MINISTÈRE DE L'ÉCONOMIE, DES FINANCES ET DE LA RELANCE
Liberté
Égalité
Fraternité

Working Papers

No. 2021/3 • August 2021

Direction générale du Trésor

WHICH INDUSTRIAL FIRMS MAKE DECARBONIZATION INVESTMENTS?

Romain Faquet

WHICH INDUSTRIAL FIRMS MAKE DECARBONIZATION INVESTMENTS?

Romain Faquet

The views expressed in this working paper are those of the authors and do not necessarily represent positions of the French Treasury. Its publication aims at stimulating debate and generating comments and criticism.

Romain Faquet is currently working at the Treasury of the French Ministry for the Economy, Finance and Recovery

romain.faquet@dgtresor.gouv.fr (+33-1-44-87-17-78)

Contents

Abstract

This working paper seeks to understand the drivers of decarbonization investments by industrial firms in France. I investigate the extent to which firms' characteristics determine the probability of engaging in green investments and the size of these investments. By mobilizing a selection model on individual panel data estimated between 2013 and 2018, I show that the adoption of decarbonization technologies increases with firm size, energy intensity, productivity, and inclusion in the ETS, and decreases with firm age. Among firms investing in decarbonization, the size of the decarbonization investments is determined by firms' energy intensity. The analysis suggests that some highly-emitting sectors have tended to invest less in decarbonization than other sectors, providing emerging evidence that decarbonization investments should be accelerated as a priority in these sectors.

Keywords: environmental investments, decarbonization, firms, manufacturing, greenhouse gas emissions.

JEL Classification Numbers: D22, Q52, Q58.

This work benefited from access to the Secure Data Access Center (SDAC).

Résumé

Ce document de travail cherche à comprendre les déterminants des investissements de décarbonation des entreprises industrielles en France. J'étudie la mesure dans laquelle les caractéristiques des entreprises déterminent la probabilité d'engager des investissements verts et l'intensité de leurs montants. En mobilisant un modèle de sélection sur données individuelles de panel estimé entre 2013 et 2018, je montre que l'adoption de technologies décarbonées croît avec la taille de l'entreprise, son intensité énergétique, sa productivité, son appartenance au SEQE, et est d'autant plus élevée que l'entreprise est jeune. Parmi les entreprises industrielles qui décarbonent, les taux d'investissement en technologies décarbonées sont déterminés par l'intensité énergétique de ces entreprises. Enfin, l'analyse met en exergue que certains secteurs très émetteurs ont moins investi en technologies de décarbonation que d'autres secteurs, ce qui suggère que les investissements de décarbonation devraient être accélérés prioritairement dans ces secteurs.

Mots-clés : investissements verts, décarbonation, entreprises, industrie, émissions de gaz à effet de serre.

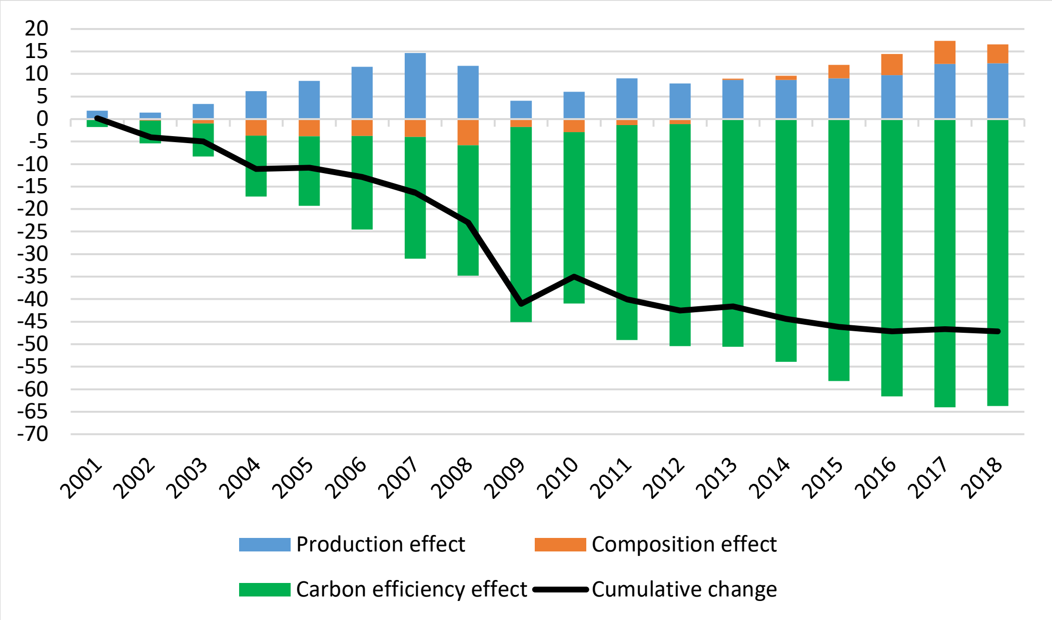
Classification JEL : D22, Q52, Q58

Introduction

French industry[1] directly[2] emitted 79.5 MtCO2e[3] of greenhouse gases (GHG) in 2018, representing 17.9% of national emissions.[4] The industrial sector is decarbonizing faster than the total economy: industrial emissions fell by 45% between 1990 and 2018 compared to 19% for all sectors combined. These emissions are linked to both energy combustion and industrial processes. Their main characteristics are that they are highly concentrated both at the sector level - three quarters of emissions are generated by three sectors: chemicals, metallurgy and minerals[5] - and at the firm level - three quarters of emissions come from the 1,200 industrial establishments covered by the EU Emissions Trading System (EU ETS).[6]

The decrease in GHG emissions from French industry was exclusively driven by the continuous improvement of carbon efficiency, i.e. by technical progress that allows for the improvement of manufacturing processes. A simple decomposition of the main forces driving the dynamics in industrial emissions shows that without a decrease in the carbon intensity of production, emissions from this sector would have increased significantly, particularly in relation to the positive growth rate of the real value added of the industrial sector and the evolution of its sectoral structure towards relatively more emitting sectors such as chemicals (see Figure 1).[7] It should be noted that this reasoning based on emissions from production on the national territory does not take into account the increasing penetration of industrial imports in household consumption.

Figure 1: Contributions to the cumulative decrease in industrial GHG emissions since 2000 (MtCO2e)



Source: DG Trésor calculations based on Citepa and Insee data.

These decarbonization investment strategies are incentivized and supported by numerous economic policy levers including the EU ETS, energy saving certificate mechanism to improve energy efficiency, the Heat Fund managed by the ADEME to support the production of renewable energy, and BPI France green loans. In response to the Covid-19 crisis, the recovery plan *France Relance* is devoting an additional €1.2 billion

[1] As defined by the National Low-Carbon Strategy, i.e. excluding the energy industry but including the building construction industry.
[2] Emissions related to the production of energy used subsequently by industrial firms are not counted.
[3] Carbon dioxide equivalent (CO2e) is a measure used to compare the emissions from various greenhouse gases based upon their global warming potential. For example, the global warming potential for methane over 100 years is 21. This means that emissions of one million tons of methane is equivalent to emissions of 21 million tons of carbon dioxide.
[4] Calculations based on Citepa data.
[5] Calculations based on Citepa data.
[6] Citepa (2020).
[7] For the method and data used, see Annex 1. The diagnosis is unchanged for manufacturing industry.

between 2020 and 2022 to help industrial firms invest in lower-emission equipment,[8] and €2.0 billion to support the development of green hydrogen, particularly in the industrial sector.

In the coming years, achieving the industrial sector's emission reduction targets will require seeking out more costly and less accessible decarbonization sources in low-carbon manufacturing processes and removing technological barriers in cement and steel in particular (e.g. reduction of iron ore by hydrogen, extraction of aluminium using electrolysis, electric cracking technologies). In this perspective, making a long-term carbon price increase trajectory credible as soon as possible will be a *sine qua non* condition for triggering these investments. Complementarily, subsidies to support green research, development and innovation (RDI) and adoption of these technologies may be required to counter the well-documented path dependence in firms' investments.[9]

The National Low-Carbon Strategy (NLCS) sets a reduction target for industry emissions of − 35% in 2030 (53 MtCO2e) and − 81% in 2050 (16 MtCO2e) compared to 2015. This necessary but ambitious trajectory assumes totally decarbonized energy, with residual emissions coming from industrial processes. The observed average rate of emissions reduction of −1.3% per year since 2013 is insufficient to achieve these objectives: −3.5% per year is needed between 2019 and 2030 to reach the mid-term objective. More generally, recent studies, which take into account not only the trend decline in emissions but also the emissions cut due to the French recovery plan, have assessed the possibility of achieving the NLCS's target of reducing emissions from the French economy by 40 % in 2030 compared to 1990. They show that the objectives will probably not be achieved without additional measures, particularly in industry.[10]

Supporting green investments requires a better understanding of their drivers. Paradoxically, while these investments are the main lever for decarbonizing the industrial sector, few econometric studies on recent data have sought to understand the characteristics of firms investing in decarbonization. The economic literature has indeed mainly focused on the causal impact of the EU ETS on GHG emissions. This literature, including on French data, has shown that the EU ETS has made it possible to reduce the emissions of energy-intensive industrial establishments without affecting employment or competitiveness.[11]

To our knowledge, two recent studies shed more light on the demographics of firms adopting green investments. Using Swedish data, Jaraité *et al.* (2014) provided evidence that firm's environmental expenditures and investments to mitigate air pollution increase with firm size, energy-intensity and EU ETS coverage. On French data, Dussaux (2020) found that the rise in energy prices drives industrial establishments to make air pollution abatement investments.

The present study contributes to the literature in three ways. First, it specifically targets decarbonization investments that is investments in technologies that aim to reduce GHG emissions, not air pollutants. Second, it considers a rich set of firm demographic determinants of such investments through a matching of several databases. Third, it seeks to explain both the adoption and intensity of decarbonization investments in manufacturing between 2013 and 2018 using a panel selection econometric model.

This paper is structured as follows: part 1 presents the data used, part 2 the econometric modelling and part 3 the results.

[8] It basically provides for support for energy efficiency and process adaptation on the one hand and support for low-carbon heat production on the other.
[9] See Aghion *et al.* (2016).
[10] BCG (2021); Rexecode (2021).
[11] For a comprehensive review of the literature, see Glachant and Mini (2020).

1. Data

1.1 Description of databases

To conduct this study, the data on decarbonization investments in the industrial sector from the ANTIPOL survey (see Box 1) are matched between 2013 and 2018 with three other databases:

- the exhaustive tax database FARE, which provides structural information on French firms, in particular their balance sheets and profit and loss accounts (but also other information such as employment);
- the EACEI survey on industrial energy consumption (see Box 2);
- the EU registry recording French industrial establishments covered by the EU ETS.

Box 1: The ANTIPOL survey

Industry's decarbonization investments are derived from the ANTIPOL survey on environmental protection expenditures. This survey, which has been conducted entirely by INSEE since 2012,[12] focuses on 11,000 industrial establishments with 20 or more employees in the industrial sector including energy (more precisely, divisions 5 to 35 of the NAF rev.2 industry classification are surveyed). The sampling method is a stratified design by size and sector of activity, with an exhaustive stratum for establishments with more than 250 employees. The change in the sampling frame in 2012 leads to a break in sampling.

This annual survey covers investments to treat, measure, control or limit the pollution generated by industrial activity.[13] Investments to protect the environment totaled €1.327 billion in 2019. This expenditure is broken down according to the areas concerned including wastewater, air protection, GHG limitation, waste, landscapes and biodiversity. We focus on decarbonization investments, i.e. investments explicitly dedicated to cut GHG emissions. In 2019, €311 million was aimed at reducing GHG emissions, the second most important destination after investments to reduce air pollution (€328 million).[14]

Another distinction is made between so-called specific investments that is equipment entirely dedicated to environmental protection (such as filters, carbon capture and storage), and so-called integrated investments, meaning the adoption of clean production technologies. The latter correspond to the additional cost generated by the choice of equipment with better environmental performance than the market standard. In 2018, decarbonization investments amounted €234 million and were divided between €171 million of specific investments and €63 million of integrated investments.[15]

The relatively modest aggregate amounts recorded can probably be explained by the fact that firms only record their most significant decarbonization actions and not the technical progress incorporated during the renewal of industrial equipment over time. Moreover, pure energy efficiency investments are not counted by definition when they have no effect on GHGs.

[12] A complementary survey was previously conducted by the statistical services of the Ministry of Agriculture to measure firms' environmental protection expenditures in the agri-food industries.
[13] Study expenses (environmental studies for investment, impact studies and regulatory audits) and current expenses (operating and maintenance expenses for anti-pollution investments, for example) are also surveyed.
[14] Other important expenses concern wastewater (212 M€), soil (163 M€), landscapes and biodiversity (145 M€).
[15] This detailed information is not yet available for 2019.

Box 2: The EACEI survey

This annual survey measures the quantities consumed and expenditure by type of energy in the industrial sector. It has been unified and conducted entirely by INSEE since 2013,[16] and is carried out among approximately 8,500 industrial establishments with more than 20 employees (excluding the energy industry, more precisely divisions 7 to 33 excluding 19 of NAF rev.2 are surveyed). The sampling method is based on a stratified design by size, sector of activity and region, with an exhaustive stratum for establishments with more than 250 employees and/or high energy consumption. The change in the sampling frame in 2013 leads to a break in sampling.

In 2018, industrial energy consumption amounted 37.4 million tonnes of oil equivalent (Mtoe)[17] for a bill of €14.5 billion. The three most energy-intensive sectors are chemicals (11.8 Mtoe), metallurgy (8.9 Mtoe) and agri-food industries (5.3 Mtoe).[18] The main energy sources consumed are gas (10.8 Mtoe), electricity (9.9 Mtoe), solid mineral fuels (6.0 Mtoe) and petroleum products (3.3 Mtoe).

1.2 Data matching and creation of the final dataset

The four databases are matched on the basis of the SIREN identifier of legal units (see Table 1, “before filtering”). Since the ANTIPOL and EACEI databases are surveyed at the establishment level, the data from establishments sharing the same identifier are aggregated. This raises the question of whether these aggregated investment and energy consumption data are representative of the legal unit they are matched to. First, all establishments of the same firm are not systematically surveyed in the same year. Second, a firm may have an industrial establishment while having a main activity in market services. That is why only groups of establishments whose aggregate employment from the EACEI and ANTIPOL databases each represents between 90% and 110% of the legal unit's employment from FARE are kept in the sample used for this study (see Table 1, “after filtering”). Negative value added observations are also removed (see Table 1, “final dataset”).

This results in an unbalanced panel dataset of 4,394 industrial firms observed in at least one year between 2013 and 2018, and 8,918 individual firm-year observations. Our dataset include both adopters (i.e. firms that make decarbonization investments in a given year) and non-adopters (i.e. firms that do not make these investments in a given year). When a firm is not observed in a given year, it cannot be known whether or not it adopted green technologies.

Table 1: Steps in the matching of the four databases

Firm-year observations	2013	2014	2015	2016	2017	2018
Before filtering	3546	3456	3531	4511	3620	3621
After filtering	1601	1528	1750	2154	1159	1034
Final dataset	8918					

Source: DG Trésor calculations.

[16] Previously, a complementary survey was conducted by the statistical services of the Ministry of Agriculture to measure the energy consumption of the agri-food industries.
[17] The tonne of oil equivalent (toe) represents the quantity of energy contained in one tonne of crude oil, or 41.87 gigajoules. This unit is used to express in a common unit the energy value of the various energy sources: 1 toe is, for example, equal to 1,616 kg of coal or 11.6 MWh of electricity.
[18] This is followed by rubber-plastic (4.6 Mtoe) and wood (4.0 Mtoe).

In our final dataset, the dependent variable is the decarbonization investment rate of each firm, constructed as the decarbonization investment expenditures divided by the value added (VA). This ratio is computed for specific investments, integrated investments and total investments.

The study considers the following set of explanatory variables, based on their potential theoretical influence on decarbonization investments.

- Firm size, measured by the number of employees. A positive impact is expected if ability to amortize fixed costs triggers investment.
- Firm performance, measured first by its labor productivity, that is the log of the ratio of real value added (deflated by sectoral value added prices from the annual national accounts) to the number of employees, and second by its profitability, that is the ratio between gross operating surplus and nominal value added. A positive impact is expected: profitability frees up room for investment, all else being equal.
- Firm age, with an ambivalent theoretical effect: a younger firm may have more difficulty in obtaining financing but may suffer less lock-in in older technologies.
- Carbonized energy use intensity, measured by the log of the ratio of carbonized energy consumption to real value added. Carbonized energy includes gas (natural gas and other gases), solid mineral fuels (coal, coke) and oil products (petroleum coke, butane/propane, heavy fuel oil and heating oil). A positive impact would be expected: energy intensity increases the need to abate more pollution all else being equal. This variable is lagged by one year to avoid potential simultaneity bias: decarbonization investments could indeed contemporaneously affect carbonized energy consumption in return.
- The average cost of carbonized energy, measured as the log of the ratio of carbonized energy expenditures to carbonized energy consumption. A firm that does not consume any carbonized energy has zero expenditure. A positive impact can be expected: since energy is a production cost, the higher the price the stronger the incentives to reduce its consumption and therefore its emissions. This variable is also lagged by one year to avoid a potential simultaneity bias.
- Inclusion in the EU ETS (phase 3). This is a dummy variable. A positive impact is expected: by putting a price on carbon, the EU ETS encourages firms to reduce pollution by investing in specific investments or in decarbonized production technologies.

1.3 Descriptive statistics

Table 2 presents the characteristics of industrial firms that make decarbonization investments in a given year, as opposed to those that do not. They appear to be much larger, more productive, more profitable and more energy-intensive, pay a lower average energy cost and are the same age. Finally, they are three times more likely to be covered by the EU ETS.

Firms that make specific investments in a given year are relatively larger, more productive, more profitable, and more energy-intensive than firms that invest in integrated technologies. They are also almost twice as likely to be subject to the EU ETS.

Table 2: Characteristics of firms making and not making decarbonization investments in a given year

	Non-adopters			Adopters, total investments			Total sample		
	P10	P50	P90	P10	P50	P90	P10	P50	P90
Size (employees)	35	128	432	80	283	1383	37	140	511
Labor productivity (K€)	36.9	58.9	106.0	42.2	69.7	127.0	37.4	60.1	109.2
Profit margin (%)	−8.1	21.0	47.5	−1.0	26.0	52.3	−7.3	21.7	48.2
Energy intensity (toe/M€)	0	17.1	177.0	1.6	30.3	366	0	18.2	196.5
Average energy cost (€/toe)	0	523.7	837.4	301.7	480.6	719.3	0	516.1	826.5
Age (years)	11	29	58	12	30	59	11	29	58
EU ETS membership (%)	5.6			15.3			6.9		
Firm-year observations									
All sectors	7770			1148			8918		
Chemical industry	743			156			899		
Equipment goods and transport equipment	1028			274			1302		
Food industry	1359			269			1628		
Metallurgy	1776			135			1911		
Non-metallic minerals	599			61			660		
Wood and paper	407			100			507		
Other industries	1858			153			2011		

Source: DG Trésor calculations.

Note: P10, P50, P90 represent respectively the first decile, the median, and the last decile. For example, the median size of firms in our total sample is 140 employees. Within this sample, firms that make green investments ("adopters") in a given year have a median size of 283 employees, compared to 128 for those that do not adopt these technologies ("non-adopters").

	Adopters, specific investments			Adopters, integrated investments		
	P10	P50	P90	P10	P50	P90
Size (employees)	113	352	1859	71	245	1081
Labor productivity (K€)	45.3	73.9	142.6	40.6	66.5	118.4
Profit margin (%)	0.0	27.0	52.9	−1.0	25.8	52.0
Energy intensity (toe/M€)	1.7	32.7	500.1	1.2	27.6	309.7
Average energy cost (€/toe)	281.7	465.4	669.5	32207	488.3	727.7
Age (years)	11	30	59	13	30	59
EU ETS membership (%)	19.5			12.2		
Firm-year observations						
All sectors	614			663		
Chemical industry	91			77		
Equipment goods and transport equipment	168			148		
Food industry	128			164		
Metallurgy	75			72		
Non-metallic minerals	32			33		
Wood and paper	59			63		
Other industries	61			106		

Source: DG Trésor calculations.

Note: P10, P50, P90 represent respectively the first decile, the median, and the last decile. For example, within the firms that decarbonize, those that make specific investments in a given year have a median size of 352 employees compared to 245 for those that make integrated investments.

Table 3 presents the frequency and distribution of decarbonization investments in our final dataset. A few observations can be made:

- A minority of industrial firms make investments to cut GHG emissions: only 12% of firms adopt decarbonizing technologies on average in a given year.
- Firms investing in integrated technologies in a given year generally do not adopt specific technologies in the same year: the total share of investing firms is close to the sum of the share of firms engaging in specific investments and in integrated ones.
- The distribution of these investments is highly skewed: the median is quantitatively modest at €33K while significant investments are concentrated in the last decile, 10 times higher than the median.
- The amounts of specific investments are higher than the integrated investments: however, as previously mentioned, these statistics are not directly comparable since integrated investments only represent the extra costs associated with clean technologies compared to market standards.

Table 3: Frequency and distribution of annual decarbonization investments

		Specific investments	Integrated investments	Total investments
Adoption (%)		6.9	7.4	12.9
Investment if adoption (K€)	Mean	169	83	138
	P10	4	2	3
	P50	43	20	33
	P90	395	200	330
	P95	730	356	601
	P99	2000 (2871)	1000 (1822)	1847 (2672)
Investment ratio if adoption (% VA)	Mean	0.8	0.6	0.8
	P10	0.01	0.01	0.01
	P50	0.1	0.1	0.2
	P90	1.1	1.1	1.2
	P95	2.2	2.5	2.5
	P99	7.3 (30.5)	9.9 (18.3)	9.9 (29.0)

Source: DG Trésor calculations.
Note: P10, P50, P90, P95 and P99 represent respectively the first decile, the median, the last decile, and the 95th and 99th percentiles of the distribution of investment variables. For the last percentile the average value of the variables is also given in brackets.

2. Econometric model

In view of the small number of industrial firms making decarbonization investments, a panel selection model is used. This method makes it possible to distinguish the factors affecting the probability of adoption (extensive margin) and those affecting the intensity of these decarbonization investments (intensive margin).

2.1 Rationale for specification of the model

The mass point of zero decarbonization investments highlighted in the previous descriptive statistics rules out the use of a linear model on the total sample. Fitting a linear equation on the sub-sample of adopters is also to be ruled out: there is no reason to assume that the adoption of decarbonization technologies is random. In other words, it is unlikely that the determinants of investment intensity are independent of the factors influencing the decision to invest, and in this case the ordinary least squares estimator would be biased.

This selection effect is solved by the joint modelling of the probability of investing and the intensity of these investments, for which Heckman (1976) derived a consistent two-step estimator for cross-sectional data. This method was extended by Wooldridge (1995) to panel data models that is including unobserved individual heterogeneity.

The model can be written as follows:

$$\begin{cases} D_{it} = 1 \; if \; D_{it}^{*} = Z_{it}\gamma + \eta_i + u_{it} > 0 \\ \\ I_{it} = X_{it}\beta + \alpha_i + \varepsilon_{it} \;\; if \; D_{it} = 1 \end{cases}$$

The first equation is a selection equation for decarbonization technologies. It describes the choice to invest or not based on a dummy variable D_{it}. D_{it}^{*} is the latent variable that models the trade-off underlying the decision to invest as a function of a vector of explanatory variables Z_{it}, an unobserved individual effect η_i potentially correlated with Z_{it}, and an idiosyncratic error term u_{it}. Z_{it} includes all the explanatory variables defined previously (see part 1.2.), plus a carbonized energy use dummy, time dummies and industry dummies.

The second equation linearly models the investment rate as a function of a vector of explanatory variables X_{it}, an unobserved individual effect α_i potentially correlated with X_{it} and an idiosyncratic error term ε_{it}. X_{it} includes the same variables as Z_{it} but the dummy for carbonized energy use, and that for two reasons. First, adopters are most likely to be carbonized energy users, which removes any variability between these two variables in the sub-sample of adopters. Second, depriving the second equation of a variable driving adoption improves parameters identification in selection models.

2.2 Estimation procedure

The Wooldridge estimator is based on the following two steps (see Annex 2 for a sketch of the proof).

First, a correlated random-effects probit estimator is used to fit a new version of the selection equation. In this equation, individual heterogeneity is no longer completely unobserved but divided into a part modelled explicitly in terms of the time averages of the explanatory variables (the so-called Mundlak hypothesis) and a purely random part.

$$D_{it} = 1 \; if \; D_{it}^{*} = Z_{it}\gamma + \overline{Z}_{\iota}\delta + v_{it} > 0 \; ; \; \overline{Z}_{\iota} = \frac{1}{T}\sum_{t=1}^{T} Z_{it} \; ; v_{it} \sim N(0, \sigma_v^2)$$

Second, a generalized least squares (GLS) estimator is used to fit a new version of the linear investment regression to the sub-sample of adopters. For $D_{it} = 1$,

$$I_{it} = X_{it}\beta + \overline{X}_{\iota}\psi + \omega\hat{\lambda}_{it} + e_{it} \; ; \; \overline{X}_{\iota} = \frac{1}{T}\sum_{t=1}^{T} X_{it} \; ; \hat{\lambda}_{it} = \frac{\varphi\left(\frac{Z_{it}\hat{\gamma} + \overline{Z}_{\iota}\hat{\delta}}{\hat{\sigma}_v}\right)}{\Phi\left(\frac{Z_{it}\hat{\gamma} + \overline{Z}_{\iota}\hat{\delta}}{\hat{\sigma}_v}\right)}$$

φ and Φ being respectively the Gaussian probability density function and cumulative distribution function. The $\hat{\lambda}_{it}$ are called inverse Mills ratios. Derived from the first stage probit, they correct for selection bias, that's that the adopters are not a random subset of industrial firms. This ratio is a monotonic decreasing function of the probability that a firm invests given its observed characteristics. For example, firms that have a low probability of investing given their observed characteristics have high inverse Mills ratio and may adopt green technologies due to unobserved factors. The coefficient ω of the inverse Mills ratio in the second stage regression assesses whether the unobserved factors predicting the probability of investing are correlated or not with unobserved factors affecting the intensity of these investments.

The GLS estimator of the model's coefficients is consistent but the standard errors are underestimated because they do not account for the uncertainty from first stage estimation of the inverse Mills ratios. That's why the standard errors are bootstrapped at the firm level. 100 new samples are created by randomly drawing firms from the original sample with replacement,[19] and then used to estimate the two-step selection model.

[19] All the observations of each firm drawn from the original sample are then taken into account. The bootstrapped standard errors are stable from 100 repetitions.

3. Results

3.1 Probability of adoption of decarbonization investments (extensive margin)

The first step concerns the estimation of the selection equation by a correlated random-effects probit estimator (see Table 4a). Individual heterogeneity (Mundlak-type) is statistically correlated with the explanatory variables meaning that a pure random-effects model is rejected by the data.[20] Three comments are in order.

Table 4a: Main results of the selection equation (first stage: extensive margin)

	Total investments		Specific investments		Integrated investments	
	Coeff.	*p-value*	Coeff.	*p-value*	Coeff.	*p-value*
Size	$3.4\ 10^{-4}$	*0.01*	$2.8\ 10^{-4}$	*0.00*	$2.1\ 10^{-4}$	*0.00*
Labor productivity	0.49	*0.00*	0.69	*0.00*	0.21	*0.06*
Profit margin	−0.26	*0.12*	−0.46	*0.02*	−0.02	*0.93*
Energy intensity	0.07	*0.00*	0.05	*0.09*	0.06	*0.02*
Average energy cost	−0.11	*0.03*	−0.04	*0.65*	−0.06	*0.30*
Energy use dummy	0.22	*0.59*	−0.20	*0.73*	−0.04	*0.93*
EU ETS membership	0.41	*0.00*	0.47	*0.00*	0.09	*0.50*
Age	−0.07	*0.01*	−0.06	*0.06*	−0.06	*0.00*
Individual heterogeneity (Mundlak-type)	yes	*0.00*	yes	*0.01*	yes	*0.00*
Sectoral dummies	yes	*0.00*	yes	*0.00*	yes	*0.00*
Time dummies	yes	*0.02*	yes	*0.06*	no	*0.24*
Rho	0.44		0.45		0.39	
Firm-year observations	8918					

Source: DG Trésor estimates.
Note: The p-values are associated with the test of the nullity of the coefficient for each explanatory variable, and the joint nullity of the coefficients for individual heterogeneity and the sectoral and time dummies. All these tests are based on robust standard errors clustered at the firm level. Rho is the proportion of the variance of the error term that is contributed by the random effect.

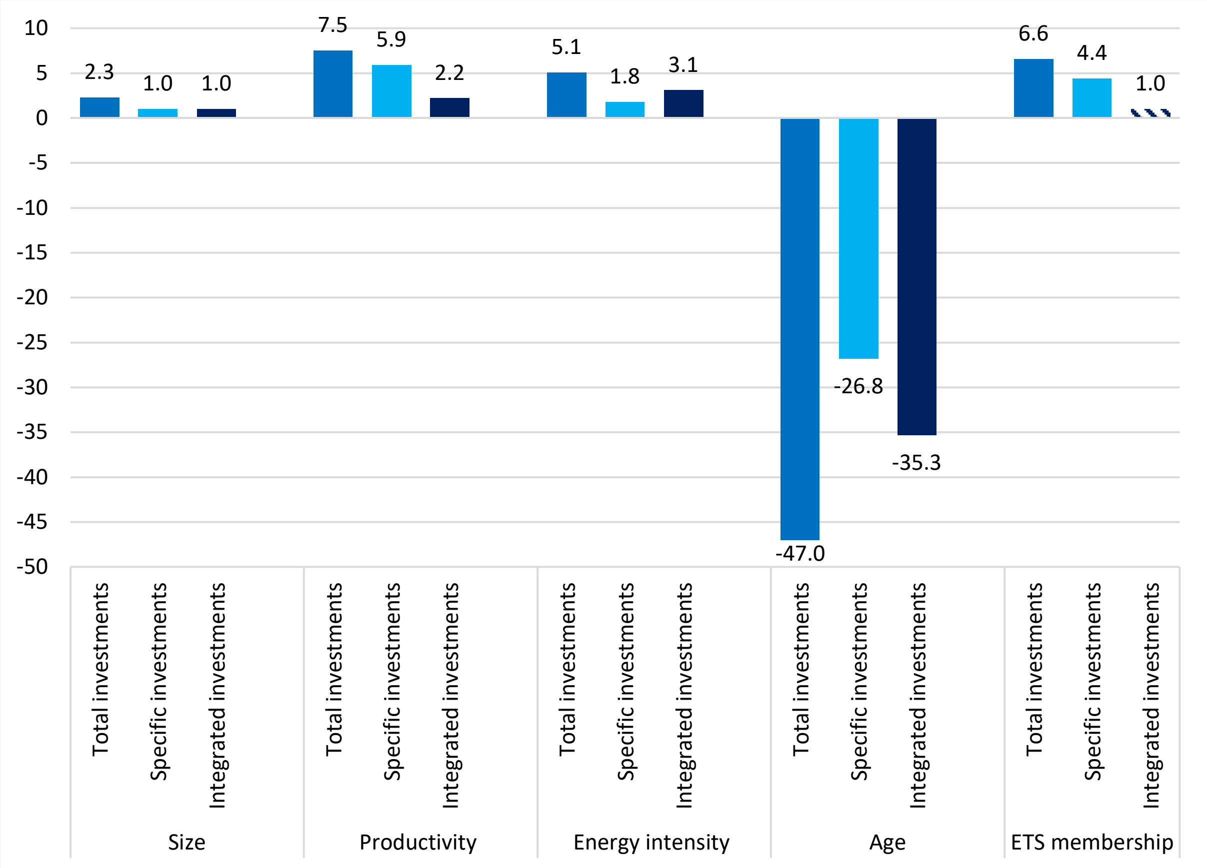
First, all other things being equal, four explanatory variables, namely firm size, productivity, energy intensity and EU ETS membership, have a positive and statistically significant effect on the adoption of decarbonization investments within industries, in accordance with theoretical intuitions. EU ETS membership does not influence integrated investments.

Second, two explanatory variables, namely firm age and its average energy cost, have a negative and statistically significant effect on the adoption rate within industries. However, the latter result is not robust since the effects are not statistically significant when the model is fitted to both sub-samples.

To illustrate the economic magnitude of these results, Figure 2 simulates the change in the probability of decarbonizing if one statistically significant predictor moves from the first decile to the last decile of its distribution (or from zero to one for discrete variables), holding all other variables in the model at their means. The age effect clearly dominates quantitatively all other variables, which may suggest that substantial lock-in effects hinder the greening of industrial investments.

[20] This joint test is equivalent to the Hausman test. Moderate unobserved heterogeneity uncorrelated with included variables remains: it explains between 39 and 45% of the variance of the error term.

Figure 2: Difference in estimated probability of adoption between first and last deciles firms according to different firm characteristics, in p.p.



Source: DG Trésor estimates.
Note: The estimated probability of adopting decarbonization technologies in high productivity firms (top decile) is 7.5 percentage points higher than in low productivity firms (bottom decile), holding all other firm characteristics at their means. Bars are hatched if the underlying coefficient is not statistically significant.

Third, a particular focus on sectoral dummies shows that, all other things being equal, the probability of investing depends statistically significantly on the sector of activity (see Table 4b).

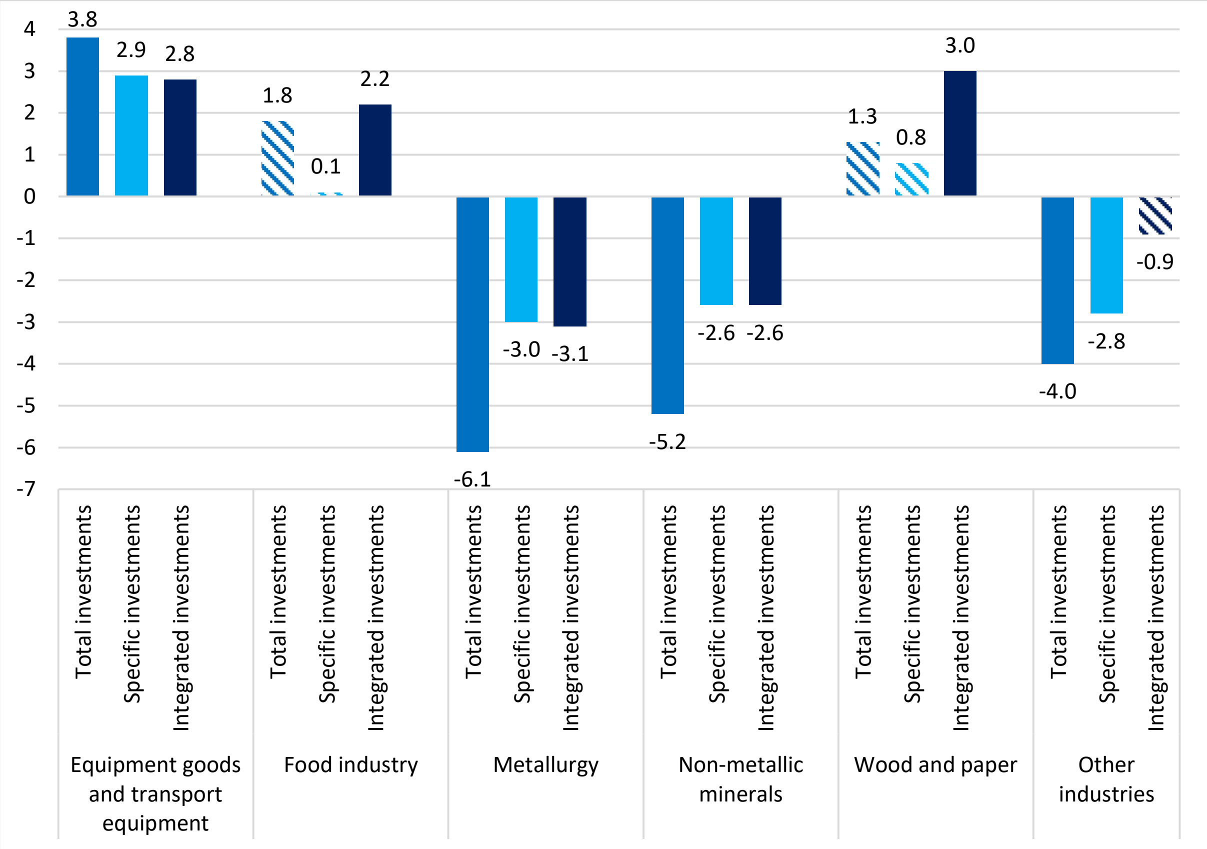
Table 4b: Sectoral dummies of the selection equation

	Total investments		Specific investments		Integrated investments	
	Coeff.	*p-value*	Coeff.	*p-value*	Coeff.	*p-value*
Equipment goods and transport	0.22	*0.05*	0.27	*0.04*	0.23	*0.06*
Food industry	0.11	*0.30*	0.01	*0.93*	0.19	*0.10*
Metallurgy	−0.49	*0.00*	−0.43	*0.00*	−0.37	*0.00*
Non-metallic minerals	−0.39	*0.00*	−0.35	*0.03*	−0.30	*0.05*
Wood and paper	0.08	*0.56*	0.08	*0.63*	0.25	*0.09*
Other industries	−0.29	*0.01*	−0.40	*0.00*	−0.10	*0.42*

Source: DG Trésor estimates.
Note: The p-values are associated with the test of the nullity of the coefficient for each explanatory variable. The chemical industry is the reference category.

To illustrate the economic magnitude of these results, Figure 3 simulates the adoption of green technologies in each industrial sector in contrast to that in the chemical industry, holding all other variables in the model at their means. After controlling for other factors, industrial firms belonging to the most polluting sectors (chemicals, metallurgy, non-metallic minerals) do not trigger these investments more than others. The study does not make it possible to say whether this is due to higher abatement costs, insufficiently targeted subsidies for these sectors, or unavailable technologies. In any case, this econometric result provides emerging evidence that decarbonization investments should be accelerated as a priority in these highly-emitting sectors.

Figure 3: Difference in estimated probability of adoption between industrial sectors, in p.p.



Source: DG Trésor estimates.
Note: The estimated probability of adopting decarbonization technologies in the minerals sector is 5.2 percentage points lower than in the chemical industry, holding all other firm characteristics at their means. Bars are hatched if the underlying coefficient is not statistically significant.

3.2 Decarbonization investment intensity (intensive margin)

The results of the second stage, which fits the decarbonization investments regression at the intensive margin using the GLS estimator, are reported in Table 5. Individual heterogeneity (Mundlak-type) is not statistically correlated with the explanatory variables, i.e. a pure random-effects model is not rejected by the data. Three comments are in order.

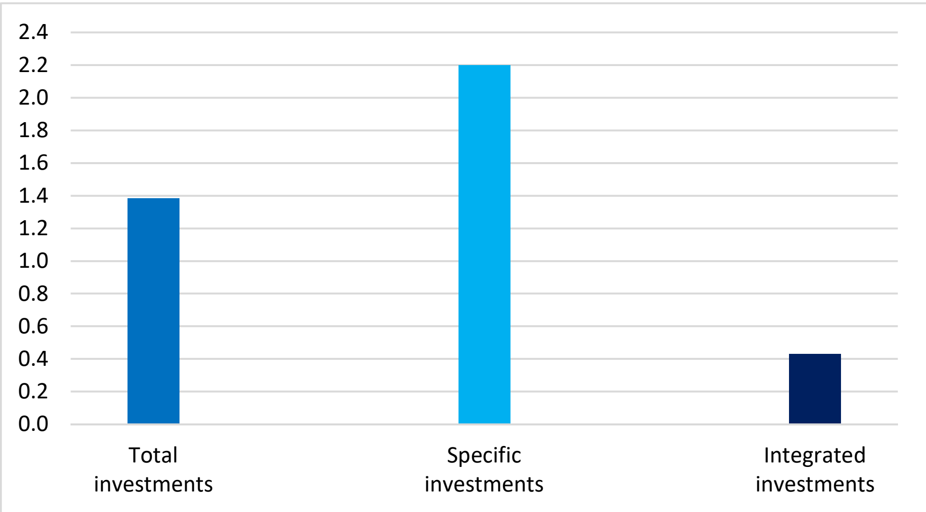
Table 5: Main results of the investment equation (second stage: intensive margin)

	Total investments		Specific investments		Integrated investments	
	Coeff.	*p-value*	Coeff.	*p-value*	Coeff.	*p-value*
Size	$2.21\ 10^{-6}$	*0.14*	$2.85\ 10^{-6}$	*0.18*	$4.2\ 10^{-6}$	*0.10*
Labor productivity	0.005	*0.28*	0.007	*0.49*	0.01	*0.10*
Profit margin	−0.006	*0.44*	−0.005	*0.57*	−0.003	*0.55*
Energy intensity	0.004	*0.01*	0.005	*0.02*	0.003	*0.04*
Average energy cost	−0.001	*0.21*	0.0008	*0.64*	−0.002	*0.05*
EU ETS membership	0.004	*0.35*	0.003	*0.61*	0.002	*0.56*
Age	$9.3\ 10^{-5}$	*0.13*	0.0001	*0.22*	$4.0\ 10^{-5}$	*0.18*
Inverse Mills ratio	0.02	*0.02*	0.02	*0.13*	0.04	*0.02*
Individual heterogeneity (Mundlak-type)	no	*0.20*	no	*0.14*	no	*0.62*
Sectoral dummies	yes	*0.01*	yes	*0.01*	yes	*0.08*
Time dummies	yes	*0.12*	yes	*0.05*	yes	*0.08*
Rho	0.97		0.99		0.88	
Firm-year observations	1148		614		663	

Source: DG Trésor estimates.
Note: The p-values are associated with the test of the nullity of the coefficient for each explanatory variable, and the joint nullity of the coefficients for individual heterogeneity and the sectoral and time dummies. All these tests are based on bootstrapped standard errors (clustered at the firm level). Rho is the proportion of the variance of the error term that is contributed by the random effect.

First, all other things being equal, the investment ratio in decarbonization technologies is influenced by carbonized energy intensity within industries. To illustrate the economic magnitude of these results, Figure 4 simulates the change in decarbonization investment intensity between firms in the first and last deciles of energy intensity.

Figure 4: Change in firm's estimated investment rate in decarbonization technologies when energy intensity moves from the first decile to the last decile of its distribution, in p.p.



Source: DG Trésor estimates.
Note: The estimated investment ratio in decarbonization technologies for high energy-intensive firms (top decile) is 1.4 percentage point higher than for low energy-intensive firms (first decile).

Second, the estimated inverse Mills ratios are statistically significant and positive for the reference sample and the subsample for integrated investments. This result validates the econometric choice: unobserved firm-level characteristics explain positively decarbonization investments both at the extensive and the intensive margins. In other words, firms that have a *low probability* of investing given their observed characteristics (that is high inverse Mills ratios firms) and do invest, also *invest more* than expected all else equal. Among these unobserved factors, the environmental sensitivity of the firm's manager might play a role.[21]

Third, a closer look at the distribution of the inverse Mills ratios shows that some firms that are most likely to invest according to their observable characteristics (that is low inverse Mills ratios firms, say in the first decile of total sample, see Table 6) do not invest.

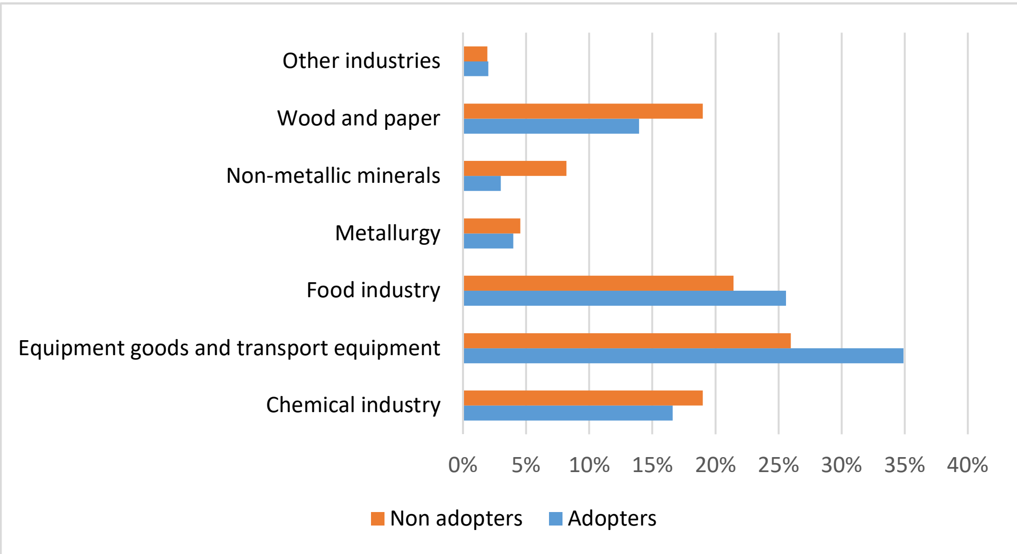
Table 6: Distribution of inverse Mills ratios

	Adopters	Non-adopters	Total sample
P1	0.18	1.10	0.94
P5	0.97	1.44	1.65
P10	1.15	1.62	1.54
P50	1.80	2.18	2.12
P90	2.39	2.64	2.62
P95	2.50	2.74	2.73
P99	2.77	2.93	2.92

Source: DG Trésor estimates.

A relatively large proportion of these low inverse Mills ratios non-adopters belong to the most polluting industrial activities (see Figure 5). The study does not make it possible to say whether this is due to higher abatement costs, insufficiently targeted subsidies for these sectors, or unavailable technologies. In any case, this econometric result provides emerging evidence that decarbonization investments should be accelerated as a priority in these highly-emitting sectors.

Figure 5: Sectoral distribution of firms having inverse Mills ratio in the first decile



Source: DG Trésor estimates.

[21] This interpretation, which I cannot prove in this work, seems plausible in the light of the emerging literature linking the top management team/board composition to firm's environmental performance. For a comprehensive review of the literature, see Nuber and Velte (2021).

4. Conclusion

This research on the drivers of decarbonization investments by French industrial firms leads to the following conclusions:

- Only a minority of industrial firms make green investments and the distribution of these investments is highly skewed.
- The adoption of decarbonization technologies increases with firm size, energy intensity, productivity, and inclusion in the ETS, and decreases with firm age.
- The probability of investing depends on the sector of activity, but firms belonging to the most emitting sectors (chemicals, metallurgy, and non-metallic minerals) do not make these investments more than others, all other things being equal. The study does not make it possible to say whether this is due to higher abatement costs in these sectors, unavailable technologies or insufficiently targeted subsidies for these sectors.

References

Aghion, P., A. Dechezleprêtre, D. Hemous, R. Martin, J. Van Reenen (2016), “Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry”, *Journal of Political Economy*, 124, 1-51.

BCG (2021), « Évaluation d’impact des mesures prises depuis 2017 sur la réduction des gaz à effet de serre en France à horizon 2030 ».

Citepa (2020), « Bilan des émissions en France de 1990 à 2018 », Rapport national d’inventaire, juin.

Dussaux, D. (2020), “The joint effects of energy prices and carbon taxes on environmental and economic performance: Evidence from the French manufacturing sector”, *OECD Environment Working Papers*, n°154.

Glachant, M., C. Mini (2020), « Quand le carbone coûtera cher. L’effet sur la compétitivité industrielle de la tarification du carbone », *Les Notes de la Fabrique*, Presses des Mines.

Heckman, J. (1976), “The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models”, *Annals of Economic and Social Measurement*, 5, 475-492.

Jaraite, J., A. Kazukauskas, T. Lundgren (2014), “The effects of climate policy on environmental expenditure and investment: evidence from Sweden”, *Journal of Environmental Economics and Policy*, 3(2), 148-166.

Nuber, C., P. Velte (2021), “Board gender diversity and carbon emissions: European evidence on curvilinear relationships and critical mass”, *Business Strategy and the Environment*, 30(4), 1958-1992.

Rexecode (2021), « Les émissions françaises de gaz à effet de serre d’ici 2030 : impact de la récession et du plan de relance ».

Trognon, A. (2003), « L'économétrie des panels en perspective », *Revue d'économie politique*, 113(6), 727-748.

Wooldridge, J. (1995), “Selection correction for panel data model under conditional mean independence assumptions”, *Journal of Econometrics*, 68, 115-132.

Annex 1: Kaya decomposition of industrial GHG emissions

1. Method

The level of GHG emissions is decomposed as a function of total industrial production P, the relative weight of each industrial sector w_i and the carbon efficiency of each sector CE_i.

$$GHG = P\frac{GHG}{P} = \sum_{i=1}^{n} P\frac{P_i}{P}\frac{GHG_i}{P_i} = \sum_{i=1}^{n} P.w_i.CE_i$$

The decomposition of industrial GHG emissions cumulative change into separate additive contributions is obtained by using Laspeyres decomposition.

$$\Delta GHG_t = GHG_t - GHG_0 = \Delta P_t \sum_{i=1}^{n} w_{i0}.CE_{i0} + P_0 \sum_{i=1}^{n} \Delta w_{it}.CE_{i0} + P_0 \sum_{i=1}^{n} \Delta CE_{it}.w_{i0} + \Delta P_t \sum_{i=1}^{n} \Delta w_{it}.CE_{i0} + \Delta P_t \sum_{i=1}^{n} w_{i0}.\Delta CE_{it} + P_0 \sum_{i=1}^{n} \Delta w_{it}.\Delta CE_{it} + \Delta P_t \sum_{i=1}^{n} \Delta w_{it}.\Delta CE_{it}$$

Laspeyres decompositions result in residual contributions (embodied in the last 4 terms of the decomposition) which are here evenly distributed between the contributions of the three variables of interest. Those three contributions are written as follows:

Production effect

$$\Delta P_t \sum_{i=1}^{n} w_{i0}.CE_{i0} + \frac{1}{2}.\Delta P_t \sum_{i=1}^{n} \Delta w_{it}.CE_{i0} + \frac{1}{2}\Delta P_t \sum_{i=1}^{n} w_{i0}.\Delta CE_{it} + \frac{1}{3}\Delta P_t \sum_{i=1}^{n} \Delta w_{it}.\Delta CE_{it}$$

Composition effect

$$P_0 \sum_{i=1}^{n} \Delta w_{it}.CE_{i0} + \frac{1}{2}.\Delta P_t \sum_{i=1}^{n} \Delta w_{it}.CE_{i0} + \frac{1}{2}P_0 \sum_{i=1}^{n} \Delta w_{it}.\Delta CE_{it} + \frac{1}{3}\Delta P_t \sum_{i=1}^{n} \Delta w_{it}.\Delta CE_{it}$$

Carbon efficiency effect

$$P_0 \sum_{i=1}^{n} \Delta CE_{it}.w_{i0} + \frac{1}{2}\Delta P_t \sum_{i=1}^{n} w_{i0}.\Delta CE_{it} + \frac{1}{2}P_0 \sum_{i=1}^{n} \Delta w_{it}.\Delta CE_{it} + \frac{1}{3}\Delta P_t \sum_{i=1}^{n} \Delta w_{it}.\Delta CE_{it}$$

2. Data

The industrial sector is considered as defined by the NLCS based on the Citepa[22] nomenclature of economic activities. Sectoral GHG emissions are taken from the Citepa database and are expressed in MtCO2e. Sectoral production is taken from the Insee annual national accounts. To remove price effects, sectoral value added are computed at constant 2000 prices.[23] A transition matrix is created to re-aggregate national accounts sectoral data so that it matches as close as possible the Citepa nomenclature (see Table 7).

[22] Centre interprofessionnel technique d'études de la pollution atmosphérique. It carries out the national inventory of GHG and air pollutant emissions.
[23] The chain-linked volumes are not considered because the sectoral shares would not be additive. Sectoral value added deflators in level A88 are not available before 2000.

Table 7: Transition matrix between Citepa and Insee nomenclatures

Citepa	INSEE (NAF rev.2 in level A88)
Chemical industry	20 21
Building construction	41 42 43
Machinery, Equipment goods and transport equipment	26 27 28 29 30 33
Agri-food industry	10 11
Ferrous metallurgy	24 25
Non-ferrous metallurgy	
Non-metallic minerals	8 23
Wood and paper	17
Other manufacturing industries	13 14 15 16 18 22 31 32

Source : DG Trésor.

Annex 2: Wooldridge's (1995) estimator

In 1995, US econometrician J. Wooldridge derived a method for extending selection models to panel data. This is an important theoretical contribution to panel data econometrics. Indeed, the handling of unobserved heterogeneity is problematic in non-linear models. First, the introduction of fixed effects cannot be solved by within or first difference estimators, unlike in linear models, and introducing individual dummies leaves insufficient degrees of freedom in the estimation. Second, assuming pure random effects, that is uncorrelated to the explanatory variables, is not credible. To avoid these pitfalls, Wooldridge relies on previous work (notably that of Y. Mundlak) which aimed at building a unified approach to deal with unobserved heterogeneity in panel data econometrics, easily transposable to non-linear frameworks.[24]

More precisely, Wooldridge proposes a set of four hypotheses allowing to extend Heckman's (1976) two-step consistent estimator to panels. Let the following selection model be used in this working paper (see part 2.1.):

$$\begin{cases} D_{it}^* = Z_{it}\gamma + \eta_i + u_{it} \; ; \; D_{it} = 1 \; if \; D_{it}^* > 0 \\ I_{it} = X_{it}\beta + \alpha_i + \varepsilon_{it} \;\; if \; D_{it} = 1 \end{cases}$$

Selection is not exogenous and individual effects can be correlated with the explanatory variables, that is

$$E(\alpha_i + \varepsilon_{it} \,|X_{it}, D_{it} = 1\,) \neq 0.$$

The methodology relies on two assumptions concerning the selection equation:

(1) There is a linear relationship between the conditional expectation of the individual effect and the time averages of the explanatory variables in the selection equation (this is the so-called Mundlak hypothesis).

$$\eta_i = \overline{Z}_\iota\delta + c_i \; ; \; \overline{Z}_\iota = \frac{1}{T}\sum_{t=1}^{T} Z_{it} \; ; \; c_i \perp (Z_{it}, u_{it})$$

$$\Rightarrow \; D_{it}^* = Z_{it}\gamma + \overline{Z}_\iota\delta + c_i + u_{it}$$

(2) The new error term $v_{it} = c_i + u_{it}$ of the selection equation is uncorrelated with the explanatory variables and follows a normal distribution. That's why this new selection equation can be fitted to the data using a random-effects probit estimator.

$$c_i \sim N(0, \sigma_c^2); \; u_{it} \sim N(0,1)$$

$$\Rightarrow \; v_{it} \sim N(0, \sigma_v^2) \; ; \; \sigma_v^2 = \sigma_c^2 + I_T$$

The methodology also relies on two assumptions specifying the relationship between the unobserved variables of the two equations that is the correction for selection bias.

(3) There is a linear relationship between the conditional expectation of the individual effect of the linear equation, the time averages of its explanatory variables and the error term of the selection equation.

$$E(\alpha_i|X_i, Z_i, v_{it}) = \overline{X}_\iota\psi + \phi v_{it} \; ; \; X_i = (X_{i1} \ldots X_{iT}) \; ; Z_i = (Z_{i1} \ldots Z_{iT}) \; ; \; \overline{X}_\iota = \frac{1}{T}\sum_{t=1}^{T} X_{it}$$

(4) There is a linear relationship between the conditional expectation of the error term of the linear equation and the error term of the selection equation.

[24] For a brief history of panel data econometrics, see Trognon (2003).

$$E(\varepsilon_{it}|X_i, Z_i, v_{it}) = E(\varepsilon_{it}|\, v_{it}) = \rho v_{it}$$

$$(3) + (4) \Rightarrow E(\alpha_i + \varepsilon_{it}\,|X_i, Z_i, D_{it} = 1\,) = E(\alpha_i\,|X_i, Z_i, D_{it} = 1\,) + E(\varepsilon_{it}\,|X_i, Z_i, D_{it} = 1\,)$$

$$= \bar{X}_i\psi + \phi E(v_{it}\,|X_i, Z_i, D_{it} = 1\,) + \rho E(v_{it}\,|X_i, Z_i, D_{it} = 1\,)$$

$$= \bar{X}_i\psi + \omega E(v_{it}\,|X_i, Z_i, D_{it} = 1\,) = \bar{X}_i\psi + \omega\lambda_{it}\,;\; \lambda_{it} = \frac{\varphi(\frac{Z_{it}\gamma + \bar{Z}_i\delta}{\sigma_v})}{\Phi(\frac{Z_{it}\gamma + \bar{Z}_i\delta}{\sigma_v})};\; \omega = \phi + \rho$$

$$D_{it} = 1 \;\Rightarrow I_{it} = X_{it}\beta + \bar{X}_i\psi + \omega\lambda_{it} + e_{it}$$

$$E(e_{it}\,|X_i, Z_i, D_{it} = 1\,) = E(\alpha_i + \varepsilon_{it} - E(\alpha_i + \varepsilon_{it}\,|X_i, Z_i, D_{it} = 1\,))|X_i, Z_i, D_{it} = 1\,) = 0$$

$$V(e_{it}\,|X_i, Z_i, D_{it} = 1\,) = \sigma_\alpha^2 + I_T\sigma_\varepsilon^2$$

That’s why this augmented version of the linear equation is fitted to the data using the GLS estimator.