Assessing the impact of support policies for energy efficient technology in Switzerland

Author(s):
Ley, Marius

Publication Date:
2012-11

Permanent Link:
https://doi.org/10.3929/ethz-a-007572162

KOF

KOF Working Papers

Assessing the Impact of Support Policies for Energy Efficient Technology in Switzerland

Marius Ley

No. 321
November 2012

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

KOF

ETH Zurich
KOF Swiss Economic Institute
WEH D 4
Weinbergstrasse 35
8092 Zurich
Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

Assessing the Impact of Support Policies for Energy Efficient Technology in Switzerland*

Marius Ley
ETH Zürich
KOF Swiss Economic Institute
marius.ley@kof.ethz.ch

November 12, 2012

Abstract

This study uses data of a representative sample of Swiss firms to assess the effectiveness of policy measures directed at the diffusion of energy efficient technology (EET). Three different outcome variables (number of EET applications implemented, CO_2 reduction, and EET related investment) are analysed using methods based on the estimated propensity score in an attempt to overcome the problem of non-random assignment. I conclude that, even after controlling for non-random assignment, diffusion support from the two institutions taken into consideration has indeed been beneficial in spurring adoption of energy saving technology and in reducing emissions of CO_2. Additionally, an estimator for Average Treatment Effects on the Treated (ATT) that directly relies on the propensity score has been found to produce better results, in terms of efficiency, than the widely used Nearest Neighbour Matching (NNM) procedure.

Keywords: Energy Efficiency, Energy Policy, Technology Diffusion, Propensity Score Matching

JEL-Classification: O30, Q42, Q48

Acknowledgements: I am grateful for fruitful comments provided by Spyros Arvanitis, Peter Egger, Michael Siegenthaler, Martin Wörter and participants at the European Network on Industrial Policy (EUNIP) International Workshop on Evaluating Innovation Policy (2011, Firenze) as well as at the KOF Brown Bag Seminar (Zürich, 2012). Any remaining errors are the author's alone.

*This study has been financed by the Swiss Federal Office of Energy (SFOE).

1 Introduction

A number of measures have been put into place in Switzerland during the last two decades to support innovation and diffusion of energy saving technologies. This study uses firm level data, collected in the spring of 2009 by means of a survey among Swiss firms belonging to both the manufacturing and service sectors, to determine the effectiveness of some of those measures directed at the diffusion stage.

Public awareness that energy systems need to become more efficient and less reliant on non-renewable sources — furthermost, fossil ones — has gradually increased over the last decades, much due to concerns that current use and technology of energy are not sustainable. From a economic perspective, government intervention destined at supporting the adoption of energy saving technologies can be justified on the grounds of *market failures* that prevent an optimal free-market outcome from resulting. It may thus be of use to briefly discuss market failures; whereby a distinction can be made between *externalities* — which in turn can be of negative or of positive nature — and market failures of other types.

Negative *externalities* arise due to external costs of use of energy that are not imposed on the user in the absence of well-designed policy instruments (such as levies or tradable permits). The most prominent (and vigorously debated) topic here, of course, is global climate change attributable to the burning of fossil fuels, mainly in the form of the release of massive amounts of carbon dioxide (CO_2) into the atmosphere. On the other hand, the technological change literature stresses the existence of positive externalities of adoption of new technology — the fact that a firm's decision to apply new technology may ultimately benefit its competitors or other enterprises (such as suppliers of the relevant technology or other potential end-users), resulting in a net benefit to the economy that actually exceeds what the initial adopter may reap in terms of enhanced productivity or market share. The positive externality argument, however, is of more relevance in the literature dealing with research and development (R&D) and innovation activities (of the private sector), rather than in the technology diffusion context the present paper is concerned with. Nevertheless, it would be unreasonable to rule out a priori the existence of such *network* or *epidemic* effects. Karshenas and Stoneman (1993) present an empirical diffusion model capable of capturing these effects, and an analysis by Arvanitis and Ley (2012) applied to the case of energy efficient technology suggests that positive externalities indeed may play a role here, too.

Optimal diffusion of technology may be hindered by a number of market failures other than the classical externality type. They are by no means only limited to the field of energy efficiency, but are thought to be of particular magnitude in this context by many authors (e.g. Battisti, 2008). Popp et al. (2009, p. 30ff) provide a comprehensive overview of theoretical and empirical findings related to various types of market failures. Worth noting are *agency problems,* which may arise particularly in landlord/tenant relationships, where investments in appliances or thermal insulation of buildings are made by the landlord, but their running costs or savings are borne by the tenant. *Information inefficiencies* in the process of technology diffusion are likely to arise, as information (e.g. as for what is the most suitable energy-efficient technology that can be implemented for a certain process or purpose within a given firm or household) has the characteristics of a public good. Finally, many studies stress that the implicit *discount rates* underlying the purchasing decisions of households firms with respect to energy-efficient investments are relatively high. Strictly speaking, this phenomenon taken by itself does not constitute a market failure; however it may be a manifestation of various underlying imperfections such as uncertainty or capital market failures.

It should be stressed that subsidising diffusion may not be (and will not be, in most cases) the optimal policy response to most of the market failures discussed. In the presence of information inefficiencies, to take just the most evident example, public policy might instead turn towards promoting the dissemination of technology-related knowledge that firms or households would otherwise find difficult gather. In the case of negative externalities, conventional economic wisdom suggests that market-based instruments like emission taxes or permits would be most efficient. However, it is far beyond the scope of the present paper to provide a comparative analysis as to how different policy instruments perform relative to each other (see section 3 for further literature on this topic). Instead, the focus will be to empirically analyse if two specific policy instruments presently in place in Switzerland are successful in fulfilling their intended purpose, which is to promote the adoption of energy efficient technologies at the firm level. While it is fully justified (and of major economic and political interest) to ask whether there are better instruments to reach this goal or whether the funding intended for this purpose is of adequate size, such questions will not be addressed by this study.

From the point of view of an economist concerned with policy evaluation, the two policy schemes considered here have the interesting feature that, following their implementation, the target population of firms is divided into two clearly distinct groups: those who have received support and those who have not. This makes it possible to use microeconometric treatment effects estimation techniques to assess whether this policy has been successful. Essentially, this means that we can find an answer to the question of what would have been, on average, the performance in terms of certain outcome measures of a firm who has received support, if this firm had not benefited from any such support. This is in contrast to policies that are not designed to provide support to a subgroup of successful applicants, but rather affect the economy or the target population broadly, i.e. in a non-specific manner, such as emission taxes or permits.[1]

Treatment effects estimation allows as to focus on the question whether there is *crowding out* of private investment by public support. This refers to a situation where

[1] Microeconometric evaluation of the effectiveness of this latter type of policies would require different techniques, such as Computable General Equilibrium (CGE) modelling.

a certain investment (e.g. in R&D or in environmental measures) would have been undertaken by the firm regardless of whether the subsidy is granted or not.[2] In nearly all circumstances, crowding out effects cannot be completely ruled out a priori when designing support policies, as the private benefits of a measure (or any non-pecuniary motives of the firm's manager to implement it) cannot be known exactly to the agency that decides to whom subsidies should be allocated. For a policy programme to be effective, it is thus crucial to avoid crowding out effects as much as is possible. In the present empirical study, the central question will therefore be if the case of complete crowding out of private investments in energy technologies by public support measures can be credibly ruled out. In the case of complete crowding out, such policies would of course be entirely ineffective. In the opposite case of no (or only partial) crowding out, the policy can be presumed to fulfil (at least partly) its intended purpose.

The remainder of this paper is organised as follows. Section 2 describes the current institutional setting of support measures for energy efficient technologies (henceforth EET) diffusion in Switzerland alongside with the availability of data for the present paper, section 3 resumes the relevant literature, section 4 provides the theoretical basis for the analytical approach, section 5 outlines the data and discusses implementational details, section 6 presents the results and section 7 concludes.

2 Institutional Background

Swiss energy policy is complicated by the fact that the three tiers of government (federal, cantonal and municipal) are all involved in designing and implementing policy measures. In particular, the Federal Constitution states that "The Cantons shall be primarily responsible for measures relating to the use of energy in buildings"[3]. As a result, Swiss Cantons have implemented a wide range of measures aimed at reducing the energy consumption in buildings. These include technical requirements for newly constructed and renovated buildings, as well as financial support schemes to real estate owners who construct or renovate objects according to specific energy-efficient standards. Since buildings use more than 40% of final energy consumption (and more than half of total consumption of fossil fuels) in Switzerland (BFE, 2011), reducing energy use related to buildings has been a long-standing priority of Swiss energy policy, and thus local governments (i.e. cantons and, to some extent, municipalities) rather than the federal government are the prime donors of financial assistance. In contrast, legislation on the use of energy of things other than buildings (in particular: installations, vehicles and appliances) belong to the domain of federal policy (Federal Constitution, Art.

[2] In other words, if the private benefits of the implemented measure are positive even in the absence of a subsidy, in a world of rational decision taking firms.

[3] *Federal Constitution of the Swiss Confederation*; Art. 89/4.
All legal texts referred to in this paper may be accessed through http://www.admin.ch/ch/e/rs/rs.html

89/3).[4] Strebel (2011) provides a detailed overview of the peculiarities of energy policy in the context of Switzerland, and explores the implications they (and inter-governmental institutions, in particular) have on the diffusion of specific policy instruments.

In the framework of the Kyoto protocol, Switzerland has committed to a 8% reduction of greenhouse gas emission for the reference period 2008–2012 relative to its 1990 emission levels. A reduction target of 10% specific to CO_2 — the single most important greenhouse gas — has subsequently been formulated in the Federal CO_2 Emission Reduction Law[5], enacted in 1999. This has prompted the creation of two new schemes promoting dedicated projects to reduce CO_2 emissions by firms: the *Climate Cent Foundation*[6]; and the possibility for firms of being exempt from the CO_2 levy conditional on implementing emission reduction measures, usually with assistance from the private sector based *Energy Agency of the Swiss Economy* (henceforth EASE).[7] Based on a mandate negotiated with the federal government, both of these private sector based institutions provide support to firms — either in the form of investment subsidies to firms implementing technology that effectively reduces CO_2 emissions (in the case of the former) or by providing expertise to firms wishing to be exempt from the CO_2 levy (in the case of the latter).

The existence of two complementary schemes is mainly due to the fact that, based largely on political considerations, the federal government has decided to impose a CO_2 levy on fossil fuels for heating purposes only, whereas no such levy has yet been introduced on fuels used to power engines for transportation or other purposes.[8] This decision has been reached following a joint effort by private sector organisations to form the *Climate Cent Foundation* in order to impose a moderate *voluntary* levy on engine fuel imports from October 2005 until August 2012,[9] and to allocate the proceeds towards selected projects directed at reducing CO_2 emissions both domestically and abroad, thus avoiding the much higher federal CO_2 levy to cover fossil engine fuels as well. Every firm or public institution is in principle eligible to apply for subsidies granted by the Climate Cent Foundation, no matter whether the proposed reduction of CO_2 emissions is achieved by cutting the use of engine fuels (on which the Foundation's proceeds are generated) or heating fuels (which, in contrast, are subject to the CO_2 levy).

In contrast, the EASE's mission is to assist firms to mitigate the potentially severe effects (particularly, loss of international competitiveness) of the federal CO_2 levy by

[4] As is the case with support for energy efficiency related R&D, which is however not of concern for this paper.

[5] *Bundesgesetz vom 8. Oktober 1999 über die Reduktion der CO_2-Emissionen* (CO_2-Gesetz)

[6] http://klimarappen.ch/

[7] http://www.enaw.ch/

[8] The rate of the levy is currently set, with effect from January 2010, at CHF36 (EUR30 at current exchange rates) per ton of CO_2.

[9] Since no fossil fuels of any kind are mined in Switzerland, all domestic consumption was affected by this voluntary import levy. It amounted to CHF 0.015 per litre of petrol and diesel imports, which is roughly a sixth of the effective rate of the federal CO_2 levy currently applied to fossil fuels for heating purposes. The timespan of this measure was limited as it was intended to be replaced by other instruments in the post-Kyoto period.

providing them expertise related to technical, reporting and legal issues required for a successful application towards the federal government for exemption from the CO_2 levy. Such an exemption is granted if the firm undertakes efforts to reduce its CO_2 emissions as appropriate in the individual case, a possibility explicitly provided for in the CO_2 Emission Reduction Law. Thus, only users of fossil fuels for heating purposes are potential beneficiaries of this scheme.

Our survey has featured questions about participation in the three schemes broadly outlined above: (a) financial support by cantons and municipalities; (b) financial support by the Climate Cent Foundation; (c) exemption from the CO_2 levy with assistance of the EASE. Despite disposing of this breadth of information, only the answers related to support from categories (a) and (b) are considered in this analysis; the effects of the EASE are thus neglected. This is due to the limited potential reach of the latter scheme: only firms with significant expenses for fossil fuels for heating purposes may participate, which would complicates the construction of the comparison group to be used in the present analysis.

In terms of their volume, the two schemes considered in the present study are roughly comparable. In 2008, cantons have spent 65 million Swiss francs on support for energy efficiency.[10] The Climate Cent Foundation supported domestic energy efficient projects with 41 million Swiss francs in 2008 and nearly twice as much in the subsequent year (Stiftung Klimarappen, 2010).

The dataset contains a number of potential *outcome variables* that can be considered as indicators of whether diffusion support has been successful:

- the *number of technological fields* in which a firm has invested in improved, i.e. more energy efficient technologies (count variable);
- whether any *reduction of CO_2 emissions* has resulted from such technology adoption (binary variable);
- the size of *investments* related to energy efficient technologies (a nonnegative numerical variable).

The variable *number of technological fields* refers to a precise list of specific energy efficient technologies, for which each firm participating in our survey was able to state whether it had invested in or not. Table 1 presents this technology list, alongside the respective number of technology users in our dataset.

Since our data stems from a non-experimental setting, the probability for a given firm of receiving any support must be expected to depend on various firm characteristics that were known to the supporting agency prior to taking a decision whether to grant any support or not. Supporting agencies usually are charged with the task of scrutinising all applications for such support in a manner to allocate funds or expertise to those potential

[10]Source: http://www.endk.ch/kantone.html. No precise numbers are available for support by Swiss municipalities.

beneficiaries where the expected benefits, in terms of enhanced energy efficiency or reduced emission of greenhouse gases, are highest. Not controlling for this non-random dependence of support on firm characteristics would result in biased estimates of the impact of diffusion support (a problem referred to as *selection bias*, or absence of *random assignment* or *random treatment*), and therefore suitable methods are required in order to rule out this source of bias.

Table 1: Number and Percentage (in Terms of Total Number) of Users of Specific EET Technologies

Technology	Number of Users	%
Electrical machines and drive systems	466	44.4
Information and communication technologies	459	43.8
Consumer electronics	174	16.6
Components of process engineering	472	45.0
Process engineering	282	26.9
Engines of motor vehicles	309	29.5
Motor vehicle bodies (weight/aerodynamics)	140	13.3
Traffic management system	111	10.6
Isolation	524	50.0
Lighting (incl. control systems)	610	58.2
Heating (incl. control systems)	602	57.4
Cooling systems	385	36.7
Air conditioning	450	42.9
Photovoltaics	71	6.8
Electricity from biomass	30	2.9
Wind energy	18	1.7
Combined heat and power (biomass)	18	1.7
Combined heat and power (fossil)	59	5.6
Solar heat	60	5.7
Heat from biomass	51	4.9
Geothermal heat	25	2.4
Heat pumps	178	17.0
Heat recuperation systems	322	30.7
District heating	130	12.4
Any of the above technologies	1049	100.0

3 Relevant Literature

On the theoretical side, a notable body of literature aims at ranking different policy instruments fostering the use of energy efficient or "clean" (if pollution reduction is the aim) technology in terms of criteria such as cost-effectiveness. Most of these exercises reach the conclusion that market-based instruments, such as taxes or emission permits, outperform direct measures (as resumed by Popp et al., 2009, page 24). A number of authors however stress the fact that a combination of several instruments might be the adequate solution, allowing to "reduce emission at a significantly lower cost than any single policy alone" (Fischer and Newell, 2008). Emission pricing in combination with subsidies for technology R&D and learning constitute a promising policy portfolio, according to this study. Similarly, in a theoretical framework of endogenous and directed technical change, Acemoglu et al. (2012) show that an optimal policy combines “carbon taxes” (i.e. pricing harmful emissions) and research subsidies.

To the best of my knowledge, no empirical study assessing the effectiveness of support policies on the diffusion of energy efficient technologies has been conducted yet. Velthuijsen (1993), rather than being concerned with specific policies in place, provides an early empirical study of the factors hindering businesses to implement energy efficiency measures and derives policy implications. In a broader context of assessing policy effectiveness, two strands of literature experiencing intensified research interest in recent years are worth mentioning: assessment of policies (a) promoting renewable sources of energy in the production of electricity (RES-E), and (b) supporting R&D and innovation in energy efficient technologies and/or for renewable energy sources.

Policies of various kinds supporting RES-E have been introduced as early as in the 1980s[11] in many countries, see e.g. Johnstone et al. (2010). As a consequence, a relatively abundant data basis exists that lends itself to empirically analysing and comparing, on the country or state level, the effects of various policy instruments that have been put into practice. Carley (2009) and Delmas and Montes-Sancho (2011) rely on data of policies implemented by states of the U.S. and their effect on electricity generation from non-traditional renewable sources. Mulder (2008) as well as del Río and Tarancón (2012) use data on wind turbine capacity in the EU to assess the effectiveness of various policy instruments.

The effects of public policy on R&D for energy efficiency and/or renewables have been analysed, e.g., by Johnstone et al. (2010). They use patent counts related to several renewable energy source technologies on a panel of 25 countries. Noailly (2012) measures the impact of various policy parameters — including, but not only governmental R&D expenditures — on innovation in energy efficient technology in buildings, again relying on patent counts across seven European countries.

Rennings and Rammer (2009) is not concerned with the effects of policy, but uses German survey data and econometric methodology similar to the present paper in order to analyse firm-level effects of energy and resource efficiency innovations. Similarly, Arvanitis et al. (2010) rely on propensity score matching estimation in order to investigate the impact of a support scheme of the Swiss government on innovation projects of any nature. They use several measures of innovation performance as outcome variables and rely on four different statistical matching methods as robustness checks. Only minor differences in the outcomes are found when comparing various types of matching methods (and the impact of this kind of support is found to be positive and significant).

[11] or prior to that date, with regards to support specifically for R&D in RES-E.

4 Statistical Foundations of Treatment Effects Estimation

Microeconometric evaluation of the performance of policy programmes has recently become an active area of research, perhaps most prominently in labour economics, where the political desire to provide ex post insights about the effectiveness of such measures has been strongest, resulting in a vast array of methods and estimators being proposed to this end (Caliendo and Hujer, 2006). Even though these estimators may differ in their interpretations and identifying assumptions, they all share as a common point of departure the definition for the *treatment effect* of unit i upon outcome variable y:

$$\tau_i \equiv y_i^1 - y_i^0$$

Where y_i^1 represents the outcome of y_i in the situation where unit i receives treatment (i.e. successfully participates in the programme and, for instance, receives a subsidy); and, similarly, y_i^0 measures the same outcome given that there is no treatment for i. To this end, it is convenient to define a binary treatment indicator T_i such that $T_i = 1$ in the case of treatment and $T_i = 0$ otherwise, so that the realised outcome can be expressed as $y_i = (1 - T_i)y_i^0 + T_i y_i^1$. Both y_i^1 and y_i^0 — and, consequently, τ_i — are random variables. An inherent feature of treatment effects analysis is that, for every unit i, only one of these two states is realised and observed, whereas the other is not — this is the Fundamental Problem of Causal Inference (Holland, 1986). Some distributional assumptions are thus required for estimation and inference of average treatment effects, as will become clearer in the following.

Of major interest in the recent evaluation literature is the *average treatment effect on the treated* (ATT). As its name suggests, it is simply the average treatment effect over all those observations in the population (or sample) that have received treatment:[12]

$$ATT \equiv \mathrm{E}(\tau_i | T_i = 1) = \mathrm{E}(y_i^1 - y_i^0 | T_i = 1)$$

In order to pave the way for the estimators used in this study, this expression will be extended in two respects: first, by conditioning on a vector of observable pre-treatment variables X_i, we introduce additional information that will later be useful — given additional assumptions to be specified below — for identification of the average treatment effect of interest. Second, by splitting the expectation of a difference into a difference of expectations, we obtain a first term that can be estimated in a straightforward manner by taking population means, since y_i^1 is realised and observed for every unit appearing

[12]Often, researchers and policymakers are interested on the (hypothetical) average effect of a programme under the assumption that *all* units in the population (or sample) are treated. The relevant parameter to estimate would then be the *average treatment effect* (ATE), defined simply as $ATE \equiv \mathrm{E}(\tau_i) = \mathrm{E}(y_i^1 - y_i^0)$, without conditioning on $T_i = 1$. As the two measures capture two different phenomena, their interpretations naturally differ; as usually do their estimated values in practice. The present study is confined to ATT, as it is deemed the more informative parameter for the policies in question, and since estimates for ATE turned out to be much less reliable in terms of estimated standard errors.

in the ATT (as $T_i = 1$, $y_i^1 = y_i$):

$$
\begin{aligned}
ATT = \mathrm{E}(y_i^1 - y_i^0|X_i; T_i = 1) = \mathrm{E}(y_i^1|X_i; T_i = 1) - \mathrm{E}(y_i^0|X_i; T_i = 1) \\
= \frac{1}{n_T} \sum_{i|T_i=1} y_i - \mathrm{E}(y_i^0|X_i; T_i = 1)
\end{aligned}
$$

Where, given a total number of n observations in the population (or sample) at hand, $n_T = \sum_{i=1}^{n} T_i$ denotes the number of treated observations . To be precise, the so-called stable unit treatment value assumption (SUTVA) needs to be satisfied for this last equality to hold, i.e. the treatment received by unit i must not affect the outcomes of any other unit $j = i$ in the sample. A sufficient condition for SUTVA to hold is random sampling (Wooldridge, 2002), which is the case in the present study.

By contrast, the second term $\mathrm{E}(y_i^0|X_i; T_i = 1)$ cannot be readily estimated from the population, as it is an expectation of unobserved hypothetical outcomes y_i^0. As is common in evaluation studies, we will rely on the conditional mean independence (CMI) assumption in order to overcome this problem (see e.g. Wooldridge, 2002, p. 607). Formally, CMI states that $\mathrm{E}(y_i^0|X_i; T_i) = \mathrm{E}(y_i^0|X_i)$ and $\mathrm{E}(y_i^1|X_i; T_i) = \mathrm{E}(y_i^1|X_i)$. In words, provided that X_i contains all relevant information related to firm characteristics prior to the treatment decision, knowing whether a unit has been treated or not does not carry any additional information about the expected values of the potential outcomes y_i^0 and y_i^1.

Assuming that CMI holds and, moreover, $\mathrm{P}(T_i = 1|X_i) < 1$ for all X_i,[13] we can estimate $\mathrm{E}(y_i^0|X_i; T_i = 1)$ on the basis of observations in the sample that have not received treatment. A very intuitive procedure to achieve this would be to construct a plausible 'ex post' control group consisting of non-treated observations such that the empirical distribution of X in this control group would match exactly the empirical distribution of X in the group of the treated. In real world applications, this is usually impractical due to the fact that a prudently chosen covariate vector X is multidimensional (and may contain several variables measured on a continuous scale) up to a degree that it is practically impossible to find exact matches in X between the two groups, given a sample of finite size. Several estimators to overcome this "curse of dimensionality" (Ho et al., 2007) problem have been proposed and broadly applied by the literature. In the present study, I apply two of them that share the common feature that they rely on the estimated *propensity score* $\hat{p}(X_i) \equiv \hat{\mathrm{P}}(T_i = 1|X_i)$, i.e. the estimated probability of receiving treatment given the covariates X_i.

The first estimator used here is a version of a matching estimator based on the propensity score. It relies on assembling a comparison group of non-treated observa-

[13] This is the common support (CS) condition required for ATT. It guarantees that, for any X_i for which treated observations exist, we can find non-treated observations having the same (or reasonably close) X_i with nonzero probability.

tions by matching them on the basis of their estimated propensity score; rather than directly on the basis of their covariate values (as suggested, but deemed impractical in the instructive example above). The underlying idea is that the propensity score is a so-called *balancing score* (Rosenbaum and Rubin, 1983), i.e. if CMI with respect to X_i holds, it equally holds with respect to $\hat{p}(X_i)$. For each observation i appearing in the group of treated observations, we thus choose one observation (or a fixed number of k observations) from the pool of non-treated observations that comes (or that come) "reasonably" close to i in terms of its (or their) estimated propensity score. Several variants to this basic idea exists; for a detailed overview, see Caliendo and Kopeinig (2008). The matching algorithm chosen for the present study is *nearest neighbour matching* (NNM): for each i having been treated, only the single non-treated observation $j(i)$ having the most similar estimated propensity score is retained: $|\hat{p}(X_i) - \hat{p}(X_{j(i)})| \leq |\hat{p}(X_i) - \hat{p}(X_k)|$ for any non-treated observation k. The NNM estimator for ATT thus becomes

$$\hat{ATT}_{NNM} = \frac{1}{n_T} \sum_{i|T_i=1} (y_i - y_{j(i)}) = \frac{1}{n_T} \sum_{i|T_i=1} y_i - \frac{1}{n_T} \sum_{i|T_i=1} y_{j(i)} \tag{1}$$

The second estimator used in the present study does not build upon any matching algorithm, but instead relies directly on the following result:

$$ATT = \frac{1}{\mathrm{P}(T=1)} \mathrm{E}\left[\frac{T - p(X)}{1 - p(X)} y\right]$$

This identity can be derived again by assuming CIA and CS (see e.g. Wooldridge, 2002, p. 615). Given a consistent estimator of the propensity score $\hat{p}(X)$, a consistent estimator for ATT is

$$\hat{ATT}_{HIR} = \frac{1}{n_T} \sum_{i=1}^{n} \frac{T_i - \hat{p}(X_i)}{1 - \hat{p}(X_i)} y_i = \frac{1}{n_T} \sum_{i|T_i=1} y_i - \frac{1}{n_T} \sum_{j|T_j=0} \frac{\hat{p}(X_j)}{1 - \hat{p}(X_j)} y_j \tag{2}$$

This estimator is termed $\hat{ATT}_{HIR}$ due to the fact that Hirano, Imbens and Ridder (2003) have shown it to be efficient under certain assumptions. As the procedure chosen by the present study does not entirely fulfil these assumptions,[14] claiming $\hat{ATT}_{HIR}$ to be efficient in our context cannot be justified on analytical grounds. However, if we limit our attention to the two estimators presented here, a glance at equations 1 and 2 can provide some interesting insights. As the component for the sample mean among treated units is identical in both estimators ($\frac{1}{n_T} \sum_{i|T_i=1} y_i$), they only differ in their calculation of the means for the counterfactual control group. Whereas $\hat{ATT}_{NNM}$ computes this value from the matched observations $y_{j(i)}$ (of which there are n_T or less), $\hat{ATT}_{HIR}$ uses the information available from all non-treated observations (which amount to $n - n_T$) by applying variable weights. Given a sufficiently large pool of non-treated observations relative to the amount of treated ones, the number of observations from

[14] In particular, that $\hat{p}(\cdot)$ be a series estimator (Wooldridge, 2002).

which information is drawn is thus much larger for the latter estimator. Provided that the weights are not too unevenly distributed, the intuition is that a lower variance will result from this. Such intuitive reasoning may of course fail; but it definitely provides a good motivation for a particular study such as the present one to include an empirical comparison of the performance of the two estimators with the data at hand by looking at their respective standard errors.

Calculating consistent standard errors for both of these estimators is, however, far from straightforward. Variance in both $A\hat{T}T_{NNM}$ and $A\hat{T}T_{HIR}$ originates not only from the fact that they are calculated on the basis of y_i's, but also from the first step of estimating $\hat{p}(X_i)$ on the basis of X_i's (and thus by relying on randomly sampled variables in both steps). In practice, taking correctly into account the two steps when producing estimates for standard errors is cumbersome to implement and computationally expensive to a degree that most applied studies refrain from doing it.[15] Instead, they either treat the first step of estimating $\hat{p}(X_i)$ as deterministic and resort to a simplified variance calculation based solely on the sampling variance of y_i, or they use bootstrapping to approximate the distribution and/or standard errors of average treatment effects estimates. See Caliendo and Kopeinig (2008) for a detailed discussion of this; and, in particular, Lechner (2002) for simulation results justifying the widespread use of the former approach. In the present study, I provide standard error estimates and tests on the statistical significance of results based both on simplified variance calculations and on bootstrapping. To this end, a non-parametric bootstrap with $B = 1000$ replications will be conducted, which includes a re-estimation of the propensity score equation at each iteration. This approach allows for a simple check of robustness regarding the reliability of the simplified procedure of variance calculation. some more stuff.

Simplified variance estimates are calculated as follows:

$$\hat{\text{Var}}(A\hat{T}T_{NNM}) = \frac{1}{n_T}\hat{\text{Var}}(y_i|T_i = 1) + \frac{\sum_{j\in\mathcal{M}} n_{j(i)}^2}{n_T^2}\hat{\text{Var}}(y_j|T_j = 0) \tag{3}$$

$$\hat{\text{Var}}(A\hat{T}T_{HIR}) = \frac{1}{n_T}\hat{\text{Var}}(y_i|T_i = 1) + \frac{1}{n_T^2}\sum_{j|T_j=0}\left(\frac{\hat{p}(X_j)}{1-\hat{p}(X_j)}\right)^2\hat{\text{Var}}(y_j|T_j = 0) \tag{4}$$

Where $\mathcal{M}$ is the set of index numbers of matched observations as determined by the nearest neighbour matching algorithm, and $n_{j(i)}$ is the number of times observations j has been matched to a treated observation by this algorithm.[16]

Since $A\hat{T}T_{HIR}$ as defined in equation 2 does not have the desirable property of linearity in the outcome variable y (this condition only holds asymptotically), a normal-

[15] Montes-Rojas (2009) provides a solution for a consistent variance estimator of $A\hat{T}T_{HIR}$.

[16] If matches are drawn from the set of non-treated observations without replacement, no multiple occurrences of matches will be present and the variance calculation for $A\hat{T}T_{NNM}$ slightly simplifies. In the present study, I draw matches with replacement.

isation of the weights used in the summation for the mean counterfactual outcome will be applied in the estimations that follow. See subsection 8.1 in the appendix for more details.

5 Data and Implementation

5.1 Data

Data used for the present paper has been collected by means of the Energy Technology Survey, a one-time survey conducted by KOF Swiss Economic Institute in 2009. The survey was addressed to members of the KOF Enterprise Panel, a representative sample of Swiss enterprises covering a wide range of economic activities — essentially the entire private sector, with the exception of agriculture, forestry and mining. Participation to the survey being voluntary, 2324 out of 5837 recipients returned valid questionnaires, resulting in a participation rate of 39.8%. Due to item nonresponse and implausible answers in those questions serving as explanatory variables for propensity score estimation, nearly 300 observations had to be rejected. In addition, 97 observations exhibiting extreme values in certain variables were excluded,[17] resulting in a number of 1966 observations retained for the econometric analysis that follows.

Table 2 gives an overview of the coverage of 29 branches and three size classes in the resulting sample. In addition, table 2 lists the numbers of firms having adopted one or more energy efficient technologies (EET), besides those having been supported to this end by each of the two (or any of the two) supporting agencies presented in section 2. 1049 firms — more than half of the sample — are users of EET. The respective probabilities of being among the group of EET adopters and of having successfully applied for support for this seem to differ only to a limited degree between the various economic branches. Looking at the size classes, it turns out that medium-sized and, even more so, large firms adopt EET more often. Interestingly, support for EET adoption is obtained much less often by small firms than medium-sized or large ones; and this finding holds not only in relation to the total number of firms in the sample, but also in relation to EET adopters.

Table 1 provides an overview of the various technology applications used by the survey to capture those that can be considered as energy efficient; and, for each of them, the number of adopters in the sample. The respective numbers of adopters reveal large differences between commonly implementable and thus widely used technologies (such as efficient lighting technologies) and rather specific technologies implemented only by a small group of firms (such as wind energy, or combined heat and power from

[17]Excluded observations: 38 firms with less than 5 as well as 7 firms with more than 5000 full time equivalent employees; plus 52 firms with a year of foundation prior to 1850. Not excluding these "extreme valued" observations for treatment effects estimation results in a somewhat less satisfying degree of balance, but does not considerably change the findings related to average treatment effects — see the subsection on robustness checks (6.3).

Table 2: Composition of Data Set by Industry and Size Class, Including Numbers of Supported and Adopting Firms

	Number of firms supported by			Total number of	
Industry (NACE Rev. 1.2)	Climate Cent Fnd.	Cantons, Munici-palities	Any of these	EET adopters	All firms
Food, beverage, tobacco (15, 16)	5	1	6	57	79
Textiles (17)	0	0	0	10	16
Clothing, leather (18, 19)	0	0	0	4	7
Wood processing (20)	0	2	2	19	33
Paper (21)	1	1	2	15	20
Printing (22)	1	0	1	25	50
Chemicals (23, 24)	6	1	7	46	73
Plastics, rubber (25)	3	0	3	31	48
Glass, stone, clay (26)	2	1	3	26	38
Metal (27)	1	1	1	13	23
Metal working (28)	3	4	6	74	155
Machinery (29)	3	3	6	100	169
Electrical machinery (31)	0	0	0	24	46
Electronics, instruments (30, 32, 331 – 334)	1	2	3	57	117
Watches (335)	0	1	1	13	31
Vehicles (34, 35)	1	1	1	13	19
Other manufacturing (36, 37)	0	1	1	15	32
Energy, water (40, 41)	2	7	7	29	39
Construction (45)	4	6	7	88	183
Wholesale trade (50. 51)	1	3	4	69	153
Retail trade (52)	1	1	2	55	115
Hotels, catering (55)	2	4	6	57	90
Transport, telecommunication (60 – 63)	2	4	6	75	121
Banks, insurance (65 – 67)	0	2	2	46	103
Real estate, leasing (70, 71)	2	4	5	9	14
Computer services (72, 73)	0	0	0	20	45
Business services (74)	1	5	5	45	126
Personal services (93)	2	1	2	10	13
Telecommunication (64)	0	0	0	4	8
Total	44	56	89	1049	1966
Size class (Number of FTE employees)					
Small (< 50)	4	10	13	409	983
Medium ($50 - 250$)	27	33	55	439	725
Large (≥ 250)	13	13	21	201	258
Total	44	56	89	1049	1966

biomass).

Descriptive statistics for variables that enter the specification of the propensity score equation (see the next subsection) are provided by table 3, whereas table 4 contains descriptive statistics for five outcome indicators for which ATT estimates will be produced. These outcome indicators are based on the three outcome variables described in section 2. As we wish to remain as flexible as possible in our choice of measurement of policy success, no a priori assumption was made e.g. as for whether success, within the targeted firms, should be measured by increases in their share of EET investments in total investments; or by increases in their per capita EET investments (i.e. EET investments divided by the number of full-time equivalent employees, or the logarithm thereof). Likewise, no clear presumption is made as to whether an increase in the breadth of the targeted firms' EET activities (as measured by the number of EET applications) is supposed to be induced in absolute terms (increasing the absolute value of the outcome variable) or in relative terms (increasing the logarithm of its value). Therefore, and since this comes at little computational expense, ATT estimates for five indicators ("variants" of the outcome variables of section 2) will be provided. The last two columns in table 4 indicate that — without conditioning on any pre-treatment covariates — for each outcome indicator, the mean value within the group of supported ("treated") firms is higher than within the group of non-supported firms. The central question of how much (if any) of this observed difference is attributable to treatment will be resolved in the next section.

Table 3: Descriptive Statistics of Explanatory Variables

Variable	Minimum	Maximum	Mean	Standard Deviation
LANG_FR (French-speaking)	0.000	1.000	0.161	0.367
LANG_IT (Italian-speaking)	0.000	1.000	0.043	0.203
LOG_EMPL	1.609	8.310	3.978	1.331
EMPL	5.000	4063.000	140.335	313.318
EMPL_SQ	25.000	16507969.000	117812.461	836167.425
AGE	1.000	158.000	56.748	35.665
LOG_AGE	0.000	5.063	3.789	0.793
FOREIGN (firm ownership)	0.000	1.000	0.141	0.349
HI_EDU (% higher edu. employees)	0.000	100.000	20.897	20.063
LOGC_HI_EDU (Logistic HI_EDU)	–3.045	3.045	–1.395	1.063

Table 4: Descriptive Statistics of Outcome Indicators

	All Observations				Treated:	Untreated:
Variable	Minimum	Maximum	Mean	Standard Deviation	Mean	Mean
EET Investments share	0.000	100.000	3.196	10.108	16.283	2.632
log per capita EET Inv.	1.411	12.226	6.693	1.711	7.471	6.570
CO_2 Reduction yes/no	0.000	1.000	0.360	0.480	0.872	0.335
log N adopted Tech.	0.000	3.178	1.506	0.722	1.927	1.467
N adopted Tech.	0.000	24.000	3.024	3.903	8.101	2.784

5.2 Choice of Propensity Score Specification

As mentioned earlier, both of the estimators for average treatment effects (ATT) used in this study crucially depend on having at our disposition a consistent estimate of the propensity score $\hat{p}(X_i)$. Given the uncertainty as for the correct specification of the propensity score equation (i.e. which parameters affect the probability of being supported, and by which functional form this influence is best captured), special care has been taken to find a specification that provides a convincing degree of balance of the covariates between the treated and matched groups. The present subsection addresses this problem.

The specification and results of the propensity score estimation ultimately chosen are given in Table 5. A Probit model has been used for this purpose. Industry affiliation for five manufacturing categories, the construction sector and two service categories is controlled for by dummy variables. Similarly, the language of the questionnaire sent to each firm is accounted for: there are German, French and Italian speaking regions in Switzerland; where German constitutes the reference category for the dummy variables constructed for this purpose. Since linear, quadratic and logarithmic formulations taken alone all led to unsatisfactory results in the case of the variables for firm age and firm size (as measured by the number of full time equivalent employees), a pragmatic approach has been chosen by letting logarithmic as well as linear terms (plus a quadratic term in the case of firm size) enter the equation side by side. FOREIGN is another binary variable indicating whether the firm is by majority non-domestically owned. To control for human capital intensity, a logistic transformation of the percentage of employees having completed higher education enters the specification (LOGC_HI_EDU).

Table 5: First Stage (Propensity Score) Estimation

	Probability of Support	
(Intercept)	–2.4563***	(0.5573)
IND_MANUF1	–0.4143	(0.3270)
IND_MANUF2	0.1217	(0.2082)
IND_MANUF3	–0.3827	(0.2388)
IND_MANUF4	–0.2196	(0.2583)
IND_CONSTR	–0.4096	(0.2597)
IND_SERV1	0.1977	(0.2257)
IND_SERV2	–0.3882	(0.2546)
LANG_FR	–0.3111*	(0.1733)
LANG_IT	–0.5403	(0.4018)
LOG_EMPL	0.3578***	(0.0904)
EMPL	–0.0006	(0.0007)
EMPL_SQ	0.0000	(0.0000)
LOG_AGE	–0.1708	(0.1486)
AGE	0.0055*	(0.0032)
FOREIGN	–0.3585**	(0.1766)
LOGC_HI_EDU	0.1070*	(0.0607)
N	1966	
Maximised log-L	–319.7	
Null log-L	–362.4	
Pseudo-R^2	0.118	

Probit estimates. Standard errors in brackets. Stars denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels respectively.

Estimates in Table 5 reveal that, controlling for other covariates, firms located in the French and Italian speaking regions of Switzerland have been less likely to receive support than firms located in the German speaking part. While raising a politically relevant question in its own right, this finding does not impair in any way the validity of my results. A potential problem might arise only with the hypothesis that this effect is driven by institutional differences between French- and Italian-speaking cantons with respect to their German-speaking counterparts, which would mean that the observed underrepresentation is confined to support from cantons and municipalities and not present in CCF support. As a consequence, the propensity score estimation used in this study may be flawed. Estimating the effect of the covariates on the probability of receiving support from either one of the two schemes separately rather than on the probability of receiving support from at least one of the two schemes, however, yields no evidence for the above hypothesis (see subsection 6.3); the estimated coefficients being similar in sign and magnitude when considering the two schemes separately. On the contrary, an ex-post evaluation of a specific Climate Cent Foundation sub-programme (BFE, 2010; *in German only*) acknowledges this apparent underrepresentation of participants from the French and Italian regions in the context of the CCF and identifies information problems as the most probable cause (in the non-German speaking parts, a smaller share of potential recipients was aware about the program's existence).

A simple check of whether $\hat{p}(X_i)$ thus obtained is acceptable as an estimate for the propensity score can be carried out by comparing the mean difference in the values of components of X_i resulting from treatment as predicted by the estimators for ATT. In other words, $A\hat{T}T$ can be calculated for a pre-treatment covariate by letting $y_i = x_i$ for any x_i contained in X_i. By definition, $ATT = \mathrm{E}(x_i^1 - x_i^0|X_i; T_i = 1) = 0$ in this case,[18] so that consistent estimates of ATT for x_i asymptotically tend to zero (and any deviations in small samples are attributable to sampling error). Tables 6 and 7 make use of this property and report estimates for such average "treatment" effects for covariates used to estimate $\hat{p}(X_i)$ (except for the industry dummies) on the basis of $A\hat{T}T_{NNI}$ and $A\hat{T}T_{HIR}$, respectively.

The sample means within the groups of treated, matched and untreated observations appear in the first three columns, while the fourth and fifth columns contain the difference between the means of treated and matched and p-Values for the associated t-Test. Systematic deviations from zero should be interpreted as a hint that estimation of $\hat{p}(X_i)$ if flawed, due to e.g. inappropriate functional form assumptions. Clearly, this is not the case here, as p-values above 0.5 are found throughout. The median p-values across the ten variables considered are 0.661 (for $A\hat{T}T_{NNM}$) and 0.855 (for $A\hat{T}T_{HIR}$), respectively.

Since good balance requires similarity in the empirical distributions of the covariates between the groups of treated and matched observations, ensuring that means of covari-

[18]In fact, for any pre-treatment covariate x_i, there can be no meaningful distinction between x_i^0 and x_i^1, as the value for x_i is fixed rather than dependent on treatment. In the case of (erroneously) putting a variable that actually is affected by treatment into the covariate vector X_i, the affirmation that ATT thus defined is zero still holds by construction (due to the fact that the ATT estimator is obtained by conditioning on X_i).

Table 6: Covariates: Means among Treated and Matched Observations (NNM)

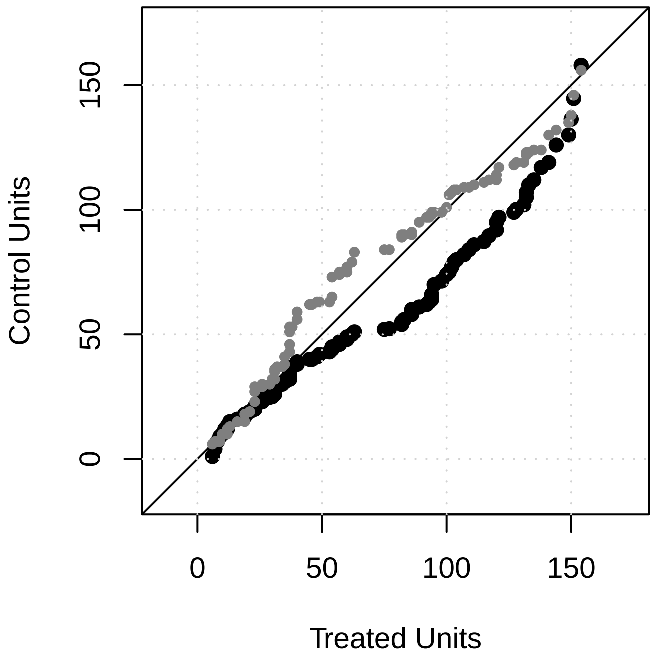
	Means			Difference	
	Treated	Matched	Untreated	Tr – Ma	p-Value
LANG_FR (French-speaking)	0.090	0.056	0.164	0.034	0.517
LANG_IT (Italian-speaking)	0.011	0.011	0.045	0.000	1.000
LOG_EMPL	4.855	4.948	3.936	–0.094	0.627
EMPL	233.640	254.448	135.911	–20.808	0.664
EMPL_SQ	144777.640	185080.583	116533.878	–40302.942	0.705
AGE	70.034	73.101	56.118	–3.067	0.618
LOG_AGE	3.972	4.036	3.780	–0.064	0.613
FOREIGN (firm ownership)	0.101	0.079	0.143	0.022	0.660
HI_EDU (% higher edu. employees)	20.584	21.870	20.912	–1.286	0.662
LOGC_HI_EDU (Logistic HI_EDU)	–1.335	–1.344	–1.398	0.009	0.955

Table 7: Covariates: Means among Treated and Matched Observations (HIR)

	Means			Difference	
	Treated	Matched	Untreated	Tr – Co	p-Value
LANG_FR (French-speaking)	0.090	0.088	0.164	0.002	0.957
LANG_IT (Italian-speaking)	0.011	0.013	0.045	–0.001	0.918
LOG_EMPL	4.855	4.862	3.936	–0.007	0.955
EMPL	233.640	229.455	135.911	4.185	0.902
EMPL_SQ	144777.640	133929.490	116533.878	10848.151	0.843
AGE	70.034	71.327	56.118	–1.293	0.790
LOG_AGE	3.972	4.009	3.780	–0.036	0.698
FOREIGN (firm ownership)	0.101	0.092	0.143	0.009	0.789
HI_EDU (% higher edu. employees)	20.584	21.676	20.912	–1.092	0.582
LOGC_HI_EDU (Logistic HI_EDU)	–1.335	–1.318	–1.398	–0.017	0.867

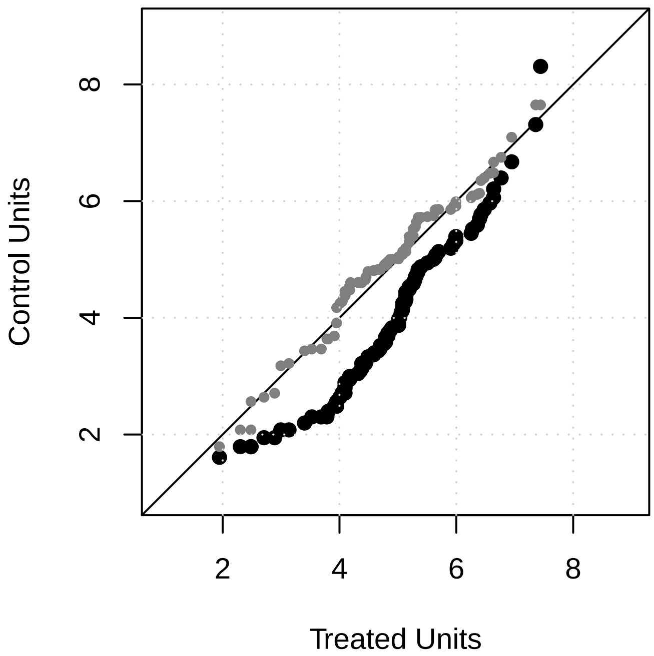
ates between the two groups do not significantly differ is a necessary but not sufficient criterion for balance. Ideally, the empirical distributions of covariates in both groups should coincide all over their respective support region. As a visual aid, QQ plots may give some indication as for how closely the empirical distributions of a variable between two samples coincide. Figures 1 to 3 plot empirical quantiles of matched (relying on the nearest neighbour matching procedure on which $\hat{ATT}_{NNM}$ is based) versus treated values with respect to the three variables AGE, LOG_EMPL and HI_EDU. For the sake of completeness, figure 4 shows coincidence of distributions for the estimated propensity score. Coincidence of distributions (which manifests itself by how close the dots come to lie to the 45 degree line) is clearly improved in the course of matching for the variables AGE and LOG_EMPL, whereas the effect is less evident for HI_EDU, but still present to some degree in the lower segment of the distribution range.

Figure 1: QQ Plot of Covariate AGE (Treated versus Raw and Matched)



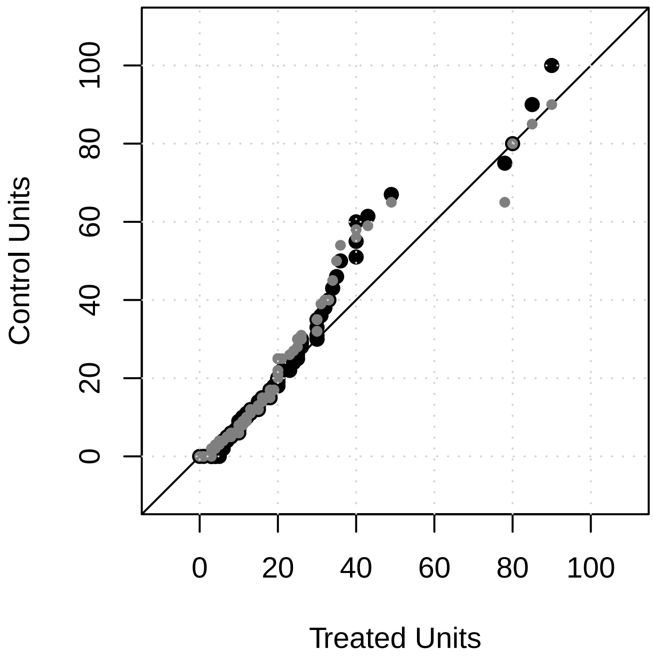
Black dots: “raw” (prior to matching) untreated units vs. treated units;
gray dots: matched untreated units vs. treated units.

Figure 2: QQ Plot of Covariate LOG_EMPL (Treated versus Raw and Matched)



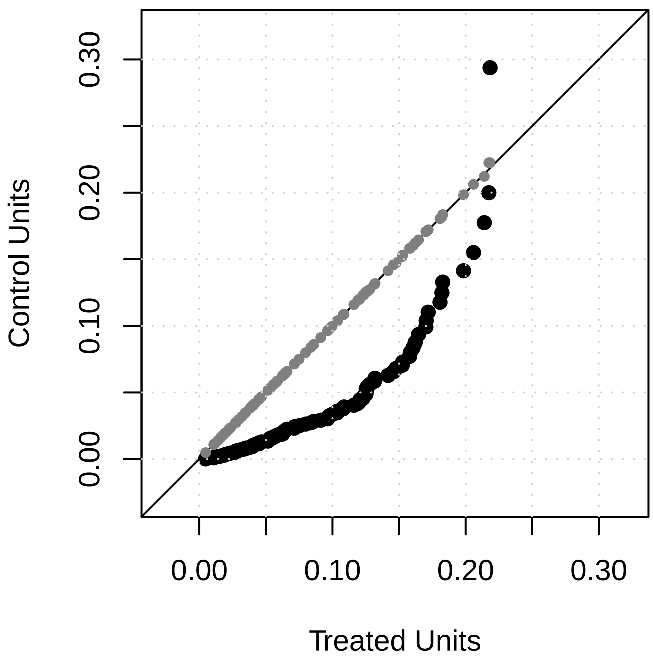
Black dots: “raw” (prior to matching) untreated units vs. treated units;
gray dots: matched untreated units vs. treated units.

Figure 3: QQ Plot of Covariate HI_EDU (Treated versus Raw and Matched)



Black dots: "raw" (prior to matching) untreated units vs. treated units;
gray dots: matched untreated units vs. treated units.

Figure 4: QQ Plot of Estimated Propensity Score (Treated versus Raw and Matched)



Black dots: "raw" (prior to matching) untreated units vs. treated units;
gray dots: matched untreated units vs. treated units.

6 Results

6.1 Baseline Results

Table 8 presents, for each of the five indicators covering the three outcome variables, mean values of the treated, matched and untreated observations according to $\hat{ATT}_{NNI}$. Table 9 likewise does for $\hat{ATT}_{HIR}$. The average treatment effect of the treated (ATT), defined as the difference in means of the treated and matched, is reported on the right-hand side of the tables, alongside the p-Value of the associated t-test for the null hypothesis of a zero treatment effect. It follows that, for each of the five indicators, a positive average treatment effect on the treated (ATT) results, no matter which of the two estimators considered here is used. The effects are statistically significant at the 1% level for all indicators and for both estimators. With the exception of log per capita EET investments and the log number of adopted technologies, the effects even pass a test of significance at a level of 0.1%, for both estimators. In other words, all of the three outcome variables appear positively affected by the fact of having been supported by either of the two institutions. In the case of EET investments, this finding is invariant as to whether such investments are measured as a percentage of total investments, or in relation to the firm's total number of employees. However, the latter formulation turns out slightly less robust in terms of statistical significance.

Table 8: Outcome Indicators: Means of Treated, Matched and Untreated (NNM)

	Means			Difference	
	Treated	Matched	Untreated	Tr – Ma (ATT)	p-Value
EET Investments share	16.283	3.594	2.632	12.688	0.000***
log per capita EET Inv.	7.471	6.408	6.570	1.063	0.003***
CO_2 Reduction yes/no	0.872	0.407	0.335	0.465	0.000***
log N adopted Tech.	1.927	1.623	1.467	0.304	0.008***
N adopted Tech.	8.101	3.483	2.784	4.618	0.000***

Standard errors and p-values according to simplified variance calculation as in equation 3.
Stars denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels respectively.

Given the difficult data situation with regards to quantitative variables — e.g. about the amount of support received by firms, or about units of energy saved or CO_2 emissions avoided — possibilities for quantifying the gains of the policy schemes considered here in more concrete terms are very limited. A valuable insight provided by the present study is, nevertheless, that crowding-out problems do not seem to appear (or, in any case, they appear to a limited degree only) in the context of current public or private sector based support schemes for EET in Switzerland. That is to say, firms have implemented projects that enhance their energy efficiency and invested funds to this end, to an extent that would not have resulted, if they had not been granted the financial support by the schemes considered here. If this were not the case (i.e. in the presence of complete

Table 9: Outcome Indicators: Means of Treated, Matched and Untreated (HIR)

	Means			Difference	
	Treated	Matched	Untreated	Tr – Ma (ATT)	p-Value
EET Investments share	16.283	4.348	2.632	11.935	0.000***
log per capita EET Inv.	7.471	6.716	6.570	0.755	0.008***
CO_2 Reduction yes/no	0.872	0.492	0.335	0.380	0.000***
log N adopted Tech.	1.927	1.552	1.467	0.375	0.000***
N adopted Tech.	8.101	3.882	2.784	4.219	0.000***

Standard errors and p-values according to simplified variance calculation as in equation 4.
Stars denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels respectively.

crowding-out), no positive impact would result in estimates of the Average Effect of Treatment on the Treated (ATT).

As has already been mentioned (and as is usual with all studies of this kind), the results found here are valid only insofar as no information having affected individual firms' probability of being supported has been absent from the data, i.e. that the set of covariates in my propensity score estimation (presented in Table 5) is complete, and that the equation has been correctly specified. This is evidently a strong assumption, however one that cannot be circumvented. In this sense, care has been taken that the present study represents the best possible effort to assess the effectiveness of EET adoption policy in Switzerland, given the data limitations inherent in most outcome analysis based on observational research.

6.2 Comparative Empirical Assessment of the Two Estimators

Table 10 compares the estimates for the two estimators already reported in tables 8 and 9 and, in addition, provides standard errors and p-values based both on simplified variance calculations and on bootstrapping, as described in subsection 4.

A first finding — already evident in tables 8 and 9 — is that estimates differ somewhat between the two procedures, but the quantitative findings (sign and statistical significance) are roughly the same, no matter which estimator is chosen. For the effect on log per capita EET investments, the average effect when estimated by the NNM procedure is 41% higher than when estimated by the HIR estimator, whereas for the other four indicators, the relative differences are less than 25%, and there is little evidence that one estimator systematically produces estimates of larger magnitude than the other.

Turning to the estimated standard errors, however, we can see that $A\hat{T}T_{HIR}$ produces more precise estimates (i.e. at a lower estimated standard error) than $A\hat{T}T_{NNM}$ for all five indicators. This finding holds for both cases of comparing standard errors; i.e. those based on simplified variance calculation, and those obtained by bootstrapping.

If we compare, for each of the two estimators individually, differences between calcu-

Table 10: Outcome Indicators: Treatment Effects, both NNM and HIR, and Bootstrapped Standard Errors

Indicator	Estimator	Estimate	Std. Err (S)	p-Value (S)	Std. Err (BS)	p-Value (BS)	p-Value (BS, emp.)
EET Inv. share	$\hat{ATT}_{HIR}$	11.935	3.005	0.000	3.165	0.000	0.000
	$\hat{ATT}_{NNM}$	12.688	3.172	0.000	3.682	0.001	0.000
log pc EET Inv.	$\hat{ATT}_{HIR}$	0.755	0.275	0.008	0.298	0.016	0.009
	$\hat{ATT}_{NNM}$	1.063	0.342	0.003	0.421	0.017	0.032
CO_2 red. yes/no	$\hat{ATT}_{HIR}$	0.380	0.040	0.000	0.044	0.000	0.000
	$\hat{ATT}_{NNM}$	0.465	0.068	0.000	0.089	0.000	0.000
log N Tech.	$\hat{ATT}_{HIR}$	0.375	0.075	0.000	0.074	0.000	0.000
	$\hat{ATT}_{NNM}$	0.304	0.112	0.008	0.128	0.023	0.002
N Tech.	$\hat{ATT}_{HIR}$	4.219	0.470	0.000	0.459	0.000	0.000
	$\hat{ATT}_{NNM}$	4.618	0.617	0.000	0.785	0.000	0.000

Standard errors (S) and p-values (S): based on simplified variance calculation as in equations 3 and 4. Standard errors (BS) and p-values (BS): based on bootstrapped estimates and normal approximations thereof, respectively. p-values (BS, emp.): based on empirical quantiles of bootstrapped estimations. Number of bootstrapping replications (B): 1000

lated and bootstrapped standard errors, it turns out that bootstrapped standard errors are always and considerably higher than calculated ones for $\hat{ATT}_{NNM}$: by a margin of 22% on average for the five indicators (with individual margins ranging between 14% and 31%). As for $\hat{ATT}_{HIR}$, they are not uniformly higher (only in three out of five cases), but remain higher on average by 4% (individual margins: from –2% to 10%). These findings suggest that we should treat standard error estimates based on the simplified procedure with caution, as the degree of underestimation might be substantial, in particular when NNM is used. It may be sensible to assume that in larger samples (or in samples with a larger number of treated units) this kind of bias may be lower, but in the present case it is far from negligible. As a practical consequence with the sample at hand, t-tests based on bootstrapped rather than calculated standard errors indicate that ATT estimates are different from zero only within a confidence bound of 95% (as opposed to 99%, when relying on calculated standard errors), for three out of the ten cases listed in table 10.

The finding of a larger discrepancy between calculated and bootstrapped variances in the case of $\hat{ATT}_{NNM}$ than in the case of $\hat{ATT}_{HIR}$ means that the efficiency advantage of HIR over NNM is more dramatic when bootstrapped standard errors are considered: relying on simplified variance calculation, standard error estimates for NNM appear larger by between 5% and 70% than for HIR; and relying on bootstrapping, they appear larger by between 16% and 102%.

There are thus two main insights from this subsection: $\hat{ATT}_{HIR}$ outperforms $\hat{ATT}_{NNM}$ in terms of efficiency; a finding that is confirmed no matter whether standard errors are produced according to simplified variance calculation or on the basis of bootstrapping, the efficiency advantage being considerably larger in the second case. And, simplified variance calculation according to equations 3 and 4 produces standard error estimates that are most likely affected by severe downward bias.

6.3 Robustness Checks

Several robustness checks have been performed in order to confirm that the results of the baseline estimations presented above are not driven by the specific sample and estimation procedure chosen.

Public vs. Private Sector Support

A separate analysis for the impacts of the *public* entities (cantons and municipalities) on the one hand, and the *private* sector (Climate Cent Foundation) on the other hand does not generate any new insights. The estimated effects remain positive and fairly similar in magnitude for both programmes, but generally at a lower statistical significance.[19] Yet, positive effects (significant at a 5% level) on CO_2 emission reduction and on the number of technologies remain for each of the two supporting bodies. As for EET investments and when looking at the NNM estimator, the null hypothesis of no treatment effect cannot be rejected in the case of EET investment share for the Climate Cent Foundation, and of log per capita EET investments for Cantons and Municipalities. Nevertheless, estimates based on the HIR estimator remain significant at 5% for both of these outcome indicators.

No Sample Restriction

As mentioned in subsection 5.1, 97 observations were excluded for calculation of baseline estimates out of concern that their extreme values in terms of size or age might negatively affect the quality of estimates (see footnote 17). In order to assess whether this might introduce additional bias through sample selection, ATT estimates were also computed using a non-restricted sample that incorporates these 97 cases. Goodness of matching slightly worsened as a result, the median p-value for covariate difference now being 0.586 for the NNM procedure (previously: 0.661), and 0.814 for the HIR procedure (previously: 0.855). Treatment effects estimates hardly were affected by this; $A\hat{T}T_{HIR}$ turned out slightly (but not dramatically) lower for all five outcomes, and $A\hat{T}T_{NNM}$ were slightly lower for some and slightly higher for the other outcome indicators. At any rate, effects remained statistically significant at comparable levels, with the exception of $A\hat{T}T_{HIR}$ for log per capita EET investments, which was significant at only 5% (baseline: 1%) in the non-restricted sample.

[19] This is not surprising, as considering only one of the two programmes at a time means an even lower number of treated units on which the estimate can be based (see table 2), resulting in a higher sampling error.

Control Group Restriction

For the estimates presented so far, the control group was assembled by drawing from the entire pool of survey respondents (minus those observations that did not provide responses to questions that were essential for the computation of estimates, and those observations exhibiting extreme values as described in footnote 17). This procedure could be criticised as being overly permissive in the following sense: not having at my disposition any pre-treatment covariates that capture a firm's inherent patterns of energy use, and given the fact that these patterns may be very heterogeneous, the CMI assumption might not hold. Put differently: treated firms may have been more likely to receive support due to the fact that they e.g. have been owners of vehicles or buildings (which are both energy-intensive), or have been relying on particularly energy-intensive processes; and characteristics such as these may not have been taken into account properly by controlling for the covariates that are available in the present study.

If a violation of the CMI of this kind is suspected, it is evidently not possible to overcome the resulting bias without including additional covariates that capture the above-mentioned parameters (which we unfortunately do not dispose of). However, insights can be provided by a robustness check of the following kind: restricting the potential control group to those observations that have, ex post, invested in at least one energy-efficient technology leaves us with a sample for which violations of the CMI of the above kind are highly unlikely; as only those observations are considered where actual opportunities for improving energy efficiency were present in the first place. The resulting estimates will evidently have a different (and somewhat problematic) interpretation than the baseline estimates of subsection 6.1; since, by construction, they neglect the possibility that there may be cases of firms who have received support and who actually would not have undertaken *any* EET investment without support. What matters here is the fact that they provide a means of obtaining conservative estimates of ATT in the sense that problems with the CMI as described above (which would lead to an upward bias in estimates) are ruled out, at the cost of incurring a different kind of bias (which is downward) as a result of reducing the pool of comparison observations in a manner that is most probably too restrictive.

The results (not shown here) of these estimations, based on 1049 observations in the sample having adopted at least one EET application, indicate that only the effects on the outcome indicators "reduction of CO_2 emissions" and "number of technological fields" are strongly affected. They remain positive, but drop in magnitude by 88% and 56%, respectively for $\hat{ATT}_{NNM}$, or by 69% and 46% for $\hat{ATT}_{HIR}$, in comparison to the baseline estimates. As a result, the effect on "reduction of CO_2 emissions" is not significantly different from zero in the NNM estimator, but it remains significant at 1% in the HIR estimator. The effect on the "number of technological fields" remains significant at 1% in both estimators. ATT estimates for the other three outcome indicators fall by no more than one third or remain roughly constant, and their statistical significance is not affected to a notable degree; except for per capita log EET investments in the NNM estimator, which is significant at 5% only (baseline: 1%). In summary, results of this robustness check suggest that, if there is any upward bias in my baseline estimates due to not sufficiently taking into account energetic characteristics prior to treatment of firms, it is only minor in the sense that the qualitative finding of a positive treatment effect remains unaffected for four out of five outcome indicators (and, for the fifth outcome indicators, is questionable on the basis of one out of the two estimators only).

7 Conclusion

Survey data covering a representative sample of Swiss firms belonging to the manufacturing, construction and service sectors has been used to assess if any public and private support received by these firms in order to adopt energy efficient technology (EET) has resulted in a favourable outcome in terms of the number of adopted technologies, CO_2 emission reduction and the amount of investment related to such technology. A total number of 89 firms in the sample have been identified as being beneficiaries of support programmes; either by the private sector based Climate Cent Foundation, or by public entities (Cantons or Municipalities). Applying nearest neighbour propensity score matching, as well as the HIR estimator (as described in Hirano, Imbens and Ridder, 2003) that directly builds on the estimated propensity score, all of the outcome variables have been found to respond positively and significantly to such support measures.

A secondary finding of the present study concerns the use of the HIR estimator as an alternative for estimation of average treatment effects on the treated (ATT). In comparison to the widely used procedure of nearest neighbour matching (NNM) based on the propensity score, this estimator is shown to perform considerably better in terms of efficiency. This can intuitively be attributed to the fact that, for computing estimates for ATT, this latter estimator uses information of a potentially much larger pool of (untreated) comparison observations than the NNM estimator, while essentially requiring the same classical assumptions for identification. Further studies using real data or Monte Carlo simulations may be helpful to determine whether this HIR procedure outperforms NNM, or others of the more traditional procedures,[20] to a similar degree in other situations.

While the analysis, due to inherent data limitations, does not bring about any conclusions regarding the magnitude of success (i.e., what has been the gain in terms of reduced energy consumption by monetary unit of means granted, and whether the funds have been allocated in the most cost-effective manner), it provides tangible evidence that no discernible crowding-out effects emerge as a result of these support schemes. This

[20]As there is a large variety of matching estimators used in practice, the Nearest Neighbour Matching (NNM) estimator just being a particular one of them, a thorough evaluation of the potential merits of the HIR estimator should confront it to several variants of matching estimators.

is an important finding that contributes to justifying ongoing efforts to facilitate diffusion of energy efficient technology, given that reducing our energy systems' reliance on non-sustainable and environmentally costly sources of energy is a long-term challenge that cannot be expected to solve itself without decisive and well-coordinated policy intervention.

8 Appendix

8.1 Linearity of $\hat{ATT}_{HIR}$

$\hat{ATT}_{HIR}$ as defined in equation 2 has the inconvenience that the weights $\frac{1}{n_T}\frac{\hat{p}(X_j)}{1-\hat{p}(X_j)}$ appearing in the summation for the mean of the (counterfactual) control group need not sum up to one. As a consequence, $\hat{ATT}_{HIR}$ does not necessarily have the property of linearity in the outcome variable y. This means that linear transformations of y may result in estimate values (and their test statistics) varying in an unexpected manner, particularly in small samples.

The requirement that the above summation of weights equal one (and thus linearity of $\hat{ATT}_{HIR}$), however, holds asymptotically. To see this, consider

$$\text{plim}_{n\to\infty}\frac{1}{n_T}\sum_{j|T_j=0}\frac{\hat{p}(X_j)}{1-\hat{p}(X_j)} = \text{plim}_{n\to\infty}\frac{1}{\sum_{i=1}^{n}T_i}\sum_{i=1}^{n}(1-T_i)\frac{\hat{p}(X_i)}{1-\hat{p}(X_i)}$$

which, for any given X_i, is indeed one (again assuming consistency in $\hat{p}(X_i)$). It follows that, for any distribution of X_i, the above probability limit is one.

We can thus modify $\hat{ATT}_{HIR}$ by normalising the weights of the second summation term as follows:

$$\hat{ATT}^{*}_{HIR} = \frac{1}{n_T}\sum_{i|T_i=1} y_i - \frac{1}{\sum_{j|T_j=0}\frac{\hat{p}(X_j)}{1-\hat{p}(X_j)}}\sum_{j|T_j=0}\frac{\hat{p}(X_j)}{1-\hat{p}(X_j)}y_j$$

in order to obtain an estimator for ATT that is linear in y, without losing consistency: as $n\to\infty$, $\hat{ATT}^{*}_{HIR}$ coincides with $\hat{ATT}_{HIR}$; a finding which directly follows from the above result. The corresponding adjustments required for the variance calculation as in equation 4 are straightforward.

References

Acemoglu, D., Aghion, P., Bursztyn, L., Hemous, D. (2012), The environment and directed technical change, *American Economic Review*, 102, 1, 131–166

Arvanitis, S., Donzé, L., Sydow, N. (2010), Impact of Swiss technology policy on firm innovation performance: An evaluation based on a matching approach, *Science and Public Policy*, 27, 1, 63–78

Arvanitis, S., Ley, M. (2012), Factors determining the adoption of energy-saving technologies in Swiss firms: An analysis based on micro data, *Environmental and Resource Economics (in press)*, DOI 10.1007/s10640-012-9599-6

Battisti, G. (2008), Innovations and the economics of new technology spreading within and across users: Gaps and way forward, *Journal of Cleaner Production*, 16, 1, 22–31

BFE (2010), Evaluation des Gebäudeprogramms der Stiftung Klimarappen, *Bundesamt für Energie (BFE)*, Bern

BFE (2011), Analyse des schweizerischen Energieverbrauchs 2000–2010 nach Verwendungszwecken, *Bundesamt für Energie (BFE)*, Bern

Caliendo, M., Hujer, R. (2006) The microeconometric estimation of treatment effects — an overview, *AStA Advances in Statistical Analysis*, 90, 1, 199–215

Caliendo, M., Kopeinig, S. (2008), Some practical guidance for the implementation of propensity score matching, *Journal of Economic Surveys*, 22, 1, 31–72

Carley S. (2009), State renewable energy electricity policies: An empirical evaluation of effectiveness, *Energy Policy*, 37, 3071–3081

Delmas, M. A., Montes-Sancho M. J. (2011), U.S. state policies for renewable energy: Context and effectiveness, *Energy Policy*, 39, 2273–2288

Fischer, C., Newell, R. G. (2008), Environmental and technology policies for climate mitigation, *Journal of Environmental Economics and Management*, 55, 142–162

Hirano, K., Imbens G. W., Ridder G. (2003), Efficient estimation of average treatment effects using the estimated propensity score, *Econometrica*, 71, 4, 1161–1189

Ho, D. E., Imai, K., King, G., Stuart, E. A. (2007), Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference, *Political Analysis*, 15, 199–236

Holland, P. W. (1986), Statistics and causal inference, *Journal of the American Statistical Association*, 81, 945–60

Johnstone, N., Haščič, I., Popp, D. (2010), Renewable energy policies and technological innovation: Evidence based on patent counts, *Environmental and Resource Economics*, 45, 133–155

Karshenas, M., Stoneman, P. L. (1993), Rank, stock, order, and epidemic effects in the diffusion of new process technologies: An empirical model, *Rand Journal of Economics*, 24, 4, 503–528

Lechner, M. (2002), Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods, *Journal of the Royal Statistical Society*, 165, 1, 59–82

Montes-Rojas, G. (2009), A note on the variance of average treatment effects estimators, *Economics Bulletin*, 29, 4, 2937–2943

Mulder, A. (2008), Do economic instruments matter? Wind turbine investments in the EU(15) *Energy Economics*, 30, 2980–2991

Noailly, J. (2012), Improving the energy efficiency of buildings: The impact of environmental policy on technological innovation, *Energy Economics*, 34, 795–806

Popp, D., Newell, R. G., Jaffe, A. B. (2009), Energy, the environment, and technological change, *NBER Working Papers*, 14832, Cambridge, MA

Rennings, K., Rammer, C. (2009), Increasing energy and resource efficiency through innovation: An explorative analysis using innovation survey data, *Czech Journal of Economics and Finance (Finance a uver)*, 59, 5, 442–459

del Río, P., Tarancón, M.-A. (2012), Analysing the determinants of on-shore wind capacity additions in the EU: An econometric study, *Applied Energy*, 95, 12–21

Rosenbaum, P. R., Rubin, D. B. (1983), The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70, 41–55

Stiftung Klimarappen (2010), Jahresbericht 2009, *Stiftung Klimarappen*, Zürich

Strebel, F. (2011), Inter-governmental institutions as promoters of energy policy diffusion in a federal setting, *Energy Policy*, 39, 467–476

Velthuijsen, J. W. (1993), Incentives for investment in energy efficiency: An econometric evaluation and policy implications, *Environmental and Resource Economics*, 3, 153–169

Wooldridge, J. M. (2002), Econometric analysis of cross section and panel data, MIT Press, Cambridge MA