UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Identifying firm-level costs in District Heating when firms can endogenously exploit variation in competitive intensity across product markets

Bachelor Thesis in Industrial and Financial Management

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Spring Semester 2016

Supervisor: Magnus Söderberg

Joakim Hiramsson 911230
Fredrik Nihlmark 901213

Abstract

Price constraints in the Swedish district heating (DH) market were removed in 1996 and since then there has been an ongoing debate about whether price regulation should be re-introduced or not. One of the problems for policy makers is that they have lacked important economic measures such as reliable data on firm-level costs. In this study, total firm-level cost is estimated controlling for that DH firms can endogenously exploit variation in competitive intensity across the DH and electricity markets. Since the electricity market is subject to more intense competition, firms that provide both DH and electricity have incentives to strategically allocate costs from the electricity operation to the DH operation. This suspicion invalidates descriptive comparisons of costs based on the statistics provided by the Swedish Energy Markets Inspectorate. With a nearest-neighbor covariate matching algorithm it is revealed that firms that provide both DH and electricity inflate their DH costs with on average 16.00 %. This strategic cost manipulation is fully passed on to consumers who therefore are forced to subsidize the firms' aggressive supply of electricity. To circumvent this problem, policy makers are advocated to price regulate firms that offer both DH and electricity.

Keywords: Competition, District heating, Price regulation, Strategic behavior, Cost allocation, Matching algorithm.

Sammanfattning

Den svenska fjärrvärmemarknaden avreglerades år 1996 och sedan dess har det pågått en diskussion huruvida prisreglering skall återinföras eller inte. Ett av problemen för beslutsfattare är att de har saknat viktiga ekonomiska värden så som trovärdig kostnadsdata på bolagsnivå. I denna studie estimeras fjärrvärmebolagens kostnadsnivåer, samtidigt som bolagens möjlighet att exploatera variation i konkurrensintensitet mellan fjärrvärme- och elektricitetsmarknaderna kontrolleras för. Eftersom elmarknaden är utsatt för högre konkurrens har bolagen som erbjuder både fjärrvärme och elektricitet incitament att strategiskt allokera kostnader från elverksamheten till fjärrvärmeverksamheten. Denna misstanke ogiltigförklarar jämförelser av kostnader baserat på beskrivande data från den svenska Energimarknadsinspektionen. Med en nearest-neighbor kovariat matchningsalgoritm avslöjas det att bolag som förser marknaden med både fjärrvärme och elektricitet blåser upp sina fjärrvärmekostnader med i genomsnitt 16,00 %. Denna strategiska kostnadsmanipulation övervältras helt på konsumenterna som tvingas att subventionera bolagens aggressiva elförsörjning. För att kringgå detta problem rekommenderas att de bolag som tillhandahåller både fjärrvärme och elektricitet underkastas prisreglering.

Nyckelord: Konkurrens, Fjärrvärme, Prisreglering, Strategiskt beteende, Kostnadsallokering, Matchningsalgoritm.

Contents

1 | Introduction

Price regulation was removed in the Swedish district heating (DH) sector in 1996 when the entire energy market was deregulated (Muren 2011; Holm 2013). Since the beginning of the 2000s there has been an ongoing debate about whether price regulation should be re-introduced or not. The firms themselves and the Government are currently of the opinion that no regulation is needed[1], whereas some recent academic investigations have pointed at the advantage of price regulation (Biggar & Söderberg 2012). However, no consensus has been reached, primarily because it has been difficult to determine how welfare was affected by the market change in 1996.[2]

The most prominent challenge has been to determine the actual cost level of DH. Some attempts have been made to shed light on this issue. For example, the cost of providing networks in low density areas has been investigated by Reidhav and Werner (2008), the distribution capital cost under different supply and demand conditions has been scrutinised by Persson and Werner (2011) and the size of the the price-cost margins has been investigated by Westin and Lagergren (2002) and by the Government (SOU 2004).

However, surprisingly little effort has been made to understand the overall, firm- or network-level, costs. One reason for this is the uncertainty about the theoretically appropriate cost function for the vertically integrated DH firms. However, even reduced form approaches are generally lacking. The studies referred to above are few and tend to focus on specific cost aspects rather than providing insights about the overall cost structure. Another reason is that several of the Swedish DH firms provide multiple utility services. This complicates the empirical identification of the pure DH cost function, and in particular since firms have incentives to allocate costs that belong to competitive sectors (e.g. electricity production) to DH where the competitive pressure is comparatively much weaker. This situation has provided policy makers with insufficient information to determine what the most appropriate design of the DH market should be. The overall purpose of this study is to investigate whether there is any need to redesign the Swedish DH market. More specific research questions are formulated at the end of this section.

[1] This question has informally been discussed with several CEOs of DH firms.

[2] The average Swedish household spends approximately 4 % of its total expenditures on electricity and other fuels (excluding fuels for transportation). This corresponds to about 15 % of the total cost of housing. This information is available at Statistics Sweden's website. Last visited: 2016-05-20. Available at: http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__HE__HE0201__HE0201A/HUTutgift1/?rxid=a9b94be2-f057-4c90-8d47-34d8b6228c81

DH firms can be categorized in three groups: 1) those that only provide DH, 2) those that provide DH and other monopolized public services (water and sewerage, waste disposal and collection, recycling handling and metropolitan networks) and 3) those that provide both DH and electricity. Firms in the third group sometimes also have supplementary services similar to those in group 2. An intuitive way to determine the cost level of DH for all groups is to estimate a cost function based on data from group 1 and to use that function to extrapolate the cost to the other two groups. However, that approach is only valid if firms in the three groups have structurally similar cost characteristics. This restriction is likely to be violated if firms endogenously determine which group they want to belong to, e.g. that larger firms or that firms that are located in warmer areas have cost advantages when it comes to coordinating multiple services. Regardless of the precise nature of the group selection mechanism, it is clear that group selection is endogenous and to simply extrapolate cost values based on the assumption of random group selection is unrealistic. To circumvent this problem a matching algorithm is used, where firms in group 2 and 3 are matched to firms in group 1 to determine their cost levels.

Conceptually, the research begins by specifying a general, reduced-form cost function for DH and this function is used to estimate a cost model for every firm group. Next, based on the average cost levels of firms only providing DH, the true average cost levels of DH for firms that offer both DH and other utility services are determined using the matching algorithm. With this information at hand the following research questions can be answered:

1. What is the relative cost efficiency between single and multiple service providers?

2. Do firms that provide both DH and electricity act strategically and allocate costs from the electricity operation to the DH operation?

3. Are the potential cost differences also reflected in the average price-cost margin - across firm groups and time?

The rest of the report is structured as follows. Chapter 2 explains how DH works and discusses in more detail what affects the cost of providing the DH service. Chapter 3 describes the relevant economic theories. Chapter 4 explains the methodology with a particular focus on the matching algorithm as a way to solve the non-random allocation of firms in the three groups. Chapter 5 includes the econometric results. Chapter 6 discusses the methodology at some length, whereas Chapter 7 summarizes the conclusions and recommendations.

2 | Background

This chapter first gives a brief introduction to the technical characteristics of DH, explaining the concepts of DH production and distribution. The second section gives an explanation about how various factors can affect the cost of DH. This knowledge is essential for the understanding of subsequent chapters, and in particular how to specify an appropriate cost model. The chapter ends with the delimitations of this study.

2.1 How District Heating Works

Figure 1 illustrates the whole DH process; from heat generation, to distribution, to usage at the consumer site.

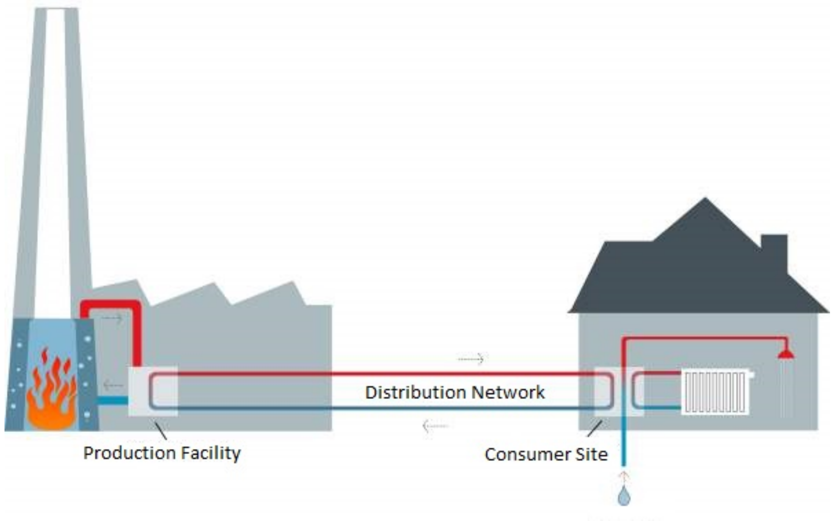

Figure 1: Illustration of the DH process.[3]

2.1.1 Production

Heat production takes place in thermal power plants where the most common way of producing heat is by burning different types of fuel, e.g. natural gas, oil, waste, coal and bio-materials (Sipilä 2015). In heat-only boiler plants, heat is the main product and in combined heat and power (CHP) plants, electricity is the main product with heat being the by-product (SOU 2005). Another way of producing heat is by using large solar panels for harvesting the sun's energy, which heats the circulating water (Pauschinger 2015). In some industrial

[3]District Heating Södertörn. Last visited 2016-04-11. Available at: http://sfab.se/Fjarrvarme/

processes, heat is produced as a by-product and if it is not possible to re-use this heat in the industrial process itself, it is sometimes fed into the DH network in the form of waste heat (SOU 2005).

2.1.2 Distribution

DH distribution systems consists of two well-isolated pipes located underground. One pipe contains the highly pressurized heated water, which has a temperature between 70 - 120 degrees Celsius. When the heated water enters a building, a heat exchanger extracts the heat from the heated water to heat the water on the consumer's side. The heated water is then fed to radiators and other systems inside the building (e.g. tap water system). The cooled water on the supplier's side is then fed back from the heat exchanger to the production plant via the other pipe to be re-heated and then re-distributed (Biggar & Söderberg 2012). On the way back to the production plant the cooled water is used for keeping pavements and football fields free of ice.[4]

2.2 Factors Leading to Cost Differences Across DH Firms

There are several factors that influence the cost of DH. These factors can vary across firms, time or both. In this section, these factors are presented in order to ease the understanding of the cost model specification.

2.2.1 Firm-Specific Factors

The different firm-specific factors are related to production and distribution, the structure of the district heating firms and geographical and social factors. Below, the most important of these factors are presented.

As described in section 2.1.1, heat can be produced in different ways in DH facilities, leading to differences in production related costs across firms. For instance, the type of fuel mix used to generate heat along with the type of power plant used are greatly affecting the cost levels of DH. Further, the age of the production facility, which is related to the technology used in that facility, is another important cost driver. The newer the technology, the higher the efficiency and, thus, the lower the costs. But older investments can be fully amortized, making them comparatively cheap to operate. Thus, the net effect of network/facility age is ambiguous. Also, the number of employees and the degree of maintenance required for the facilities and the length of the distribution network are important cost drivers of production and distribution.

[4]Swedish District Heating Association. Last visited 2016-04-12. Available at: http://www.svenskfjarrvarme.se/Fjarrvarme/Sa-funkar-fjarrvarme/

Furthermore, if the production facility is close to consumers and the consumers live close to each other, the cost structure is allowed to be more advantageous (less capital investment needed and less heat lost during distribution). Related to this is also the size of the population, i.e. the consumer base. The higher the number of consumers, the lower the size of fixed initial investment per consumer. Another firm-specific factor related to production and distribution of DH is policy initiatives. Costs are greatly dependent on fuel price trends, taxes and economic control instruments, e.g. feed-in tariffs and emission rights among others. This will affect the firms differently due to different usage of fuel mixes.[5]

The structure of the district heating firms is a very important factor that can lead to cost differences across firms and because of its relative importance, it is the cost aspect that is focused most on in this study. A firm producing DH is often included in a corporate energy group with production and sales of electricity. The energy group often has additional utility services, typically including some or all of the following: water and sewerage, waste disposal and collection, recycling handling and metropolitan networks providing broadband, television and telephony via fiber optic networks (Westin & Lagergren 2002).

This study divides DH firms into three groups based on the composition of provided services:

- DH_{only} - a firm only providing DH and no other services. These firms represent the pure cost of DH and serve as the reference group.
- DH_{multi} - a firm providing DH and some or all of the additional services described above. These other services are technically separated from DH. Therefore, it should be easier to separate costs from DH and costs from such other services.
- $DH_{multiEL}$ - a firm providing DH, electricity production and additional services like the DH_{multi}-group. The production techniques of DH and electricity are often coupled because heat is produced as a by-product of electricity, as described in section 2.1.1 (SOU 2005). This greatly complicates the allocation of cost components to the right part (DH or electricity). This directly also hampers the possibility to evaluate the market conduct (SOU 2003).

[5] Vattenfall. Last visited 2016-04-12. Available at: http://corporate.vattenfall.com/about-energy/energy-distribution/district-heating/pricing/

There are several reasons to believe that the costs of DH differs across these three groups. Such differences can occur because of, for example, coordination costs, shared administrative costs, easier access to some fuel types (e.g. waste) and strategic allocation of costs. Cost allocation and strategic behavior are described in closer detail in section 3.3.

Knowledge about the consumer's situation in the market is necessary to understand how the firms' strategic behavior might affect the consumer. When the consumer is about to choose a heating technology, the consumer faces a large initial investment. At this stage, DH along with other heating technologies such as thermal heating and electric heating are competing with each other. Once the consumer has chosen DH as heating technology and the investment is sunk, the consumer is locked-in to the technology for a considerable amount of time, often up to 20 years (Hepbasil & Kalinci in press; Shah, Col Debella, & Ries 2008).[6] During this time the consumer cannot change from DH to another heating technology and has no option but to pay the price of heating via DH. The consumer is thus exposed to the risk of increasing prices, caused by strategic behavior of the DH firm. So, even though DH might look like an attractive heating alternative at the time of the investment, it is subject to the risk of increasing prices over time (SOU 2003).

The above mentioned is thus related to a social factor that affects the way the DH firms can operate, therefore indirectly affecting the costs. Another example of a social factor is the attitudes towards the DH system (strong support for the green party can lead to favors for DH), which can also affect the way the DH firms can operate. There are different factors related to the geographical location of the production facility that create cost differences across firms. The production facilities throughout the country are exposed to different topography along with soil and climate conditions, such as temperature and height above sea level.[7]

2.2.2 Year-Specific Factors

Year-specific factors are factors that affect the costs of DH due to an event that was specific at a certain year. Elections, storms, introduction of a new law among other things are examples of year-specific factors.

[6] If the consumer was to choose another heating alternative, the consumer would be locked to that technology as well.

[7] Swedish Energy Market Inspectorate. Last visited 2016-04-15. Available at: http://ei.se/en/District-heating/Factors-that-affect-the-price-of-district-heating1/

2.3 Delimitations

In this study it has been assumed that DH firms minimize their costs across all business areas within the firm, meaning that the effect of ownership structure of the firms is disregarded. Furthermore, the cost function (specified in 4.2) only take the delivered heat (Q) and the total length of the distribution network (L) into consideration. The rest of the cost affecting factors are included in the firm- and year-specific fixed effects. Finally, complete data for the Swedish DH firms was only available for the time period of 2009 to 2014.

3 | Theory

This chapter first gives a quick introduction to why the Swedish DH market is regarded as a typical example of a natural monopoly. The first section also describes the pricing process and how natural monopolies can be regulated. The following section describes some economic cost theory. The last section describes how providing multiple utility services in a DH firm may give rise to strategic behavior among firms in the DH industry.

3.1 Natural Monopolies

Providing certain services require large fixed costs. Such fixed costs occur, for example, when large distribution or other infrastructure networks are needed to provide the service. Every firm that wants to offer such a service to consumers must construct its own network and pay these fixed costs. From an economic point of view, it is not feasible to build multiple parallel networks since each network requires payment of the fixed costs. Therefore, it is more efficient if one supplier supplies the whole market demand. By doing so, one firm can spread its large fixed costs over the whole demanded market quantity. The supplier is therefore able to supply the market demand at a decreasing average cost. This market situation is referred to as a natural monopoly (Currier 2004).

The DH market is often referred to as a natural monopoly because of the large fixed costs needed to build the production and distribution network (Westin & Lagergren 2002; Sjödin & Henning 2004).[8]

The monopolist has superior control over the market price. A profit maximizing monopolist will produce less output at a higher price compared to a situation where the firms are competing with each other. The monopolist is able to take advantage of its dominant market position and makes an excessive economic profit by setting the price at the monopoly price. This monopoly price is inefficient and causes losses in welfare, known as dead weight losses. Natural monopolies are often subject to some kind of price regulation, that puts an upper limit on the price. This reduces the dead weight loss in society and it also reduces the monopolist's economic profit. As for the DH market in Sweden, it has been difficult to determine whether the DH firms set their price of DH at the monopoly price, since the true cost levels of DH is not known for firms providing multiple utility services. For a further explanation regarding monopoly

[8] Tschirhart (1996) mentions electricity and natural gas utilities as other examples of natural monopolies.

pricing and price regulation, the reader is referred to Mankiw (2011, ch. 15) and Hirschey (2003, ch. 11,13).

3.2 Cost Theory

This section puts relevant economic cost theory into the context of DH, which is important for the motivation and understanding of the cost function that is used in subsequent sections.

3.2.1 Economies of Scale

Assuming that a DH firm only provides heat that is denoted Q. The firm's average cost is calculated as

$$AC = \frac{FC + VC \cdot Q}{Q} \tag{1}$$

where FC is total fixed cost and VC is variable cost. If the average cost decreases as the output quantity increases, the firm enjoys economies of scale. For further information the reader is referred to Mankiw (2011, pp. 272-273).

3.2.2 Economies of Scope

Consider a DH firm that provides two services, DH and electricity. The cost of providing DH and electricity are denoted $C(DH)$ and $C(EL)$, respectively. The cost of providing DH and electricity jointly is denoted $C(DH, EL)$. According to Kwoka (2002), economies of scope is experienced if the following inequality holds

$$C(DH, EL) < C(DH) + C(EL) \tag{2}$$

Here only DH and electricity are considered, but of course, Equation 2 can be generalized into jointly providing N different services together with DH. The generalized form of Equation 2 is

$$C(DH, S_1, S_2, ..., S_N) < C(DH) + C(S_1) + ... + C(S_N) \tag{3}$$

where $C(S_1) + ... + C(S_N)$ denotes the cost of providing services $S_1, ..., S_N$ separately. Coordination costs and shared administrative costs are both factors of economies of scope.

3.2.3 Economies of Vertical Integration

Economies of vertical integration is similar to the concept of economies of scope, but instead considers the integration or separation of production and distribution. A firm only producing DH experiences certain production costs, denoted $C(P)$. A second, separate, firm responsible for distribution of DH experiences certain distribution costs, denoted $C(D)$. In the case of DH firms, integrating production and distribution results in costs denoted $C(P, D)$. Economies of vertical integration is experienced if the following inequality holds (Kwoka 1996)

$$C(P, D) < C(P) + C(D) \tag{4}$$

3.2.4 Cost Efficiency

Consider three firms X, Y and Z. Firm X only provides DH and firms Y and Z provides DH and additional services. Figure 2 is used to illustrate the differences in cost efficiency among the firms. Firm X provides DH at an average cost of AC_X. Firm Y provides DH at an average cost AC_Y and firm Z provides DH at an average cost AC_Z.

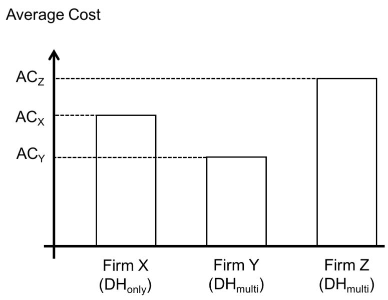

Figure 2: Illustration of cost efficiency. Firm Y is more cost efficient than firm X and firm Z is less cost efficient than firm X.

In Figure 2 firm X provides DH at a higher average cost than firm Y. Thus, firm Y is more cost efficient than firm X. Since firm Y is a multiple utility services company, jointly providing multiple services reduces the average cost of DH. Therefore, firm Y experiences economies of scope. For firm Z the opposite is true, i.e. firm X is more cost efficient than the multiple utility service firm Z. Hence, firm Z experiences diseconomies of scope.

Cost efficiency of firm Y is calculated according to

$$\text{Cost Efficiency of Firm Y (\%)} = \frac{AC_X - AC_Y}{AC_X} \cdot 100 \tag{5}$$

Similarly, the cost efficiency of firm Z is calculated by the use of Equation 6

$$\text{Cost Efficiency of Firm Z (\%)} = \frac{AC_X - AC_Z}{AC_X} \cdot 100 \tag{6}$$

3.3 Cost Allocation and Strategic Behavior

In CHP plants, electricity and heat are produced at the same time. This implies that the production processes of electricity and DH are linked closely together. Therefore, it is hard to distinguish which costs are caused by which production process. This greatly reduces the cost transparency of firms utilizing the CHP technology for production of electricity and heat. This is a problem in itself because it greatly complicates the identification of the pure cost function of DH. Further, firms owning CHP plants are often members of a corporate energy group, which allows for the allocation of costs among different business areas within the energy group. This enhances the problem of insufficient cost transparency. A production technology with low cost transparency in combination with the possibility to allocate costs between different business areas might give the DH firms incentives to act strategically in order to maximize its total profits (SOU 2003).

The electricity market is highly competitive since the consumers are free to choose any supplier they want. An electricity supplier therefore needs to offer competitive prices in order to be able to sell the electricity it produces. Compared to the electricity market, the DH market is much less competitive. The possibility of allocating costs between business areas within the firm's energy group provide firms with great opportunities to allocate costs from the electricity operation to the DH operation. By doing so, firms that normally have an inefficient and non-competitive electricity production would be able to offer a competitive electricity price. The DH consumers would then subsidize inefficient electricity production, or in other words, firms selling both electricity and heat produced in CHP plants would be competitive on the electricity market by using its dominant position in the DH market. The problem for the DH consumers is that they are locked-in to the DH technology once they have invested in it. This leaves them unprotected to future price increases (SOU 2003).

The above described allocation process might also be possible in cases when firms provide DH together with additional services. But because other services are not coupled to the production of DH in the same way as electricity is in CHP plants, the concept is easier to grasp in the context of DH and electricity.

Figure 3 illustrates how strategic behavior in the form of cost allocation affects the average cost of DH. In Figure 3, firm U currently provides DH at an average cost of AC_U. Firm U then decides to lower its electricity price and in order to do so, it allocates some of the costs of electricity production to DH. As illustrated in Figure 3, this will increase the average cost of DH and will also be reflected in a higher DH price.

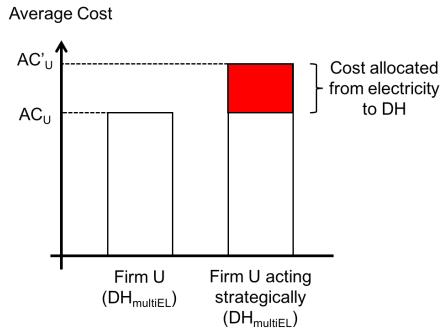

Figure 3: Illustration of the allocation process.

In order to investigate strategic behavior among DH firms, their reported average cost is compared to an imputed average cost, reflecting the true cost of DH. The imputation process is further described in section 4.4. If the reported average cost is higher than the imputed average cost, there is a reason to believe that firms engage in this type of strategic behavior.

Strategic behavior is measured according to

$$\text{Strategic Behavior (\%)} = \frac{AC'_U - AC_U}{AC_U} \cdot 100 \tag{7}$$

4 | Methodology

This section describes the methodology. Figure 4 illustrates the different steps of the workflow in which this study has been carried out. In section 4.1 the data extraction and treatment is explained. A quadratic generic cost function for DH is specified in the following section. In section 4.3 the implementation of a multiple regression analysis in order to establish cost models is described. Further, a brief explanation of how a matching algorithm was implemented to impute average cost levels and conceptually how it works along with a motivation of choice is presented in section 4.4. The last section 4.5, provides information on how firm behavior is estimated based on cost efficiencies, strategic behavior and price-cost margins.

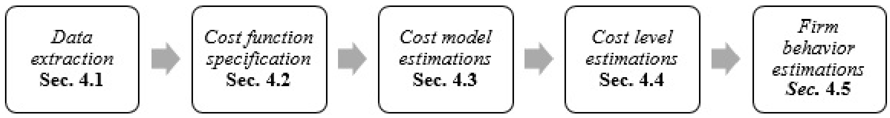

Figure 4: The workflow of this study.

4.1 Data Extraction

This study is based on a data set that was extracted from the Swedish Energy Market Inspectorate's (SEMI) website.[9] The data consists of observations between year 2009 and 2014. Each observation corresponds to one DH firm in one year. The relevant variables used are:

- Total heat delivered by each firm and year, Q $[MWh]$.
- Total length of the DH distribution network for each firm and year, L $[km]$.
- Average cost for each firm and year, AC $[tSEK/MWh]$. The average cost is defined as the total cost divided by Q, adjusted for inflation.[10]
- The average price for all firms at each year, P $[tSEK/MWh]$, which is defined as total operating revenues divided by Q, adjused for inflation.

The first three variables are important since they are included in the cost function that is specified in the following section. To categorize firms into the three different groups that were described in section 2.2.1, two dummy variables are introduced:

[9]The extracted data is available at the Swedish Energy Market Inspectorate's website. Last visited: 2016-04-08. Available at: http://ei.se/.sv/Fjarrvarme/inrapporterad-data/

[10]Information about CPI is available through Statistics Sweden's website. Last visited 2016-04-08. Available at: http://www.scb.se/sv_/Hitta-statistik/Statistik-efter-amne/Priser-och-konsumtion/Konsumentprisindex/Konsumentprisindex-KPI/

- No_elec_sales, which has a value of 1 if the company only produces DH in the production facility, otherwise 0.
- Multi_serv, which has a value of 1 if the firm offers any other services than DH, such as water and sewerage, waste disposal and collection, recycling handling and metropolitan networks. If a firm only provides DH, the value of Multi_serv is 0. If a firm produces both DH and electricity, No_elec_sales will be 0 and Multi_serv will be assigned a value of 1.

Consequently, these two variables allow an identification of three groups of firms. First are the firms that only produce DH and offer no other service. This group is referred to as the DH_{only}-group. Second are the firms that produce DH and offer other services but not electricity (the DH_{multi}-group) and third is the group that produces both DH and electricity, and possibly other services, (the $DH_{multiEL}$-group).

The information about what services firms offer was collected from SEMI's database, firm websites and annual reports. Further, all observations were checked manually and eliminated if it had clearly unreasonable values. Lastly, the 5 % smallest and the 5 % largest DH firms in terms of yearly average delivered heat between 2009 and 2014, were eliminated from the data set. The 5 % smallest firms produced on average 5.6 GWh/year, while the 5 % largest produced 2500 GWh/year. They are considered to be too different from the sample average, which is 252 GWh/year. This leaves a total of 942 complete observations, where the three different firm groups consist of 382 observations (DH_{only}-group), 325 observations (DH_{multi}-group) and 235 observations ($DH_{multiEL}$-group), respectively.

4.2 District Heating Cost Function Specification

One cost function that performs well in relation to the general cost function criteria[11] is the quadratic cost function. Equation 8 states the quadratic cost function when both Q and L are considered to be production outputs.

$$TC = \alpha_0 + \alpha_1 Q + 0.5\alpha_2 Q^2 + \alpha_3 L + 0.5\alpha_4 L^2 + 0.5\alpha_5 QL \tag{8}$$

The corresponding average cost can be specified by dividing Equation 8 by Q, yielding

$$AC = \beta_0 + \beta_1 \frac{1}{Q} + 0.5\beta_2 Q + \beta_3 \frac{L}{Q} + 0.5\beta_4 \frac{L^2}{Q} + 0.5\beta_5 L \tag{9}$$

[11] For more information about the cost function criteria, the reader can consult Baumol, Panzar and Willig (1982).

The quadratic cost function in Equation 9 generally outperforms most other cost functions, including the commonly used translog[12] function (Färe, Martins-Filho, & Vardanyan 2009; Shaffer 1998). The translog function is unsuitable because it is unable to fulfill many of the cost function criteria defined in Baumol, Panzar, and Willig (1982). One particular problem with the translog cost function is that it involves the logarithm, which is undefined for zero. Therefore, the translog cost function is unable to handle zero output in one of multiple production outputs (Kwoka 1996; Shaffer 1998). This makes the translog cost function unsuitable to use for modeling specialized production, which is required in order to estimate economies of scope or product specific economies of scale (Pulley & Braunstein 1992). Hence, the quadratic cost function is better suited to model multiple production outputs and multiple production technologies (Pulley & Braunstein 1992; Kwoka 1996; Röller 1990). Also, the fact that the translog cost function involves the logarithm makes it less suitable to handle certain fixed effects, where, for some observations, the fixed effects variables must equal zero (Kwoka 2005).

Looking at the cost function in Equation 9, the average cost of DH is assumed to be determined by the costs of production and distribution, proxied with the output factors Q and L, respectively. It is expected that the costs of production and distribution are affected by the input prices of fuel, capital and labour but these input prices are not considered in Equation 9, for various reasons. Including the input price of fuel in the cost function requires complete information about the type of fuel and the fuel mix used by each individual firm. Even though it is assumed that there is a national price for each type of fuel, it would be very hard to find information about the price of bio-materials and waste since they are not traded on a market in the same way as, for example, oil is. Besides, even if this information was available it would still require firms to provide complete information about current fuel mixes. Since DH firms were not required to provide this information throughout the sample period, it is not possible to include the average input fuel price in the cost function. Excluding the price of capital and labour is less problematic. Sweden has a national capital market and therefore the price of capital does not vary across firms. The cost of capital will therefore be captured by the year-specific fixed effects described in section 4.3. This is also the case with the price of labour, since it does not vary considerably across the short time period considered in this study. Thus, it will be captured by the firm-specific fixed effects (Biggar & Söderberg 2012).

[12] For a brief introduction to the translog cost function, the reader is referred to Ch. 4 in Kwoka (1996).

4.3 District Heating Cost Model Estimations

With a complete data set and a specified cost function, the third step of the workflow could be processed. The statistical package STATA was used to perform all analyses.[13]

To account for unobserved, firm-specific fixed effects, a fixed effects linear model was applied when using multiple regression analysis.[14] Further, to cope with the year-specific factors described in section 2.2.2, the effects that are similar for all firms during the same year are controlled for using year-specific effects.[15] The STATA code used to estimate the cost models for all three groups of DH is included in Appendix A.1.

4.4 District Heating Cost Levels Estimations

Since there is reason to believe that DH firms act strategically as discussed in section 3.3, the average cost levels reported to SEMI by the DH firms providing multiple utility services are not considered reliable. Therefore, it is necessary to identify the true average cost of DH for these firms, i.e. what the actual average cost of DH would be if the firms only provided DH. This can easily be done by extrapolating these costs based on the cost model for the firms only providing DH. This produces, however, unreasonable results since this extrapolation is only valid if firms in the three groups have structurally similar cost characteristics.[16] This restriction is likely to be violated if firms endogenously determine which group they want to belong to. In this case it is clear that group selection is endogenous because firms themselves decide which services they provide. To simply extrapolate cost values based on the assumption of random group selection is therefore unrealistic. To predict more reliable results, it is desirable to base the true average costs of DH for the firms offering multiple services on the average costs of DH for firms only providing DH that are similar to each other in terms of Q and L. This can be accomplished by the use of a matching algorithm. In particular, STATA was used in this study to apply a nearest-neighbor covariate matching algorithm to impute the average cost levels of DH for multiple service firms (DH_{multi}- and $DH_{multiEL}$-group) based on the average cost

[13]For further information, visit STATA's website. Last visited: 2016-04-18. Available at: http://www.stata.com/why-use-stata/.

[14]For deeper knowledge about fixed effects, the reader is referred to Blumenstock, J. Last visited: 2016-04-18. Available at: http://www.jblumenstock.com/files/courses/econ174/FEModels.pdf.

[15]For deeper knowledge about year effects, the reader is referred to Darthmouth College. Last visited: 2016-04-22. Available at: https://www.dartmouth.edu/ ethang/Lectures/Class17/Always%20Control%20for%20Year%20Effects%20in%20Panel%20Regressions.pdf.

[16]The extrapolation gives negative average costs for some firms and unreasonably high costs for other firms. Therefore, these results are considerable unreliable.

levels of DH for firms only providing DH (DH_{only}-group). A brief explanation of the concept of this matching algorithm and why it was chosen is provided below.

This study consists of a sample of DH firms, where some receive a treatment and some do not. In this case two different treatments could be observed, namely (1) the treatment of not producing electricity and (2) the treatment of offering multiple services. For the first treatment, the treated group (that received the treatment) comprises of the DH_{only}-firms and the control group (that did not receive the treatment) comprises of the $DH_{multiEL}$-firms. Similarly, for the second treatment, the treated group comprises of the DH_{multi}-firms and the control group comprises of the DH_{only}-firms. It would be interesting to determine what the actual average cost levels of DH would be for the $DH_{multiEL}$-firms if it received the treatment (did not produce electricity) and for the DH_{multi}-firms if it did not receive the treatment (did not offer multiple services). These average cost levels are in general referred to as treatment effects, or the potential outcome for each firm group. Matching algorithms can be used for this type of analysis and there are numerous algorithms available. As mentioned above, the nearest-neighbor covariate matching algorithm was used in this investigation to impute these average cost levels.

Nearest-neighbor covariate matching is a type of matching algorithm that estimates the treatment effect from observational data by imputing the missing potential outcome (i.e. in this case what the average cost would be if it did or did not receive the treatment) for each unit (firm in this case) by using an average of the outcomes of similar units in the opposite group. The similarity between units is based on a distance metric known as the Mahalanobis distance, which is given as a weighted function of the covariates.[17] The covariates in this case are defined as the variables affecting the average cost, which are the ones included in the quadratic cost function corresponding to Equation 9. The matching is done with replacement, which means that an untreated subject can be used as a match more than once. The reason for this is to increase the matching quality and decrease the estimator bias (Caliendo & Kopeinig 2003). Further, the matching is done only with one nearest-neighbor (one-to-one matching), which means that the most similar unit in the treatment group will directly determine the potential outcome of the unit in the control group and vice versa. According to Caliendo and Kopeinig (2003) it is advisable to match to more than one nearest-neighbor, but this involves a trade-off between variance and bias.

[17]To read more about treatment effects and nearest-neighbor matching in STATA, the reader is referred to the STATA manual. Last visited: 2016-05-04. Available at: https://www.stata.com/manuals13/te.pdf.

Matching with one single nearest-neighbor is associated with a decreased bias at the expense of an increased variance. This is however desirable in this case, since there is a risk due to a relatively small sample size that the second or even third closest neighbor is too far away to be reasonably classified as a similar firm. Consequently matching multiple neighbors would more likely increase the bias, thus lowering the quality and reliability of the result. In Appendix A.2 the STATA code is provided for how to impute the potential outcomes, i.e. average cost levels, for the DH firms providing multiple utility services.

Methods based on matching algorithms have become widely used to estimate casual treatment effects. There are several known studies where it has been applied. Deheijia and Wahba (1999) uses a propensity score matching algorithm to estimate the treatment effect on a labour training program on postintervention earnings. Gerfin and Lechner (2002) also use a propensity score matching algorithm to evaluate whether an introduced labour market policy entailing several programmes in Switzerland had any effects on the individual employment probability. However, in this study, a covariate matching algorithm is used instead. The difference between these two algorithms is that the covariate matching algorithm uses a distance metric based on covariate values to determine the similarity of two units and matching is based on the distance between those two units. The propensity score matching algorithm, on the other hand, uses estimated predicted probabilities to establish propensity scores and units with similar propensity scores are matched (Zhao 2003). The reason why the covariate matching algorithm was chosen was because it was crucial that the matched firms are as similar as possible in terms of the covariates.[18] This way the firms with similar total lengths of the distribution network and delivered heat is matched. Abadie and Imbens (2004) have shown that nearest-neighbor covariate matching algorithms are not consistent when matching on two or more continuous covariates, but this can be mitigated by using a bias-corrected estimator.[19] However, in this investigation, using a bias-corrected estimator produced illogical and unreasonable average cost levels, why this correction was not applied.

[18] The propensity score matching algorithm was tried once, but produced highly unlikely results. According to Zhao (2003) the most suitable matching algorithm depends on the characteristics of the data. In this case, the characteristics of this data were judged to be more appropriate for a covariate matching algorithm.

[19] For a detailed mathematical exposition of the bias-corrected estimator, see Abadie and Imbens (2011).

4.5 Firm Behavior Estimations

With two complete data sets of average cost levels of DH, namely SEMI's data and the imputed data calculated according to the previous section, the last step of the working process could be processed. This involves calculating the net cost effect, considering cost efficiencies, strategic behavior and the development of price-cost margins over time. In the following sections, the estimation of these values are explained.

4.5.1 Estimation of Cost Efficiencies

To determine cost efficiencies, the yearly average of the average costs for each firm group between 2009 and 2014 were calculated based on SEMI's data. Then a total average of all the years for each firm group was calculated. It is this total average cost for each firm group that is used in Equations 5 or 6 to estimate the cost efficiencies for DH_{multi}- and $DH_{multiEL}$-firms. Cost efficiencies were also calculated based on the imputed data.

4.5.2 Estimation of Strategic Behavior

For the imputed data, the exact same calculations were made as for the SEMI-data when making calculations for the cost efficiency. From this the strategic behavior was calculated by the use of Equation 7 for each of the firms providing multiple utility services.

4.5.3 Estimation of Price-Cost Margins

As for the price-cost margins, they were established in the same way as the yearly average of the average costs for each firm group and for both data sets. From this the development of the price-cost margin between 2009 and 2014 based on the two data sets could be plotted. By comparing these plots, it is possible to determine whether some firms use their position to act strategically and increase their margins. It also shows to what extent cost differences are absorbed by the firms or passed on to consumers. Furthermore, the price-cost margin was calculated for each observation to enable a distribution plot, which illustrates the spread for the different firm groups.

5 | Results

This chapter presents the results of the study and it starts by presenting the results from the multiple regression analyses, one for each group of DH firm. The second section of the chapter presents the results from the matching algorithm. The chapter ends with an estimation of firm behavior with regards to cost efficiency, strategic behavior and price-cost margins.

5.1 District Heating Cost Model Estimations

Table 1 shows the results from the estimations of Equation 9 for the three different groups of DH firms.

Table 1: Output from the multiple regression analysis of each DH firm group.

Variable	DH_{only} Mean (S.E.)		DH_{multi} Mean (S.E.)		$DH_{multiEL}$ Mean (S.E.)	
$Const.$	0.4174 (0.0327)	***	0.5450 (0.0364)	***	0.5437 (0.0483)	***
$\frac{1}{Q}$	4471.2490 (783.327)	***	1136.967 (1074.294)		14625.25 (5739.401)	**
$0.5Q$	-1.55e-6 (6.28e-7)	**	-1.12e-6 (4.99e-7)	**	-1.36e-7 (2.95e-7)	
$\frac{L}{Q}$	-1.0482 (1.3201)		31.918 (23.565)		-13.3318 (31.1366)	
$0.5\frac{L^2}{Q}$	1.865e-4 (3.065e-4)		0.0343 (0.2729)		0.4257 (0.2479)	*
$0.5L$	5.75e-5 (5.79e-5)		-5.625e-4 (7.624e-4)		-2.433e-4 (2.975e-4)	
R^2	0.2121		0.2100		0.1680	
No. of obs.	382		325		235	

Notes: Significance levels *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Table gives mean values of each coefficient with standard error in parenthesis. Year- and firm-specific fixed effects were included in each regression.

The results in Table 1 reveal that the cost models for the three DH firm groups are different. This implies, for example, that a simple extrapolation based on DH_{only} would be invalid. For all DH firm groups, the quadratic cost function specified in Equation 9 is able to capture a sizeable part of the total variation of the average cost - the within R^2 ranges from 16.80 % to 21.21 %. It is assumed that the remaining cost drivers discussed in section 2.2 are controlled for by the firm- and year-specific fixed effects. This assumption appears to be acceptable given that many demand and supply factors are driven by long-term decisions and persistent behavior that only changes slowly over time, or when new policies are introduced at the national level.

5.2 District Heating Cost Level Estimations

The results from the treatment effects matching algorithm are shown in Table 2.

Table 2: The treatment effects after applying the matching algorithm.

Treatment Effect	*No elec. sales*		*Multi. serv.*
Treatment Group	DH_{only}		DH_{multi}
Control Group	$DH_{multiEL}$		DH_{only}
	Coeff. (S.E.)		Coeff. (S.E.)
Average Cost	-0.1086 *** (0.0127)		0.0122 (0.0083)
No. of obs.	622		707
No. of Nearest Neighb.	1		1

Notes: Significance level *** $p < 0.01$. Table gives the change in mean average cost, in $tSEK/MWh$, with standard error in parenthesis. Matching is done with replacement.

The results show that $DH_{multiEL}$-firms report on average 0.1086 $tSEK/MWh$ higher average DH cost than the DH_{only}-firms. This result is significant at the 1 % level.[20] It was expected, and found, that joint production of electricity and DH lowers the actual average cost of DH, which means that this firm group enjoys economies of scope. Hence, this result raises suspicions about cost allocation and strategic behavior among firms providing both electricity and DH. The treatment of adding multiple services to firms producing only DH (i.e. with no electricity production) increases the average cost level of DH by 0.0122 $tSEK/MWh$, on average. But this result is not significant at the 10 % level. Thus, based on this result, there is no indication of DH_{multi}-firms being, on average, engaged in cost allocation and strategic behavior. This result is reasonable since, compared to electricity, production of other utility services are not closely linked to the production of DH, which increases cost transparency of

[20]To ensure that it is the electricity production that contributes to this treatment effect, the DH_{multi}- and $DH_{multiEL}$-firms are matched with regard to the treatment of not producing electricity in Table B.1 in Appendix B.

firms providing multiple utility services. Also, services provided by the DH_{multi}-firms are not traded on markets that are clearly more competitive than the DH market. Their incentives to act strategically are thus weaker than what they are for $DH_{multiEL}$-firms. From these treatment effect estimations, the imputed true average cost levels for all three DH firm groups are calculated. These costs are summarized in Table 4.

5.3 Firm Behavior Estimations

This section provides results about firm behavior estimation by calculating cost efficiencies between firms only providing DH and firms providing multiple utility services, strategic behavior of the firms providing multiple utility services and price-cost margins for all DH firm groups.

5.3.1 Estimation of Cost Efficiencies

Table 3 presents the yearly average cost levels of DH for each firm group. It also gives the total average cost per year from 2009 to 2014 for each DH firm group. The average cost levels in Table 3 are based on the costs collected by SEMI.

Table 3: The average cost levels of DH for every DH firm group based on SEMI's data.

Year	DH_{only}	DH_{multi}	$DH_{multiEL}$
2009	0.5323	0.5077	0.5938
2010	0.5186	0.5168	0.5788
2011	0.5457	0.5403	0.6094
2012	0.5175	0.5359	0.5921
2013	0.5353	0.5604	0.5932
2014	0.5530	0.5494	0.6176
TOT.AVG	0.5337	0.5351	0.5975

Notes: Table gives average cost in $tSEK/MWh$. The last row gives the average of years 2009 to 2014.

Based on Table 3, cost inefficiencies of the DH_{multi}- and $DH_{multiEL}$-firms are 0.26 % and 11.95 %, respectively. This reveals that DH_{multi}- and $DH_{multiEL}$-firms provide DH more inefficiently than DH_{only}-firms, which is associated with diseconomies of scope. This result gives further indications of $DH_{multiEL}$-firms being engaged in strategic cost allocation.

Table 4 presents the yearly average of the average cost of DH for each firm group per year from 2009 to 2014 based on the imputed average cost levels given from the treatment effects from the matching algorithm in section 5.2. Table 4 also includes the total average of the yearly average cost for each DH firm group. Based on the imputed cost levels, DH_{multi}-firms are 0.34 % more cost efficient than the DH_{only}-firms, meanwhile the $DH_{multiEL}$-firms are 3.50 % more cost efficient. This implies that these DH firm groups are enjoying economies of scope. However, only the result for the $DH_{multiEL}$-firms is significant.

Table 4: The average cost levels of DH for every DH firm group based on imputed cost levels.

Year	DH_{only}	DH_{multi}	$DH_{multiEL}$
2009	0.5323	0.5162	0.4923
2010	0.5186	0.5288	0.5177
2011	0.5457	0.5258	0.5218
2012	0.5175	0.5304	0.5234
2013	0.5353	0.5421	0.5120
2014	0.5530	0.5481	0.5236
TOT.AVG	0.5337	0.5319	0.5151

Notes: Table gives average cost in $tSEK/MWh$. The last row gives the average of years 2009 to 2014.

5.3.2 Estimation of Strategic Behavior

The imputed data represents the true average costs for providing DH and to provide further arguments for or against cost allocation, it is desirable to compare reported and imputed average cost. Figure 5 provides such an illustrative comparison between reported and imputed average cost levels for all three firm groups. The bars in Figure 5 are based on values given by Tables 3 and 4.

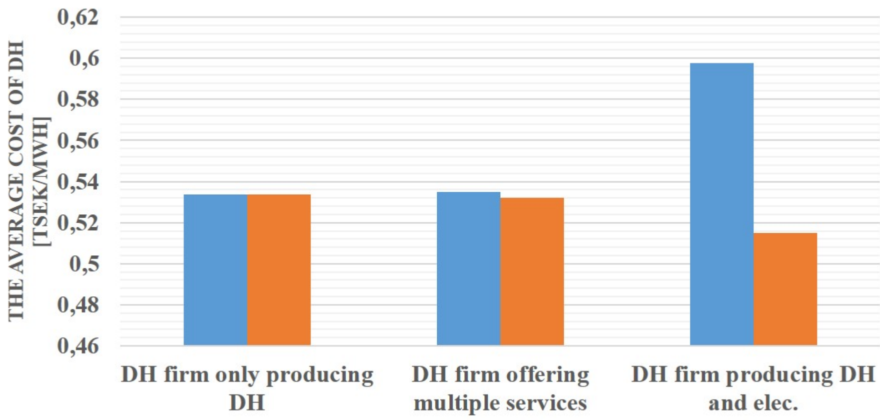

Figure 5: Comparison between SEMI's data (blue) and imputed data (red) of the total average cost levels of DH for each DH firm group.

In Figure 5, it is possible to see that, based on reported average cost values, DH_{multi}- and $DH_{multiEL}$-firms are less cost efficient than firms only providing DH. It also illustrates that the opposite is true for the imputed average cost levels; DH_{multi}- and $DH_{multiEL}$-firms are more cost efficient than DH_{only}-firms. This difference clearly reveals that firms providing multiple services are reporting average costs of DH that are, on average, higher than the true average cost of DH. This difference between reported and imputed average cost levels is likely to come from strategic cost allocation. The strategic behavior of $DH_{multiEL}$-firms is 16.00 %, meaning that $DH_{multiEL}$-firms, on average, strategically report 16.00 % higher average costs to SEMI and to their consumers. For DH_{multi}-firms, the strategic behavior is 0.60 %. However, only the result for $DH_{multiEL}$-firms is significant. There is no difference between reported and imputed average cost levels for DH_{only}-firms. This is due to that the imputation is based on the reported average cost levels for DH_{only}-firms. Thus, reported and imputed average cost levels are the same for DH firms only providing DH.

5.3.3 Estimation of Price-Cost Margins

To further investigate the strategic behavior of DH firms, the average price-cost margins are calculated and compared over time for both reported and imputed data, for all firm groups. Table 5 shows the average price-cost margin for all firms and for each firm group per year, from 2009 to 2014 based on SEMI's data.

Table 5: The average price-cost margins for all DH firms and for each DH firm group based on SEMI's data.

Year	*All Firms*	DH_{only}	DH_{multi}	$DH_{multiEL}$
2009	1.0726	1.0695	1.0682	1.0858
2010	1.0653	1.0695	1.0461	1.0877
2011	1.0407	1.0457	1.0291	1.0491
2012	1.0795	1.1051	1.0493	1.0796
2013	1.0772	1.1022	1.0428	1.0819
2014	1.0934	1.1069	1.0861	1.0838

Notes: Table gives price-cost margin in percent in decimal form.

With the values in Table 5, the development of the reported average price-cost margin can be illustrated. Figure 6 illustrates this development for all firms and for each firm group, respectively. The price-cost margins for the different firm groups follow the same trend and the margins are roughly of the same magnitude. On average, the DH_{only}-firms have a price-cost margin of 8.32 %, the DH_{multi}- and $DH_{multiEL}$-firms have a margin of 5.36 % and 7.79 %, respectively.

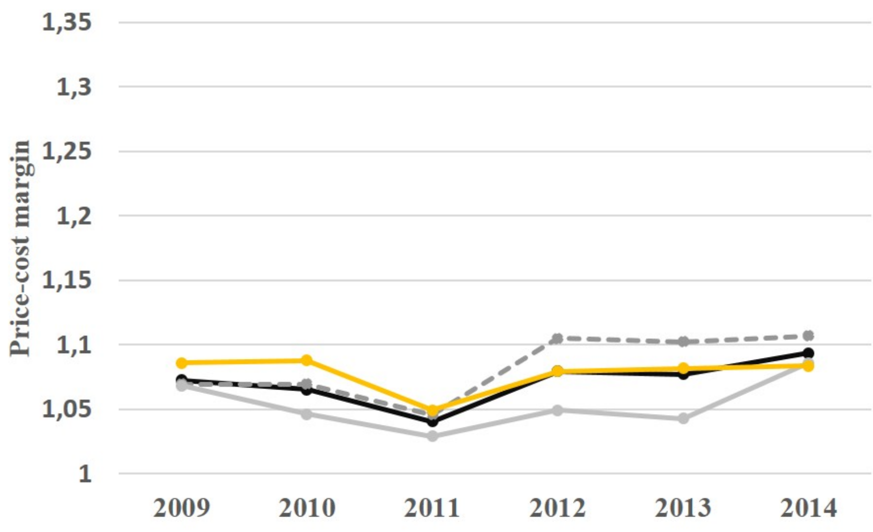

Figure 6: The price-cost margin development between 2009 and 2014 based on SEMI's data. The solid black graph gives values for all DH firms, the dashed grey graph gives values for DH_{only}-firms, the solid grey graph gives values for DH_{multi}-firms and the yellow graph gives values for $DH_{multiEL}$-firms.

Table 6 shows the average price-cost margin for all firms and for each firm group per year, from 2009 to 2014, based on imputed data.

Table 6: The average price-cost margins for all DH firms and for each DH firm group based on imputed data.

Year	*All Firms*	DH_{only}	DH_{multi}	$DH_{multiEL}$
2009	1.1188	1.0695	1.0563	1.3160
2010	1.0973	1.0695	1.0442	1.2307
2011	1.0990	1.0457	1.0634	1.2416
2012	1.1239	1.1051	1.0617	1.2294
2013	1.1385	1.1022	1.0808	1.2561
2014	1.1536	1.1069	1.1040	1.2845

Notes: Table gives price-cost margin in percent in decimal form.

With the values in Table 6, the development of the imputed average price-cost margin can be illustrated. Figure 7 illustrates this development for all firms and for each firm group, respectively.

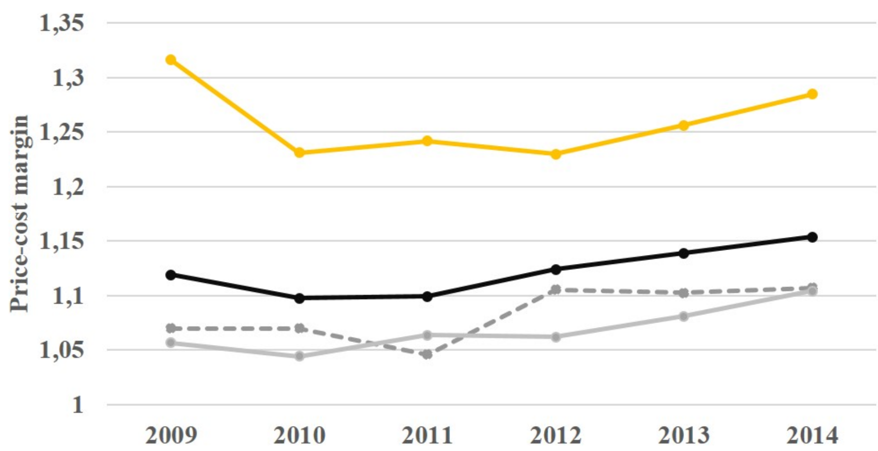

Figure 7: The price-cost margin development between 2009 and 2014 based on imputed data. The solid black graph gives values for all DH firms, the dashed grey graph gives values for DH_{only}-firms, the solid grey graph gives values for DH_{multi}-firms and the yellow graph gives values for $DH_{multiEL}$-firms.

Figure 7 illustrates that the two firm groups DH_{only} and DH_{multi} have, on average, imputed average price-cost margins of 8.32 % and 6.84 % respectively, with minor deviations over time. But the result for the DH_{multi}-firms is not statistically significant. Firms providing DH and electricity have a coherently higher imputed average price-cost margin than the two other groups of DH firms. $DH_{multiEL}$- firms have, on average, an imputed average price-cost margin of 25.90 %, with larger fluctuations over time. It is also noteworthy that the price-cost margin for all firms has increased every year since 2011. Figure B.1 in Appendix B illustrates the distribution of price-cost margins for all firms both for reported and imputed data.

A comparison between the imputed average price-cost margins and the average price-cost margins reported by the firms reveals that there is a major difference between the imputed and reported average price-cost margin for the $DH_{multiEL}$-firms. The imputed average price-cost margin is 25.90 % compared to 7.79 % for the reported average price-cost margin. This shows that the average price-cost margin for pure DH is well above the price-cost margin based on reported values. This is also illustrated by Figures 6 and 7. Clearly, this further supports the fact that $DH_{multiEL}$-firms allocate costs from the competitive electricity operation to the comparatively much less intensively competitive DH operation. This strategic cost allocation is fully passed on to consumers, who therefore are forced to subsidize the firms' aggressive supply of electricity. As for the DH_{multi}-firms, there is virtually no difference between imputed and reported average price-cost margin and the imputations are not significant for this firm group. There is no difference between the imputed and reported average price-cost margin for DH_{only}-firms because they serve as reference values during the imputation. Figure B.2 in Appendix B illustrates the distributions of price-cost margins for the three DH firm groups.

6 | Discussion

6.1 Evaluation of the Cost Function and Cost Model

The data extracted consists of observations between 2009 and 2014. After filtration and removing of extreme observations, 942 observations with complete data remained. To enable even more accurate analyses, it would be desirable to have data from earlier years. Unfortunately SEMI did not begin to collect data until 2007 and the data for the first two years was of relatively low quality, why it was not included in the analyses.

To try to improve the cost models described in Table 1, it would be desirable to include some of the factors described in section 2.2. However, including more factors in the cost function will result in a more extensive grouping of DH firms. Since the data set used in this study includes a limited amount of observations, this would imply fewer observations in each group. Also, the observations might be spaced far apart with respect to the input factors used in the cost function. Few observations spaced far apart might be problematic when performing the regression; the model might not be able to explain the observations. To circumvent this problem and to test the effect of one specific grouping of DH firms, multiple regressions should be performed, where each regression includes only one new variable, i.e. one regression for geographic grouping, one for owner structure grouping and so on. However, in this study, the factors described in section 2.2 are assumed to be controlled for by the firm- and year-specific fixed effects. For some of these factors, it might be better to include some of them directly in the cost function. The most interesting of these factors are presented below.

Differences in production techniques across firms can be captured, in full or partly, by the differences in fuel mix across firms. Therefore, it would be desirable to include the DH firm's individual fuel mixes in the cost function. Today, however, DH firms are not required to report their fuel mix to SEMI. Therefore, many observations lack information about fuel mix and including this in the cost function would reduce the number of observations included in the analyses. Thus, it would bring uncertainty to the result presented in Table 1.

In this study, firms are not divided into different groups based on their geographic position. Firms in the north of Sweden experience differences with respect to climate conditions, that can warrant more detailed investigations. However, this grouping was beyond the time scope of this study.

DH firms have different ownership structures. Many DH firms are privately owned and in some municipalities the DH business is under direct control of the local municipal political assembly. However, during recent years there has been a trend of moving DH, together with other utility services, into separate, municipally owned firms. Regardless of owner structure, it is assumed that all DH firms aim at minimizing their costs. This assumption holds for private companies, since cost minimization is a natural way to achieve profit maximization. Many municipalities have a clear policy with respect to pricing of utility services. They explicitly state that their utility services should have a price that is competitive, to keep current inhabitants satisfied and to attract new inhabitants.[21] This desire is consistent with cost minimization. The tendency to move DH and other utility services into separate companies under municipal ownership makes it harder for municipalities to cross-subsidize businesses that are not directly linked to DH by over-investing in DH. This implies that forming corporations makes DH more likely to only carry its own cost and represent a kind of cost minimization in itself. An additional argument supporting the assumption of cost minimization is that DH consumers has the ability to formally complain and to have the National District Heating Committee reviewing the prices if they are unable to reach an agreement with its local DH supplier. Even though the DH firm is not legally obliged to follow the committee's recommendation, ignoring the committee can lead to bad publicity. Bad publicity can make it harder for the DH firm to attract new consumers. These are the arguments on which the assumption of all firms minimizing their costs are based. However, Muren (2011) show that owner structure does affect the price of DH and, therefore, it would be desirable to include owner structure in the cost function specification to be able to see if this has any effect on the empirical work.

6.2 Evaluation of the Matching Algorithm

The matching algorithm applied in this study to impute the average cost levels of DH for firms providing multiple utility services is a nearest-neighbor covariate matching algorithm, as mentioned in section 4.4. The matching is done based only on the similarity of the heat delivered by the firms and the firms' length of their distribution network. This means that firms with different owner structures, fuel mixes, geographical and social conditions can be matched, i.e. a privately owned firm in northern Sweden that mainly uses oil and waste heat to produce and provide electricity and DH can have its true cost of DH based on a

[21] A survey conducted by Sandoff (2008) concludes that offering a low price of DH is one of the major arguments that municipalities use for motivating them into owning DH.

municipality owned firm in southern Sweden that mainly uses waste to produce heat, as long as their Q and L are similar. The reasonableness of the imputed average cost levels can thus be discussed. Since the sample size is relatively small, it is, however, hard to tell how the bias would be affected if, for instance, the firms could be matched geographically. How would the bias change if Piteå is matched with a geographically similar firm that is quite different in Q and L compared to when Piteå is matched to a firm that is similar in terms of Q and L but different geographically? This depends on how significant each variable is with respect to the average cost.

It is desirable to match on the most significant variables, but this study was delimited to only consider Q and L. Two advantages gained by including as many significant variables in the matching algorithm as possible are 1) probably more reliable results given a sufficient sample size so that suitable matches can be guaranteed and 2) more treatment effects could be studied. However, if the sample size is not large enough, like in this case, inclusion of more variables would jeopardize the current bias, since it would be hard to find suitable matches. If, for instance, the fuel mixes and the geographical location would be included in the matching algorithm, it would be very hard to find a similar firm in terms of all these variables in the opposite treatment group given the current sample size. Consequently, one would risk an increased bias. At the same time there is a risk that the interpretation of the results that the $DH_{multiEL}$-firms act strategically is partly explained by the fact the matches are made on only Q and L. However, from the multiple regression analysis, the delivered heat has been shown to capture a sizeable part of the total variations of the average cost of DH. Therefore the current results have to be considered relatively reliable.

7 | Conclusions and Recommendations

From the cost model estimation, it is concluded that the delivered heat captures a sizeable part of the total variation of the average cost of DH for all firm groups. The matching algorithm used in this study works well for its intended task. It successfully imputes reasonable average costs levels of DH for the firms providing multiple utility services. However, the matching can be extended in future work to include more observations and additional covariates to improve its performance even further.

The difference between the imputed true average cost levels and the corresponding values reported by the firms supports the hypothesis that firms providing multiple utility services act strategically. This is particularly true for the firms providing both DH and electricity. These firms are cost inefficient compared to the firms only providing DH based on SEMI data, but they are comparatively cost efficient based on imputed data. This reveals a strategic behavior of 16.00 %, meaning that the $DH_{multiEL}$-firms report, on average, 16.00 % higher average costs of DH than what true costs of DH are. By comparing the price-cost margins based on their reported average costs and the imputed true cost, this strategic behavior is further demonstrated. Also, the price-cost margin for the $DH_{multiEL}$-firms reveals that the strategic cost allocation is fully passed on to consumers, who therefore are forced to subsidize the firms' aggressive supply of electricity. As for the DH_{multi}-firms, there is no statistical support that they engage in such behavior and to the extent that they do, it is of a much smaller magnitude.

The reasonableness of these results has been discussed and further work with inclusion of more variables in both the cost function and in the matching algorithm along with a more thorough data gathering should be carried out to see if it could further decrease the bias and thereby further strengthen the current arguments that the firms providing DH, electricity and other multiple utility services act strategically.

In the case of strategic behavior, the textbook advocates two alternative solutions: 1) to separate the production of DH and electricity into two separate firms, which is not a technical and economical feasible alternative in this situation, since the production processes are coupled; or 2) to regulate the prices set by the firms that provide both DH and electricity. Thus, the only feasible option policy makers have in this situation is to price regulate the firms that provide DH and electricity.

References

Abadie, A. & Imbens, G. (2004). *Large Sample Properties of Matching Estimators for Average Treatment Effects (previous version: Simple and Bias- Corrected Matching Estimators for Average Treatment Effects)*. Harvard University: Working paper.

Abadie, A. & Imbens, G. (2011). "Bias-Corrected Matching Estimators for Average Treatment Effects". In: *Journal of Business & Economic Statistics* 29.1, pp. 1–11.

Baumol, W.J., Panzar, J.C., & Willig, R.D. (1982). *Contestable Markets and the Theory of Industry Structure*. 1st ed. New York, NY: Harcourt College Pub.

Biggar, D. & Söderberg, M. (2012). *The effect of future price expectations on customers' willingness to make sunk investments in relaince on a monopoly service*.

Caliendo, M. & Kopeinig, S. (2003). *Some Pratical Guidance for the Implementation of Propensity Score Matching*. Peking University: China Centre for Economic Research.

Currier, K.M. (2004). "Natural monopoly regulation in the presence of cost misreporting". In: *Atlantic Economic Journal* 32.1, pp. 49–61.

Deheijia, R.H. & Wahba, S. (1999). "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs". In: *Journal of the American Statistical Association* 94.448, pp. 1053–1062.

Färe, R., Martins-Filho, C., & Vardanyan, M. (2009). "On Functional Form Representation of Multi-output Production Technologies". In: *Journal of Productivity Analysis* 33.2, pp. 81–96.

Gerfin, M. & Lechner, M. (2002). "A Microeconometroc Evaluation of the Active Labour Market Policy in Switzerland". In: *The Economic Journal* 112, pp. 854–893.

Hepbasil, A. & Kalinci, Y. (in press). "A review of heat pump heating systems". In: *Renewable and Sustainable Energy Reviews*.

Hirschey, M. (2003). *Managerial Economics*. 10th ed. Mason, OH: South Western.

Holm, J. (2013). "Inför prisreglering av fjärrvärme - Vilka lärdomar kan dras från ekonomisk teori och empiri?" In: *Konkurrensverkets uppdragsforskningsrapport*, p. 68.

Kwoka, J.E. (1996). *Power Structure: Ownership, Integration and Competition in the U.S. Electricity Industry*. 1st ed. Boston, MA: Kluwer.

Kwoka, J.E. (2002). "Vertical economies in electric power: evidence on integration and its alternatives". In: *International Journal of Industrial Organization* 20.5, pp. 653–671.

Kwoka, J.E. (2005). "Electric Power Distribution: Economies of Scale, Mergers and Restructuring". In: *Applied Economics* 37.20, pp. 2373–2386.

Mankiw, N.G. (2011). *Principles of Economics*. 6th ed. Mason, OH: South Western.

Muren, A. (2011). *Exploatering eller reglering av naturliga monopol - Exemplet fjärrvärme*.

Pauschinger, T. (2015). "Solar thermal energy for district heating". In: *Advanced District Heating and Cooling (DHC) Systems*. Ed. by Wiltshire, R. Saint Louis: Woodhead Publishing, pp. 99–120.

Persson, U. & Werner, S. (2011). "Heat distribution and future competitiveness of district heating". In: *Applied Energy* 88.3, pp. 568–576.

Pulley, L.B. & Braunstein, Y.M. (1992). "A Composite Cost Function for Multiproduct Firms with an Application to Economies of Scope in Banking". In: *The Review of Economics and Statistics* 74.2, pp. 222–230.

Reidhav, C. & Werner, S. (2008). "Profitability of sparse district heating". In: *Applied Energy* 85.9, pp. 867–877.

Röller, L-H. (1990). "Quadratic Cost Functions with an Application to the Bell System". In: *The Review of Economics and Statistics* 72.2, pp. 202–210.

Sandoff, A. (2008). *Ägar- och företagsstyrning i kommunala energibolag*.

Shaffer, S. (1998). "Functional Forms and Declining Average Costs". In: *Journal of Financial Services Research* 14.2, pp. 91–115.

Shah, V., Col Debella, D., & Ries, R.J. (2008). "Life cycle assessment of residential heating and cooling systems in four regions in the United States". In: *Energy and Buildings* 40.4, pp. 503–513.

Sipilä, K. (2015). "Cogeneration, biomass, waste to energy and industrial waste heat for district heating". In: *Advanced District Heating and Cooling (DHC) Systems*. Ed. by Wiltshire, R. Saint Louis: Woodhead Publishing, pp. 45–73.

Sjödin, J. & Henning, D. (2004). "Calculating the marginal costs of a district-heating utility". In: *Applied Energy* 78.1, p. 2.

SOU (2003). *Tryggare fjärrvärmekunder - ökad transparens och åtskillnad mellan el- och fjärrvärmeverksamhet (SOU 2003:115)*. Stockholm: Ministry of the Environment and Energy.

SOU (2004). *Skäligt pris på fjärrvärme (SOU 2004:136)*. Stockholm: Ministry of the Environment and Energy.

SOU (2005). *Fjärrvärme och kraftvärme i framtiden (SOU 2005:33)*. Stockholm: Ministry of the Environment and Energy.

Tschirhart, J. (1996). "Monopsony power and the existence of natural monopoly in energy utilities". In: *Fuel and Energy Abstracts* 37.2, pp. 327–340.

Westin, P. & Lagergren, F. (2002). "Re-regulating district heating in Sweden". In: *Energy Policy* 30.7, pp. 583–596.

Zhao, Z. (2003). *Data Issues of Using Matching Methods to Estimate Treatment Effects: An Illustration with NSW Data Set*. 1st ed. Peking: China Centre for Economic Research.

Electronic References

Blumenstock, J. Last visited: 2016-04-18. Available at: http://www.jblumen stock.com/files/courses/econ174/FEModels.pdf.

Darthmouth College. Last visited: 2016-04-22. Available at: https://www.dartmou th. edu/ ethang/Lectures/Class17/Always%20Control%20for%20Year%20Ef fects%20i n%20Panel%20Regressions.pdf.

District Heating Södertörn. Last visited 2016-04-11. Available at: http://sfab.se/Fjarrvarme/.

MATLAB. Last visited 2016-05-06. Available at: http://se.mathworks.com/products/ matlab/.

STATA. Last visited 2016-04-18. Available at: http://www.stata.com/why-use-stata/.

STATA manual - Teffects Nnmatch & Teffects Intro. Last visited 2016-05-04. Available at: https://www.stata.com/manuals13/te.pdf.

Statistics Sweden. Last visited 2016-04-08. Available at: http://www.scb.se/sv_/Hitta-statistik/Statistik-efter-amne/Priser-och-konsumtion/Konsumentprisindex/Konsumentprisindex-KPI/.

Statistics Sweden. Last visited 2016-05-20. Available at: http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__HE__HE0201__HE0201A/HUTutgift1/?rxid=a9b94be2-f057-4c90-8d47-34d8b6228c81.

Swedish District Heating Association. Last visited 2016-04-12. Available at: http://www.svenskfjar rvarme.se/Fjarrvarme/Sa-funkar-fjarrvarme/.

Swedish Energy Market Inspectorate. Last visited 2016-04-08. Available at: http://ei.se/ sv/Fjarrvarme/inrapporterad-data/.

Swedish Energy Market Inspectorate. Last visited 2016-04-15. Available at: http://ei.se/ en/District-heating/Factors-that-affect-the-price-of-district-heating1/.

Vattenfall. Last visited 2016-04-12. Available at: http://corporate.vatten fall.com/about-energy/energy-distribution/district-heating/pricing/.

A | Appendix

Here the STATA syntax is given for the two cases of STATA-analysis. In section A.1 the STATA syntax for cost model estimations is presented, whereas in A.2 the STATA syntax for cost level estimations is given.

A.1 STATA Syntax for Cost Model Estimations

The following syntax is used to generate the necessary variables according to the cost function that the data is fitted to by the help of multiple regression analysis:

gen q_heat_div = 1 / q_heat

gen q_heat_half = 0.5 * q_heat

gen q_heat_length_div = length / q_heat

gen q_heat_length_sq = 0.5 * length * length / q_heat

gen length_half = 0.5 * length

To define the data as panel data and to order the panel members (in this case the firms) within the panel by time (in this case year), the following syntax is used:

xtset case year

To fit the panel data to the conventional quadratic generic cost function by the use of multiple regression analysis and to include firm- and year-specific fixed effects, the following syntax is applied:

For firms that only produce DH and offer no other service (DH_{only}-group):
xtreg cost_avg q_heat_div q_heat_half q_heat_length_div q_heat_length_sq length_half i.year if sample == 1 & no_elec_sales == 1 & multi_serv == 0 , fe

For firms that only produce DH, but offer other services (DH_{multi}-group):
xtreg cost_avg q_heat_div q_heat_half q_heat_length_div q_heat_length_sq length_half i.year if sample == 1 & no_elec_sales == 1 & multi_serv == 1 , fe

For firms that produce both DH and electricity and offer other services ($DH_{multiEL}$-group):

xtreg cost_avg q_heat_div q_heat_half q_heat_length_div q_heat_length_sq length_half i.year if sample == 1 & no_elec_sales == 0 & multi_serv == 1 , fe

Here the **fe**, at the end of each command, is telling STATA to apply a fixed effects linear model. The syntax **i.year** is telling STATA to account for the year-specific fixed effects. The results from applying these syntax can be found in section 5.1 Table 1.

A.2 STATA Syntax for Cost Level Estimations

The following syntax is used to apply a nearest-neighbor covariate matching algorithm to estimate the average treatment effect of introducing multiple services to DH firms only producing and offering DH:

teffects nnmatch (cost_avg q_heat_div q_heat_half q_heat_length_div q_heat_length_sq length_half) (multi_serv) if no_elec_sales==1 & sample==1, gen(_nn)

Note that the covariates correspond to the variables in Equation 9 presented in section 4.2. **gen(_nn)** generates a variable **_nn1** containing the observation number of the nearest-neighbor which it is matched with. This variable is necessary to generate to be able to perform postestimations. By default only one nearest-neighbor is matched and the matching is done with replacement.

The following syntax is applied in order to determine the potential outcomes, i.e. the imputed average cost levels:

predict po0 po1, po

where **po0** is the potential outcome (average cost) for the treatment group (DH_{multi}-firms) when not receiving the treatment (multi_serv) and it is the potential outcome for the control group (DH_{only}-firms) when not receiving the treatment. **po1** is the potential outcome for the treatment group when receiving the treatment and it is the potential outcome for the control group when receiving the treatment. The interesting parameter here is **po0** for the treatment group.

For the other treatment effect of not producing electricity to DH firms producing and offering both DH and electricity, the following syntax was used:

teffects nnmatch (cost_avg q_heat_div q_heat_half q_heat_length_div q_heat_length_sq length_half) (no_elec_sales) if for_teffects==1 & sample==1, gen(_nn2)

Similarly performed as for the first treatment. However, a new variable that was called **gen(_nn2)** had to be introduced in this case, but for the same reason as for the first treatment. Also a variable **for_teffects** was introduced in order to separate the DH_{multi}-firms from the DH_{only}- and $DH_{multiEL}$-firms, which were matched here.

Just like for the first treatment, the potential outcomes were estimated from the following syntax:

predict po01 po11, po

where **po01** in this case is the potential outcome for the treatment group (DH_{only}-firms) when not receiving the treatment (no_elec_sales) and it is the potential outcome for the control group ($DH_{multiEL}$-firms) when not receiving the treatment. **po11**, however, is the potential outcome for the treatment group when receiving the treatment and it is the potential outcome for the control group when receiving the treatment. Consequently the interesting variable here is **po11** for the control group, since it corresponds to what the average cost levels would be for the $DH_{multiEL}$-firms when not producing electricity.

The average treatment effect for both treatments can be found in Table 2 under section 5.2.

B | Appendix

Various plots and one table to support or further illustrate some results are presented here. MATLAB, which is a software used for computations and image processing, has been used diligently in this study for image processing.[22]

Table B.1 shows the result of the treatment effect between DH_{multi}- and $DH_{multiEL}$-firms when receiving/not receiving the treatment of not producing electricity.

Table B.1: The treatment effect after applying the matching algorithm.

Treatment Effect	$No\ elec.\ sales$
Treatment Group	DH_{multi}
Control Group	$DH_{multiEL}$
	Coeff. (S.E.)
Average Cost	-0.096 *** (0.0100)
No. of obs.	560
No. of Nearest Neighb.	1

Notes: Significance level *** p < 0.01. Table gives the change in mean average cost, in $tSEK/MWh$, with standard error in parenthesis. Matching is done with replacement.

[22] For more information, the reader is referred to the homepage of MATLAB. Last visited: 2016-05-06. Available at: http://se.mathworks.com/products/matlab/

Figure B.1 shows the price-cost margin distributions for both reported and imputed data, for all DH firms.

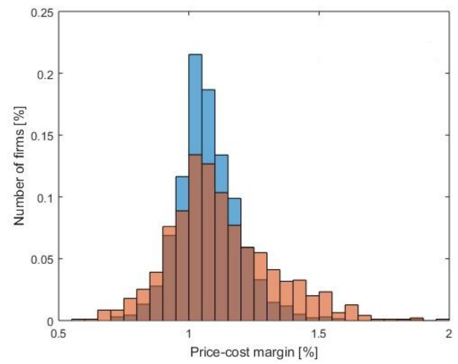

Figure B.1: The distributions of price-cost margins for all firms based on SEMI's data (blue) and imputed data (red). Note: There are only two distributions in the Figure, the distributions become darker in the middle because the colors are imposed on each other.

Figure B.2 illustrates the distribution of price-cost margins based on the different DH firm groups. The blue distribution shows price-cost margins for DH firms only producing DH and is based on data from SEMI. The red and yellow distributions shows imputed data. More specifically, the red distribution shows price-cost margins for firms providing multiple utility services and the yellow distribution shows price-cost margins for firms producing both DH and electricity and offering other utility services.

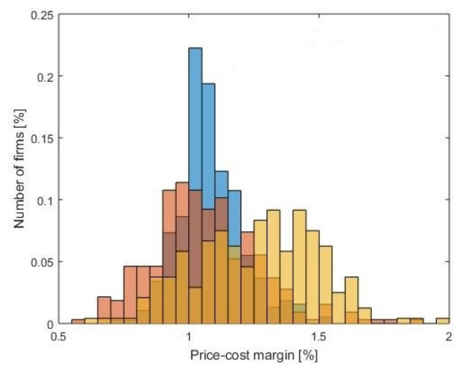

Figure B.2: The distributions price-cost margins for the three different firm groups. For DH_{only}-firms (blue), data is based on SEMI's values. The two other distributions are based on the imputed data. The red distribution represent DH_{multi}-firms and yellow distribution represents the $DH_{multiEL}$-firms. Note: There are only three distributions in the Figure, the distributions become darker in the middle because the colors are imposed on each other.