Monopoly Regulation When Consumers Need to Make Sunk Investments: Theory and Evidence from the Swedish District Heating Sector

Darryl Biggar a, Matthieu Glachant b and Magnus Söderberg b,1

a Australian Competition and Consumer Commission (ACCC) and Australian Energy Regulator (AER)2

b CERNA, Mines ParisTech, 60 Boulevard St Michel, 75006 Paris, France

26 February 2016

ABSTRACT

Regulatory agencies routinely seek to promote price stability. Such practices have no clear rationale under the neoclassical approach to natural monopoly regulation. An alternative view is that regulation exists to protect customers’ relationship-specific sunk investments. In a theory model we show that when consumers need to make sunk investments, welfare increases when a regulator provides assurance of future rigid prices. We use data from the Swedish district heating sector during the 1998-2007 period to explore the impact of monopoly pricing decisions on the take-up rate of district heating and to identify the gap with a counterfactual scenario featuring price regulation.

Keywords: Forward-looking consumers, pricing behaviour, regulation, district heating, Sweden

JEL Classifications: D42, L51, L52, L97

1 Corresponding author: IFEL B5, School of Business Economics and Law, University of Gothenburg, PO Box 610, S-405 30 Gothenburg, SWEDEN. magnus.soderberg@gu.se

2 The views expressed in this paper are the views of the authors and do not necessarily reflect the views of the ACCC or the AER.
1. INTRODUCTION

This paper explores an issue at the foundations of natural monopoly regulation. According to neoclassical theory, the primary economic objective for natural monopoly regulation is to reduce or eliminate deadweight loss (e.g. Crew and Kleindorfer, 2006). However, many economists have observed that this approach leads to policy recommendations (such as the embrace of Ramsey pricing, peak load pricing, or other forms of price discrimination) that regulators tend to eschew in practice. In addition, regulators frequently pursue policies, such as price stability, or incremental cost-based pricing, that have no clear rationale in traditional neoclassical theory.3

As an alternative, it has been suggested that a key rationale of natural monopoly regulation is to protect and promote sunk complementary investments made by potential and existing customers of the regulated firm (Goldberg, 1976; Biggar, 2009). According to this view, a distinguishing feature of natural monopoly industries is the need for customers to make sunk, relationship-specific investments in order to extract the most value from the monopoly service.4 These could be investments in specialised customer-premises equipment, human capital, or investment in a particular location, such as investments by an aluminium smelter relying on local electricity supply, household investment in gas appliances relying on the local gas supply, or investment in a coal mine close to a major rail spur. Once these investments are sunk, they are subject to the risk that the monopoly will increase its prices and expropriate the value of the investments. When this hold-up problem cannot be solved through private vertical integration or long-term contracting arrangements, regulation may play a role. According to this view, a key rationale for natural monopoly regulation is to protect and thereby promote customers’ sunk investments to ensure that they take up and get the most value out of the monopoly service.

To date, no empirical test has been carried out to determine whether there is economic substance in this alternative justification for price control regulation. But what would such a test look like? The key hypothesis is that price regulation, by protecting customers from price increases, promotes customers’ sunk investments and therefore promotes take-up of the monopoly service. Perhaps the ideal test would be a randomised controlled trial in which customers must invest to take up the monopoly

---

3 Crew and Kleindorfer (2006) suggest that regulators should emphasise “…some degree of price stability” (p. 72), but they do not elaborate on why that would be advantageous.

4 The need to make sunk investments to extract the most value from a transaction commonly arises in competitive markets. However, in the case of monopoly markets, these sunk investments are also inevitably relationship-specific. The monopoly problem concerns protecting and thereby promoting these sunk investments.
service and where service providers are randomly subject to price regulation. We could then observe whether there is a material difference in the take-up rate of the monopoly service in regulated versus unregulated markets. If a material difference was found, we could infer that a key objective of price regulation, through controlling the path of prices, is to promote the take-up of the monopoly service.

Such a test is not feasible. However, the Swedish district heating market has a number of desirable features which come close. In Sweden, district heating is provided by a large number of small, independent monopoly firms and, during the period of our study, prices were unregulated. In addition, their networks have expanded considerably since the end of the 1990s. This stands in contrast to most monopolistic network sectors in the developed world, where customers have already made the most of the required sunk investments, and where prices are subject to price control regulation by an independent regulatory agency. Furthermore, customers who wish to make use of district heating must purchase and install customer-premises equipment, which can cost more than ten times the annual consumption expenditure on heating. In Sweden, this has given rise to concerns that customers who have made this investment are locked into the district heating service (EI, 2007; Henning, 2006) and that, fearing future increases in prices, they might be reluctant to make the necessary investments – potentially choosing environmentally or economically inferior heating alternatives (SOU 2004, p. 15; EI, 2007, p.41).

In this paper, we use the natural variation in pricing policies of these utilities to explore the impact of pricing decisions on the take-up rate of district heating and to identify the gap with a counterfactual scenario featuring price regulation. The analysis involves three steps. We first develop a model featuring a private monopoly providing a district heating service to consumers. The objective is to go deeper into the motivation of the research question by characterizing monopoly pricing behaviour, the hold-up problem, and the potential role of regulation as a commitment mechanism to increase consumer incentives to connect to the network. We show that the possibility to commit increases both consumer surplus and firms’ profits relative to a scenario without commitment, and that the monopoly then chooses to commit to rigid prices. This commitment is however not credible without third party enforcement.

---

5 As a result there have been calls for price controls by the Swedish Competition Authority and the Swedish Energy Markets Inspectorate (SCA, 2009; EI, 2007). The Energy Markets Inspectorate, in arguing the need for regulation, explicitly emphasises that district heating customers are in a “weak position” with respect to their suppliers and that regulation would “build long-term confidence in district heating as a product” (EI, 2007, p. 66). This may have created a regulatory threat established during our study period and that partially disciplined district heating firms. We come back to this when interpreting the results in Section 6.
These findings decisively hinge upon the assumptions that consumers are forward-looking and risk averse when adopting district heating. In the second step, using data from the Swedish district heating sector from 1998-2007, we estimate the fixed-effects demand model developed by Berry (1994) to test these assumptions. The dependent variable is the ratio between residential property owners that invest in district heating and those that are in a position to invest in a new heating system in each local market. Assuming that customers form expectations about future prices based on recent past price outcomes, we construct two variables that measure (i) the slope of the long-term price trend and (ii) variations around this long-term trend. In line with the theory, we find that the rate of connections to the district heating network decreases when these variables increase. Since all of the price variables must be considered endogenous, we collect the share of fossil fuel, which is strongly related to the cost of providing district heating, directly from the firms and use that as an instrument.

The econometric estimation thus confirms that consumer investments are influenced by future price levels and volatility. In the last stage, we use the estimates to simulate a counterfactual scenario with an optimal price regulation. In this scenario, the price is set to the unit production cost averaged over time so that the price is held constant. We find that the market share is 11.3%, which is remarkably close to 9.3%, the predicted market share under the business-as-usual (unregulated) scenario. A look at descriptive statistics suggests that this limited gap is driven by both low mark-ups and a low inter-annual price volatility. One possible reason for this pricing behaviour is the threat of price regulation during the study period. Since the market price deregulation of 1996, regulation has been a subject of continuous political debate. Public dissatisfaction has brought the market to the brink of regulation several times. The height of the debate was the passing of the District Heating Act (DHA) in 2008, which created a formal procedure through which consumers could complain to a national committee about the prices they were charged.

This paper primarily contributes to the economic literature on natural monopoly regulation pioneered by Goldberg (1976), which stresses the problem of incomplete contracting between buyers and suppliers. A recent example is the work by Chisari and Kessides (2009), who examine the case of developing countries where natural monopoly sectors with low coverage ratios face significant network expansion opportunities. This literature is essentially theoretical and we propose an empirical study to assess the importance of the concern. Our results lend support to the common regulatory practice of promoting price stability – a practice which has no clear rationale under the neoclassical approach to public utility regulation. They support the view of natural monopoly regulation as a form of long-term contract seeking to protect and promote the sunk relationship-specific investments of the monopoly service provider and its customers. This has important implications for the analysis of regulatory policy and the design of regulatory institutions.
The paper is related to several studies which have found that when there is scope for inter-temporal substitution of purchases, past purchases, together with expectations about future price paths, have an effect on present demand. This literature has focused on both the case where goods are storable and consumers can hold inventories (Erdem, Keane and Imai, 2003; Sun, Neslin and Srinivasan, 2003; Hendel and Nevo, 2004, 2006a, 2006b; and Su, 2010), and the case where goods are durable (Chah, Ramey and Starr, 1995; Nair, 2007; Chevalier and Goolsbee, 2009; and Gowrisankaran and Rysman, 2009). In the next section we review the more specific literature on consumer behaviour in Sweden with respect to investments in district heating, and show that it too strongly supports our fundamental assumption that consumers are forward-looking.

Our analysis could also be related to the broader literature on ‘price stickiness’ and ‘price rigidity’. An example is the study of consumer habit formation by Nakamura and Steinsson (2011). Nakamura and Steinsson emphasise that if the supplier of a habit-forming product cannot commit to a price path, customers face a time-inconsistency problem – the decision to consume in the first period increases their subsequent demand, which the supplying firm may exploit with a higher price in subsequent periods.

The paper continues with a description of the Swedish district heating sector. In Section 3 we lay out our modelling approach. Section 4 describes the data. Section 5 contains the empirical evaluation where the two key assumptions are tested. Section 6 provides some simulation results where different pricing scenarios are evaluated. Section 7 concludes.

2. DISTRICT HEATING IN SWEDEN

Swedish district heating firms are vertically integrated local monopolies that produce heat in a heat centre and distribute it to customers’ properties through a network of underground pipelines carrying hot water or steam. At the customer’s property, a heat exchanger extracts heat energy and the cooler water is returned to the heat centre to be re-heated and re-distributed. District heating is only economical in densely populated areas, which is why networks rarely cross municipality borders. In 2007, district heating met approximately 50% (or 47 TWh) of the total heat demand in Sweden. In the same year, it was the most common form of heating for multi-dwelling houses in 234 of the 290 Swedish municipalities (market share around 75%), whereas in detached dwellings (i.e. one- and two-}

---

6 We use the words ‘firm’ and ‘municipality’ synonymously in this paper.
family dwellings) it had a market share of around 10% (SCA, 2009; SCB, 2009; SEA, 2010). The only widespread substitutes for district heating in densely populated areas are electricity-based technologies (e.g. radiant/convection/fan heating and air/ground heat pumps). Wood-fired and ground-based technologies are mostly used in rural areas as they are emission- and land-intensive. Cities have a shortage of ground space, which is required for drilling, and wood fire causes unacceptable levels of smoke exhaustion. In our empirical study we arbitrarily consider that consumers have four heating alternatives: district heating, electricity, remaining technologies (including renewables and various technology combinations), and the outside option, which is when the consumer opts out from investing in a fixed device (e.g. passive heating that only relies on heat generated by humans and portable devices).

Apart from differences in price level and structure, the Swedish heating market is subject to both horizontal and vertical differentiation. For example, district heating is considered more environmentally friendly than electricity (around 40% of electricity is generated from nuclear power plants) and district heating has higher service reliability than electricity, although it also requires more indoor space.

Heating equipment installed in residential dwellings has an expected lifetime of 20 years (Hepbasil and Kalinci, in press; Shah et al., 2008) and, to access a particular technology, consumers have to make a large upfront investment that can amount to ten times the annual consumption cost. In principle, consumers can switch to other technologies before the end of their device’s lifetime, but qualitative and anecdotal evidence suggests that they tend not to. Rydén et al. (2013) conduct interviews with representatives from district heating firms, and ascertain that only a negligible fraction of customers have switched prematurely from district heating to other technologies. In fact, many district heating firms report that no single customer has switched from district heating in recent decades. The fact that consumers do not revise their investment decisions has important implications for our modelling strategy that we explain in Section 3.

---

7 This section reports statistics from a number of sources. When not explicitly stated, data can be accessed from either ‘Nils Holgersson’s annual price comparisons’ (www.nilsholgersson.nu), Statistics Sweden (www.scb.se), the Energy Markets Inspectorate (www.ei.se), or the Swedish District Heating Association (www.svenskfjarrvarme.se).

8 E.g. Värnamo Energi, a mid-sized district heating firm in the centre of Sweden, states on its homepage that no customer has switched from its DH service since it was established in 1984 (www.varnamoenergi.se, visited on the 12th February 2016).
The price of all heating technologies consists of three components. Apart from the installation cost, consumers pay a fixed annual price and a variable price per kWh of heat. District heating tariffs were exempted from sector-specific price regulation in 1996, on the basis that electricity is a competing source of energy for heating purposes. Prices have risen steadily since 1996. ‘Nils Holgersson’s annual price comparisons’, which has published prices for all public utility services in Swedish municipalities since 1996, report a real average increase of the list price (total consumption price for a representative consumer) of approximately 12% over the ten years from 1998 to 2007. Detailed plant-level statistics collected by Statistics Sweden confirm this increase in the average consumption price. This rise can be compared with the regulated electricity distribution price, which only increased by 1% in real terms during the same ten-year period. Customers and media have expressed concerns that the price rise of district heating is driven by an increasing mass of locked-in customers. The number of district heating customers at the national level went up from 149,000 in 1998 (SCB, 2001) to 289,000 in 2007 (SCB, 2009) and the average network expansion increased from 4-5 km of lines per annum at the end of the 1990s to 7-10 km in 2006-07. Hence, despite price increases, there is no sign that district heating firms are reaching a slowdown in the demand for connections. Firms could have reduced the hold-up problem by subsidising the cost of purchasing a heat exchange device for new customers, but the lack of competition in densely populated areas may be the reason why district heating firms have not done so in practice.

The scant literature on consumer heating choices in Sweden supports the assumptions made in Section 4, i.e. that consumers form expectations about future prices based on past prices, that consumers are risk averse, and that demand for heating is inelastic in the short run. For example, Mahapatra and Gustavsson (2007) found that the annual consumption price was the single most important factor to influence property owners when they chose a heating technology. Based on interviews with representatives from Swedish district heating firms, Jörgensen (2009) concludes that consumers form expectations about future prices based on past prices. Moreover, Forsaeus Nilsson et al. (2008) and Isaksson (2005) found that consumers are negatively affected by uncertain contractual conditions when making decisions about the choice of heating technology. Lastly, there is ample empirical evidence to support the assumption that the short-run elasticity of space heating is very low in Scandinavia. Brännlund et al. (2007) estimate the short-run price elasticity of heating to be -0.05 in Sweden based on quarterly national data for the period 1980-1997. Leth-Petersen and Togeby (2001)"

---

9 Electricity distribution tariffs are regulated by the Swedish Energy Markets Inspectorate, which is required to take into account customers’ interest in low and stable electricity prices (SOU, 1995). See Jamasb and Söderberg (2010) for further details on the Swedish electricity market and how it is regulated.

10 These reports are available online (www.nilsholgersson.nu). This means that past list prices are available to all consumers considering connecting to the network.
report a price elasticity based on annual panel data close to zero for residential customers in Denmark. This study is interesting because it is restricted to existing customers (i.e. it excludes entrance of new customers) and it therefore shows that the price sensitivity for heating is very low in the short run.

3. CONCEPTUAL FRAMEWORK

In Section 3.1, we develop a simple illustrative model featuring a private monopoly providing a district heating service. In Section 3.2 we use the theoretical framework and expand the demand side to derive an estimable model.

3.1 A model of monopoly pricing

In a first version of the model, we introduce simplifying assumptions that allow us to concentrate the theoretical analysis on the interplay between consumer choice and monopoly pricing behaviour. A representative consumer\textsuperscript{11} chooses between district heating and another heating technology at time $t$. If she opts for district heating, she first needs to make an investment $I_t$ to connect to the network and, once connected, she makes a binary consumption decision, i.e. she either decides to purchase a fixed amount of heating service, or to switch to the outside heating option.\textsuperscript{12} The lifetime of the investment is $K$.

For ease of presentation, consumer utility of the outside heating option is normalized to zero in this first version of the model. The indirect utility for the heating service in period $t + k$ of a consumer who is connected is given by

$$u_{t+k} \equiv u(v_{t+k} - p_{t+k})$$

where $v_{t+k}$ captures the value of the heating service and $p_{t+k}$ is its price. The function $u$ is a utility function that exhibits standard properties: $u(0) = 0$, $u' > 0$ and $u'' < 0$. The concavity of $u$ implies that the consumer is risk averse. Thus, price volatility will reduce her propensity to invest.

\textsuperscript{11} Assuming a representative consumer is sufficient to characterize the hold-up problem and its welfare consequences. However, in a more refined version of the model that is introduced in Section 3.2, we assume that consumers are heterogeneous.

\textsuperscript{12} This assumption is realistic. As shown in Section 2, the short-run price elasticity for heating is close to 0.
District heating is provided by a single profit-maximizing firm\textsuperscript{13} at cost $\theta_{t+k}$. $\theta_{t+k}$ is the realization in period $t + k$ of stochastic variable $\theta$ which is i.i.d. distributed in the interval $[\underline{\theta}, \overline{\theta}]$. We assume that $\nu_{t+k} \geq \overline{\theta}$, i.e. the gain from trade, is positive at any cost level.\textsuperscript{14} It follows that, once the consumer is connected, district heating Pareto dominates the alternative heating option. Accordingly, the socially optimal price schedule should maximize the incentives to connect. Under our assumptions, the highest incentives are obtained when the firm makes no profit and the price is fixed (because the consumer is risk averse). More precisely we have:

**Lemma 1.** The socially efficient price is constant over time and equal to the expected cost $E[\theta]$.

Proof. The firm’s expected discounted profit at time $t$ is $\sum_{k=0}^{K} \delta^k E[p_{t+k} - \theta_{t+k}]$ where $E[\ ]$ is the expectation operator and $\delta$, a discount factor ($0 < \delta < 1$). This simplifies to $\sum_{k=0}^{K} \delta^k (p_{t+k} - E[\theta])$ as $\theta$ is i.i.d. □

We now examine different pricing scenarios with and without price commitment and explore their impact on consumer investment incentives. If the firm is unable to commit to future prices, it chooses the price that maximizes its profit at each period. Reasoning backward, this leads the firm to charge the limit price at which the consumer is indifferent to purchasing the service or not once she is connected. That is, the price such that $\nu_{t+k} = p_{t+k}$. Under this scenario, the consumer keeps purchasing district heating in any period, but the firm fully extracts the gains from trade.

Moving backward to the connection decision at time $t$, the consumer looks forward and maximizes her expected lifetime utility:

$$U_t \equiv -l_t + \sum_{k=0}^{K} \delta^k E[u(\nu_{t+k} - p_{t+k})]$$

(1)

As the consumer anticipates $\nu_{t+k} = 0$ in any future period, $U_t = -l_t < 0$; she never invests. This leads to a second Lemma.

\textsuperscript{13} In practice, district heating firms may not always maximise profit, but here we assume they behave as if they do since our primary objective is to highlight the hold-up problem generated by the pricing behaviour of profit-maximizing firms. This will not affect the econometric analysis since we will focus on the demand-side of the market.

\textsuperscript{14} In the worst case where the production cost is $\overline{\theta}$, the limit price leaving a firm with zero profit = yields higher consumer surplus than switching to the alternative heating option.
Lemma 2. If the firm is not able to commit to a price schedule at time $t$, the consumer always opts for the outside heating option. The consumer utility and firm’s profit are zero.

Note that the result of no connection is extreme because we have assumed a unique representative consumer who purchases a fixed quantity of service. The monopoly can thus fully extract the gain from trade. With consumer heterogeneity and/or continuous quantities, infra-marginal consumers would derive a positive surplus, and equilibria with partial connections would be feasible. Introducing these complexities would however not qualitatively alter the general insight resulting from the following propositions.

Assume now that at time $t$ the firm is able to commit to a price schedule for all future periods $t + k$. Importantly, to be credible, this commitment requires an outside party (e.g. a regulator) since it is always profitable for the firm to renege on the deal ex post in each period.\(^{15}\) Again, our objective is to characterize the price schedule chosen by the monopoly in this scenario. We examine two classes of price scheme. The first is a stochastic price that fully transfers cost shocks to the consumer: $p_{t+k} = \theta_{t+k} + m$, where $m$ is a constant mark-up. The second is a constant price $p_{t+k} = E[\theta] + m$. The relevant difference is that the latter allocates all the risk to the firm, while the former fully allocates the risk to the consumer.\(^{16}\) We characterize the equilibrium value of $m$ in each case and then compare the profits with the two price schedules in order to identify which will be used by the monopoly firm.

Under the stochastic price scheme, if the consumer is connected to the district heating network, her utility in period $t + k$ is

$$u_{t+k} = u(v_{t+k} - \theta_{t+k} + m).$$

Note that the consumer always purchases in the case where $m = 0$ since we have assumed that $v_{t+k} - \bar{\theta} \geq 0$. It follows that the firm will always set a value of $m$ which leads the consumer to keep using district heating in all future periods, as the worst it can do ($m = 0$) yields a non-negative profit. Moving backward at the time of the investment decision $t$, the consumer adopts district heating if

\(^{15}\) Implicit enforcement could also come from the threat of regulation or loss of votes in a political economy setting.\(^{16}\) Schedules with intermediate risk allocations are feasible, but the analysis of these two polar cases is sufficient to make our point.
\[ U_t = -I_t + \sum_{k=0}^{K} \delta^k E[u(v_{t+k} - p_{t+k})] \geq 0 \]

Since \( \theta \) is i.i.d., this condition is equivalent to

\[ I_t \leq \sum_{k=0}^{K} \delta^k E[u(v_{t+k} - \theta - m)] \quad (2) \]

Given this demand equation, the identification of the mark-up \( m \) requires us to consider two cases. The first is when the consumer participation constraint (2) is satisfied in the particular case where \( m = 0 \). Depending \textit{inter alia} on the level of \( I_t \), the firm will have some room to increase \( m \) above 0 without inducing the exit of the consumer. The equilibrium value \( m^* \) of the mark-up is then given by the indifference condition:

\[ I_t = \sum_{k=0}^{K} \delta^k E[u(v_{t+k} - \theta - m^*)]. \quad (3) \]

Importantly, the corresponding price is lower than the price without commitment, since inducing entry requires a positive surplus for the consumer in order to recoup the initial investment. This is a critical point: the ability to commit \textit{ex ante} allows the firm to charge a lower price that induces entry, and thus increases future profits. The second case is where eq. (2) is not satisfied with \( m = 0 \). In this case, the firm has simply no room to find a positive mark-up that will induce the consumer to adopt district heating. The consumer chooses the outside heating technology and profit is zero, the same as without commitment.

Consider now the fixed tariff \( p_{t+k} = E[\theta] + m \). Expected lifetime expected utility becomes:

\[ U_t = -I_t + \sum_{k=0}^{K} \delta^k u(v_{t+k} - E[\theta] - m). \]

The comparison with stochastic prices is straightforward. For the same \( m \), and because \( u'' < 0 \), \( U_t \) is higher than under the stochastic price since \( u(v_{t+k} - E[\theta] - m) > E[u(v_{t+k} - \theta - m)] \). Thus, the consumer is more likely to connect as the adoption condition is weaker than in (2). In contrast, whether prices are stochastic or rigid makes no difference for the firm if the consumer adopts the
technology since the firm is risk neutral.\textsuperscript{17} That is, the discounted flows of profits under the two schemes are identical:

\[
\sum_{k=0}^{\infty} \delta^k (E[\theta] + m - \theta_{t+k}) = \sum_{k=0}^{\infty} \delta^k m
\]

From these properties it follows that the firm will be able to find a value for \( m \) that keeps the consumer connected to the district heating network and generates a positive profit if the following constraint is satisfied

\[
l_t \leq \sum_{k=0}^{K} \delta^k u(v_{t+k} - E[\theta])
\]

In that case, \( m^* \) is higher than the value under the stochastic schedule. We can now compare the two price schedules:

1. If \( l_t \leq \sum_{k=0}^{K} \delta^k u(v_{t+k} - E[\theta]) \), the consumer adopts district heating under both price schedules. However, profit is higher with the fixed price while consumer surplus is identical.

2. If \( \sum_{k=0}^{K} \delta^k u(v_{t+k} - E[\theta]) \leq l_t \leq \sum_{k=0}^{K} \delta^k E[u(v_{t+k} - \theta)] \), the consumer only adopts district heating when the price is fixed. Both parties enjoy a positive surplus whereas their surplus would be zero with the stochastic price.

3. If \( l_t \geq \sum_{k=0}^{K} \delta^k E[u(v_{t+k} - \theta)] \), the consumer does not adopt district heating. Utility and profit are zero under both price schedules.

We are now able to write a proposition that collects the main findings.

**Proposition.** Under commitment, the firm chooses a constant price which leads the representative consumer to always opt for district heating. This equilibrium Pareto dominates the equilibrium without commitment.

This analysis yields a simple message. If risk adverse consumers need to make sunk investments to use district heating, the commitment to future rigid prices is Pareto improving. Where long-term contracts are not feasible (or if they are renegotiable), price regulation can offer the necessary mechanism to achieve this outcome. This result decisively hinges upon the assumptions that consumers are forward-looking and risk averse when adopting district heating. The empirical

\textsuperscript{17} These results would hold under the more general hypothesis that the firm is less averse to risk than the consumer.
investigation will test these two assumptions. It will also provide estimates used in subsequent simulations to assess the economic importance of these issues.

3.2 An estimable demand equation

We now introduce several complexities on the demand-side of the model in order to have a more realistic representation of how consumers choose heating services. The first extension is that consumers have the choice between four technologies (as explained in Section 2): (i) district heating, (ii) electricity heating, (iii) remaining technologies and (iv) the outside option. An additional extension is that we assume that consumers are heterogeneous. More specifically, the lifetime utility of a consumer $i$ who resides in local jurisdiction $m$, and who chooses technology $j$ in year $t$ is now

$$U_{m,j,t} + \omega_{i,m,j,t}$$

where $U_{m,j,t}$ is the average lifetime utility and $\omega_{i,m,j,t}$ is a consumer-specific deviation from the average that follows an extreme value distribution. We will come back to the specification of $U_{m,j,t}$ shortly. We employ Berry’s (1994) generalisation of McFadden’s (1973) discrete-choice demand model by transforming the multinomial logit model into a linear model. In Berry’s framework the probability that good $j$ is purchased asymptotically corresponds to its market share at time $t$. Hence, one can write: $s_{m,j,t} \equiv e^{U_{m,j,t}} / \sum_k e^{U_{m,k,t}}$. Normalizing the utility of the outside option $U_{m,0,t}$ to zero and taking the natural logarithm implies that

$$\ln s_{m,j,t} - \ln s_{m,0,t} = U_{m,j,t}.$$ 

A concern with this equation is that district heating and electricity heating are closer substitutes for each other than district heating is for other technologies (see Section 2). To relax the hypothesis of irrelevance of independent alternatives (IIA) and to account for the substitutability between district heating and electricity, we use a nested choice structure where district heating and electricity are in one nest and remaining technologies are in the other one. Since we are only interested in consumers’ preferences when they invest in district heating, we drop the $j$ subscript. In this nested-logit framework, the equation describing the market share of district heating becomes:

$$\ln (s_{m,t}) = U_{m,t} + \sigma \ln(s_{m\mid g,t}) + \ln(s_{0,t}).$$ (4)
where $s_{m|g,t}$ is the market share of district heating in a segment consisting of district heating and electricity and $\sigma \in [0,1]$ is a scalar that parameterizes the within-nest correlations. Note that the model collapses to the standard logit when $\sigma = 0$.

One crucial advantage of Berry's approach is to derive a demand equation for each product that does not depend on the characteristics of the other products (except the within-nest market share $s_{m|g,t}$). This restriction to the estimation of a single product is justified on two grounds: our research question only concerns district heating, and we do not have sufficient price data on electric heating and other heating options.

**Econometric specification**

We now need to specify the function $U_{m,t}$ that will be inserted into eq. (4). Adapting the notations from Section 3.1, we re-write eq. (1) as:

$$U_{m,t} = -I_{m,t} + \sum_{k=0}^{K} \delta^k u(p_{m,t+k} - p_{m,t+k}).$$

We will estimate the following linear transformation of this equation:

$$U_{m,t} = \alpha_1 p_{m,t-1} + \alpha_2 p_{m,t-1}^{\text{slope}} + \alpha_3 p_{m,t-1}^{sd} + \alpha_4 \Delta \ln(\ell_{m,t-1}) + \eta_m + \theta_t + \xi_{m,t} \tag{5}$$

We now justify this specification. To start with, the linear combination $\alpha_1 p_{m,t-1} + \alpha_2 p_{m,t-1}^{\text{slope}} + \alpha_3 p_{m,t-1}^{sd}$ captures the size of the utility loss associated with the payment of the district heating service. By definition, $p_{m,t-1}$ is the observed price in period $t-1$, $p_{m,t-1}^{\text{slope}}$ and $p_{m,t-1}^{sd}$ are respectively the slope of the long-term price trend and standard deviation of prices around that trend. In this expression, the term $\alpha_4 \Delta \ln(\ell_{m,t-1})$ represents consumer expectations about future prices. Similar to Branch (2004) and Evans and Honkapohja (2001), we assume that consumers make a straight line through available prices and extrapolate to form their expectations. The slope coefficient of this line indicates what price levels consumers expect in the future, relative to observed prices. The prices are lagged by one period because of the delay between the time when consumers decide to connect and when the actual connection takes place. The role of $\alpha_3 p_{m,t-1}^{sd}$ is different: when

---

18 Using these two price terms together with year and firm fixed effects, the prediction error varies from 0.18% to 0.34%. Thus, the errors are very small and the two price terms closely approximate rational expectations.
the average consumer is risk averse it serves as a proxy for utility losses associated with future price volatility. It thus captures the fact that the utility function $u$ is concave. We expect $\alpha_3 < 0$.

It is necessary to determine how many past prices consumers take into consideration when they calculate $p_{tslope}$ and $p_{trend}$. We set this number to three, i.e. that consumers use $p_t$, $p_{t-1}$ and $p_{t-2}$. This procedure assumes that consumers attach equal weights to the three prices and ignore more historical prices. There are both practical and conceptual justifications for why we choose three prices. The practical one is that two prices ($p_t$ and $p_{t-1}$) would create a perfect fit between the line and the observations, resulting in no price variation, and that four prices reduce the estimable sample by over 50% compared to three prices. The conceptual justification is that many firms publish annual reports, including prices for the last 2-3 years, on their homepages, making these easily accessible, while more historical prices are costlier to access.

To deal with the problem that we do not observe connection costs, $l_{m,t}$, and values of usage, $v_{m,t+k}$, in the data, we introduce in eq. (5) $\eta_m$ which is a vector of firm fixed effects that control for variation across local jurisdiction, $\theta_t$, a vector of time-specific effects and $\xi_{m,t}$, a random noise.\footnote{The Energy Markets Inspectorate has developed principles to determine connection prices for physical energy networks and these principles are largely used by district heating utilities. Two primary cost components are: material/equipment and the distance between the network and the connecting property. The market for material/equipment is national (i.e. covered by year fixed effects) and the distance is strongly correlated to population density, which exhibits very little within-variation. The inclusion of density in the empirical evaluation shows that it is not even vaguely related to the number of connections. Hence, we assume that connection prices have been constant.}

Another challenge is that every household does not have access to district heating. In reality, only households located at a certain distance from the network have the opportunity to connect. The number of properties that can be connected in each time period is thus related to the expansion of the network. This leads us to include $\Delta \ln(\ell_t)$ as a control variable. This variable is the network expansion between period $t-1$ and $t$.

Substituting eq.(5) in eq.(4) leads to the base equation that will be estimated below\footnote{This baseline specification assumes that consumers have the same preferences for falling and rising price trends. In the case of investments with long lifetimes, one cannot form strong expectations about the influential direction of past price decreases since a fall in the price may merely be interpreted by the consumer as indicating that there is room to increase prices in the future (e.g. Adeyemi and Hunt (2007) find a negative effect on demand from cumulative price decreases in one of their models). We estimate models where the effects from...}:
\[
\ln(s_{m,t}) - \ln(s_{m,0,t}) = \alpha_1 p_{m,t-1} + \alpha_2 p_{m,t-1}^{\text{slope}} + \alpha_3 p_{m,t-1}^{\text{sd}} + \\
+ \alpha_4 \Delta \ln(\ell_{m,t-1}) + \sigma \ln(s_{m,\text{lag},t}) + \eta_m + \theta_t + \xi_{m,t}
\] (6)

The whole framework hinges on the assumption that the consumer never revises her heating choice once she is connected. This is debatable. In theory, when making her investment decision, the consumer can anticipate the possibility of switching to other heating options. If this is the case, future prices and characteristics of other technologies matter and should be included in \(U_{m,t}\). However, we have seen in Section 2 that such premature revisions of investment decisions do not occur in practice (essentially because investments are sunk). In any case, we could not do differently because of a lack of data on the price of alternative heating options.

4. DATA

The data set includes aggregate annual market data from 126 firms from 1997-2007. The firms included are those that have reported complete data for all variables of interest during at least two years and those that have had district heating networks for at least five years. The average firm has 5.2 observations and the data is unbalanced.

A property owner can either occupy her dwelling or rent it out, and it is possible that an owner’s incentives to invest in a heating device may differ depending on whether or not the property is rented out.\(^{21}\) It is not possible to directly separate connected dwellings occupied by their owner from rented-out properties. However, dwellings that are rented out are almost exclusively multi-dwelling houses (i.e. apartment blocks) owned by large firms or the local council. Those are often connected to the network in the first wave (1-3 years after the production facility is completed) to reach the critical mass that is needed to make district heating financially viable. To control for this, and other non-equilibrium behaviours at the start-up of networks, we exclude the first five years of operation for all networks.

\(^{21}\) About 75% of Swedish households live in detached houses. Practically all of these own their properties since there is no tradition of renting detached houses in Sweden and no established marketplace for such properties. A proportion of the remaining 25% also occupy their own properties, but a non-negligible share of households nevertheless rent the dwelling they occupy.
In Table A1 in the Appendix we provide detailed statistics of municipalities that are at different stages of their district heating programmes. In this and the next paragraph we summarise the most important conclusions based on that Table. Of a total of 290 municipalities, 250 had district heating in 2007. Of those 250, 16 were created between 2001 and 2007 and are therefore excluded to ensure that all observations represent networks older than five years. Missing data results in the exclusion of a further 108 municipalities. This group is rather large because it is difficult to access data on heating technology market shares and the share of fossil fuel. Such data are only partially available through public sources, and so these variables were primarily collected directly from municipalities and firms.

The existence of district heating networks tends to be positively related with population and population density. Income is similar across all groups, suggesting that consumers’ willingness to pay is not an important determinant. Consistent with expectations, Table A1 reveals that (i) the number of multi-dwelling houses is substantially higher in municipalities with district heating (higher connection density reduces distribution costs), and (ii) high demand for heating (high HDD) increases the likelihood of a district heating network. The firms in the sample operate in municipalities with a slightly lower number of inhabitants and density compared to the population of municipalities with a district heating network (see Table A1). However, the characteristics of price variables in the sample and population are generally very similar, both in terms of mean and range.

District heating prices are at the centre of this study and it can be informative to take a closer look at their characteristics. Table 1 displays characteristics for: (i) price averaged over time (denoted \( \bar{p}_m \)), (ii) price slope \( p_{m,t}^{sd} \), and (iii) the percentage change from the first to the last reported prices for each firm. The statistics indicate noticeable heterogeneity across firms, with the highest average price being twice as high as the lowest price. However, the distribution is negatively skewed, with fewer firms in the low price tail compared to the high tail. The average standard deviation around the long-term price trend is 0.7, but the positively skewed distribution varies from practically zero, i.e. a situation where the price series closely follows the long-term trend, to above 3. Firms increased their real prices over the period by approximately 10% on average. The distribution of the price difference is relatively symmetric, with the largest price decrease amounting to almost 30% and the largest price increase to almost 60%.

### Table 1. Price characteristics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min</th>
<th>Max</th>
<th>Shapiro-Wilk test, Prob(&gt;z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price averaged over time: ( \bar{p}<em>m = T^{-1} \sum</em>{t=1}^{T} p_{m,t} )</td>
<td>54.79</td>
<td>5.91</td>
<td>34.67</td>
<td>67.18</td>
<td>0.01</td>
</tr>
<tr>
<td>Standard deviation of de-trended</td>
<td>0.67</td>
<td>0.50</td>
<td>0.13</td>
<td>3.03</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Empirical identification of price characteristics relies on variation within firms. It is possible that firms adopt a certain price policy and then apply that policy over an extended period of time. For example, firms may decide that prices should follow the consumer price index and that cost shocks should be smoothed over a certain number of years. Such policies can imply that within-variations of $p_t^{slope}$ and $p_t^{sd}$ are small and that it is hard to empirically identify their effects. However, Panel A in Table 2 reveal that the within-variations of these variables are larger than the between-variations.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$p_{m,t}$</th>
<th>$p_{m,t}^{slope}$</th>
<th>$p_{m,t}^{sd}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Firms included in sample</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between-variation</td>
<td>5.88</td>
<td>0.96</td>
<td>0.46</td>
</tr>
<tr>
<td>Within-variation</td>
<td>3.68</td>
<td>1.57</td>
<td>0.54</td>
</tr>
<tr>
<td>Panel B: Firms not included in sample</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between-variation</td>
<td>6.54</td>
<td>1.66</td>
<td>0.66</td>
</tr>
<tr>
<td>Within-variation</td>
<td>3.63</td>
<td>1.61</td>
<td>0.61</td>
</tr>
</tbody>
</table>

When comparing price statistics for firms included in the sample with those that are not included, we find very similar values. Panel B in Table 2 reveals that both within- and between-variations are similar for all three price variables. Thus, there is no evidence that pricing behaviours are different for those municipalities that are not included in the estimable sample.

Is there any sign that these variables have affected property-owners’ inclination to connect to the district heating network? We begin by calculating the market share of district heating for each firm and year as the ratio between ‘the number of property owners installing district heating’, and ‘the total number of property owners in a position to invest in a new heating technology’. Those in a position to invest in a new heating device are the sum of all newly constructed dwellings from year $t-1$ to $t$ and 5% of the existing stock of dwellings. Similar market share calculations are made for electricity and the outside option (see Table 3), since these are used in the analysis in Section 4. Next we compute the correlations between the district heating market share and the three price characteristics displayed in Table 1. It turns out that the correlation between market share and the price, both averaged over time, is -0.23. The other two correlations are very close to zero. Hence, this descriptive analysis suggests that a high price reduces consumers’ propensity to make sunk investments in district heating,
whereas the long-term price trend and price variability have no clear impact on consumers. However, these findings ignore within-firm variation and other controls.

Table 3. Descriptive statistics for variables used to estimate (6).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description (measurement unit)</th>
<th>n</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p )</td>
<td>Unit price of district heating (öre/kWh) (^a)</td>
<td>659</td>
<td>55.527</td>
<td>7.391</td>
<td>32.990</td>
<td>72.892</td>
</tr>
<tr>
<td>( p_{\text{sld}} )</td>
<td>Slope of line fitted through three most recent prices</td>
<td>659</td>
<td>0.8052</td>
<td>1.8618</td>
<td>-9.3079</td>
<td>7.2469</td>
</tr>
<tr>
<td>( p_{\text{sd}} )</td>
<td>Sum of squared deviations of the de-trended price based on three most recent prices</td>
<td>659</td>
<td>0.6318</td>
<td>0.7491</td>
<td>0.0007</td>
<td>8.5111</td>
</tr>
<tr>
<td>( \Delta \ell )</td>
<td>Network expansion from period ( t-1 ) to ( t ) (km) (^c)</td>
<td>659</td>
<td>8.0086</td>
<td>14.631</td>
<td>0</td>
<td>120</td>
</tr>
<tr>
<td>( s_0 )</td>
<td>(Residential properties not investing in a fixed device) / (Residential properties in a position to choose heating technology) (^b,e)</td>
<td>659</td>
<td>0.1007</td>
<td>0.0072</td>
<td>0.0776</td>
<td>0.1193</td>
</tr>
<tr>
<td>( s_{DH} )</td>
<td>(Residential properties connected to district heating network) / (Residential properties in a position to choose heating technology) (^b,e)</td>
<td>659</td>
<td>0.1146</td>
<td>0.1169</td>
<td>0</td>
<td>0.8911</td>
</tr>
<tr>
<td>( s_{EL} )</td>
<td>(Residential properties investing in an electricity based device) / (Residential properties in a position to choose heating technology) (^b,e)</td>
<td>659</td>
<td>0.1691</td>
<td>0.1998</td>
<td>0.0008</td>
<td>0.9779</td>
</tr>
<tr>
<td>( \text{hdd} )</td>
<td>No of heating degree days in previous year (^a)</td>
<td>659</td>
<td>3640</td>
<td>739.8</td>
<td>1613</td>
<td>7309</td>
</tr>
<tr>
<td>( \text{inco} )</td>
<td>Average individual income net of taxes (kSEK) (^a)</td>
<td>659</td>
<td>190.2</td>
<td>14.737</td>
<td>149.95</td>
<td>254.91</td>
</tr>
<tr>
<td>( \text{fos} )</td>
<td>Share of fossil fuel used in production (^e)</td>
<td>659</td>
<td>0.0692</td>
<td>0.0640</td>
<td>0.0176</td>
<td>0.5230</td>
</tr>
<tr>
<td>( \text{pop16} )</td>
<td>Share of inhabitants aged from 0 to 16 (^a)</td>
<td>659</td>
<td>0.1986</td>
<td>0.0168</td>
<td>0.1469</td>
<td>0.2590</td>
</tr>
<tr>
<td>( \text{homes} )</td>
<td>Number of household dwellings (^a)</td>
<td>659</td>
<td>25 253</td>
<td>45 123</td>
<td>1 869</td>
<td>428 163</td>
</tr>
</tbody>
</table>

\(^a\) Source: Statistics Sweden. 
\(^b\) Source: Statistics Sweden. Number of consumers who are in a position to choose heating technology in year \( t \) consists of all new dwellings completed in \( t \) plus 5% of the existing stock of dwellings. This follows from the assumption that the average lifetime of a heating device is 20 years.
\(^c\) Source: The Swedish District Heating Association.
\(^d\) Source: Annual reports; homepages.
\(^e\) Information collected from municipalities and firms directly.

5. ESTIMATION

5.1 Econometric issues

Endogeneity is a major concern when estimating eq. (5). This firstly concerns the within-firm market share, which is endogenous by construction as it is the market share for district heating divided by the sum of the market shares for district heating and electricity. To mitigate the problem, we adopt a 2SLS model and instrument the within-firm market share with heating degree days (HDD). HDD is designed to reflect the demand for heating. The precise definition is the number of degrees that a day’s average temperature is below 18°C, the temperature below which buildings need to be heated. The logic for using HDD as an instrument is that the electricity price is strongly correlated with temperature, since temperature directly influences the production cost of electricity, and this is transmitted to the many consumers who have a contract with time-varying electricity prices. In contrast, the price of district heating is set annually and in advance and thus is not affected by short-term changes in temperature.
The price variables can also be endogenous, but for a different reason. The price that consumers observe when they decide to make their sunk investment is set in the previous period, but the price can still be endogenous if it is subject to autocorrelation. For instance, if firms use a stochastic price scheme (described in Proposition 1), autocorrelation can occur if costs are autocorrelated, and a fixed price scheme (described in Proposition 2) will result in perfect autocorrelation. This can also occur if price reviews are costly. As \( p_{m,t-1} \), \( p_{m,t-2} \), and \( p_{m,t-3} \) are used to calculate \( p_{m,t-1}^{slope} \) and \( p_{m,t-1}^{sd} \), all three price variables in eq. (5) are potentially endogenous.

In order to parsimoniously select the instruments, we push the identification of the autocorrelation structure by estimating the following dynamic model using the standard within-estimator and clustered SEs: \( p_{m,t} = \alpha p_{m,t-1} + \eta_m + \epsilon_{m,t} \). Here \( \alpha = 0.76 \) and it is highly significant, confirming that \( p_{m,t} \) and \( p_{m,t-1} \) are highly correlated. To investigate the correlation between \( p_{m,t} \) and \( p_{m,t-2} \), we then run the following regression: \( \hat{\epsilon}_{m,t} = \alpha' p_{m,t-2} + \eta'_m + \epsilon'_{m,t} \) where \( \hat{\epsilon}_{m,t} \) is the residuals from the first model. Note that the AR(1) component identified in the first model is excluded to reduce multicollinearity between \( p_{m,t-1} \) and \( p_{m,t-2} \). Here we estimate that \( \alpha' = -0.06 \) and it is not significant at the 10% level. Thus, in the following we consider \( p_{m,t-1} \) to be endogenous but that deeper lags of \( p_{m,t-1} \) are exogenous.

We use Wooldridge’s (2002, pp. 623-625) approach where an instrument is used to generate predicted values of the endogenous variable. These predictions are then used as an instrument in a standard 2SLS estimation. As eq. (5) describes demand, we need to instrument with cost shifters. We take the share of fossil fuel used in the production of district heating, in practice, mostly oil, in municipality \( m \) in year \( t-1 \) (\( fos_{m,t-1} \)). Oil is an expensive fuel type, and its share of the total amount of fuel used strongly influences the total cost of providing district heating. In contrast, it has no impact on the price of the alternative heating technologies. In particular, oil is a negligible part of the Swedish electricity production mix, which primarily consists of hydropower and nuclear. Thus, the share of fossil fuel does not influence consumers’ relative utility for district heating.

The estimation procedure involves the following stages:

Stage 1. Regress \( p_{m,t} \) on \( fos_{m,t-1} \), and allow the effect of \( fos \) to be firm-specific:

22 However, a one-stage procedure where both and are included as explanatory variables in the same model gives the same qualitative results as the two-stage procedure.
\[ p_{m,t} = \text{os}_{m,t-1} \times \text{Firm FE}_{m,t} + \text{Firm FE} + \text{Year FE} + \varepsilon_{m,t} \]

**Stage 2.** Calculate and save the predicted values: \( \hat{p}_{m,t} \).

**Stage 3.** Use the predicted values \( (\hat{p}_{m,t}) \) together with the actual values lagged by one and two periods to construct \( \hat{p}^{s\text{slope}}_{m,t-1} \) and \( \hat{p}^{s\text{sd}}_{m,t-1} \).

**Stage 4.** Use \( \hat{p}_{m,t-1} \), \( \hat{p}^{s\text{slope}}_{m,t-1} \) and \( \hat{p}^{s\text{sd}}_{m,t-1} \) as instruments for \( p_{m,t-1} \), \( p^{s\text{slope}}_{m,t-1} \) and \( p^{s\text{sd}}_{m,t-1} \) when estimating (5).

Since two of the prices that make up \( \hat{p}^{s\text{slope}}_{m,t-1} \) and \( \hat{p}^{s\text{sd}}_{m,t-1} \) are also included in \( p^{s\text{slope}}_{m,t-1} \) and \( p^{s\text{sd}}_{m,t-1} \), the instruments turn out to be strong in all estimations displayed in Table 4. The F-statistics for each of the three price variables in all of the first stage regressions are well above the conventional threshold of 10.

Finally, standard errors are clustered over firms in all estimations to deal with potential within-firm correlation. Within-correlation can occur if there are spill-over effects across neighbours in the same geographical area, e.g., a connected consumer shares its experiences with unconnected neighbours.\(^{23}\)

### 5.2 Results

We begin by estimating eq. (5) using OLS. The results are displayed in column (1) in Table 4 and our expectations are confirmed. In particular, consumers are less inclined to take up district heating as the price slope and price volatility increase. **Size of the effect?** For instance, market share if a one standard deviation increase of \( p^{s\text{slope}}_{m,t-1} \) and \( p^{s\text{sd}}_{m,t-1} = 0 \); District heating also acquires a larger share of the market when more consumers access the district heating network, i.e. as the network expands. Furthermore, district heating and electricity partly substitute a within-nest correlation of 0.54. Note, however, that the coefficient of the current price level is not significant. This is probably due to a combination of (i) limited within-variation of \( p_{m,t-1} \) that means the price level is partly captured by the firm FE, and (ii) the co-variation of \( p_{m,t-1} \) and \( p^{s\text{slope}}_{m,t-1} \). The correlation between two variables is 0.40.

The results when the within-nest market share is instrumented are displayed in column (2) of Table 4. The \( \sigma \) coefficient reduces from 0.54 to 0.24, suggesting a weaker degree of substitutability between

---

\(^{23}\) A dynamic panel data specification, which provides a more precise within-correlation structure, was evaluated but it did not increase estimation efficiency. See column (3) in Tables A2 and A3 in Appendix 2 for details.
district heating and electricity. However, the coefficients of the price variables are practically unchanged.

Finally we estimate a model where $\ln(s_{m|g})$, $p_{m,t-1}$, $p_{m,t-1}^{slope}$ and $p_{m,t-1}^{sd}$ are all treated as endogenous. This is the base model that will be used for the simulations. Results are presented in column (3). Most coefficients remain unchanged except that of $p_{t-1}^{sd}$, which decreases by around 50%. This justifies the instrumentation of the price variables. The nest and price coefficients all display the expected signs and are significant, at least at the 10% level.

Table 4. Estimation output of eq. (5).

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{m,t-1}$</td>
<td>0.0014</td>
<td>0.0089</td>
<td>0.0014</td>
</tr>
<tr>
<td>$p_{m,t-1}^{slope}$</td>
<td>-0.0922</td>
<td>-0.1061</td>
<td>-0.1006</td>
</tr>
<tr>
<td>$p_{m,t-1}^{sd}$</td>
<td>-0.4001</td>
<td>-0.4300</td>
<td>-0.1947</td>
</tr>
<tr>
<td>$\Delta \ln(f_{m,t-1})$</td>
<td>0.1467</td>
<td>0.1388</td>
<td>0.1384</td>
</tr>
<tr>
<td>$\ln(s_{m</td>
<td>g,t})$</td>
<td>0.5386</td>
<td>0.2441</td>
</tr>
</tbody>
</table>

| Year FE | Yes | Yes | Yes |
| Firm FE | Yes | Yes | Yes |
| $R^2$ (within) | 0.301 | 0.260 | 0.244 |

Notes. Standard errors clustered over firms in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable: $\ln(s_{m,t}) − \ln(s_{m,0,t})$. (1) OLS, entire sample; (2) 2SLS with instrumentation of the within-nest market share, entire sample; (3) 2SLS with instrumentation of the within-nest market share and the price variables, entire sample.

Results are robust to a number of different specifications and sample conditions. To test the robustness of the specification we estimate a model where income is included as an additional control and in another model we allow consumers to respond differently to positive and negative price slopes. To test the robustness of the sample, we first exclude the three municipalities with the largest populations (Stockholm, Gothenburg and Malmoe). These municipalities, which are substantially larger than all other municipalities, have scale and scope opportunities that may lead to different

---

24 Eq. (5) assumes that consumers have the same preferences for falling and rising price trends. In the case of investments with long lifetimes, we cannot form strong expectations about the influential direction of past price decreases, since a fall in price may merely be interpreted by the consumer as the existence of room to increase prices in the future (e.g. Adeyemi and Hunt (2007) find a negative effect on demand from cumulative price decreases in one of their models). See Table A2 in Appendix 2 for details.
supply conditions. They also have different demand characteristics due to less indoor space on average. In another estimation we exclude municipalities with urban population density below 8 individuals per hectare. In contrast to the largest municipalities, these networks have unfavourable supply conditions. In a third test we exclude the most northern municipalities. These have different demand for heating since they are located in the coldest region. Finally, we exclude the 10% of municipalities with the highest shares of rented dwellings. This is because owners who occupy their dwellings and those who rent them out can have different incentives for making long-term investments. Further details and estimation results are included in Appendix 2, Tables A2-4.

6. SIMULATIONS

We have shown in the preceding section that \( p^{slope} \) and \( p^{sd} \) affect the take-up of the service. In this section, we simulate a counterfactual scenario with a regulated price in order to estimate the economic significance of this effect.

Under the optimal regulation, the price is fixed – as consumers are risk averse while the district heating firm is not – and is equal to the unit cost averaged over time (see Lemma 1). A practical problem is that cost statistics are only available from 2007 onwards, which is posterior to our sample period (1998-2007). Our strategy is then to extrapolate the values for the pre-2007 period using post-2007 data, assuming that market conditions were stable from 1998-2014. To check whether this assumption is valid, we take the 2007-2014 data, deflate it to the same base (year 2000) and compare descriptive statistics for \( p \) and \( p^{slope} \) across the two sample periods, i.e. 1998-2007 and 2007-2014. As displayed in Table 5, it turns out that the two price variables have very similar characteristics in the two samples. There is no evidence that market conditions changed substantially across the periods.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Period: 1998-2007</th>
<th>Mean</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td></td>
<td>55.527</td>
<td>7.391</td>
</tr>
<tr>
<td>$p^{slope}$</td>
<td></td>
<td>0.805</td>
<td>1.862</td>
</tr>
<tr>
<td>$p$</td>
<td></td>
<td>0.791</td>
<td>2.212</td>
</tr>
</tbody>
</table>

The simulation procedure is then as follows. We consider an average municipality where the independent variables in eq.(6) are set to the sample average (including the price variables $p$, $p^{slope}$ and $p^{sd}$). Using the coefficients displayed in column (3) in Table (4), the predicted district heating market share in this average municipality is 9.3%. This is the Business-As-Usual (BAU) benchmark against which we will compare the scenario with optimal regulation. Next, we calculate the average cost across all municipalities and over the 2007-2014 period. This average is $c = 52.743$. Last, we calibrate by assuming that $p_{m,t-1} = c$, $p^{slope} = 0$ and $p^{sd} = 0$.

The predicted market share with regulation is 11.3%, which is remarkably close to the BAU market share of 9.3%. In theory, this limited gap can have two nonexclusive sources of explanation: low mark-ups and/or a low inter-annual price volatility. A look at descriptive statistics in the post-study period 2007-2014 suggests that both factors are at work: The mark-up is around 6.4%, which is arguably quite low while the volatility, which is mostly described by $p^{sd}$, is much lower than the cost volatility $c^{sd}$ (the mean is 6.379 compared to 11.832 as shown in Table 6).

Table 5. A comparison of cost and price variables in the 2007-2014 period.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>56.356</td>
</tr>
<tr>
<td>$p^{slope}$</td>
<td>0.775</td>
</tr>
<tr>
<td>$p^{sd}$</td>
<td>6.379</td>
</tr>
<tr>
<td>$c$</td>
<td>52.895</td>
</tr>
<tr>
<td>$c^{slope}$</td>
<td>0.505</td>
</tr>
<tr>
<td>$c^{sd}$</td>
<td>11.832</td>
</tr>
</tbody>
</table>
Since the market price deregulation of 1996, regulation has been a subject of continuous political debate. Public dissatisfaction has brought the market to the brink of regulation several times. The height of the debate was the passing of the District Heating Act (DHA) in 2008, which created a formal procedure through which consumers can complain about the prices they were charged to a national committee.

6. CONCLUSIONS

We focus on decisions where consumers make long-term sunk investments in a service while subject to uncertainty about future usage prices. Given that consumers are forward-looking and risk averse, we show that welfare increases if the firm can commit to a price scheme. As this commitment is not credible without third-party enforcement, this suggests that price regulation could play a role. Specifically, we show that the lower the future price level and the lower the uncertainty of future prices, the higher the propensity to make sunk investments. In the empirical part of the paper, we test whether the assumptions that lead to these conceptual predictions hold, i.e. whether consumers are forward-looking and risk averse.

We use data from the Swedish district heating market where consumers need to make a material long-term sunk investment in a heat exchange device in order to take up the district heating service. Faced with this investment decision, customers form a view about the future path of prices through the slope of the long-term price trend, and their (average) preference for risk is measured through the volatility of prices around the long-term price trend. The empirical results are robust and consistent with the assumptions. Specifically, we find that both the price slope and volatility are negatively related to sunk investments in district heating.

These findings support the view that one role of natural monopoly regulation is to provide potential customers with some assurance as to the stability of the likely future path of prices, thus encouraging them to make sunk complementary investments and take up the public utility service. This is potentially an important difference in perspective with important implications for regulatory policy. To explore this point further, we have simulated a counterfactual scenario with an optimal price regulation. The predicted social optimum would be 11.3%, which is only slightly higher than 9.3%, the predicted market share in the business-as-usual (unregulated) scenario. These results are driven by both low mark-ups and a low price volatility (compared to the cost volatility). This suggests that Swedish district heating firms self-regulated to a large extent during the study period. The indication is that there was a credible threat of price regulation: a draft law on district heating was discussed.
during the study period (and finally adopted in 2008). Assessing empirically the stringency of this regulatory threat is beyond the scope of the present paper, but it clearly deserves further research.
REFERENCES


Biggar, D., (2009), Is protecting sunk investment by consumers a key rationale for natural monopoly regulation?. Review of Network Economics 8, 128-152.


Erdem, T., M. Keane and S. Imai, (2003), Consumer Price and Promotion Expectations: Capturing Consumer Brand and Quantity Choice Dynamics under Price Uncertainty. Quantitative Marketing and Economics 1, 5-64.

Evans, G.W. and Honkapohja, (2001), Learning and Expectations in Macroeconomics, Princeton University Press.


SCB, (2009), Electricity supply, district heating and supply of natural and gasworks gas 2007, EN 11 SM 0901.


APPENDIX 1.

Table A1. Characteristics of Swedish municipalities in 2007 for different stages of district heating adaptation.

<table>
<thead>
<tr>
<th></th>
<th>Municipalities with no district heating network in 2007</th>
<th>Municipalities with district heating network in 2007</th>
<th>Municipalities with district heating networks, but where networks were too young to be included in sample (i.e. created after 2000)</th>
<th>Municipalities with district heating created 2000 or earlier but incomplete data</th>
<th>Municipalities with district heating created 2000 or earlier and complete data (i.e. municipalities included in sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of municipalities</td>
<td>40</td>
<td>250</td>
<td>16</td>
<td>108</td>
<td>126</td>
</tr>
<tr>
<td>Population</td>
<td>14 009 (10 797)</td>
<td>34 490 (65 600)</td>
<td>17 955 (18 993)</td>
<td>22 260 (23 440)</td>
<td>47 065 (87 950)</td>
</tr>
<tr>
<td>Population density</td>
<td>1 163 (452)</td>
<td>1 445 (688)</td>
<td>1 535 (842)</td>
<td>1 409 (737)</td>
<td>1 464 (625)</td>
</tr>
<tr>
<td>Income</td>
<td>226 000 (34 000)</td>
<td>225 000 (28 000)</td>
<td>234 000 (63 000)</td>
<td>227 000 (31 000)</td>
<td>222 000 (17 000)</td>
</tr>
<tr>
<td>No of detached houses</td>
<td>4 727 (2 623)</td>
<td>7 364 (6 556)</td>
<td>4 925 (2 959)</td>
<td>5 642 (4 287)</td>
<td>9 150 (7 893)</td>
</tr>
<tr>
<td>No of apartment houses</td>
<td>1 636 (1 919)</td>
<td>9 497 (29 112)</td>
<td>3 193 (5 022)</td>
<td>4 719 (7 013)</td>
<td>14 392 (39 927)</td>
</tr>
<tr>
<td>HDD (Heating degree days)</td>
<td>3 855 (669)</td>
<td>4 112 (729)</td>
<td>3 687 (366)</td>
<td>4 096 (704)</td>
<td>4 179 (768)</td>
</tr>
</tbody>
</table>

APPENDIX 2.

The Appendix provides robustness checks of eq. (5) when nest and prices are assumed to be exogenous (Table A2), when the nest is endogenous and prices exogenous (Table A3), and when both the nest and prices are endogenous (Table A4).

Table A2. Estimation output of eq. (5) when nest and all price variables are exogenous.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ln(s_{m,t-1}) - \ln(s_{m,0,t-1}) )</td>
<td></td>
<td></td>
<td>0.0191</td>
<td>0.0032</td>
</tr>
<tr>
<td>( p_{t-1} )</td>
<td>0.0017 (0.0193)</td>
<td>-0.0040 (0.0193)</td>
<td>0.0070 (0.0218)</td>
<td>0.0032 (0.0204)</td>
</tr>
<tr>
<td>( p_{t-1}^{\text{slope}} )</td>
<td>-0.0929** (0.0403)</td>
<td>-0.1464** (0.0646)</td>
<td>-0.0921** (0.0428)</td>
<td>-0.1060** (0.0433)</td>
</tr>
<tr>
<td>( p_{t-1}^{\text{slope}} D_{p_{t-1}^{\text{slope}} &gt; 0} )</td>
<td></td>
<td>0.0958 (0.0836)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p_t^{\text{rd}} )</td>
<td>-0.4013*** (0.0832)</td>
<td>-0.4323*** (0.0795)</td>
<td>-0.4018*** (0.0871)</td>
<td>-0.3568*** (0.0841)</td>
</tr>
<tr>
<td>( \Delta \ln(s_{t-1}) )</td>
<td>0.1476*** (0.0444)</td>
<td>0.1442*** (0.0436)</td>
<td>0.1603*** (0.0533)</td>
<td>0.1329*** (0.0509)</td>
</tr>
<tr>
<td>( \ln(s_{mlg,t}) )</td>
<td>0.5400*** (0.0568)</td>
<td>0.5421*** (0.0569)</td>
<td>0.5497*** (0.0441)</td>
<td>0.5463*** (0.0552)</td>
</tr>
<tr>
<td>( \ln(c_{t-1}) )</td>
<td>-0.0086 (0.0252)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Year FE | Yes | Yes | Yes | Yes |
Firm FE | Yes | Yes | Yes | Yes |

A-B AR(2) P>z | 0.301 | 0.303 | 0.311 |
R² (within) | 659 | 506 | 576 |
No obs | 659 | 659 | 506 |

Notes. Standard errors clustered over utilities are reported in brackets. * \( p < 0.10 \), ** \( p < 0.05 \), *** \( p < 0.01 \). Dependent variable: \( \ln(s_{m,t}) - \ln(s_{m,0,t}) \). (1) OLS, with income added as explanatory variable, entire sample; (2) OLS, control for different responses to positive and negative price slopes, entire sample; (3) GMM, dynamic panel data model, entire sample; (4) OLS, municipalities with 10% highest share of rented dwellings excluded (m=14). Eq. (3) uses the first difference of the dependent variable in period t-2 as instrument for the lagged dependent variable.
Table A3. Estimation output of eq. (5) when nest is endogenous and price variables exogenous.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \ln(s_{m,t-1}) - \ln(s_{m,0,t-1}) )</td>
<td>0.0088 (0.0196)</td>
<td>0.0049 (0.0202)</td>
<td>0.0032 (0.1524)</td>
<td>0.0149 (0.0205)</td>
</tr>
<tr>
<td>( p_{t-1} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p_{t-1}^{slope} )</td>
<td>-0.1058 (0.0432)</td>
<td>-0.1461 (0.0648)</td>
<td>-0.0944 (0.0414)</td>
<td>-0.1194 (0.0451)</td>
</tr>
<tr>
<td>( p_{t-1}^{sd} )</td>
<td>-0.4295 (0.0940)</td>
<td>-0.4537 (0.0899)</td>
<td>-0.4118 (0.0872)</td>
<td>-0.3951 (0.0937)</td>
</tr>
<tr>
<td>( \Delta \ln(\ell_{t-1}) )</td>
<td>0.1364 (0.0423)</td>
<td>0.1370 (0.0416)</td>
<td>0.1590 (0.0509)</td>
<td>0.1269 (0.0468)</td>
</tr>
<tr>
<td>( \ln(s_{m[t,g,t]}) )</td>
<td>0.2435 (0.1400)</td>
<td>0.2486 (0.1372)</td>
<td>0.4674 (0.1680)</td>
<td>0.2872 (0.1428)</td>
</tr>
<tr>
<td>( \ln(co_{t-1}) )</td>
<td>0.0038 (0.0269)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes. Standard errors clustered over utilities are reported in brackets. *\( p < 0.10 \), **\( p < 0.05 \), ***\( p < 0.01 \). Dependent variable: \( \ln(s_{m,t}) - \ln(s_{m,0,t}) \). (1) 2SLS, Nest endogenous, with income added as explanatory variable, entire sample; (2) 2SLS, Nest endogenous, control for different responses to positive and negative price slopes, entire sample; (3) GMM, Nest endogenous, dynamic panel data model, entire sample; (4) 2SLS, Nest endogenous, municipalities with 10% highest share of rented dwellings excluded (\( m=14 \)). Eq. (3) uses the first difference of the dependent variable in period \( t-2 \) as instrument for the lagged dependent variable.

Table A4. Estimation output of eq. (5) when both nest and price variables are endogenous.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( P_{t-1} )</td>
<td>-0.0026 (0.0215)</td>
<td>0.0079 (0.0233)</td>
<td>-0.0176 (0.0218)</td>
<td>0.0108 (0.0215)</td>
</tr>
<tr>
<td>( p_{t-1}^{slope} )</td>
<td>-0.1030 (0.0409)</td>
<td>-0.1235 (0.0416)</td>
<td>-0.0943 (0.0466)</td>
<td>-0.1098 (0.0421)</td>
</tr>
<tr>
<td>( p_{t-1}^{sd} )</td>
<td>-0.1938 (0.1033)</td>
<td>-0.1781 (0.1002)</td>
<td>-0.1916 (0.1138)</td>
<td>-0.1846 (0.1015)</td>
</tr>
<tr>
<td>( \Delta \ln(\ell_{t-1}) )</td>
<td>0.1454 (0.0447)</td>
<td>0.1151 (0.0477)</td>
<td>0.1320 (0.0489)</td>
<td>0.1271 (0.0475)</td>
</tr>
<tr>
<td>( \ln(s_{m[t,g,t]}) )</td>
<td>0.2534 (0.1358)</td>
<td>0.2825 (0.1231)</td>
<td>0.2996 (0.1133)</td>
<td>0.2892 (0.1395)</td>
</tr>
</tbody>
</table>

Notes. Standard errors clustered over firms in brackets. *\( p < 0.10 \), **\( p < 0.05 \), ***\( p < 0.01 \). Dependent variable: \( \ln(s_{m,t}) - \ln(s_{m,0,t}) \). (1) 2SLS with instrumentation of the within-nest market share and the price variables, and the three municipalities with the highest populations excluded; (2) 2SLS, with instrumentation of the within-nest market share and the price variables, low density municipalities excluded (those with fewer than 8 inhabitants per hectare, \( m=40 \)); (3) 2SLS, with instrumentation of the within-nest market share and the price variables, most northern municipalities excluded (municipalities located in the four most northern counties, \( m=26 \)); (4) 2SLS, with instrumentation of the within-nest market share and the price variables, municipalities with 10% highest share of rented dwellings excluded (\( m=13 \)).