Adverse Selection in Carbon Offset Markets in China

Method Note by FU Zixuan

May 21, 2025

1 Empirical Evidence

The empirical section reports two type of regression relationships. The first one is on the treatment effect of **project registation** on **firm's emission**. The other presents evidence on how board's **approval decision** is affected by the the **rate of return** of the propsoed project.

1.1 Emission outcomes

Main specification The regression specification is as follows:

$$Y_{it} = \alpha_i + \alpha_{jt} + \sum_{\tau=-5}^{4} \beta_{1\tau} 1\left\{t - \text{Start}_i = \tau\right\} \text{Proposed}_i + \beta_{2\tau} 1\left\{t - \text{Start}_i = \tau\right\} \text{Registered}_i + \epsilon_{it} \tag{1}$$

As nicely discussed in Borusyak et al. (2024), an event study is comprised of two parts: (1) an estimating equation and (2) data structure. Looking at the **data structure**, we have a panel data with staggered adoption/varying event data and there are never-treated units which are non-proposing firms. In terms of the **estimating equation**, we have two treatment groups: (1) proposed but non-registered (β_1) and (2) registered (β_2), and dynamic (w.r.t event date) treatment effects (τ). The coefficients of interests are $\beta_{1\tau}$ and $\beta_{2\tau}$ for $\tau = -5, \ldots, 4$.

It is well-documented in the literature that the ususal n-way fixed effect estimator is biased in the case of heterogenous treatment effects across units i and across time (calendar time t or event time τ).[1] Therefore, the paper follows Borusyak et al. (2024) to tackle the issue by following a three-step procedure.

To summarize the entire estimation of the equation above 1, the paper proceeds as follows:

1. Select a candidate control firm such that the distance of the emission outcome between the treated and the control firm is minimized.

[1] Since DiD is a special case of event study, the same problem (heterogenous across units i and calendar time t) applies to DiD as well.

2. After the set of control firms is constructed, the paper employs a three-step procedure.
   (a) Estimate the fixed effects α_i, α_{jt} using **the untreated observations.**
   (b) Use the estimated fixed effects $\hat{alpha_i}, \hat{\alpha}_{jt}$ to construct the counterfactual untreated outcome $\hat{Y}_{it}(0)$ for **the treated observations.** Then compute $\hat{T}_{it} = Y_{it} - \hat{Y}_{it}(0)$.
   (c) Use $\hat{T}_{it}$ to estimate any functions of $\hat{T}_{it}$, such as the mean value of $\hat{T}_{it}$ (average over each event year τ) for each firm i.

Alternative specification There are some variation of the regression specification 1 for robustness check/comparison.

- Restrict sample to proposed firms only. The estimating equation is the same except without the $\beta_{1\tau}$ term.
- Pool the post-period indicator into one post-event indicator, therefore estimating the average change in emission after the event.
- Varying the number of control firms for a treated firm from 5 to 10.
- Varying the outcome variable Y_{it}, such as using the projected emission from project registrion instead of actual emission, or using the log of emission.

Table 1: Data Structures for Event Study Estimation

	Only Ever-Treated Units	*There are Never-Treated Units*
Common Event Date	N/A	DiD-type
Varying Event Date	Timing-based	Hybrid

Note: Author's proposed labels for event study data structures, based on whether the analysis data sample uses never treated units or not, and on whether treated units have a common event date or varying event dates. "DiD-type" = "Difference in Difference type."

1.2 Registration outcomes

The treatment effect estimation is contaminated by 3 factors: (1) self-selection of the firms, (2) external screening of the board, and (3) productivity effect. This subsection presents some evidence on the board's screening practices. The paper specifies the following linear probability model:

$$\Pr(\text{Registerd}_i = 1|X) = \alpha_t + \alpha_k + \alpha_c + \alpha_l + \text{InternalRateReturn}_i\beta_1 + X_i'\beta_2 + \epsilon_i \quad (2)$$

A linear probability model has latent variable formulation where

$$Y_i^* = b_0 + b_1' X_i + \epsilon_i, \quad \epsilon_i \sim \text{Uniform}(-a, a)$$

where $Y_i^* = 1$ if $Y_i^* > 0$ and $Y_i^* = 0$ otherwise. Whether ϵ_i is uniform or normal or logistic does not matter much, the sign of the coefficient β_1 is of importance here. A negative sign indicates that the board tends to reject projects with high IRR, that is, the non-additional projects.

2 Theoretical Model

The model adopts a fairly simple structure, but incorporates the key features from empirical evidence. Yet it is also restrictive in the sense that everything takes an explicit functional form, e.g., production function, signal distribution. It is of another question whether the model can be generalized and whether it is robust to other functional forms. This section only walks through the model presented in this paper in detail.

2.1 Firm side

Production function Firm produces two outputs: (1) the products y and (2) the emissions e, using the same input v. The two production functions are:

$$y = (1-a)zv \tag{3}$$

$$e = \left(\frac{1-a}{z_e}\right)^{\frac{1}{\alpha_e}} zv \tag{4}$$

Here, a is the abatement effort. Intuitively, the higher the effort, the lower the emissions as well as lower the output. For efficiency, z is the production efficiency while z_e is the abatement efficiency. α_e denotes the elasticity of emissions with respect to abatement effort $1-a$. That is

$$\frac{\partial e}{\partial(1-a)} \frac{(1-a)}{e} = \frac{1}{\alpha_e} \frac{e(1-a)}{(1-a)e} = \frac{1}{\alpha_e}$$

By combinging the two production functions 3, we can express one output in terms of another. Thus, we effectively treat e as an input. substitute $(1-a)$ in y:

$$y = \frac{e}{zv}^{\alpha_e} z_e zv = e^{\alpha_e} v^{1-\alpha_e} z^{1-\alpha_e} z_e \tag{5}$$

This is a Cobb-Douglas production function with e, v, as input and $\tilde{z} = z^{1-\alpha_e} z_e$ as the productivity parameter. This rewritten function is another perspective of the production procedure.

Production decision We simply assume the inverse demand curve follows $p = y^{-\frac{1}{\eta}}$. We also assume that the cost of input v is c and the price of emission is t. Then given the demand, cost, production function 5, we can solve for the optimal input v^* and e^*. This is a standard problem in intermdediate microeconomics. Our object of interests are the closed form expression of emission level e^* and emission per output level $\frac{e^*}{y^*}$.

Now let's derive it step by step :)

First, there is fixed ratio of the two inputs v and e in the equilbrium production.

$$\frac{e^*}{v^*} = \frac{\alpha_e}{1-\alpha_e}\frac{c}{t} = A$$

Then there is a fixed ratio between the input and the output $y = \tilde{z}v^{1-\alpha_e}e^{\alpha_e}$.

$$\frac{e^*}{y^*} = \frac{e^*}{\tilde{z}v^{1-\alpha_e}e^{\alpha_e}} = \frac{1}{\tilde{z}}\frac{e^*}{v^*}^{1-\alpha_e} = \frac{1}{\tilde{z}}\left(\frac{\alpha_e}{1-\alpha_e}\frac{c}{t}\right)^{1-\alpha_e} = \frac{1}{\tilde{z}}A^{1-\alpha_e}$$

Similarly,

$$\frac{v^*}{y^*} = \frac{1}{\tilde{z}}\frac{v^*}{e^*}^{\alpha_e} = \frac{1}{\tilde{z}}\left(\frac{1-\alpha_e}{\alpha_e}\frac{t}{c}\right)^{\alpha_e} = \frac{1}{\tilde{z}}\frac{1}{A^{\alpha_e}}$$

Second, we look at firm's problem by choosing y^* directly instead of choosing the two inputs. The firm chooses y^* to maximize the profit:

$$\max_y \pi = y^{1-\frac{1}{\eta}} - c\cdot v - t\cdot e = y^{1-\frac{1}{\eta}} - y\left(c\frac{1}{\tilde{z}}A^{1-\alpha_e} - t\frac{1}{\tilde{z}}A^{\alpha_e}\right)$$

Then the first order condition of $y^{-\frac{1}{\eta}} - y\frac{C_w}{\tilde{z}}$ is

$$(1-\frac{1}{\eta})y^{-\frac{1}{\eta}} - \frac{C_w}{\tilde{z}} = 0$$

Then

$$y^* = \left(\frac{C_w}{\tilde{z}}\frac{\eta}{\eta-1}\right)^{-\eta} \tag{6}$$

Consequently, e^* and v^* are solved as a fixed ratio of y^*.

$$e^* = \frac{1}{\tilde{z}}A^{1-\alpha_e}\left(\frac{C_w}{\tilde{z}}\frac{\eta}{\eta-1}\right)^{-\eta} \propto \tilde{z}^{\eta-1} \tag{7}$$

Abatement decision The firm can improve the abatement efficiency z_e by a factor of Δ_e via investing in an abatement project. At the same time, the production efficiency z grows by a factor of Δ_z. From equation **??**, we can see that the emission level $e^* \propto (z^{1-\alpha_e}z_e)^{\eta-1}$.

Thus, following a change in efficiency z_e and z, the emission level changes as follows:

$$\frac{e_1}{e_0} = \delta_z^{(1-\alpha_e)(\eta-1)}\delta_e^{(\eta-1)} \tag{8}$$

The optimal level of emission changes following a change in the efficiency level. The firm investment decision depends on the profit from the investment. Therefore we need to go from **emission change** to **profit change**. Yet this is pretty straightforward because optimality condition dictates that the profit is linear in output y^* or input and emission. That is

$$\pi \tag{9}$$

Therefore, the profit change is

$$\frac{\pi_1}{\pi_0} = \frac{1}{\eta-1}\frac{t_e}{\alpha_e}(\Delta_e^{\eta-1})\Delta_z^{(1-\alpha_e)(\eta-1)} \tag{10}$$

The **additional profit change following an investment** is therefore $\frac{1}{\eta-1}\frac{t_e}{\alpha_e}(\Delta_e^{\eta-1} - 1)\Delta_z^{(1-\alpha_e)(\eta-1)}e_0 := b(\Delta_e, \Delta_z)e_0$ Then the firm will compare this additional profit change with the **cost of investment**, which is assumed to be $F(\Delta_e, e_0)\varepsilon$

Without the CDM project, the firm will invest if $b(\Delta_e, \Delta_z)e_0 > F(\Delta_e, e_0)\varepsilon$.

CDM act Now introducting the CDM project, it will induce some firms that origianlly do not invest to invest because there is additional benefit to the baseline $b(\Delta_e, \Delta_z)e_0$ from selling carbon offset credits granted by the board. Those firms are the **additional** firms. The board would like to only grant credits to this type of firms rather than the **non-additional** firms whose private benefit is already larger than the investment cost.

Before moving to the board's side, we need to quantify the benefits from the CDM registration (additional profit from investment), we need to define how the board grants credits. The number of CER credits granted to the firm is calculated based on the emission level change in producing **the same amount of output** as before. Therefore, a project that increases the abatement efficiency to increase by a factor of Δ_e will be granted

$$\left[1 - \left(\frac{1}{\Delta_e}\right)^{\alpha_e}\right] e_0$$

Given a fixed price of CER p, the firm's additional profit from registrating a CDM project is

$$\left[1 - \left(\frac{1}{\Delta_e}\right)^{\alpha_e}\right] e_0 p \tag{11}$$

2.2 Board side

Because of asymmetric information, the board does not know the firm's productivity growth δ_z and assumes it to be 1. Second, it does not observe the cost shock ϵ and assumes it to be from a distribution F_ε. Therefore, the board is comparing the the **expected profit** from the project with the **expected cost** of the project to evaluate whether the project is additonal or not. Indeed, the firm has perfect private information on δ_z and ε. Therefore, a natural discrepancy between true additional and board's screening additional arises.

To summarize, firm is additional if

$$b(\Delta_e, \Delta_z)e_0 < F(\Delta_e, e_0)\varepsilon \quad \text{and} \left(b(\Delta_e, \Delta_z) + \left[1 - \frac{1}{{\Delta_e}^{\alpha_e}}\right] p\right) e_0 > F(\Delta_e, e_0)\varepsilon$$

The board decide whether it's additional

$$b(\Delta_e, 1)e_0 < F(\Delta_e, e_0)\varepsilon \Leftrightarrow \frac{b(\Delta_e, 1)e_0}{F(\Delta_e, e_0)} < 1$$

But since it only has a noisy signal ε^s, it sets a different threshold than $\bar{R} = 1$. We denote the threshold as $\bar{R}$.

2.3 Game

The application decision of the firm and the approval decision of the board are two sequential moves in the game. There are two types of firms that will apply (1) the additional firms, which could not earn a profit without from the investment without the registration, and (2) the non-additional firms, which would invest anyway. Denote the probability of a firm being registered conditional on the cost shock ε as $\Pr(R = 1 \mid \varepsilon)$. We denote the fixed cost of applying to be Ae_0 [2].

1. Additional firms: $\pi_A = \Pr(R = 1 \mid \varepsilon)\left[b + p\delta e_0 - F(\Delta_e, e_0)\epsilon\right]$. It applies when $\pi_A > Ae_0$.

2. Non-additional firms: $\pi_{NA} = \Pr(R = 1 \mid \varepsilon)p\delta_e e_0$ applies when $\pi_{NA} > Ae_0$.

The second stage is the board's decision. The board set a threshold $\bar{R}$, which can be equivalent to setting a threshold on the ε^s, denoted by $[\varepsilon^s]$.

In equilibrium, the $\Pr(R = 1 \mid \varepsilon) = \Pr(\varepsilon^s > \bar{\varepsilon}^s \mid \varepsilon) = 1 - F_{\varepsilon^s|\varepsilon}(\bar{\varepsilon}^s)$

The cdf $F_{\varepsilon^s|\varepsilon}$ is public knowledge, as well as the threshold $\bar{R}$ and $\bar{\varepsilon}^s$.

Now we can compare the difference in emission growth between applying and non-applying firms.

$$\begin{aligned}
&E\left[\log(g_e) \mid \text{apply}, \varepsilon\right] - E\left[\log(g_e) \mid \text{not apply}, \varepsilon\right] \\
&= (\eta - 1)\log \Delta_e + E((1-\alpha_e)(\eta - 1)\log \Delta_z \mid \text{apply}, \varepsilon) \\
&- E((1-\alpha_e)(\eta - 1)\log \Delta_z \mid \text{not apply}, \varepsilon)
\end{aligned}$$

As well as the difference between the registered and proposing-only firms.

[2] Why is the fixed cost proportional to e_0?

3 Estimation

The estimation proceeds in 4 steps. There are 5 main elements to estimate (1) production function, (2) abatement cost $F(\Delta_e, e_0)$ (3) abatement efficiency growth Δ_e , (4) pruduction efficiency growth Δ_z, (5) signal structure $F_{\varepsilon^s|\varepsilon}$ and screening threshold $\bar{\varepsilon}^s$.

Production function The paper makes the folloiwng Cobb-Douglas assumption on the production function such that when taking the log of ouput, we have

$$\log y_{it} = \log z_i^e + (1-\alpha_e)[\log z_{it} + \alpha_l \log l_{it} + \alpha_k \log k_{it}] + \alpha_e \log e_{it} \tag{12}$$

For simplicity in the exposition, I myself take output y to be observed. Following the now classical approach first proposed in Ackerberg et al. (2015), the paper assumes that there is an intermediate input m that is monotonically increasing in the productivity z_{it}.

$$\log y_{it} = \phi(l_{it}, k_{it}, e_{it}, m_{it}) + \log z_i^e + \varepsilon_{it}^m$$

In the first stage, we estimate the function

$$\phi(l_{it}, k_{it}, e_{it}, m_{it}) \equiv \alpha_l \log l_{it} + \alpha_k \log k_{it} + \alpha_e \log e_{it} + m^{-1}(l_{it}, k_{it}, m_{it})$$

In the second stage, we use the assumption that log productivity $\log(z_{it})$ follows a Markov process, such that we can write $\log z_{it} = g(\mathcal{P}_{it-1}, \log z_{it-1}) + \varepsilon_{it}^z$ where $\mathcal{P}_{it-1}$ is the survival probability of the firm i at time $t-1$. [3] Thus, we can rewrite function ϕ as

$$\hat{\phi}_{it} = \alpha_l \log l_{it} + \alpha_k \log k_{it} + \alpha_e \log e_{it} + g\big(\mathcal{P}_{it-1}, \hat{\phi}_{it-1} - \alpha_l \log l_{it-1} - \alpha_k \log k_{it-1} - \alpha_e \log e_{it-1}\big) + \varepsilon_{it}^z$$

In the equation, all the variables are either directly observed (l_{it}, k_{it}, e_{it}) or estimated from the first stage $(\hat{\phi}_{it}, \hat{\phi}_{it-1}, \hat{\mathcal{P}}_{it-1})$. Similar to how we approxiamte the ϕ, we approximate the function g by some 2nd or 3rd order polynomial function. Then we can estimate all the parameters in the model by GMM.

Abatement cost $F(\Delta_e, e_0)$ The parametric assumption is that

$$\log(F) = \log(\gamma_0) + \gamma_1 \log(\delta_e e_0) + \varepsilon$$

Since F and $\delta_e e_0$ are both observed, estimate the linear equation is straightforward.

Abatement efficiency growth Δ_e Recall the credits is calculated by

$$\left[1 - \left(\frac{1}{\Delta_e}\right)^{\alpha_e}\right] e_0$$

[3] The survival probability is not present in the paper perhaps due to the fact that the sample is not subjected to selection bias.

Since e_0 is observed, α_e is estimated in step 1. Then one can get an estimate of Δ_e for each project and get a weighted average.

$\Delta_z, F_{\varepsilon^s|\varepsilon}$ and $\bar{\varepsilon}^s$ The following parametric assumptions are imposed:

1. $\Delta_z \sim N(\mu_{\Delta_z}, \sigma^2_{\Delta_z})$
2. $\text{corr}(\varepsilon^s, \varepsilon) = \rho$

Then there are 4 parameters to estimate: $\mu_{\Delta_z}, \sigma^2_{\Delta_z}, \rho, \bar{\varepsilon}^s$. The estimation is done by GMM where 4 moment conditions are chosen with (1-3) the emissions growth rates of registered, proposed and non-applicant firms (4) the registration rate. [4]

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285.

[4] I wouldn't consider the identification argument here as complete.