplus (mi R$)		−3.48	1.99	−2.08
Δ% Infant Deaths		0.00	−0.32	−0.08
Mean Income Zips Sewer (R$)	4447.46	3681.89	4447.46	3924.27
Expansion: firm choice	Yes	Yes	Yes	Yes
Expansion: all zips with sewer	No	No	No	No
Sewer Availability Charge	No	Yes	No	Yes
Water Price	Current	Current	Current	Current
Sewer Price	Current	Current	40% discount	10% discount

Notes: The share of households connected in the simulations is given by the predicted take-up estimates in the zip codes where the firm installed the services. The differences in variable profit, consumer surplus, and infant deaths were computed with respect to the baseline. The variable profit and consumer surplus are computed over 5 years, consistently with our cost estimates. The sunk cost appears only when there is expansion. The difference in infant deaths is computed using the change in the share of households connected, and estimates from Gamper-Rabindran et al. (2010). There are 216 infant deaths in the baseline. In 2017, the exchange rate was about 3.3R$ to 1 U$.

to 66% due to the increased take up in the zip codes that already had the service. It also increases consumer surplus and reduces infant deaths. In this case, the entire cost of the policy would lie on the firm as a reduction in their variable profit when compared with the situation without the discount.

Lastly, we simulate a policy that combines charging consumers for sewer availability and giving a discount for the sewer price while the firm can choose where to expand. Column (4) of table 9 displays the outcomes with a discount of 10% on the price of sewer. In appendix H.1, we show how the firm profits and the share of households connected to sewer would vary with different levels of price discount. In this case, the firm increases the share of zip code with sewer by 11 percentage points when compared to the baseline, and the share of households connected increases to 76%.

We also investigate alternative pricing strategies involving targeted discounts for low-income groups, an availability charge, and endogenous firm choice. In appendix H.1.1, we show that with a 10% discount restricted to households with incomes in the bottom quartile of the distribution, the firm would expand water and sewer to 65% of the zip codes, as opposed to 56% when the discount is given to everyone. However, giving higher discounts to this group disincentives the expansion and would translate into a lower share of households connected.

Additionally, we discuss providing subsidies to connect sewer pipes with endogenous firm expansion in appendix H.1.2. Again, the policy is expensive, and the firm does not expand to additional zip codes, reaching only 42% of the households with sewer.

7.3 Policy Implications

Among the two sets of simulations, no one-size-fits-all policy excels in every aspect. If the primary goal is to achieve the highest percentage of households connected to water and sewer while ensuring profitability for the firm, the best approach would involve mandating the firm to expand to all zip codes, implementing a sewer availability charge, and offering a 10% discount on sewer prices. However, if mandating full expansion is not feasible, the second best alternative would be implementing the sewer availability tariff without additional price discounts.

Nonetheless, it is important to acknowledge that since the feasibility of any policy hinges on firm profitability, these policies place a heavier financial burden on consumers. Under the scenarios we explored, there is an opportunity for government intervention to provide increased price discounts for consumers. This intervention could further boost connections and enhance consumer surplus, all while maintaining the viability of the projects for the firms. To delve deeper into this alternative, it would be necessary to estimate the potential savings the government could achieve in public health and other expenses by increasing the percentage of households connected to piped water and sewer. However, this falls beyond the scope of this paper and would require a separate analysis.

8 Conclusion

This paper assesses policies for improving connections to water and sanitation services through private providers. The lack of piped water and sewage collection is still a pressing issue in numerous developing countries, including Brazil. We analyze the impact of the 2020 New Sanitation Regulatory Framework, which encourages private sector involvement to achieve universal coverage by 2033. Using a structural model, we examine household decisions on service connection and consumption and firm decisions on expansion. Our analysis is based on novel billing data from a major private provider in Brazil.

Our policy analysis reveals several key findings. Firstly, even with universal piped sewer access, only approximately 43% of households would opt for this service. Secondly, if the firm expands everywhere any of the following policies would attain the 90% sewer service coverage goal: charging for sewer upon availability, offering a 40% sewer price discount, or combining sewer availability charges with a 10% discount on sewer prices. However, these policies differ regarding who bears the burden. When we consider the endogenous expansion of firms, which mirrors market dynamics more accurately, achieving the water and sewer coverage targets set by the New Regulatory Framework using the previously described policies becomes unfeasible. In this case, the availability charge achieves 81% sewer connection rate but places the entire financial burden on consumers.

Our simulation results highlight the challenges in addressing wastewater collection deficiencies. While Brazil outperforms many other developing nations in ensuring safe drinking water, it falls short in providing well-managed sanitation services, as indicated by WHO (2023). The primary obstacles include a lack of incentives and difficulty perceiving and internalizing the societal benefits of connecting to the piped sewage network. Our results support policymakers' consideration of charging for sewer services based on availability rather than upon actual connection. This approach has shown promise as a means to move closer to achieving near-universal coverage. Combining availability-based charges with price discounts may not significantly alter access rates but can address equity concerns and help to compensate for the costs borne by consumers. However, to ensure profitability for private providers without additional government incentives, the financial burden would inevitably fall on consumers. .

This study is dedicated to examining the effectiveness of policies currently under consideration by policymakers and suggesting potential enhancements to achieve their connection targets. Further research is warranted to explore other alternative policies, including significant revisions to pricing structures, utility auctions, and innovative solutions to infrastructure development. Moreover, future research could also explore the effects of these policies on water quality which, along with access, is an important aspect of 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.

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.

Coury, M., T. Kitagawa, A. Shertzer, and M. Turner (2022). The value of piped water and sewers: Evidence from 19th century chicago. Technical report, National Bureau of Economic Research.

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 S. J. Zhu (2022). Spillovers without social interactions in urban sanitation. *Available at SSRN 3790865*.

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.

Galiani, S., P. Gertler, and E. Schargrodsky (2005). Water for life: The impact of the privatization of water services on child mortality. *Journal of political economy 113*(1), 83–120.

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.

Granja, J. (2021). Regulation and service provision in dynamic oligopoly: Evidence from mobile telecommunications.

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.

Holmes, T. J. (2011). The diffusion of wal-mart and economies of density. *Econometrica 79*(1), 253–302.

Ito, K. (2013). How do consumers respond to nonlinear pricing? evidence from household water demand. *Stanford Institute for Economic Policy Research. Retrieved from http://people. bu. edu/ito/Ito_Water_Irvine. pdf*.

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.

Jia, P. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica 76*(6), 1263–1316.

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. (2021). The Price of Power: Costs of Political Corruption in Indian Electricity.

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.

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

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

TrataBrasil (2015). Ociosidade das redes de esgoto–2015.

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.

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

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

Wichman, C. J., L. O. Taylor, and R. H. von Haefen (2016). Conservation policies: Who responds to price and who responds to prescription? *Journal of Environmental Economics and Management 79*, 114–134.

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

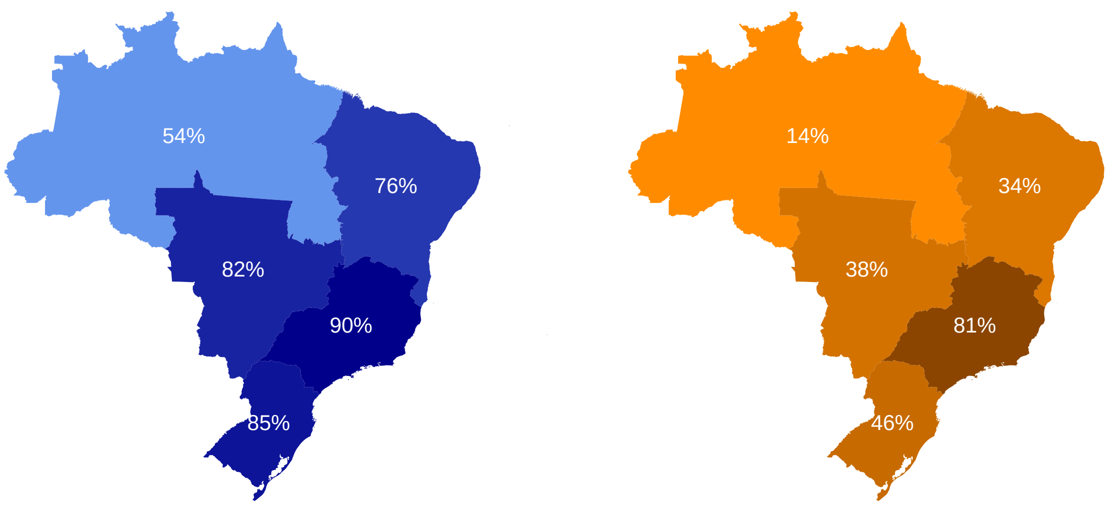
9 Appendix

A Introduction figures

Figure 6 shows the share of people connected to piped water and sewer collection in the country's regions in 2017.

FIGURE 6. SHARE OF THE POPULATION CONNECTED BY REGION IN BRAZIL (2017)

(a) Piped Water (b) Piped Sewer



Notes: Data from national sanitation survey - Pesquisa Nacional de Saneamento Básico (PNSB) from IBGE.

B Data

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 certain 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.

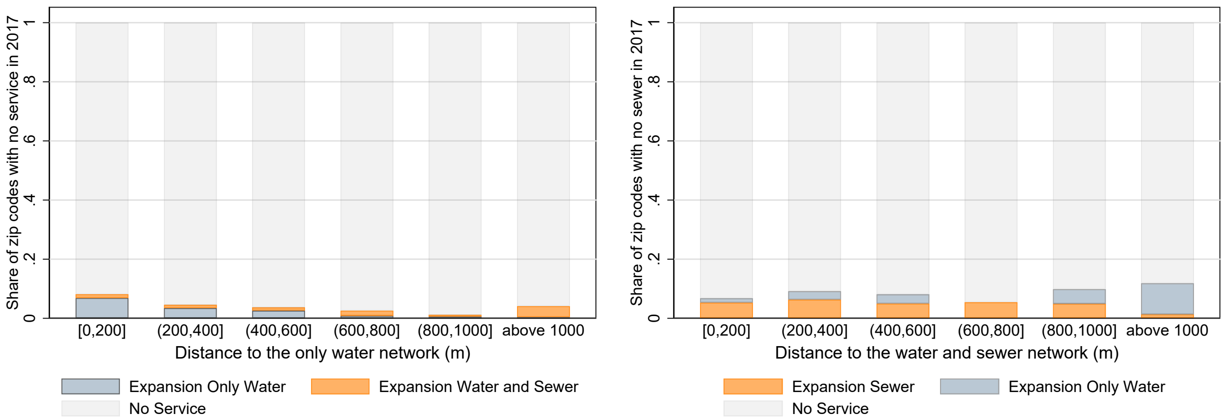
C Descriptive

C.1 Firm expansion

As shown in Figure 7a, water-only expansions were more prevalent in zip codes near the already established water pipes. Conversely, Figure 7b suggests that sewer expansions, whether combined with water expansion or not, occurred in areas where water and sewer infrastructure were already present. In both cases, there were minimal to no water or sewer expansions carried out beyond a distance of 1km from the respective network.

Figure 7. Distances to the installed network

(a) Only water (b) Water and sewer



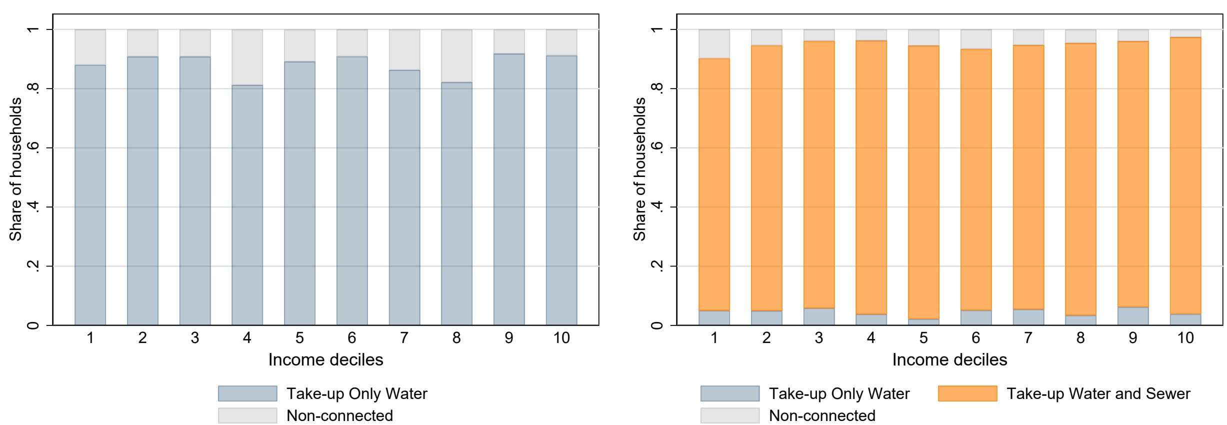
Notes: These graphs include all zip codes within the firm concession areas that did not have any service in 2017.

C.2 Service take-up

In figure 8, we show the take-up of the services in zip codes by income percentiles. The income percentiles are computed considering all zip codes within the company's concession areas. In figure 8a we show the take-up for water in zip codes with only water pipes but no sewer. In figure 8b, we show the take-up for only water and for water and sewer together in zip codes where pipes for both services are available.

Figure 8. Zip codes with installed network

(a) Only Water (b) Water and Sewer

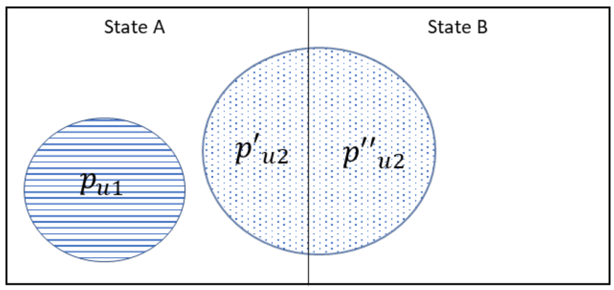


Notes: The income deciles are defined considering all zip codes within the firm's concession areas. The graph on the left side includes zip codes with only water pipes. The graph on the right side includes zip codes with water and sewer pipes.

C.3 Consumers reacting to average price

Here we illustrate the two concessions (u_1 and u_2) that allow us to isolate the effects of changes in the fixed fees from other confounding factors. In figure 9 we show that the municipalities in concession u_1 are located in two different states (States A and B), each facing different price schedules. On the other hand, all municipalities in concession u_2 are situated in State B.

FIGURE 9. ILLUSTRATION: CONCESSIONS



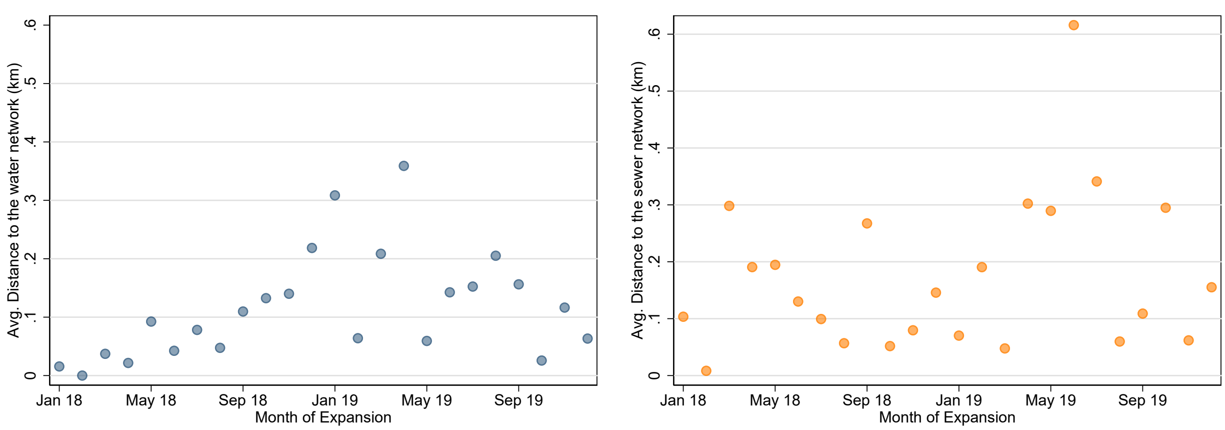
Notes: Illustrative figure of the variation used to identify consumers' reaction to changes in the fixed fee. Given that we cannot display any real maps with the areas where the firm is located, we illustrate two concessions u_1 and u_2 located in states A and B. In particular u_2 faces different price schedules depending on the state the municipalities are located in.

D Expansion model

In figure 10 we show the average distance to the installed network of the zip codes that received the service in each month of our sample. In figure 10a we show the zip codes that received only water and their average distance to the closer zip code that had water in 2017. In figure 10b we show the zip codes that received both water and sewer and their average distance to the closer zip code that had both services in 2017.

Figure 10. Distance of new zip codes per month

(a) Only water (b) Water and sewer



Notes: The graph on the left side includes zip codes that received expansion of only water pipes and the graph on the right side includes zip codes that received expansion of water and sewer pipes.

E Estimation

E.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_{zj}, 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_{zj} for each service in each zip are drawn from a Beta prior distribution with parameters λ_{1zj} and λ_{2zj}. Such that $K_{zj} \sim Binomial(N_z, S^0_{zj})$ and $S^0_{zj} \sim Beta(\lambda_{1zj}, \lambda_{2zj})$. The posterior distribution of the take-up is also a Beta distribution

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

with posterior mean

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

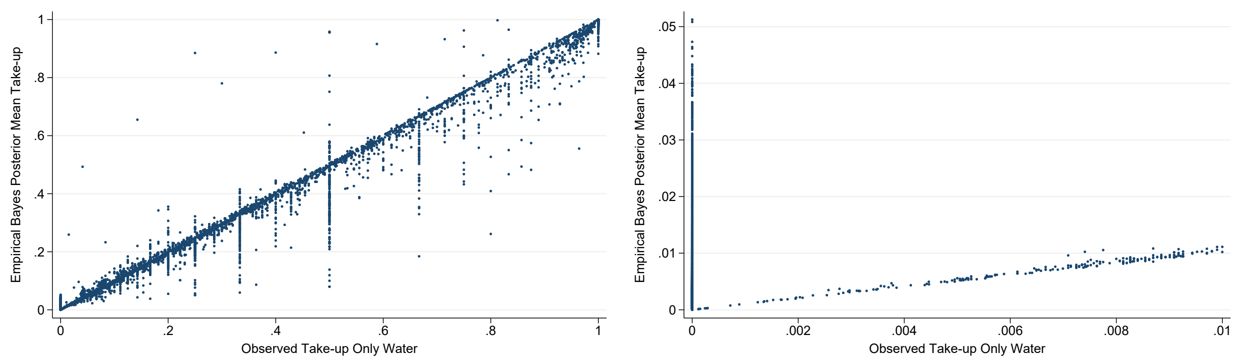
For each zip code z and service j the Beta prior is formed sing the 100 closest in income per capita that also have pipes for j, $l \in \zeta_z$, where l is a zip code from the set of similar zip codes ζ_z. The parameters of the beta prior distribution λ_{1zj} and λ_{2zj} are estimated from maximizing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

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

With the estimated parameters, we construct the posterior mean of the take-up probabilities for each zip and service $\hat{S}^P_{zj} = \frac{\hat{\lambda}_{1zj} + K_{zj}}{N_z + \hat{\lambda}_{1zj} + \hat{\lambda}_{2zj}}$, 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 11. Take-up only water: empirical Bayes Posterior vs. observed

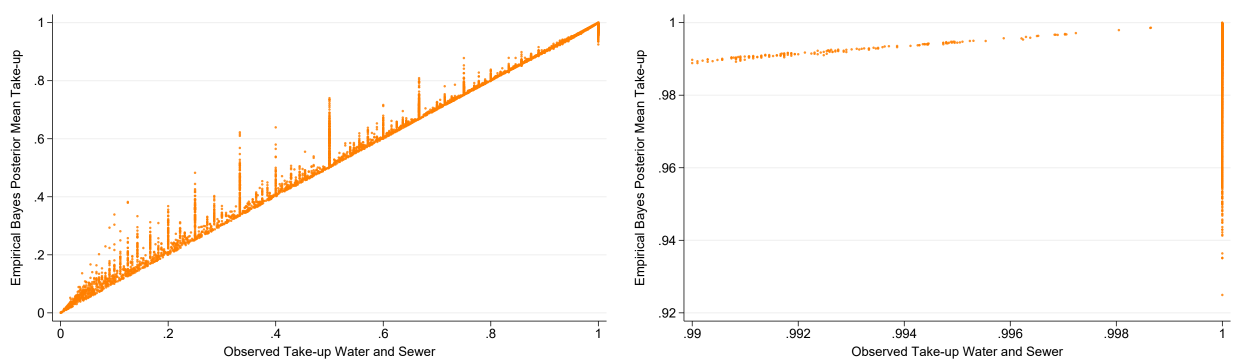
(a) All take-ups (b) Zooming in on observed take-ups below 0.1



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

Figure 12. Take-up water and sewer: Empirical Bayes Posterior vs. Observed

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

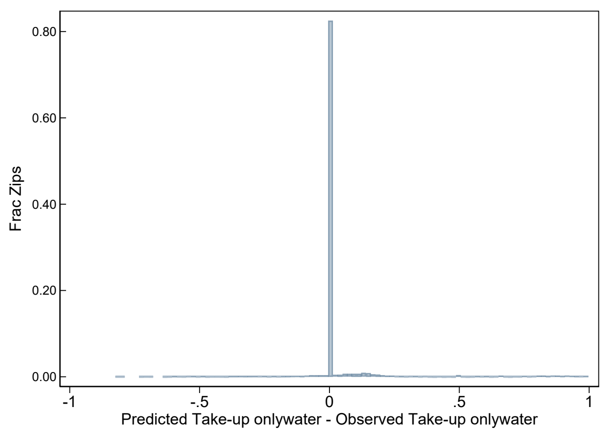


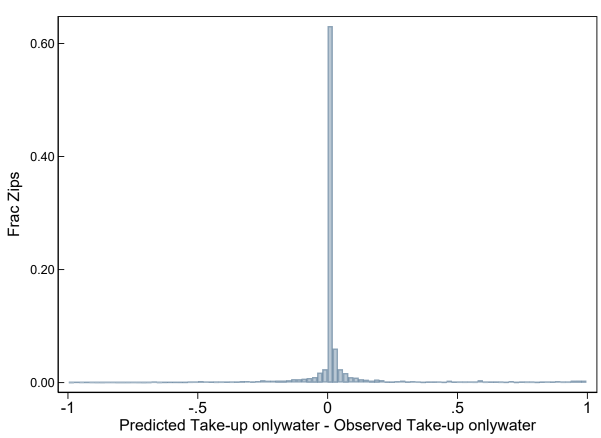
Notes: These graphs show the empirical Bayes Posterior Mean for each observed take-up of water and sewer services.

E.2 Fit take-up

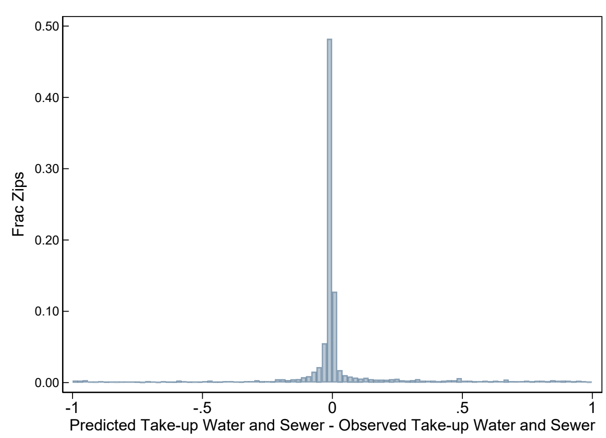
Here we plot the differences between the predicted take-up and the observed take-ups in the data. Figure 13 shows the fit of our estimates for the take-up of water services in zips where only water pipes are available. Figures 14a and 14b show, respectively, the fit of our estimates for the take-up of only water and for the take-up of water and sewer services in zips where both services are available.

Figure 13. Predicted vs. Observed Service Take-Up





(a) Only Water when there is sewer

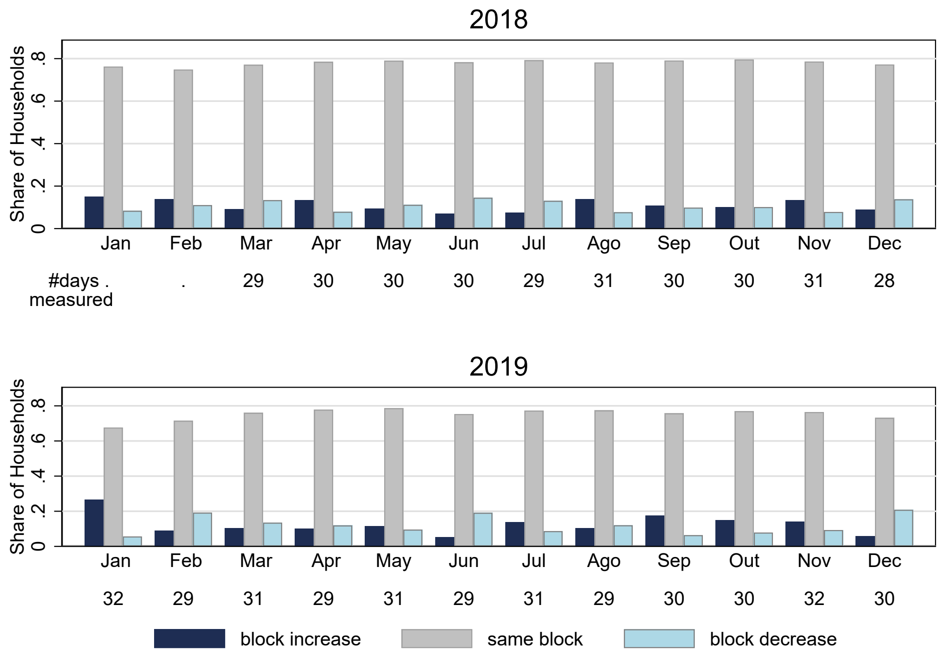


(b) Water and Sewer when there is sewer

E.3 Continuous Demand for Water

Figure 14 shows that most households do not change brackets across billing periods, and that block increases are associated with longer billing periods.

Figure 14. Bracket Change



Notes: The figure includes households already connected in 2017 and had water bills for all months of 2017 and 2018. It shows that most households do not change brackets from one billing month to the next.

Tables 10 and 11 show, respectively, the first stage and the reduced form from the continuous demand estimation with the simulated instrument for the average price and controlling for selection.

Table 10 – First Stage

	(1) First Stage Simulated IV with Selection
$\ln(AvgP_{sim})$	0.745***
	(0.009)
ln(Income)	0.044***
	(0.002)
Piped Sewer	0.140***
	(0.007)
Household size	−0.021***
	(0.003)
Urban	0.216***
	(0.015)
Municipality-Month FE	yes
Selection Correction	yes
F-statistic	6,500
Observations	6,685,016

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.Standard errors clustered at the address level. Includes only addresses with water bills for all periods in the sample and did not switch the type of service (just water/ water and sewer). The simulated average price is constructed by dividing the households into 16 groups based on income quartile, service type (only water or water and sewer), and urban or rural area, calculating the average consumption of households in the same group but located in different concessions, then computing the average price that the household would pay under the price schedule of their own concession and time if they consumed the average consumption of households in the same group but located in other concessions.

Table 11 – Reduced Form

	(1) Reduced Form Simulated IV with Selection
$\ln(AvgP_{sim})$	–0.165***
	(0.010)
ln(Income)	0.128***
	(0.004)
Piped Sewer	0.060***
	(0.008)
Household size	0.054***
	(0.005)
Urban	–0.203***
	(0.026)
Municipality-Month FE	yes
Selection Correction	yes
F-statistic	259
Observations	6,685,016

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.Standard errors clustered at the address level. Includes only addresses with water bills for all periods in the sample and did not switch the type of service (just water/ water and sewer). The simulated average price is constructed by dividing the households into 16 groups based on income quartile, service type (only water or water and sewer), and urban or rural area, calculating the average consumption of households in the same group but located in different concessions, then computing the average price that the household would pay under the price schedule of their own concession and time if they consumed the average consumption of households in the same group but located in other concessions.

F Costs

Here, we reestimate the costs assuming that managers of the firms consider the variable profits they would obtain over the course of the next 10 years or 30 from the decision to expand to each zip code, instead of 5 years as shown in the main text.

Table 12 – Cost Estimation: 10 years

MgC W (m^3)	−7.384***
	(0.399)
MgC WS (m^3)	−10.143***
	(2.669)
Cost per Distance W (km)	−17187.052*
	(9719.319)
Cost per Distance WS (km)	−56472.521*
	(33912.699)
Cost per Pop W	−1136.857*
	(601.323)
Cost per Pop WS	−2281.849**
	(1077.026)
Number of Zips	708

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.All estimated costs are in Brazilian Reais (R$). In 2017, the exchange rate was about 3.3R$ to 1 U$. The estimation considers zip codes within the firm's concessions in the north and northeast regions of the country that had no service in 2017. Here we consider that the managers consider the profits they will make over the following 10 years when making decisions.

Table 13 – Cost Estimation: 30 years

MgC W (m^3)	−9.232***
	(0.384)
MgC WS (m^3)	−10.326***
	(2.960)
Cost per Distance W (km)	−30165.897***
	(7426.441)
Cost per Distance WS (km)	−88680.417**
	(42074.368)
Cost per Pop W	−1437.771*
	(780.295)
Cost per Pop WS	−3189.050**
	(1466.231)
Number of Zips	708

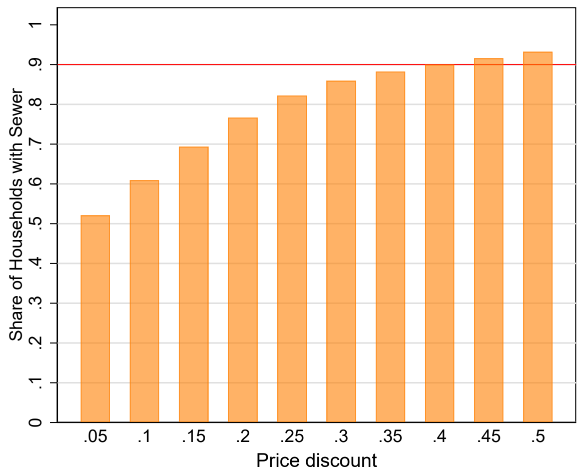
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.All estimated costs are in Brazilian Reais (R$). In 2017, the exchange rate was about 3.3R$ to 1 U$. The estimation considers zip codes within the firm's concessions in the north and northeast regions of the country that had no service in 2017. Here we consider that the managers consider the profits they will make over the following 30 years when making decisions.

G Simulations with full expansion

G.1 Price discounts

In figure 15, we show how the share of households connected to sewer varies with different levels of price discounts, considering the case where the firm has pipes for water and sewer in all zip codes within its concession. In the main text, we further explore the outcomes of setting the discount at 10%.

FIGURE 15. SEWER CONNECTIONS FULL EXPANSION AND PRICE DISCOUNTS



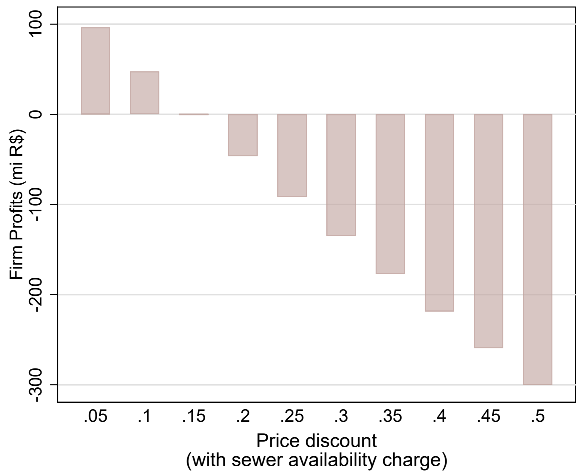
Notes: The graph displays the simulated share of sewer connections for different levels of price discount in the sewer prices in the case where the firm expands water and sewer to all zip codes.

In figure 16 we show the firm's profits, considering that it expanded to all zips with water and sewer, consumers are charged for sewer availability and different discount levels. In particular, we further discuss in the paper the case with a 10% discount since the expansion will only be viable if the provider is able to make profits with the installed network.

1. Vertical price differentiation with full expansion and availability charge

In this simulation, we explore alternative pricing strategies involving exclusive discounts for low-income groups, coupled with an availability charge for sewer in the context of full expansion. In particular, we introduce different levels of price discount for sewer to households in zip codes with average incomes within the lowest quartile of the income distribution within the areas under the firm concessions. Figure 17a shows that for any level of discount,

Figure 16. Firm profits with full expansion, sewer availability tariff, and price discounts



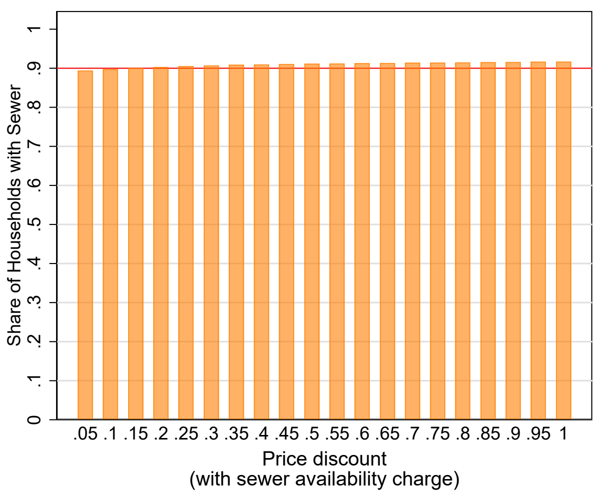
Notes: The graph displays the simulated profits of the firm for different levels of price discount in the sewer prices when the firm expands water and sewer to all zip codes and consumers are charged for sewer availability.

the overall share of households connected to sewer consistently exceeds the targeted threshold of 90%. While figure 17b shows 17b that with discounts below 70%, the full expansion with availability charge is profitable for the provider.

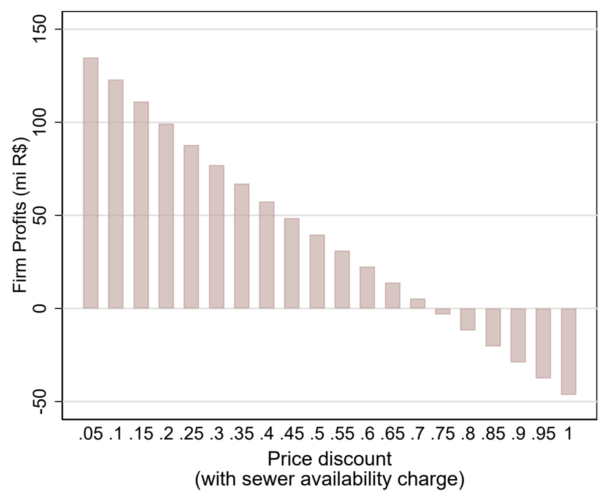
Table 14 shows the main outcomes comparing our baseline with the case where the firm expands to all zip codes with water and sewer, consumers are charged for sewer availability, and low-income consumers receive a discount of 70% on the sewer pricing.

Figure 17. Price discounts to low-income consumers with availability charge

(a) Share connected to sewer



(b) Provider profits



Notes: Here, all households pay sewer availability charges, but consumers in zip codes where the mean income is in the first quartile receive discounts for sewer. The left graph displays the share of households connected to sewer and the right graph displays the provider's profits for varying levels of discount to low-income households.

Table 14 – Simulations with full expansion, availability change, and price discounts to low-income

	(1) Baseline	(2) Full Expansion Availability Charge Price Discount to low income
Share of Zips with Water and Sewer	0.44	1.00
Share of Zips with only Water	0.44	0.00
Share Connected to Water	0.90	0.96
Share Connected to Sewer	0.34	0.91
Δ Variable Profit (mi R$)		154.12
Sunk Cost (mi R$)		−148.79
Δ Consumer Surplus (mi R$)		−0.50
Δ% Infant Deaths		−8.12
Mean Income Zips Sewer (R$)	4447.46	3531.33
Expansion: firm choice	Yes	No
Expansion: all zips with sewer	No	Yes
Sewer Availability Charge	No	Yes
Water Price	Current	Current
Sewer Price	Current	70% discount (low income)

Notes: On column (2) all households pay sewer availability charges, but consumers in zip codes where the mean income is in the first quartile receive discounts for sewer. Full expansion means that the firm installs water and sewer in all zips. The predicted take-up estimates give the share of households connected in the simulations. The differences in variable profit, consumer surplus, and infant deaths were computed with respect to the baseline. The sunk cost appears only when there is expansion. The variable profit and consumer surplus are computed over 5 years, consistently with our cost estimates. The difference in infant deaths is computed using the change in the share of households connected, and estimates from Gamper-Rabindran et al. (2010). There are 216 infant deaths in the baseline. In 2017, the exchange rate was about 3.3 R$ to 1 U$.

2. Subsidies to consumers with full expansion

In this simulation, we introduce subsidies for consumers to facilitate their connection to sewer pipes while assuming a full expansion by the firm. As previously discussed, consumers are responsible for connecting their residences to the street-level pipes once they become available, and the associated construction costs range between R$1000 and R$8000. Specifically, we offer an R$6000 subsidy to households in the lowest income quartile located in zip codes either lacking sewer service prior to expansion or exhibiting less than 90% adoption rates. The results are displayed in column (1) of Table 15. This specific scenario is presented due to the coverage ceasing to grow beyond a subsidy level of R$6000. Additional variations of these simulations are explored in Figure (18), exploring different subsidy amounts applicable to zip codes with varying income levels.

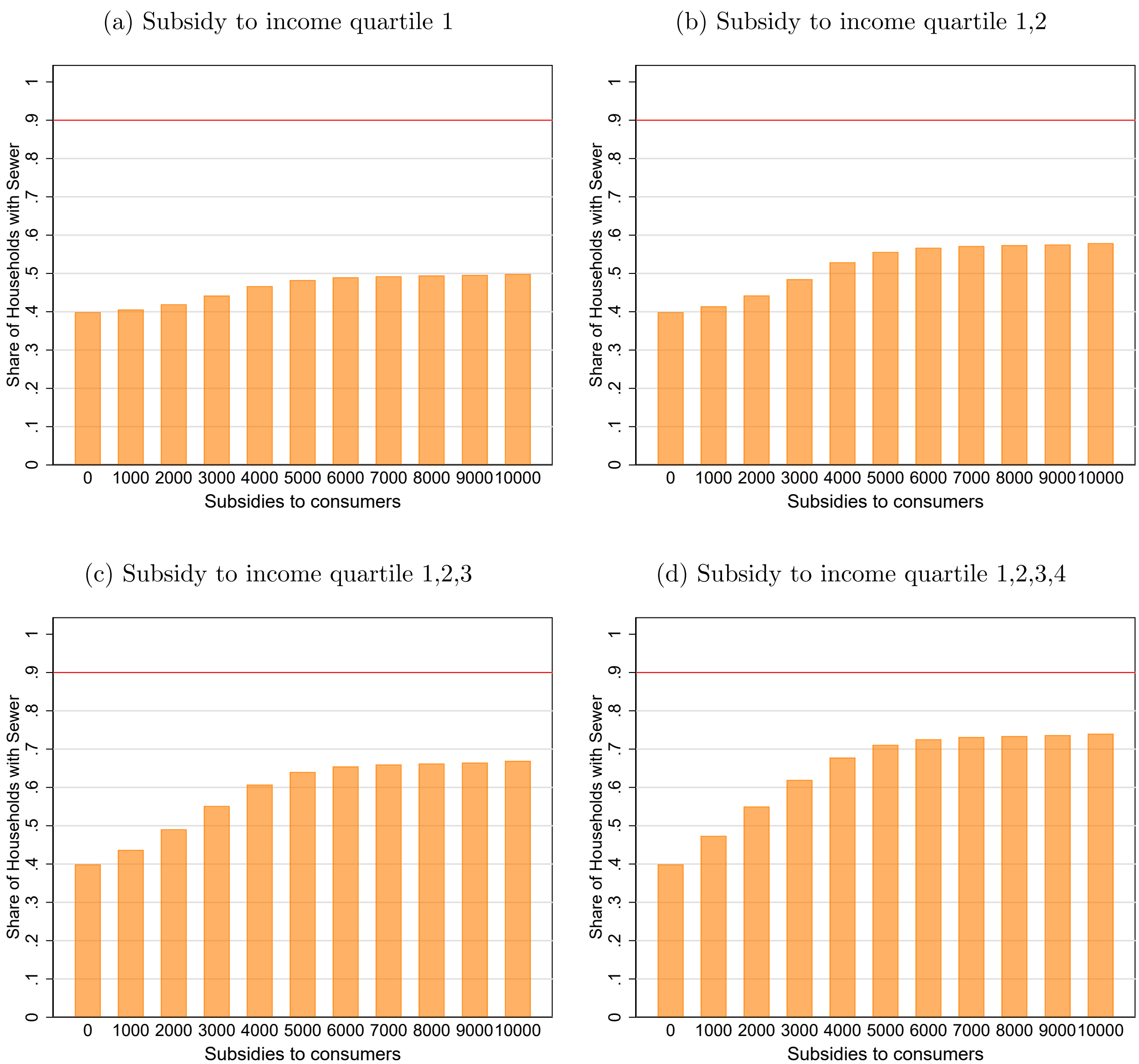
The subsidy is an expensive policy, costing more than the total sunk cost incurred by the firm in the expansion. Despite the large cost, sewer coverage only reaches 52%, still falling well short of the mandated 90% target. Given the substantial expense, the subsidy burden would likely fall on the government, as imposing additional costs on a firm already facing unprofitable expansion would not be feasible. A more focused strategy could offer subsidies only to unconnected households, reducing the costs of the policy, however, our take-up model has a limitation that any subsidy can only be offered to all households within a zip code.

Table 15 – Simulations with full expansion

	(1) Baseline	(2) Full Expansion Connection Subsidy
Share of zips with Water and Sewer	0.44	1.00
Share of Zips with only Water	0.44	0.00
Share Connected to Water	0.90	0.96
Share Connected to Sewer	0.34	0.52
Δ Variable Profit (mi R$)		−22.39
Sunk Cost (mi R$)		−148.79
Δ Consumer Surplus (mi R$)		1.38
$\text{cost}_s ubs_m i$		218.62
Δ% Infant Deaths		−7.67
Mean Income Zips Sewer (R$)	4447.46	3531.33
Expansion: firm choice	Yes	Yes
Expansion: all zips with sewer	No	No
Sewer Availability Tariff	No	No
Water Price	Current	Current
Sewer Price	Current	Current
Connection Subsidy	R$0	R$6000

Notes: Full expansion means the firm installs water and sewer in all zips. In column (2) all households receive a subsidy of R$6000 if they connect to sewer. The predicted take-up estimates give the share of households connected in the simulations. The differences in variable profit, consumer surplus, and infant deaths were computed with respect to the baseline. The sunk cost appears only when there is expansion. The variable profit and consumer surplus are computed over 5 years, consistently with our cost estimates. The difference in infant deaths is computed using the change in the share of households connected, and estimates from Gamper-Rabindran et al. (2010). There are 216 infant deaths in the baseline. In 2017, the exchange rate was about 3.3 R$ to 1 U$.

Figure 18. Sewer Subsidy by Income Quartile: With full expansion



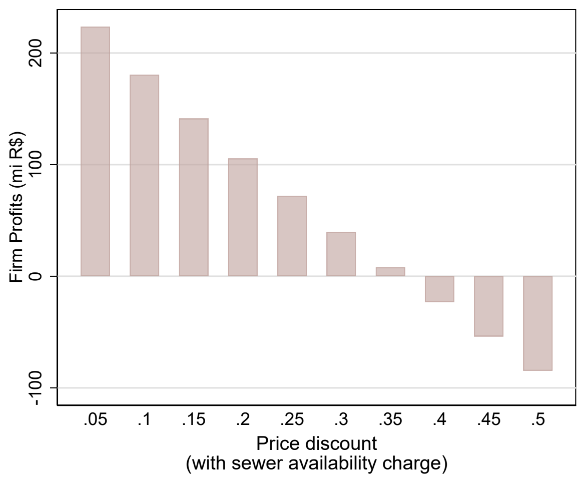
Notes: The graphs display the share of households connected to sewage for varying subsidy levels. In panel (a), only households in zips with income in the first quartile of the income distribution receive the subsidy. In panels (b), (c) we increase the number of households receiving the subsidies. In panel (d) all households receive the subsidy.

H Endogenous Expansion

H.1 Price discounts

In figure 19 we show the firm's profits with different levels of discount, considering that consumers are charged for sewer availability and the firm chose where to expand. In the main text, we further explore the outcomes of setting the discount at 10%.

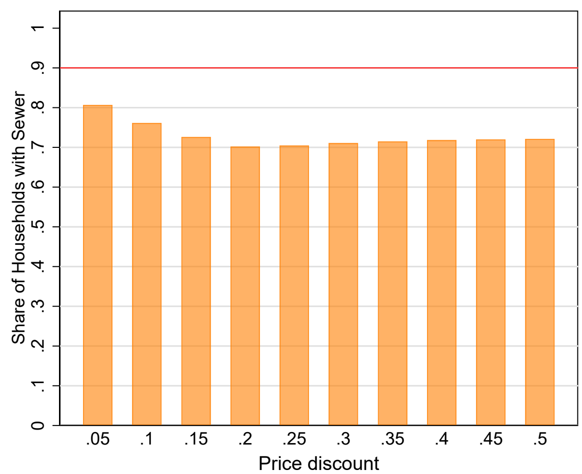
Figure 19. Firm profits with availability tariff, price discount With endogenous expansion



Notes: The graph displays the simulated profits of the firm for different levels of price discount in the sewer prices when the firm chooses where to expand and consumers are charged for sewer availability.

In figure 20 we show how the share of households connected to sewer varies with different price discounts, given that they are charged for sewer availability and the firm chooses where to expand. Here it is noteworthy that for certain levels of discount, the firm does not expand.

Figure 20. Sewer connections with price discount, sewer availability tariff and endogenous expansion

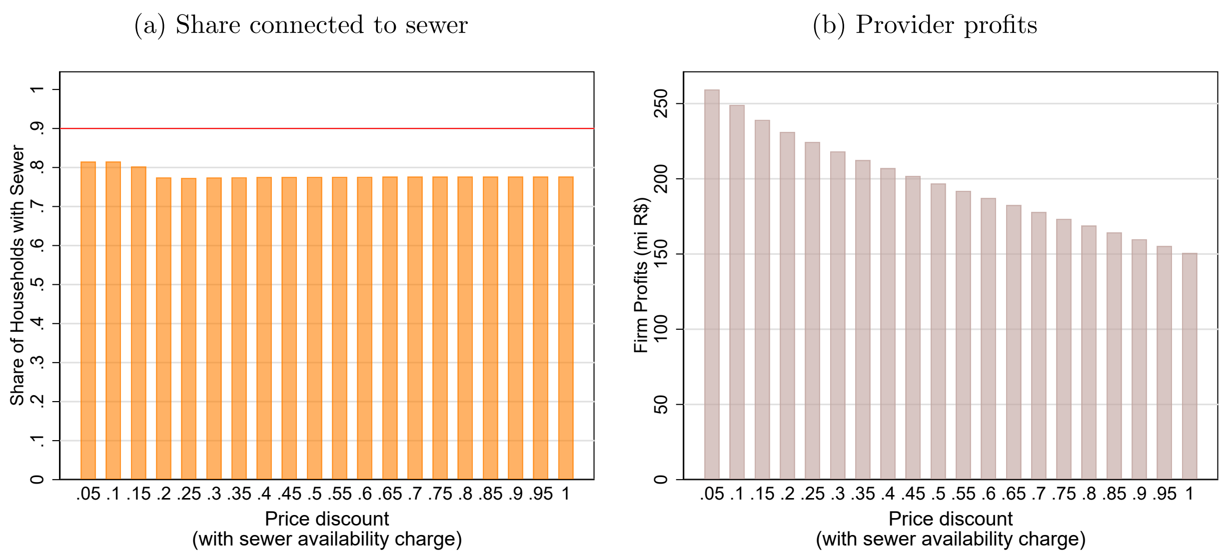


Notes: The graph displays the simulated share of households connected to sewer for different levels of price discount in the sewer prices when the firm chooses where to expand and consumers are charged for sewer availability.

1. Vertical price differentiation with endogenous expansion and availability charge

In this simulation, we investigate alternative pricing strategies that involve targeted discounts for low-income groups, alongside an availability charge and endogenous firm choice. As previously described, we introduce different levels of price discount for sewer to households in zip codes with average incomes within the lowest quartile of the income distribution within the areas under the firm concessions. Figure 21a shows that introducing any discount exceeding 10% for low-income households would lead to a decrease in the overall percentage of households connected to sewer. Figure 21b shows that with any level of discounts for low-income consumers and availability charges the firm can still make positive profits, however, it reduces the expansion to cover only profitable zip codes.

FIGURE 21. PRICE DISCOUNTS TO LOW-INCOME CONSUMERS WITH AVAILABILITY CHARGE



Notes: Here, all households in zip codes where sewer pipes are installed pay sewer availability charges, but consumers in zip codes where the mean income is in the first quartile receive discounts for sewer. The firm endogenously chooses which zip codes to expand.

Table 16 shows the main outcomes comparing our baseline with cases where consumers are charged for sewer availability and low-income consumers receive a discount of 10% or 100% on the sewer pricing. We highlight that the discounts affect the firm expansion decision; in particular, with a 10% discount to low-income households, the firm would cover 69% of the zip codes with both water and sewer, while with a 100% discount, this share reduces to 65%.

Table 16 – Simulations with full expansion, availability change, and price discounts to low-income

	(1) Baseline	(2) Endogenous Expansion Availability Charge Price Discount to Low Income	(3) Endogenous Expansion Availability Charge Price Discount to Low Income
Share of Zips with Water and Sewer	0.44	0.69	0.65
Share of Zips with only Water	0.44	0.18	0.22
Share Connected to Water	0.90	0.89	0.90
Share Connected to Sewer	0.34	0.80	0.78
Δ Variable Profit (mi R$)		275.57	178.43
Sunk Cost (mi R$)		−36.36	−27.76
Δ Consumer Surplus (mi R$)		−3.11	−1.52
Δ% Infant Deaths		−0.01	−0.05
Mean Income Zips Sewer (R$)	4447.46	3820.54	3974.78
Expansion: firm choice	Yes	Yes	Yes
Expansion: all zips with sewer	No	No	No
Sewer Availability Tariff	No	Yes	Yes
Water Price	Current	Current	Current
Sewer Price	Current	15% discount (low income)	100% discount (low income)

Notes: Endogenous expansion means the firm installs water and sewer in zip codes where it is profitable to do so. In columns (2) and (3) all households in zip codes where sewer pipes are installed pay sewer availability charges, but consumers in zip codes where the mean income is in the first quartile receive discounts for sewer. The differences in variable profit, consumer surplus, and infant deaths were computed with respect to the baseline. The sunk cost appears only when there is expansion. The variable profit and consumer surplus are computed over 5 years, consistently with our cost estimates. The difference in infant deaths is computed using the change in the share of households connected, and estimates from Gamper-Rabindran et al. (2010). There are 216 infant deaths in the baseline. In 2017, the exchange rate was about 3.3 R$ to 1 U$.

2. Subsidies to consumers with endogenous expansion

We provide the same subsidy to consumers to connect to the sewer as previously discussed, but now the firm chooses its expansion. The outcomes are presented in column (2) of table 17. The subsidy consists of R$6000 to households in the lowest income quartile located in zip codes either lacking sewer service before expansion or exhibiting less than 90% adoption rates. In this case, the firm does not expand water and sewer in any zip code, as it can be seen that the share of zips with water and sewer and only water remains the same in the first two rows of the baseline and the first simulation. However, the share of households connected to sewer increases to 42%, increasing the consumer surplus and the variable profit of the firm. Again, we find that the policy is extremely costly, and the share of holds connected to the sewer is far from the targeted mandate.

Table 17 – Simulations with endogenous expansion

	(1) Baseline	(2) Full Expansion Connection Subsidy
Share of Zips with Water and Sewer	0.44	0.44
Share of Zips with only Water	0.44	0.44
Share of Households with Water	0.90	0.90
Share of Households with Sewer	0.34	0.42
Δ Variable Profit (mi R$)		7.46
Sunk Cost (mi R$)		0.00
Δ Consumer Surplus (mi R$)		0.06
Cost Subsidy (mi R$)		163.77
Δ% Infant Deaths		−0.03
Mean Income Zips Sewer (R$)	4447.46	4447.46
Expansion: firm choice	Yes	Yes
Expansion: all zips with sewer	No	No
Sewer Availability Tariff	No	No
Water Price	Current	Current
Sewer Price	Current	Current
Connection Subsidy	R$0	R$6000

Notes: Endogenous expansion means the firm installs water and sewer in zip codes where it is profitable to do so. In column (2) all households receive a subsidy of R$6000 if they connect to sewer. The share of households connected in the simulations is given by the predicted take-up estimates in the zip codes where the firm installed the services. The differences in variable profit, consumer surplus, and infant deaths were computed with respect to the baseline. The sunk cost appears only when there is expansion. The variable profit and consumer surplus are computed over 5 years, consistently with our cost estimates. The difference in infant deaths is computed using the change in the share of households connected, and estimates from Gamper-Rabindran et al. (2010). There are 216 infant deaths in the baseline. In 2017, the exchange rate was about 3.3 R$ to 1 U$.