Working Papers | 114

19.1.2024

Least-cost Decarbonization Pathways for Electricity Generation in Finland

A CONVEX QUANTILE REGRESSION APPROACH

Natalia Kuosmanen (Corresponding author)
ETLA Economic Research, Finland
natalia.kuosmanen@etla.fi

Timo Kuosmanen
Turku School of Economics, University of Turku, Finland

Terhi Maczulskij
ETLA Economic Research, Finland

Xun Zhou
Surrey Business School, University of Surrey, UK; Department of Environment and Geography, University of York, UK

Suggested citation:
Kuosmanen, Natalia, Kuosmanen, Timo, Maczulskij, Terhi & Zhou, Xun (19.1.2024). "Least-cost Decarbonization Pathways for Electricity Generation in Finland: A Convex Quantile Regression Approach". ETLA Working Papers No 114.
http://pub.etla.fi/ETLA-Working-Papers-114.pdf

Abstract

This study investigates the least-cost decarbonization pathways in the Finnish electricity generation industry in order to achieve the national carbon neutrality goal by 2035. Various abatement measures, such as downscaling production, capital investment, increasing labor and intermediate inputs are considered. The marginal abatement costs (MACs) of greenhouse gas emissions are estimated using the convex quantile regression method and applied to unique register-based firm-level greenhouse gas emission data merged with financial statement data. We adjust the MAC estimates for the sample selection bias caused by zero-emission firms by applying the two-stage Heckman correction. Our empirical findings reveal that the median MAC ranges from 0.1 to 3.5 euros per tonne of CO_2 equivalent. The projected economic cost of a 90% reduction in emissions is 62 million euros, while the estimated cost of achieving zero emissions is 83 million euros.

Tiivistelmä

Kustannustehokkaat polut sähköntuotannon hiilidioksidipäästöjen vähentämiseen: Konveksiin kvantiiliregressioon perustuva analyysi

Tässä tutkimuksessa tarkastellaan kustannustehokkaimpia keinoja hiilidioksidipäästöjen vähentämiseksi sähköntuotannon toimialalla Suomen kansallisen hiilineutraaliustavoitteen saavuttamiseksi vuoteen 2035 mennessä. Tutkimuksessa verrataan erilaisten päästövähennysstrategioiden vaihtoehtoiskustannuksia huomioiden mahdollisuudet vähentää energiankulutusta, investoida puhtaampaan teknologiaan sekä lisätä työvoiman tai välituotteiden käyttöä. Päästövähennysten rajakustannuksia arvioidaan konveksiin kvantiiliregressioon perustuvan tilastollisen menetelmän avulla. Tutkimusaineistona hyödynnetään ainutlaatuisia rekisteripohjaisia kasvihuonekaasupäästötietoja, jotka yhdistetään yritysten tilinpäätöstietoihin. Koska päästöaineisto ei kata kaikkia yrityksiä, sovellamme kaksivaiheista Heckman-korjausta mahdollisen valikoitumisharhan korjaamiseen. Empiiristen tulostemme mukaan päästövähennysten rajakustannusten mediaani vaihtelee välillä 0,1–3,5 euroa hiilidioksiditonnilta. Arvioitu taloudellinen kustannus 90 prosentin päästövähennykselle nykyisen tasoon verrattuna on vähintään 62 miljoonaa euroa, kun taas nollapäästöjen saavuttamisen arvioitu minimikustannus on 83 miljoonaa euroa.

Ph.D. **Natalia Kuosmanen** is a Chief Research Scientist at Etla Economic Research.

Ph.D. (Econ.) **Timo Kuosmanen** is a Professor of Economics at Turku School of Economics, University of Turku.

Ph.D. **Terhi Maczulskij** is a Chief Research Scientist at Etla Economic Research and a Research Fellow at IZA.

Ph.D. **Xun Zhou** is a Senior Lecturer in Business Analytics, Surrey Business School, University of Surrey.

MMT **Natalia Kuosmanen** on Elinkeinoelämän tutkimuslaitoksen tutkimuspäällikkö.

KTT, YTM **Timo Kuosmanen** on taloustieteen professori Turun yliopistoon kuuluvassa Turun kauppakorkeakoulussa.

KTT, FM, dosentti **Terhi Maczulskij** on Elinkeinoelämän tutkimuslaitoksen tutkimuspäällikkö sekä IZA:n Research Fellow.

KTT **Xun Zhou** on vanhempi yliopistonlehtori, Surrey Business School, University of Surrey.

Acknowledgements: The financial support provided by Business Finland as part of the larger project “Innovation-led Pathways to Economic Growth” is gratefully acknowledged.

Kiitokset: Kiitämme Business Finlandia taloudellisesta tuesta osana laajempaa hankkeetta ”Innovation-led Pathways to Economic Growth”.

Keywords: Abatement cost, Convex quantile regression, Forward-looking assessment, Climate policy, Decarbonization pathways

Asiasanat: Ilmastopolitiikka, Hiilidioksidipäästöt, Konveksi kvantiiliregressio, Tulevaisuuteen suuntautuva arviointi

JEL: O44, Q43, Q51, Q52, Q54

1 Introduction

Efficient carbon abatement depends on the ability of firm managers and policymakers to identify least-cost abatement options. This requires an understanding of the role of marginal abatement cost (MAC), the cost associated with reducing the emission of one additional unit of pollutant or greenhouse gas (GHG). The MAC is a key concept in environmental economics and climate change mitigation and plays an essential role in pricing pollutants and guiding environmental policies.

Recent studies on the empirical assessment of MAC and the identification of least-cost pathways for emission reduction using the convex quantile regression (CQR) approach proposed by Kuosmanen and Zhou (2021) have gained increased popularity. This approach introduces a data-driven procedure that explicitly incorporates multiple abatement options, inefficiency, and stochastic noise, thereby providing a robust framework for estimating shadow prices and MACs. Kuosmanen et al. (2020) pioneered the application of the CQR approach in a cross-country analysis of OECD countries, revealing that actual abatement costs were too modest than predicted in the late 1990s. The EU countries bore a greater burden than their OECD counterparts in adhering to the initial Kyoto commitments. Building on these developments and findings, subsequent theoretical advancements and extensions by Dai et al. (2023c,d) further refine the CQR methodology. These contributions include an extension of properties to shape-constrained nonparametric functions and the introduction of a penalized CQR method to address quantile crossing.

Recent empirical studies demonstrated the versatility of the CQR approach. Dai et al. (2020) evaluated emissions reduction targets of Chinese provinces, revealing substantial cost variations and potential savings resulting from diverse abatement options and the adoption of more efficient technologies. Zhao and Qiao (2022) examined US coal-fired power plants, estimated shadow prices, and emphasized the regulatory impacts on the market prices of pollutants. Wen et al. (2022) assessed soil erosion abatement costs in Shaanxi Province, China, emphasizing the potential for cost-efficient solutions and highlighting the need for effective strategies that consider external variables and temporal-spatial distribution. Quinn et al. (2023) analyzed 125 countries during the Kyoto Protocol period and found that countries with set CO_2 emission targets experienced a higher MAC than prevailing emission pricing mechanisms. This highlights the importance of shadow price estimates in an emission trading system (ETS) regulation and the consequences of policy decisions. Dai et al. (2023a) utilized CQR-based quantile allocation models to evaluate resource allocation efficiency in centralized decision-making systems, revealing significant gains in Finland's business sector. Finally, Dai et al. (2023b) addressed the issue of secular stagnation in productivity growth by exploring the impact of a low-carbon transition on OECD countries. Their findings, based on a quantile shadow-price Fisher index using a penalized CQR approach, showed that accounting for GHG emissions significantly increases measured productivity growth, particularly in countries that actively reduce their emissions.

In addition to CQR, closely related estimation approaches in this field include convex nonparametric least-squares (CNLS; Kuosmanen, 2008) and stochastic nonparametric envelopment of data (StoNED; Kuosmanen and Kortelainen, 2012). Mekaroonreung and Johnson (2012) were the first to apply CNLS and StoNED to estimate shadow prices of SO_2 and NO_x emissions of US coal power plants. They find that applying the weak disposability StoNED method provides consistent estimates of the emission market prices. Xian et al. (2022) utilized the StoNED method to estimate the least MAC of CO_2 for Chinese iron and steel enterprises. Their findings show that increasing labor is the most cost-effective abatement measure for most enterprises, proposing policy implications for reducing carbon abatement costs in the industry. Recently, Rødseth (2023) applied CNLS to estimate CO_2 shadow prices, highlighting the importance of incorporating the material balance principle into shadow price estimation. In the present context, the key difference between the CQR and CNLS/StoNED approaches is that the latter approach evaluates MACs by projecting all observations to a single production frontier that represents the average practice (CNLS) or the best practice (StoNED), in the CQR approach one estimates multiple quantile frontiers to evaluate MAC locally at the current level of efficiency.

This study offers two contributions to the growing body of literature. First, we make use of unique register-based firm-level GHG emission data merged with financial statement data to empirically assess the least-cost decarbonization pathways in the Finnish electricity generation industry. Our empirical analysis not only considers the historical development of abatement costs and the least-cost abatement strategies at present, but also presents forward-looking projections to assess the economic cost of achieving Finland's carbon neutrality targets by 2035. Second, we propose a simple practical remedy for the potential sample selection bias due to zero-emission firms. Zero-emission firms refer to firms that do not report any GHG emissions. This subset of firms includes a growing number of renewable energy producers that do not emit any CO_2; however, there are also conventional firms that fail to report their emissions for various reasons. For example, the EU ETS regulation requires that all power plants with a net heat excess of 20 MW report their GHG emissions; however, the regulation does not concern smaller plants. Building on the insights of Kuosmanen et al. (2023), we adjust the MAC estimates to account for zero-emission firms by applying a two-stage method known as the Heckman correction (Heckman, 1979).

The remainder of this paper is organized as follows. Section 2 provides an overview of the Finnish electricity generation industry. Section 3 outlines the methodological framework used to estimate MAC and assesses the economic costs of the decarbonization pathways. Section 4 presents the data used in this study. Section 5 presents the empirical analysis findings. Finally, conclusions are presented in Section 6.

2 Electricity generation industry in Finland

Finland's energy sector is a significant contributor to national GHG emissions (Statistics Finland, 2022). To align with the EU's targets for reducing GHG emissions, Finland has made progress in its energy transition process, resulting in structural changes, particularly in the electricity generation industry. Currently, nuclear energy dominates, accounting for over one-third of the total electricity generation, while bioenergy and hydroelectricity follow closely, each contributing approximately 19% to the power mix.[1] Increasing renewable energy is crucial for phasing out the use of fossil fuels. This shift is expected to double industrial electricity consumption and increase the nation's total electricity use by 50% by 2050 (Paloneva and Takamäki, 2021). Therefore, the success of energy transformation depends on ensuring the availability of affordable, reliable, and low-emission electricity, with a primary focus on reducing emissions from electricity production.

[1] Statista, Electricity generation in Finland: https://www.statista.com/statistics/1391223/finland-electricity-production-by-source/.

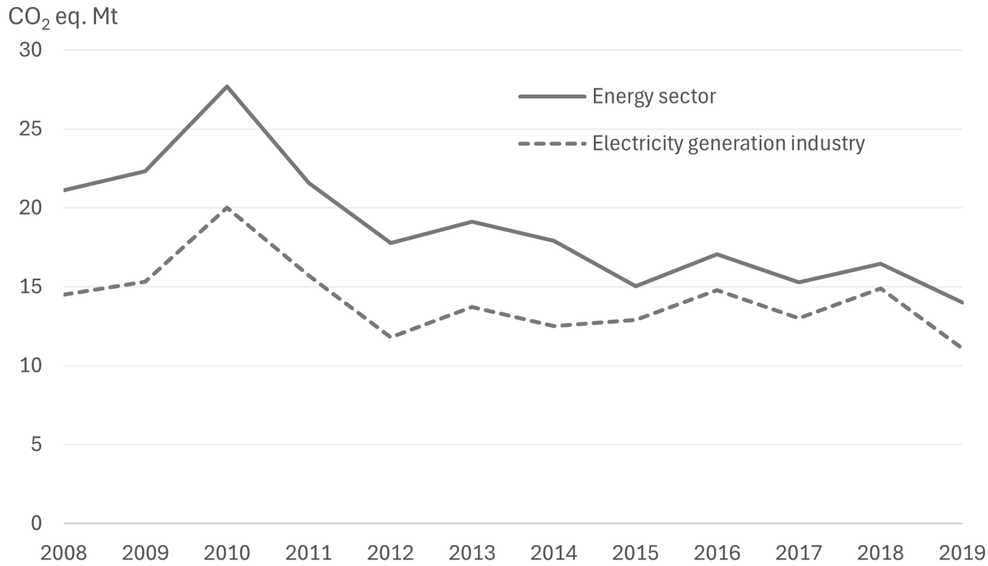

Fig. 1. GHG emissions of Finland's energy sector (NACE code 35 *Electricity, gas, steam, and air conditioning supply*; solid grey line) and electricity generation industry (TOL 2002 code 4011 *Production of electricity*; broken grey line) in 2008–2019, measured in Mt of CO_2 eq. Data sources: Eurostat air emissions accounts (solid line) and the national Greenhouse Gas Inventory of Statistics Finland (broken line).

Fig. 1 shows the GHG emissions in million tonnes of CO_2 equivalent (CO_2 eq.) for the entire energy sector (represented by the 2-digit NACE Rev. 2 code 35 *Electricity, gas, steam and air conditioning supply*) based on industry-level data from Eurostat, and specifically for the electricity generation industry (represented by the 4-digit TOL 2002[2] code 4011 *Production of electricity*) based on firm-level data from Statistics Finland for the period 2008–2019. The electricity generation industry is the largest emitter of GHG within the energy sector, accounting for approximately 70-80% of the total GHG emissions of the entire energy sector. Although emissions from the electricity generation industry (represented by the broken gray line in Fig. 1) decreased from 20 to approximately 12 million tonnes of CO_2 eq. between 2010 and 2012, the emission levels remained relatively stable thereafter. Given that this industry is expected to provide low-emission electricity in the near future to ensure low-emission energy, these trends provide a strong empirical motivation to investigate the MAC for electricity generation firms and identify the least-cost pathways for emission reduction.

[2]Statistics Finland, Standard Industrial Classification 2002: `https://www.stat.fi/en/luokitukset/toimiala/toimiala_1_20020101/?code=0201`.

Table 1. Number of electricity generation firms in Finland in 2000 and 2019 and the number of firms with reported GHG emissions.

Subdivision	2000		2019	
	All firms	With reported emissions	All firms	With reported emissions
Production of electricity with hydropower and wind power	34	-	202	2
Separate production of electricity with thermal power	4	1	10	2
Combined heat and power production	36	18	58	30
Production of electricity with nuclear power	2	1	7	2
Combined heat and power production for industry	2	1	7	2
Other production of electricity	9	3	-	-
Total	101	28	299	47

Source: Greenhouse Gas Inventory and the Business Register Database of Statistics Finland.

The Finnish electricity generation industry (the 4-digit TOL 2002 code 4011 *Production of electricity*) is further subdivided into five distinct 5-digit industries, as outlined in Table 1. This table provides an overview of electricity generation firms in Finland in 2000 and 2019, including those with reported emissions. The data were sourced from both the Greenhouse Gas Inventory[3] and Business Register[4] of Statistics Finland. The latter covers all Finnish enterprises, including those with no reported emissions. This table reveals a substantial transformation in the number of electricity generation firms in Finland between 2000 and 2019, reflecting the diversity of the industry. Noteworthy trends include an increase in the number of firms specializing in hydropower and wind power production coupled with an increase in combined heat and power generation. In 2000, the electricity generation industry comprised 101 firms, 28 of which reported their emissions. By 2019, the industry had experienced substantial growth, reaching 299 firms, of which 47 had reported emissions. Specific categories, such as hydropower and wind power production, thermal power, and

[3]Greenhouse Gas Inventory: `https://www.stat.fi/tup/khkinv/index_en.html`.
[4]Financial Statement Data Panel: `https://taika.stat.fi/en/aineistokuvaus.html`.

combined heat and power have shown variations in the number of firms over the past two decades.

However, the data in Table 1 also reveal a notable aspect: emissions are not reported for a substantial number of firms. This may be attributed to the fact that zero-emission firms refrain from emitting. Examples of zero-emission firms include those that utilize renewable energy sources, such as solar or wind power. Additionally, reliance on statistical authorities to estimate GHG emissions using energy consumption data may contribute to non-disclosure, as these data may be incomplete or missing for certain businesses.

3 Methodology

3.1 Convex quantile regression

In this study, we employ convex quantile regression (CQR) to estimate the marginal cost of abatement of GHG emissions. CQR is a data-driven method introduced by Kuosmanen and Zhou (2021) that builds upon previous studies by Wang et al. (2014) and Kuosmanen et al. (2015). Compared to previous approaches using convex regression and stochastic nonparametric envelopment of data (StoNED), where MAC estimation requires additional parametric distributional assumptions to identify a single frontier, the main advantage of CQR is that it employs multiple quantiles without making any parametric distributional assumptions. This approach is fully nonparametric and adheres to standard economic theory axioms such as monotonicity and convexity, without depending on arbitrary functional assumptions.[5] CQR addresses biases in the estimation of MAC by considering factors that are frequently overlooked or inadequately addressed in traditional shadow pricing methods, such as inefficiency, the direction vector, random noise in the data, and heteroscedasticity, which is a common issue in which the variability of a variable varies across its predicted range. CQR provides valuable information on the prices of abatement measures aligned with local efficiency levels.

Further, CQR incorporates a broader range of abatement options, including input-side options, such as fuel switching and clean technology investments, rather than simply assuming downscaling of production as the sole option, which has rarely been considered in

[5] The number of quantiles can be specified based on the sample size and desired precision. However, it is recommended to use an equidistant grid of 10 quantiles for most applications (Kuosmanen and Zhou, 2021).

previous studies (e.g., Lee, 2005). The present study recognizes the importance of considering capital investments and increased input utilization as potential strategies for emission reduction in addition to output scale reductions.

3.2 Estimation

Consider a generic semi-nonparametric production model

$$y_{it} = f(K_{it}, L_{it}, M_{it}, E_{it}) + \delta' z_{it} + \varepsilon_{it}, \tag{1}$$

where y_{it} and E_{it} represent the economic output (revenue) and bad output (GHG emissions) of firm i in period t, respectively; K, L, and M refer to capital, labor, and intermediate inputs, respectively; f is a nonparametric production function assumed to be monotonically increasing, concave, satisfying constant returns to scale (CRS); z_{it} is the contextual variable (to be discussed in more detail in the next sub-section); and ε_{it} is a composite error term that encompasses potential inefficiency and random noise.

Conditional quantile production function Q_y is defined as follows:

$$Q_y\left[\tau \mid (K, L, M, E)\right] = f(K, L, M, E) + \delta' z_{it} + \left(F_\varepsilon^{-1}(\tau)\right), \tag{2}$$

where τ $(0 \leq \tau \leq 1)$ indicates the order of the quantile, and F_ε is the cumulative distribution function of the composite error term ε. For a given quantile τ, the CQR estimator of Q_y is obtained by solving the quadratic programming (QP) problem for quantile τ:[6]

$$\min_{(\beta, \varepsilon^-, \varepsilon^+)} (1-\tau) \sum_{t=1}^{T} \sum_{i=1}^{n} (\varepsilon_{it}^-)^2 + \tau \sum_{t=1}^{T} \sum_{i=1}^{n} (\varepsilon_{it}^+)^2, \tag{3}$$

subject to

$$\begin{aligned}
& y_{it} = \beta_{it}^K K_{it} + \beta_{it}^L L_{it} + \beta_{it}^M M_{it} + \beta_{it}^E E_{it} + \delta' z_{it} - \varepsilon_{it}^- + \varepsilon_{it}^+, \quad \forall i, \forall t \\
& \beta_{it}^K K_{it} + \beta_{it}^L L_{it} + \beta_{it}^M M_{it} + \beta_{it}^E E_{it} \leq \beta_{js}^K K_{it} + \beta_{js}^L L_{it} + \beta_{js}^M M_{it} + \beta_{js}^E E_{it}, \quad \forall i, \forall t \\
& \beta_{it}^K \geq 0, \quad \beta_{it}^L \geq 0, \quad \beta_{it}^M \geq 0, \quad \forall i, \forall t \\
& \varepsilon_{it}^- \geq 0, \quad \varepsilon_{it}^+ \geq 0, \quad \forall i, \forall t.
\end{aligned}$$

[6]In the empirical analysis, the open-source Python package `pyStoNED` with the Mosek solver was utilized. `pyStoNED` can be accessed at `https://github.com/ds2010/StoNED-Python` and `https://pypi.org/project/pystoned/`.

Our main interest is in the coefficients β_{it}^K, β_{it}^L, β_{it}^M, β_{it}^E, which are the estimated subgradients of the quantile production function $Q_y(\tau|K, L, M, E)$. The non-negative variables, ε_{it}^- and ε_{it}^+, represent the negative and positive deviations, respectively, from the quantile frontier. The asymmetric loss function ensures that $100 \cdot \tau\%$ of the observations fall below that performance level τ. While ε_{it}^- and ε_{it}^+ encompass inefficiency (u) and noise (v) captured by the error term ε, our study does not explicitly identify or isolate these sources of deviation.

Following Kuosmanen and Zhou (2021), we solve Problem (3) ten times, varying parameter $\tau = \{0.05, 0.15, \ldots, 0.95\}$. Thus, we obtain ten sets of subgradient estimates $\{\beta_{it}^K, \beta_{it}^L, \beta_{it}^M, \beta_{it}^E\}$ for each firm i in year t. To obtain the unique shadow prices for each observation, we take the weighted average of the coefficients for the two quantiles closest to the observed data point. However, for observations that fall below the quantile $\tau = 0.05$ or above the quantile $\tau = 0.95$, we utilize the shadow prices associated with the nearest quantile.

Conventionally, the shadow price β_{it}^E is directly interpreted as the MAC of emissions. However, this interpretation implicitly assumes that downscaling production is the only way to reduce emissions. Alternatively, a firm could invest in cleaner technology, which usually requires additional capital investment and labor resources, or switch to cleaner fuels, which would increase intermediate inputs. To account for a broader set of abatement strategies, Kuosmanen and Zhou (2021) defined MAC as the least-cost abatement alternative:

$$\mathrm{MAC}_{it} = \min\left\{ r_{it}\frac{\beta_{it}^E}{\beta_{it}^K}, w_{it}\frac{\beta_{it}^E}{\beta_{it}^L}, \frac{\beta_{it}^E}{\beta_{it}^M}, \beta_{it}^E \right\}, \tag{4}$$

where r and w refer to the capital rents and wage rate, respectively.[7] Note that $\mathrm{MAC}_{it} \leq \beta_{it}^E$ by construction, taking a broader set of abatement strategies into account, will always yield a lower MAC estimate.

3.3 Heckman correction of zero-valued observations

The estimation of MAC in Equation (4) critically relies on the shadow price of emissions β_{it}^E. Unfortunately, the shadow prices are unidentified for zero-emission firms for which emissions

[7] In the empirical part of this study, the capital rents are estimated by the ratio of the operating profit and capital stock, and the wage rate by the ratio of the total payroll costs and the number of employees (full time equivalent). By construction, the prices of value added and intermediate inputs are equal to one.

$E_{it} = 0$. As stressed in the introduction, the subset of zero-emission firms includes renewable producers that do not emit any GHG emissions, as well as small conventional producers that emit GHG emissions but are not required to report their emissions according to the EU directive.

Since the subset of zero-emission firms is relatively large (see Section 2) and subject to endogenous selection (e.g., the use of renewable resources), simply excluding zero-emission firms from the estimation would likely cause sample selection bias. To mitigate this bias, we employ the two-step procedure introduced by Heckman (1976, 1979) to model selection in microeconometrics, such as in the context of wage equations or consumer expenditure. Recently, Kuosmanen et al. (2023) applied the Heckman correction in the nonparametric setting of convex expectile regression when the output variable y had zero-valued observations. In this study, we apply a similar approach, in which the most critical variable is E_{it}, which has a large share of zero-valued observations.

In the first step, we define a binary variable, Y, indicating whether the emission of firm i in period t, E_{it}, is greater than zero or not, as $Y_{it} = \{1 \text{ if } E_{it} > 0, \text{ and } 0 \text{ otherwise}\}$. We then use standard probit regression to estimate the likelihood of a firm having positive emissions based on the predictor variables $\boldsymbol{x}$:

$$Y_{it} = \Phi(\mathbf{x}'_{it}\gamma) + \varepsilon_{it}. \tag{5}$$

In Equation (5), Φ denotes the cumulative distribution function of the standard normal distribution, $N(0,1)$, and variables $\boldsymbol{x}$ include predictors, which consist of variables such as the number of employees, firm value added, firm age, and dummy variables for sub-industry and year. Given the parameter estimates $\hat{\gamma}$, the inverse Mills ratios are calculated as:

$$\mathrm{IM}_{it} = \frac{\phi(\mathbf{x}'_{it}\hat{\gamma})}{\Phi(\mathbf{x}'_{it}\hat{\gamma})}, \tag{6}$$

where ϕ and Φ are the density function and the cumulative distribution function of the standard normal distribution $N(0,1)$, respectively.

In the second step, we apply the CQR estimator (3) to the subsample of firms with positive emissions $E_{it} > 0$, taking the inverse Mills ratios (IM_{it}) as a contextual variable z. While the shadow prices of zero-emission firms remain unidentified, the inverse Mills ratio alleviates sample selection bias caused by the exclusion of zero-emission firms. Note that renewable energy producers with zero emissions cannot decrease their own emissions; thus, their MAC becomes infinite. While the market share of renewable producers needs to increase to achieve the policy targets, actual abatement must take place in those firms that currently generate GHG emissions.

4 Data and variables

To evaluate the MAC for reducing GHG emissions, we utilize two data sources from Statistics Finland, the national statistical authority. The first source is register-based firm-level data on GHG emissions from the National Greenhouse Gas Inventory.[8] These yearly panel data cover the period from 2000 to 2019 and include all electricity-generating firms that participated in the EU ETS, the first large-scale GHG emissions trading scheme in the world. According to EU directives, all power plants with a net heat excess of 20 MW must participate in the EU ETS. The firm-level data of the national greenhouse gas inventory used in this study are based on plant-level monitoring information submitted to the Finnish Energy Authority, which covers emissions at both the establishment and firm levels, reported as CO_2 and GHG emissions in CO_2 eq. This dataset serves as a foundation for climate policy planning and monitoring and is managed by Statistics Finland under the United Nations Framework Convention on Climate Change (UNFCCC), EU regulations, and the Kyoto Protocol. This dataset originates from official registers maintained by Statistics Finland, ensuring the reliability and precision of our analysis.

The second source is Financial Statement panel data, which provide information on all independent businesses across various industries in Finland. These panel data encompass essential firm-level details from income statements and balance sheets, including industry classification, employee count, value added, and financial metrics such as sales and fixed assets. For enterprises with at least 20 employees, data are collected directly, while information for smaller businesses is sourced from administrative records such as business taxation registers.

[8] Further information on the Greenhouse Gas Inventory is available at: https://www.tilastokeskus.fi/tup/khkinv/index_en.html.

Table 2. Descriptive statistics of the key variables.

	Revenue, M€	Emissions, 10^3t of CO_2 eq.	Labor, full-time eq.	Capital, M€	Intermediate inputs, M€
All firms					
Mean	45.00	251.17	60.00	166.01	34.78
Median	25.43	90.22	11.78	31.74	19.55
Std. Dev.	70.74	436.97	155.03	727.29	53.86
Hydropower and wind power					
Mean	9.01	0.78	21.48	20.79	6.25
Median	11.77	0.18	27.00	25.51	8.25
Std. Dev.	5.19	1.13	10.89	11.56	3.83
Separate production of electricity with thermal power					
Mean	45.27	688.39	27.63	62.92	42.07
Median	36.77	348.21	12.36	46.11	28.79
Std. Dev.	40.48	871.60	45.09	51.13	38.01
Combined heat and power production					
Mean	42.22	256.68	51.34	64.56	32.70
Median	25.61	82.91	11.06	28.90	18.87
Std. Dev.	56.10	413.36	118.11	83.54	44.66
Production of electricity with nuclear power					
Mean	335.97	193.16	779.85	4197.78	235.60
Median	296.71	0.48	814.30	4843.38	214.47
Std. Dev.	137.21	861.14	123.76	1816.86	121.80
Combined heat and power production for industry					
Mean	25.43	187.01	18.43	38.12	20.87
Median	24.89	145.16	3.00	29.21	20.66
Std. Dev.	14.49	253.34	31.16	29.05	11.38
Other production of electricity					
Mean	8.72	16.88	1.25	42.20	4.21
Median	7.61	8.73	0.30	25.58	4.97
Std. Dev.	5.28	16.81	2.13	30.69	2.01

By merging these two datasets using firm identification codes, we obtain a unique dataset that combines firm-level emission records with business register data, allowing us to investigate the cost of GHG abatement and alternative pathways of emission reduction. The merger produced a sample of 3,628 observations (523 firms) for the period 2000–2019. After excluding firms with unreported GHG emissions, the sample consisted of 798 firm-year observations (85 firms).

To estimate MAC, we use the following variables: revenue (desirable output), GHG emissions (undesirable output), labor (measured in full-time equivalent units), capital (represented by fixed assets), and intermediate inputs (derived as the difference between revenue and value added). Table 2 presents descriptive statistics for these variables across the six sub-industries according to the Finnish TOL 2002 classification. This table reveals interesting patterns across the sub-industries. For example, firms specializing in hydropower and wind power may also generate electricity from fossil fuels, but, on average, they have lower emissions and revenues than firms in other sub-industries. By contrast, firms that mainly focus on nuclear power but also have fossil fuel-powered plants are characterized by significant capital intensity, along with higher labor and intermediate inputs. Table 2 shows that many firms classified as renewable or nuclear electricity producers also have conventional plants that use fossil fuels to generate GHG emissions.

5 Results

5.1 Probit regression

We first employ a standard probit regression to predict the probability of firms having GHG emissions greater than zero over the period 2000–2019 for a sample of 523 firms. The binary outcome variable takes the value of one if a firm's emissions are greater than zero and zero otherwise. The model includes essential predictor variables, including the number of employees (measured in full-time equivalents), firm value added, firm age, and dummy variables representing sub-industry and year effects, as control variables. Table 3 reports the estimated coefficients for each predictor, along with their robust standard errors, calculated using the Stata software.

Regression analysis reveals that, on average, the probability of firms having positive GHG emissions is influenced by several key factors. The coefficient of firm size, measured by the number of employees, suggests that an increase in employees is generally associated with a lower probability of positive emissions. This relationship may be strongly influenced by nuclear power plants, as they contribute significantly to the observed patterns. By contrast, firm age has a positive impact, indicating that older firms tend to have a higher probability of positive emissions. The coefficient for firm value added is positive but has a minimal impact on the probability of positive emissions.

Table 3. Probit estimates.

Variable	*Coefficient*	*Robust st. error*
Intercept	-3.333***	0.237
Employees	-0.003***	0.001
Value added	0.000***	0.000
Firm age	0.014***	0.003
Control variables for sub-industry and year	Yes	
Log likelihood	-933.636	

Note: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

The results presented in Table 3 were used to calculate the inverse Mills ratio for a subset of 798 observations from 85 electricity generation firms for which emissions were greater than zero. By incorporating the inverse Mills ratio as an explanatory variable in the subsequent analysis, this study aims to address the truncation bias that arises from excluding observations with zero emissions.

5.2 Decarbonization of the Finnish electricity generation

After excluding zero-valued observations and considering the inverse Mills ratio as a contextual variable, we estimate the CQR for a subset of 798 observations from 85 electricity generation firms that have GHG emissions greater than zero. This analysis aims to determine MACs for GHG emissions and cost-effective abatement alternatives for Finnish electricity generation firms. We evaluate several options, including (i) downscaling production, (ii) investing in capital for carbon reduction or cleaner production technologies, (iii) increasing labor input by hiring additional technicians, and (iv) expanding intermediate inputs by increasing the use of renewable energy. The most cost-effective MAC for GHG emissions is determined by identifying the least-cost option.

Table 4 provides an overview of the MAC estimates for emission abatement in the industry, presenting the median, mean, and standard deviation values for each efficiency tier. This information offers insights into variations in abatement costs across different firm segments. Notably, higher-efficiency firms in the upper quantiles generally exhibit higher MAC values, which aligns with the economic principle that as firms optimize their processes and attain higher efficiency levels, the cost of additional emissions reductions increases.

The median MACs range from 0.05 to 3.46 euros per tonne of CO_2 eq., while the average values span from 9.63 to 3,435 euros per tonne of CO_2 eq. The large disparity between the median and mean arises from a few firms with exceptionally high MAC values, resulting in positive skewness in the distribution. This skewness may be influenced by factors such as technological constraints or industry-specific conditions of these firms. Consequently, relying on the average MAC may present a misleading picture of abatement costs, emphasizing the use of median values. Notably, most firms can abate their emissions at a very low cost.

Table 4. Marginal abatement cost (MAC) for Finnish electricity generation (2000–2019), €/t of CO_2 eq.

Efficiency tier, %	Median	Mean	Std. Dev.
0-5	0.05	18.23	80.62
5-15	0.41	19.55	94.09
15-25	0.28	430.89	3613.60
25-35	0.60	9.63	21.28
35-45	0.96	20.93	106.07
45-55	0.26	15.86	52.15
55-65	2.09	13.07	34.99
65-75	0.69	34.13	152.66
75-85	2.33	3435.79	28213.00
85-95	1.27	35.51	106.26
95-100	3.46	52.79	101.06

Table 5 presents the distributions of the least-cost abatement options across different efficiency tiers. Specifically, the figures in each row represent the share of firms within an efficiency tier with a specific least-cost strategy to reduce emissions. For instance, within the 0-5% efficiency tier, 4.9% of firms are recommended to decrease production as the most cost-effective approach, while 32.8% should make capital investments, 59.0% should increase labor, and 3.3% should increase intermediate inputs as the most cost-effective options.

Table 5. Distribution of least-cost abatement options for Finnish electricity generation firms.

Efficiency tier, %	Downscale output (y), %	Capital investment, %	Increase labor, %	Increase intermediate inputs, %
0-5	4.92	32.79	59.02	3.28
5-15	3.53	29.41	62.35	4.71
15-25	1.27	22.78	73.44	2.53
25-35	5.48	16.44	73.97	4.11
35-45	0.00	10.39	85.71	3.90
45-55	2.63	7.89	88.16	1.32
55-65	2.56	5.13	85.90	6.41
65-75	3.61	3.61	80.72	12.05
75-85	2.94	8.82	79.41	8.82
85-95	1.47	2.94	75.00	20.59
95-100	6.67	6.67	66.67	20.00

Downscaling production is the least-cost alternative for a limited subset of firms, suggesting that reducing the scale of production is not the most economical strategy for all efficiency levels. For less efficient firms in the lower tiers (0-5% to 35-45%), the least-cost strategies for reducing emissions involve more focus on capital investment and an increase in labor input. This highlights that at lower efficiency levels, investing in technology and human resources is more cost-effective for emission reduction. By contrast, more efficient firms in higher quantiles (45-55% to 95-100%) identify increasing labor and intermediate inputs as the most cost-effective approaches for abatement. These findings suggest that as firms become more efficient, optimizing labor and utilizing intermediate inputs become key strategies for achieving cost-effective emission reduction.

5.3 Forward-looking assessment of the GHG abatement cost

This part of our study is based on the work of (Dai et al., 2020), who advanced the CQR to a forward-looking assessment based on MAC estimates. Utilizing MAC estimates and aligning them with Finland's carbon emissions reduction targets, this section focuses on a forward-looking assessment of abatement costs within the country's electricity generation industry for 2021–2035. Specifically, we examine the projected economic cost for a 90% reduction and the cost of attaining zero emissions. To achieve this, we first fit an exponential trend line using nonlinear regression applied to the observed GHG emissions and the estimated MAC values spanning the period 2000–2019. The equation is as follows:

$$\mathrm{MAC}_t = A \cdot e^{b \cdot E_t} \tag{7}$$

where MAC_t is the average abatement cost in year t, A is a constant, b is the slope, and E_t is total emissions in year t. Table 6 presents the regression coefficients.

Table 6. Nonlinear regression results (2000–2019).

	Least cost	Downscale output	Capital investment	Increase labor	Increase intermediate inputs
Constant	33.22	22.33	42.22	105.18	33.77
Slope	-0.38	-0.21	-0.09	-0.45	-0.23

Note: Authors' calculations are based on Statistics Finland's data. All the coefficients are statistically significant at the 1% significance level.

Using the predicted trend, we next extrapolate the MAC of future GHG abatement and estimate the associated abatement costs. To provide insights into the potential economic costs associated with different strategies for reducing GHG emissions, Table 7 provides estimates of the abatement costs for two distinct scenarios: achieving a 90% reduction in current GHG emissions and achieving complete decarbonization (reducing emissions to zero). The abatement cost options considered in the analysis include the least-cost option, downsizing production (output reduction), investing in capital, increasing labor input, and increasing intermediate inputs.

Table 7. Estimated economic cost of abatement (M€).

	Least cost	Downscale output	Capital investment	Increase labor	Increase intermediate inputs
90% reduction	62.16	69.22	193.35	162.01	97.64
To zero	82.61	83.77	221.97	225.21	119.49

For the scenario aimed at a 90% reduction in emissions, the projected least cost is estimated to be 62.16 million euros (M€). The other abatement cost options for this scenario include 69.22 M€ for downsizing output, 193.35 M€ for capital investment, 162.01 M€ for increasing labor input, and 97.64 M€ for increasing intermediate inputs. In the case of complete decarbonization, the estimated least cost rises to 82.61 M€. The corresponding costs for the other abatement options are 83.77 M€ for downsizing output, 221.97 M€ for capital investment, 225.21 M€ for increasing labor input, and 119.49 M€ for increasing intermediate inputs.

To further explore abatement alternatives, we extend our analysis to predict the marginal cost of abating GHG emissions using either input or output. This prediction is illustrated in Fig. 2, which offers a visual representation of the evolving dynamics of abatement costs. Specifically, the figure illustrates the MAC of GHG emissions using the same four strategies: downsizing output, investing in capital, increasing labor, and increasing intermediate inputs. The horizontal axis represents GHG emissions, which decrease to zero over time, and the vertical axis represents the MAC.

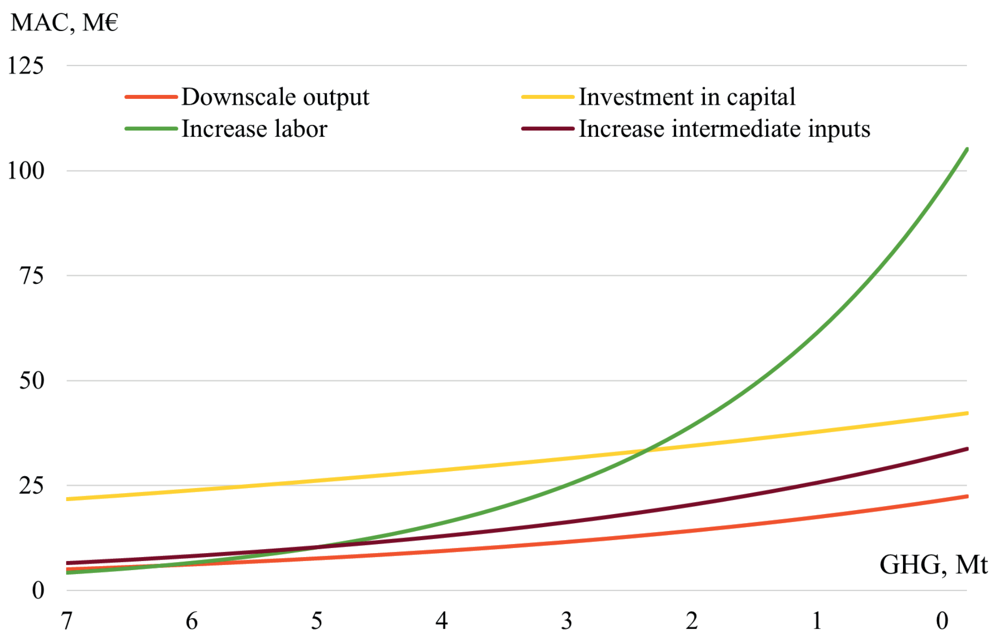

Fig. 2. MAC of abating GHG emissions using four options: downsizing output, investing in capital, increasing labor, and increasing intermediate inputs.

As discussed above, currently increasing labor input is the least-cost alternative for most firms. However, as GHG emissions continue to decrease (moving right on the x-axis), the MAC of increasing labor rises sharply. Downsizing production remains a cost-effective strategy until a certain point at which the MAC increases over time. Increasing intermediate inputs is another cost-effective option to lower GHG emissions; for instance, increasing the proportion of renewable energy consumption is the next least-cost alternative. Finally, the MAC of investing in capital has an initially higher value than the other options and is not the most cost-effective strategy for GHG abatement. However, as GHG emissions decrease, the MAC of investing in capital moderately increases and becomes a better alternative to increasing labor input. Finally, while Fig. 2 provides a useful tool for predicting the MAC of abating GHG emissions using either input or output, it is important to note that the graph is based on prediction and not actual data; therefore, it should be used as a guide rather than a definitive source of information.

6 Conclusions

Finland's electricity generation industry plays an essential role in achieving its ambitious carbon neutrality objectives and the need to provide low-emission electricity to other industries transitioning away from fossil fuels. To establish effective emission reduction targets and policies, it is necessary to determine the costs of reducing GHG emissions. To address this issue, this study examines the MACs of GHG emission reduction for the Finnish electricity generation industry using unique firm-level GHG emission data merged with register-based financial statement data. Because the starting dataset of electricity generation firms included a large number of firms for which GHG emissions were not available, we first used the Heckman correction to address the selection bias caused by excluding observations with zero emissions. Then, by employing convex quantile regression, we identify the least-cost options for each firm in our sample to reduce its emissions. Finally, we examine the least-cost decarbonization pathways in relation to Finland's carbon neutrality goal by 2035.

The empirical findings reveal that the median MACs of GHG emissions span from 0.1 to 3.5 euros per tonne of CO_2 eq., indicating substantial cost variability across different efficiency tiers. Notably, more efficient firms in higher quantiles exhibit higher MACs, consistent with the principle that achieving greater environmental efficiency often involves higher marginal costs for emission reductions. For less efficient firms in lower tiers, focusing on capital investment and increasing labor input has emerged as a least-cost strategy to reduce emissions. This emphasizes the potential benefits of technological upgrades and investments in human resources for emission reduction in the early stages of efficiency improvement. In contrast, more efficient firms in higher tiers find increasing labor and intermediate inputs to be the least-cost strategies, suggesting a shift towards optimizing labor and utilizing intermediate inputs for efficient emission reduction. The estimated costs of achieving a 90% reduction in carbon emissions and complete decarbonization were 62.1 and 82.6 million euros, respectively. These projections demonstrate that the feasibility of emission reduction strategies depends on the stringency of the set target. Although the least-cost option is financially attractive for a 90% reduction, the cost increases significantly for zero emissions. In the context of existing research, our study aligns with the growing body of literature on MAC assessments (Xian et al., 2022).

The application of least-cost abatement strategies has yielded new insights into developing effective environmental policies. By examining both efficient and inefficient firms, our study provides a better understanding of this subject matter. However, a significant drawback of our study is the inability to determine the exact reason for the absence of emissions from zero-emission firms. It is uncertain whether these firms truly do not emit emissions or if there are inconsistencies in the data.

Acknowledgments

The financial support provided by Business Finland as part of the larger project “Innovation-led Pathways to Economic Growth” is gratefully acknowledged.

References

Dai, S., Kuosmanen, N., Kuosmanen, T., and Liesiö, J. (2023a). Optimal resource allocation: Convex quantile regression approach. *arXiv preprint arXiv:2311.06590*.

Dai, S., Kuosmanen, T., and Zhou, X. (2023b). Can omitted carbon abatement explain productivity stagnation? quantile shadow-price fisher index applied to oecd countries. *Quantile Shadow-Price Fisher Index Applied to OECD Countries (June 23, 2023)*.

Dai, S., Kuosmanen, T., and Zhou, X. (2023c). Generalized quantile and expectile properties for shape constrained nonparametric estimation. *European Journal of Operational Research*, 310(2):914–927.

Dai, S., Kuosmanen, T., and Zhou, X. (2023d). Non-crossing convex quantile regression. *Economics Letters*, 233:111396.

Dai, S., Zhou, X., and Kuosmanen, T. (2020). Forward-looking assessment of the ghg abatement cost: Application to china. *Energy Economics*, 88:104758.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In *Annals of economic and social measurement, volume 5, number 4*, pages 475–492. NBER.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, pages 153–161.

Kuosmanen, T. (2008). Representation theorem for convex nonparametric least squares. *The Econometrics Journal*, 11(2):308–325.

Kuosmanen, T., Johnson, A., and Saastamoinen, A. (2015). Stochastic nonparametric approach to efficiency analysis: A unified framework. *Data Envelopment Analysis: A Handbook of Models and Methods*, pages 191–244.

Kuosmanen, T. and Kortelainen, M. (2012). Stochastic non-smooth envelopment of data: semi-parametric frontier estimation subject to shape constraints. *Journal of productivity analysis*, 38:11–28.

Kuosmanen, T., Tan, Y., and Dai, S. (2023). Performance analysis of english hospitals during the first and second waves of the coronavirus pandemic. *Health Care Management Science*, pages 1–14.

Kuosmanen, T. and Zhou, X. (2021). Shadow prices and marginal abatement costs: Convex quantile regression approach. *European Journal of Operational Research*, 289(2):666–675.

Kuosmanen, T., Zhou, X., and Dai, S. (2020). How much climate policy has cost for oecd countries? *World Development*, 125:104681.

Lee, M. (2005). The shadow price of substitutable sulfur in the us electric power plant: A distance function approach. *Journal of Environmental Management*, 77(2):104–110.

Mekaroonreung, M. and Johnson, A. L. (2012). Estimating the shadow prices of so2 and nox for us coal power plants: a convex nonparametric least squares approach. *Energy Economics*, 34(3):723–732.

Paloneva, M. and Takamäki, S. (2021). Summary of sector-specific low-carbon roadmaps. Publications of the Ministry of Economic Affairs and Employment 2021:9, Ministry of Economic Affairs and Employment of Finland.

Quinn, B., Gallagher, R., and Kuosmanen, T. (2023). Lurking in the shadows: The impact of co2 emissions target setting on carbon pricing in the kyoto agreement period. *Energy Economics*, 118:106338.

Rødseth, K. L. (2023). Shadow pricing of electricity generation using stochastic and deterministic materials balance models. *Applied Energy*, 341:121095.

Statistics Finland (2022). Finland's fifth biennial report under the unfccc. Technical report, Ministry of the Environment.

Wang, Y., Wang, S., Dang, C., and Ge, W. (2014). Nonparametric quantile frontier estimation under shape restriction. *European Journal of Operational Research*, 232(3):671–678.

Wen, X., Yao, S., and Sauer, J. (2022). Shadow prices and abatement cost of soil erosion in shaanxi province, china: Convex expectile regression approach. *Ecological Economics*, 201:107569.

Xian, Y., Yu, D., Wang, K., Yu, J., and Huang, Z. (2022). Capturing the least costly measure of co2 emission abatement: evidence from the iron and steel industry in china. *Energy Economics*, 106:105812.

Zhao, S. and Qiao, G. (2022). The shadow prices of co2, so2 and nox for us coal power industry 2010–2017: a convex quantile regression method. *Journal of Productivity Analysis*, 57(3):243–253.

Elinkeinoelämän tutkimuslaitos

ETLA Economic Research

ISSN-L 2323-2420
ISSN 2323-2420 (print)
ISSN 2323-2439 (pdf)

Publisher: Taloustieto Oy

Tel. +358-9-609 900
www.etla.fi
firstname.lastname@etla.fi

Arkadiankatu 23 B
FIN-00100 Helsinki