



**Utrecht
University**

The Innovation Impact of CO₂PL-Based Public Procurement

Jihwan Ryu (j.ryu@uu.nl)

Prof. dr. ir. Fredo Schotanus (f.schotanus@uu.nl)

Dr. Vitezslav Titl (V.Titl@uu.nl)

Table of Contents

1	Introduction	3
2	Data	5
2.1	Public Procurement Data	5
2.2	Firm Certification	7
2.3	Firm-Level Innovation Data from Orbis	10
3	Methodology	14
3.1	Treatment Definitions	15
3.2	Staggered Difference-in-Differences Framework	15
3.3	Event-Study Specification	16
3.4	Estimation sample	18
4	Results	18
4.1	Main Effects: CO ₂ PL Tender Wins with CO ₂ PL Certification	19
4.2	Comparison Groups: Certification and Procurement Heterogeneity	19
4.3	Heterogeneity by Contract Value in CO ₂ PL-Applied Procurement	22
4.4	Heterogeneity by Sectoral R&D Intensity	25
4.5	Additional Heterogeneity Analyses	28
5	Robustness Checks and Alternative Specifications	29
5.1	Placebo Test: Green Public Procurement Without CO ₂ PL Application	29
5.2	Alternative Innovation Metrics	30
5.3	Population-Level Test: Full Dutch Firm Sample	30
6	Discussion and Conclusion	31
1	Appendix	36
1.1	Heterogeneity by Firm Size	36
1.2	Heterogeneity by Tender Type	37
1.3	Robustness: Innovation Effects of GPP-Labelled Tender Wins	38
1.4	Robustness: Alternative Innovation Outcome Measures	39
1.5	CO ₂ PL-Related Tender Wins in the Full Dutch Firm Sample	40

1 Introduction

Public procurement is increasingly recognised as a potential lever for accelerating the transition toward a low-carbon economy. In the Netherlands, the CO₂ Performance Ladder (CO₂PL) has become a central instrument for encouraging climate-conscious behaviour among firms competing for public contracts. While previous evaluations have highlighted the Ladder's environmental benefits far less is known about its potential to stimulate *innovation* among participating firms.

Innovation is essential for long-run decarbonisation. Firms may develop new technologies, adopt cleaner production processes, or strengthen their environmental performance. CO₂PL-based procurement may contribute to innovation by encouraging firms to engage in continuous improvement, as certification requires systematic measurement, target setting, and periodic evaluation of CO₂ reduction efforts. However, such investments are costly, and it remains unclear whether the incentives created by CO₂PL-based procurement are sufficiently strong to induce measurable increases in innovative activity. Moreover, observed emission reductions may often reflect relatively incremental adjustments—such as logistical optimisation or changes in energy use—rather than substantial technological or organisational innovation, since the ladder is not primarily designed as an innovation policy instrument.

A substantial and largely positive body of literature has already examined the relationship between innovation and public procurement. Among others, Appelt and Galindo-Rueda (2016) show that public procurement can act as an important driver of firm-level innovation, but countries lack consistent measurement frameworks and data to properly quantify this link. Using innovation surveys and linked US administrative datasets, the authors find that firms supplying the public sector are more likely to engage in innovation and R&D, while highlighting the need for better procurement-innovation data for reliable policy evaluation. Blind et al. (2020) found firms that innovate and participate in standardization bodies are more likely to win public contracts. At the same time, winning public contracts further stimulates firms' future innovation activities, demonstrating a mutually reinforcing cycle between procurement participation and innovative performance. Although much of the literature suggests that public procurement may promote firm-level innovation, this evidence is subject to important limitations. Several studies emphasize selection effects (Uyarra and Flanagan, 2010), and, in certain contexts, the potential for innovation can also be discouraged by, for instance, risk-averse tenders (Uyarra et al., 2014).

It remains unclear whether CO₂PL-based procurement provides sufficient incentives to stimulate innovation. This uncertainty motivates the central question of this report:

To what extent does the use of the CO₂ Performance Ladder in public procurement encourage firms to innovate?

To answer this question, we combine Dutch public procurement data (2010–2024) with firm-level financial and intellectual property (IP) data from the Orbis database. Innovation is measured using a rich set of indicators, including patent counts, patent quality metrics, patent relevance and efficiency, and portfolio-level IP quality scores. In addition, we incorporate information on ISO 14001 environmental management system certification to benchmark the innovation effects of CO₂PL against a widely used, firm-level environmental standard that is not directly tied to public procurement.

Our empirical strategy exploits the timing of a firm's first win of a tender that applied CO₂PL criteria. While firms deliberately decide whether to obtain certification, the precise moment when a firm first wins a CO₂PL-using contract is plausibly quasi-random, conditional on certification and sectoral competition. This feature allows us to employ a staggered difference-in-differences design in which treatment begins in the year of the first CO₂PL tender win. Untreated observations consist of non-certified firms, firms that never win CO₂PL tenders, and firms not yet treated in the staggered rollout.

This study contributes to the literature in four ways. First, to the best of our knowledge, it provides the first causal evidence on how green public procurement affects firm-level innovation outcomes. Second, it offers the first large-scale empirical assessment linking CO₂PL tender wins to patent-based innovation indicators for Dutch firms. Third, it provides practical insights for Foundation for Climate-Friendly Procurement and Business (Stichting Klimaatvriendelijk Aanbesteden & Ondernemen, SKAO), contracting authorities, and policymakers regarding whether the CO₂ Performance Ladder not only reduces emissions but also promotes the technological development necessary for long-term sustainability transitions. Finally, while most innovation-related papers used the Community Innovation Survey (CIS), we use Orbis data to study innovation.

The remainder of this report is structured as follows. Section 2 describes the data sources and variable construction. Section 3 outlines the empirical strategy. Section 4 presents the main empirical findings, and Section 5 reports a series of robustness and sensitivity analyses. Section 6 concludes by discussing the implications of the results, their limitations, and their relevance for procurement policy.

2 Data

2.1 Public Procurement Data

The analysis draws on Dutch public tenders obtained from *TenderNed*, complemented with structured contract-level information from *TenderGuide* that identifies CO₂PL usage using search terms.

The raw tender data are recorded at the lot-winner level, as tenders may include multiple lots and multiple winning firms. To build firm-level treatment histories, we collapse the data to the firm-year level using the Dutch Chamber of Commerce identifier (KvK number). From the collapsed procurement data, we obtain the following firm-year level information:

- whether and when the firm won any tender,
- whether and when the firm won any tender in which CO₂PL was applied,
- the total annual contract value of all tenders won (irrespective of CO₂PL use),
- the total annual contract value of tenders won that applied CO₂PL, and
- the distribution of contract types (works, services, or supplies), including both total annual contract values by type and the corresponding values for tenders that applied CO₂PL.

To characterise the procurement dataset, we report several descriptive statistics in Table 1. The merged tender-firm panel contains 20,223 unique firms and 44,960 firm-year observations between 2007 and 2027, with the most of observations concentrated in the period 2013–2024. The average annual contract value of tenders won is €7.29 million, though the distribution is highly right-skewed, with a maximum of €2.67 billion.

CO₂PL adoption is present but not dominant: 3,557 firm-years (7.9%) correspond to winning at least one CO₂PL-applying tender. Using a search-term-based method to identify tenders mentioning COPL from 2012 onward, we observe only very limited usage in the early years, with two contracts in 2013 and one in 2014. CO₂PL use appears from 2015 onward and increases consistently through 2024, peaking at 598 treated firms in 2022. Among treated firm-years, the mean annual value of CO₂PL tenders is €10.9 million.

Contract types are unevenly distributed across firms: 5,184 firm-years won works contracts, 33,151 won services contracts, and 10,013 won supply contracts. This indicates that services dominate procurement activity, while works and supplies represent smaller but economically still significant portions. Notably, although works contracts account for only around one-eighth of all contract-award

firm-years (5,184 out of 44,960), they constitute a much larger share of CO₂PL-applying procurement: 1,285 out of 3,557 treated firm-years (approximately one-third). This contrast suggests that the propensity to apply CO₂PL is substantially higher in works contracts than in services or supplies, consistent with the Ladder's strong uptake in construction and infrastructure projects.

Contract values also vary substantially by type. In all tenders, works contracts average €13.1 million per year, compared to €6.14 million for services and €5.40 million for supplies. CO₂PL tenders display a similar ranking but at higher magnitudes: CO₂PL works contracts average €17.0 million, services €6.76 million, and supplies €5.25 million. These patterns suggest that CO₂PL is more commonly applied in higher-value works and service contracts.

Firms also differ substantially in the number of contracts won per year. Across all procurement, firms win an average of 2.10 contracts annually, though the distribution is highly skewed, with some firms winning more than 180 contracts in a single year. Conditional on participating in a given category, firms win on average 1.78 works contracts, 2.00 services contracts, and 2.00 supplies contracts per year. For CO₂PL-applying tenders, firms win fewer contracts on average but still display notable variation. Treated firm-years receive on average 1.59 CO₂PL contracts, consisting of 1.76 works, 1.42 services, and 1.20 supplies contracts per year. CO₂PL supply contracts rarely exceed five per firm-year, whereas CO₂PL services contracts may reach up to 23. Overall, these patterns indicate that CO₂PL exposure typically arises through a relatively small number of contracts, but with heterogeneity across contract types.

Table 1: Contract values and number of contracts per firm–year, by contract type

Contract Type	Obs	Mean Value	Median Value	Mean # Contracts	Max # Contracts
<i>All public tenders</i>					
Works	3,748	€13.1 million	€2.82 million	1.78	49
Services	19,647	€6.14 million	€0.40 million	1.99	187
Supplies	5,718	€5.40 million	€0.63 million	1.99	75
All types	44,960	€7.30 million	€0.58 million	2.10	187
<i>CO₂PL-applying tenders</i>					
Works	985	€17.0 million	€3.40 million	1.76	14
Services	1,443	€6.76 million	€1.00 million	1.42	23
Supplies	314	€5.25 million	€0.67 million	1.20	5
All types	3,557	€10.9 million	€1.53 million	1.59	23

2.2 Firm Certification

To enrich the procurement dataset, we incorporate information on firms' environmental management certifications. This includes CO₂PL certification data obtained from SKAO and ISO 14001 certification data retrieved from SCCM. These sources allow us to observe, for each firm and year, whether a firm holds a CO₂PL certificate, its certification level, and whether it is certified under the ISO 14001 environmental management standard.

Environmental management systems (EMSs) are closely linked to firms' innovation activities because they require firms to build internal capabilities, such as systematic environmental monitoring and improved information flows, that can subsequently facilitate technological and process innovation. Prior research shows that environmental management practices can stimulate innovation by reducing uncertainty, strengthening coordination, and incentivising firms to identify and implement cleaner production technologies (Wagner, 2007). At the same time, environmentally proactive firms are also more likely to adopt formal EMSs such as ISO 14001, suggesting that EMS adoption both reflects and reinforces firms' underlying innovation capacity (Ziegler and Nogareda, 2009).

The CO₂PL is a carbon-focused EMS and procurement instrument requiring firms to systematically measure emissions, set reduction targets, report transparently, and engage with supply-chain part-

ners (SKAO, 2024). Importantly, CO₂PL certification directly influences procurement outcomes: theoretically, higher certification levels yield award advantages in Dutch public tenders. This creates explicit incentives for firms to invest in energy-efficient technologies, innovate in carbon reduction processes, and strengthen organisational climate strategies. These incentives resonate with evidence showing that involvement in standards-driven procurement can reinforce innovation cycles—innovative firms tend to win more tenders, and tender participation can itself stimulate further innovation (Blind et al., 2020).

Our analysis focuses on firm-level CO₂PL certification status, measured using certificates that are valid for three years. Project-level declarations submitted in the context of individual contracts are not observed in our data and are therefore not included in the analysis. As a result, the estimates capture the effects of formal, organisation-wide certification rather than project-specific implementation of CO₂PL requirements.

To accommodate this feature of the institutional setting, we classify a firm as CO₂PL-certified for treatment purposes if it holds a valid CO₂PL certificate either in the award year or within the subsequent two years. Tender documents commonly permit bidders to commit to a CO₂PL ambition level at the time of award and to demonstrate compliance through certification within one to two years after contract award, during the initial phase of contract execution. This two-year window ensures that firms that obtain certification shortly after winning a tender, in line with contractual requirements, are correctly captured as CO₂PL participants.

ISO 14001 is a voluntary international EMS standard focused on continuous improvement and strategic integration of environmental performance (Camilleri, 2022). While it is not a carbon-specific instrument, ISO 14001 is widely recognised in Dutch public procurement and can directly influence contract outcomes when environmental management is relevant to contract execution. Official guidelines allow contracting authorities to require ISO 14001 certification when environmental performance is contractually relevant (Boot Advocaten, 2015). Moreover, EU-level guidance recognises ISO 14001 as acceptable evidence of technical and professional ability in green public procurement (European Commission, 2016). In practice, the standard has appeared in environmentally intensive tenders—such as ICT procurement and energy-performance contracts—as a qualification requirement or as supporting evidence within award criteria (European Commission, 2025).

Including ISO 14001 certification in the dataset therefore serves two purposes. First, it provides a broader measure of firms' general environmental management practices, enabling comparison between firms with different EMS approaches. Second, it allows us to contextualise whether any observed relationship between CO₂PL-related procurement exposure and innovation reflects CO₂PL's unique procurement-linked incentives, rather than generic effects of EMS adoption. By incorporat-

ing both certifications, we can determine whether CO₂PL is associated with stronger innovation outcomes than a more general EMS like ISO 14001, or whether the patterns reflect underlying environmental management capabilities shared by EMS-adopting firms.

The merged dataset is organised at the firm–year level: each observation represents a firm in a given calendar year and records its procurement activity and certification status. Since firms may appear in the data across multiple years, certification patterns differ depending on whether they are measured at the firm level (whether a firm was ever certified) or at the firm–year level (the number of years in which certification is observed). Table 2 summarises these patterns.

A further consideration concerns how certification is linked to procurement data. For the CO₂PL, tender rules in many contracting authorities allow bidders to obtain CO₂PL certification after contract award, provided that the firm commits to acquiring the certificate within a specified time-frame. To accommodate this feature of the institutional setting, we classify a firm as CO₂PL-certified for treatment purposes if it holds a valid CO₂PL certificate either in the award year or within the subsequent two years. Tender documents commonly permit bidders to commit to a CO₂PL ambition level at the time of award and to demonstrate compliance through certification within one to two years after contract award, during the initial phase of contract execution. This two-year window ensures that firms that strategically adopt the Ladder immediately after winning a tender are correctly captured as CO₂PL participants.

In contrast, ISO 14001 certification does not function as a tender-specific commitment mechanism and can be applied across a wide range of activities unrelated to public procurement. For this reason, we classify ISO 14001 status strictly based on whether the firm is certified in the award year itself. This distinction ensures comparability across certification types while respecting differences in how CO₂PL and ISO 14001 interact with procurement processes.

At the firm level, out of 20,223 firms in the procurement dataset, 1,927 firms (9.5%) were ever CO₂PL certified and 1,807 firms (8.9%) were ever ISO 14001 certified, with 828 firms (4.1%) holding both certifications at some point. At the firm–year level, the procurement panel contains 7,289 CO₂PL-certified firm–years and 6,090 ISO 14001-certified firm–years. These statistics document the prevalence and overlap of EMS adoption among firms participating in Dutch public procurement.

Table 2: CO₂PL and ISO 14001 certification: firm-level and firm-year-level summary

Panel A. Firm-level certification (20,223 firms)	Count	Percent
Firms ever CO ₂ PL certified	1,927	9.5%
Firms ever ISO 14001 certified	1,807	8.9%
Firms certified under both systems	828	4.1%
Panel B. Firm-year certification (44,960 firm-years)		
Firm-years with CO ₂ PL certification	7,289	16.2%
Firm-years with ISO 14001 certification	6,090	13.5%

2.3 Firm-Level Innovation Data from Orbis

To measure firms' innovative performance, the procurement-certification panel is merged with firm-level intellectual property (IP) indicators from *Orbis* and *Orbis Intellectual Property*. Because Orbis reports IP metrics only for firms with identifiable patent portfolios, the innovation sample is first restricted to all Dutch firms with non-missing IP information, yielding 22,558 firms.

However, only a subset of these firms ever participate in public procurement. When merging the Orbis IP dataset with the procurement panel, only 599 firms appear in both datasets. This overlap reflects the structure of the Dutch economy: many procurement-active firms—particularly in construction and services—tend not to patent, while many patenting firms rarely participate in public procurement. As a result, innovative outcomes can be analysed only for this intersection of firms.

The Orbis IP module provides several portfolio-level indicators capturing distinct dimensions of patent-based innovation (Elegbede et al., 2022):

- **Total IP Quality:** a composite score summarising overall patent quality based on more than 25 technological, legal, and financial indicators.
- **IP Relevance:** the economic weight of a firm's patent portfolio, measured as the ratio of IP asset value to turnover.
- **Technical Quality:** an assessment of the technological strength and novelty of the underlying inventions, benchmarked within their technology field.
- **Market Attractiveness:** a measure of expected commercial potential, informed by acquisi-

tion patterns, royalty rates, and litigation outcomes.

These indicators capture complementary aspects of innovation—technological content, economic relevance, and commercialisation potential (Elegbede et al., 2022). Our analysis focuses specifically on the **Technical Quality (peer group)** score as the primary innovation outcome. This decision follows from both theoretical considerations and supporting empirical results.

Environmental and EMS-driven innovations tend to be long-horizon, process-oriented, and capability-building rather than immediately market-oriented. Prior studies show that eco-innovations do not necessarily improve short-run financial or commercial outcomes, with profitability impacts depending heavily on the regulatory domain, market structure, and innovation type (Andries and Stephan, 2019; Hizarci-Payne et al., 2021). Moreover, Rennings and Rammer (2011) find that some environmental regulations stimulate market-novel innovations while others lead primarily to adaptive innovations with limited commercial effects. Because indicators such as Total IP Quality, IP Relevance, and Market Attractiveness embed financial, legal, or demand-side components, they are not always sensitive to the technological improvements induced by sustainability-oriented procurement or EMS adoption.

Empirically, technical quality is also the only Orbis IP indicator that shows a robust and statistically significant relationship with exposure to CO₂PL-related procurement in our data. Other indicators appear more sluggish, reflecting commercialisation or valuation processes whose relationship with environmental or EMS-driven innovation is theoretically ambiguous and not expected to respond systematically to procurement incentives.

Finally, we do not use patent counts as an innovation outcome. Prior evidence shows that patent quantity does not reliably capture environmental or process-oriented innovation responses (Ziegler and Nogareda, 2009). Consistently, Wagner (2007) finds that adopting an EMS increases environmental process innovation but is negatively associated with overall patenting, and that patented environmental innovations represent only a small fraction of firms' actual environmental improvements. Similarly, Appelt and Galindo-Rueda (2016) show that receiving public procurement contracts does not significantly raise firms' patent filings. These findings indicate that raw patent counts would provide a narrow and potentially misleading signal of innovation adjustments following CO₂PL exposure.

Although Crespi and Guarascio (2019) document positive effects of both regular and innovation-oriented procurement on patent stocks, their analysis was conducted at the sector-year level, capturing broad industry-level dynamics rather than firm-level behavioural responses.

Taken together, this evidence underscores the limitations of patent counts for detecting the types of organisational, technical, and process innovations that sustainability-driven procurement may induce. For this reason, we use richer intellectual-property quality indicators, especially peer-normalised technical patent quality score, which offer a more theoretically grounded, sector-comparable, and empirically sensitive measure of firm-level innovation performance.

Table 3 summarises their sectoral composition, firm size, and innovation profiles across three groups (CO₂PL only, Tender only, and Tender + CO₂PL). . We have a total of 1,017 firms with 11,201 firm-year observations. Firms with no public-procurement activity are concentrated in finance and professional services, whereas tender participants are more prevalent in manufacturing and wholesale, and CO₂PL-certified tender winners are overwhelmingly active in construction, a sector where the Ladder is widely used. Tendered firms are markedly large. On average, they generate nearly three times the operating revenue of non-tendered firms and hold an order of magnitude more assets. In contrast, CO₂PL firms tend to be smaller, reflecting the dominance of small- and medium-sized construction contractors. Innovation indicators show only modest variation in technology-based quality scores across groups, but CO₂PL firms exhibit substantially lower IP relevance and patenting intensity, consistent with their sectoral concentration in industries where environmental management and operational improvements play a greater role than patent-based innovation.

Table 3: Summary statistics by CO₂PL status and tender participation

	CO₂PL only	Tender only	Tender + CO₂PL
Number of firms	236	600	181
Number of obs.	2,563	6,494	2,144
<i>Sector shares (%)</i>			
Manufacturing	12.8	31.8	18.9
Construction	20.1	3.1	27.9
Wholesale and retail	9.3	20.6	9.5
Professional, scientific services	15.6	15.8	18.2
Financial services	30.4	10.3	7.0
<i>Financial indicators</i>			
Operating revenue (€k)	102,774	300,476	88,948
Total assets (€k)	112,993	3,656,365	193,267
Employees	324	974	324
Age (years)	30.9	40.7	38.3
<i>Innovation indicators</i>			
Patents per year	14.70	81.70	23.10
Inventions per year	6.94	33.19	7.89
Technical quality (peer)	69.74	69.37	69.81
Total IP quality (peer)	39.53	39.59	39.53
Market attractiveness (peer)	54.44	53.39	51.79
# firms in corporate group	1.61	1.74	1.38

Notes: The table reports summary statistics for three mutually exclusive groups of Dutch firms: (1) *CO₂PL only* includes firms that have been certified under the CO₂PL at least once but have never won a public tender; (2) *Tender only* includes firms that have won at least one public tender but have never held CO₂PL certification; (3) *Tender + CO₂PL* includes firms that have both held CO₂PL certification and won at least one public tender.

3 Methodology

This section presents the empirical framework used to estimate the causal effect of winning a CO₂PL-applying public tender on firm-level innovation performance. In this study, treatment is defined as the year in which a firm first wins a public tender that applies CO₂PL criteria. Because firms win such tenders in different years, treatment timing is staggered. We exploit this variation using the difference-in-differences (DiD) estimator of Callaway and Sant'Anna (2021), which compares each cohort of newly treated firms with those that have not yet been treated in the same year. This approach provides unbiased treatment-effect estimates in settings with heterogeneous treatment timing.

A central identifying assumption is that the timing of a firm's first CO₂PL-applying contract win is plausibly exogenous to its underlying innovation trajectory. While firms may endogenously decide whether or when to obtain CO₂PL certification, the treatment event, winning a tender in which the Ladder is applied, is determined by competitive tenders rather than by firms' internal innovation cycles.

This interpretation is supported by institutional evidence from Chiappinelli et al. (2025), who show in the context of U.S. federal green procurement that firms cannot meaningfully influence or anticipate the exact timing of their first environmentally relevant contract award. Because tenders are evaluated on multiple criteria and subject to substantial competition, "the award is unpredictable, making it difficult for firms to perfectly anticipate winning specific contracts" (p. 25).

The Dutch TenderNed system operates under analogous conditions. Award decisions are based on price or price-quality scoring, project-specific requirements, and competition among bidders. Although certification may improve a firm's relative competitiveness, certification alone does not determine award outcomes, nor can firms strategically choose the timing of contract availability. Conditional on (i) bidding in CO₂PL-applying tenders, (ii) sectoral procurement cycles, and (iii) certification status, the year in which a firm first wins a CO₂PL-applying contract is therefore plausibly as-good-as-random with respect to its innovation trajectory.

Empirical diagnostics are consistent with this interpretation. Event-study estimates reveal no statistically detectable pre-treatment trends in technical patent quality, indicating that treated and not-yet-treated firms were on parallel innovation trajectories prior to first exposure. This strengthens the case for interpreting the timing of first CO₂PL tender wins as a quasi-random shock, thereby justifying the staggered DiD design and the use of not-yet-treated firms as the control group.

3.1 Treatment Definitions

Firms can experience their first CO₂PL-related procurement exposure under different certification conditions. Our main treatment specification is defined as:

1. winning a CO₂PL-applying tender while holding CO₂PL certification within 2 years after contract award.

To benchmark and interpret this main effect, we construct three comparison treatments that reflect alternative certification states at the moment a firm wins its first CO₂PL-applying tender:

2. winning a CO₂PL-applying tender without CO₂PL certification;
3. winning a CO₂PL-applying tender while certified under ISO 14001 only;
4. winning a CO₂PL-applying tender while holding both CO₂PL within 2 years after contract award and ISO 14001 certification in the year of award.

These comparison groups allow us to distinguish the innovation effects of CO₂PL-based procurement from those associated with general procurement participation or broader environmental management practices. In particular, they reveal whether CO₂PL certification is necessary, complementary, or irrelevant for innovation responses to CO₂PL applied procurement.

For each treatment definition, the treatment year is the first year in which the relevant condition is satisfied. Observations prior to are coded as untreated, and firms that never meet the condition remain in the not-yet-treated group throughout the panel. Each definition generates a distinct treatment cohort, enabling separate estimation of heterogeneous innovation effects across policy pathways.

3.2 Staggered Difference-in-Differences Framework

To estimate causal effects, we employ the estimator of Callaway and Sant'Anna (2021), which is specifically designed for settings with heterogeneous treatment timing and avoids the well-known biases of traditional two-way fixed effects (TWFE) models when treatment adoption is staggered. For each treatment definition, we estimate:

$$Y_{it} = \alpha_i + \lambda_t + X'_{it}\beta + \delta_{g,t}, \quad (1)$$

where Y_{it} denotes the innovation outcome (peer-normalised technical patent quality score). The term α_i captures firm fixed effects, while λ_t captures common year shocks. The vector X_{it} contains covariates that may affect both innovation outcomes and a firm's probability of winning a CO₂PL-applying tender.

A subset of these covariates is time-invariant. *Size class* (small, medium, large, very large) proxies for structural differences in organisational capacity, while the sector-specific R&D-intensity level is based on the OECD classification of industries by R&D intensity (measured as the ratio of R&D expenditure to value added) (Galindo-Rueda and Verger, 2016). Although these characteristics do not vary over time, they are included because the Callaway–Sant'Anna estimator uses them to refine comparisons between treated and untreated firms. Intuitively, the procedure relies on these variables twice: first, to adjust for systematic differences in innovation outcomes through the regression step, and second, to construct statistical weights that ensure treated firms are compared with “similar” not-yet-treated firms.

The time-varying controls include *firm age* and the *number of affiliated firms within the same ownership group*. These capture lifecycle dynamics and organisational complexity that may influence a firm's innovation capacity and investment behaviour.

The parameter $\delta_{g,t}$ is the group–time average treatment effect ($ATT_{g,t}$), comparing firms first treated in cohort g with firms that have not yet been treated in the same year t . Estimation uses the doubly-robust inverse-probability-weighted (DR-IPW) procedure. This approach ensures that treated firms are compared with genuinely comparable not-yet-treated firms, even when firms differ in scale, sector, age, or innovation intensity. Standard errors are clustered at the firm level to account for serial correlation. The DR-IPW approach provides additional robustness by yielding consistent estimates as long as either the outcome model or the treatment model is correctly specified.

3.3 Event-Study Specification

To analyse how firms' outcomes evolve before and after they first win a CO₂PL-applying contract, we use an event-study design. This approach compares firms' performance in the years surrounding their first exposure to CO₂PL-based procurement.

For each firm i , we define the treatment year G_i as the first year in which the firm wins a contract that applies the CO₂PL. We then measure time relative to this event as

$$k = t - G_i, \quad k \in [-5, 5].$$

where $k = 0$ corresponds to the year of first exposure, negative values indicate years before treatment, and positive values indicate years after treatment. This allows us to observe firms up to five years before and after their first CO₂PL-related contract.

For each treatment definition, we estimate following effects:

- the average effect of CO₂PL exposure on treated firms,
- year-by-year effects showing how outcomes change over time relative to the first treated year.

This design serves three main purposes. First, by comparing firms' trends before treatment, we assess whether treated and untreated firms followed similar paths prior to exposure. This provides evidence on whether later differences can plausibly be attributed to CO₂PL-based procurement rather than pre-existing trends. Second, the estimates around $k = 0$ capture the short-term impact of winning a first CO₂PL-applying contract. Third, the estimates for $k > 0$ trace how firms' responses develop over time, indicating whether potential innovation effects emerge gradually, persist, or fade.

More generally, the event-study approach can be interpreted as a "before-and-after" comparison that uses untreated firms as a reference group. These include firms that have never won a CO₂PL-applying contract, as well as firms that have not yet been exposed to CO₂PL-based procurement by a given year. Comparing treated firms to these otherwise similar never-treated and not-yet-treated firms allows us to examine not only whether CO₂PL exposure matters on average, but also when its effects begin and how long they last.

The choice of a $[-5, +5]$ event window reflects existing evidence on the timing of innovation responses to demand-side interventions. Prior studies show that such effects rarely materialise immediately. For instance, Bastianin et al. (2022) report that suppliers to CERN, the European Organization for Nuclear Research, begin to exhibit higher patenting activity only 5–8 years after receiving a contract, consistent with the lengthy development cycles associated with advanced technological work. OECD analyses similarly highlight that many procurement-induced innovations fall outside the standard three-year survey horizon, suggesting that short windows risk understating the true adjustment process (Appelt and Galindo-Rueda, 2016). In the context of procurement, Crespi and Guarascio (2019) find that innovation-oriented procurement raises sector-level patent stocks with a one-year lag, pointing to adjustments that begin relatively soon after exposure but may continue accumulating over time.

This body of evidence suggests that procurement-related innovation may start to appear within a few years but generally unfolds over a longer horizon. A $[-5, +5]$ window offers a reasonable com-

promise as it is broad enough to capture both the early and medium-run dynamics while remaining narrow enough to avoid confounding longer-term structural shifts.

3.4 Estimation sample

The estimation sample is restricted to Dutch firms that either (i) have ever won a public procurement contract or (ii) have ever obtained CO₂PL certification. This restriction reflects the relevant comparison set for studying the innovation effects of CO₂PL in procurement. Because the Ladder is a procurement instrument, firms that choose to certify regardless of whether they ultimately win a contract are natural participants in public procurement markets. In contrast, firms that neither certify nor appear in public procurement are unlikely to be exposed to CO₂PL-related incentives and therefore provide a poor counterfactual.

A second reason for this sampling choice is a data limitation. TenderNed records identify only the winners of procurement procedures, not the full bidder set. This means we cannot observe firms that participated and lost. By including firms that have ever certified with CO₂PL, we retain a group of potential bidders who are plausibly comparable to winning firms in terms of procurement interest and environmental engagement.

In the difference-in-differences framework, firms contribute to the not-yet-treated comparison group in all years prior to their treatment year G_i . Firms that never experience the corresponding treatment remain in the not-yet-treated group throughout the observation window. This ensures that treated firms are always compared to the appropriate set of eligible but not-yet-treated firms under staggered adoption.

4 Results

This section presents the estimated innovation effects of exposure to the CO₂PL in public procurement. We begin with the main specification, winning a CO₂PL-applying tender while holding CO₂PL certification, and then compare this to the alternative treatment definitions. We next analyse whether acquiring CO₂PL certification has any measurable effect on innovation outcomes. Then we explore broader green procurement channels, test alternative innovation indicators, and conclude with population-wide robustness checks.

4.1 Main Effects: CO₂PL Tender Wins with CO₂PL Certification

Table 4 reports the estimated treatment effects for firms whose first exposure to a CO₂PL-applied public contract occurs under different certification conditions. Column 1 presents the main specification, focusing on firms that win a CO₂PL-applying tender while already holding CO₂PL certification or while committing to obtain certification within the prescribed post-award compliance period

The simple ATT is positive (0.426) and marginally significant at the 10% level ($p \approx 0.096$), suggesting a positive but somewhat noisy average effect across all event years. The aggregated effect across the [-5,+5] window is more precise: the ATTC of 0.28 is statistically significant at the 5% level, a statistically significant and substantively meaningful improvement in firms' peer-normalised technical patent quality. The dynamic event-study profile (Figure 5) shows no pre-treatment differences and a steady post-treatment rise, suggesting that CO₂PL-based procurement triggers a gradual upgrading in firms' technological capabilities.

Although the technical quality measure spans a 0–100 scale, its empirical distribution is extremely compressed: most firms fall between 68 and 72, and within-firm year-to-year variation is typically under 0.1 points. In this context, estimated effects in the range of 0.3–0.7 points represent meaningful shifts—on the order of half to almost one standard deviation of the within-firm distribution. The magnitude of the estimated treatment effects therefore, reflects meaningful improvements in the technological depth of firms' innovation portfolios rather than minor numerical fluctuations.

4.2 Comparison Groups: Certification and Procurement Heterogeneity

Columns 2–4 of Table 4 compare the effects of a first CO₂PL-applied contract win across alternative certification states to examine whether environmental management capabilities shape the innovation response to CO₂PL-related procurement. The corresponding dynamic treatment profiles are illustrated in Figures 5.

Winning without CO₂PL certification. Firms that win a CO₂PL-applying tender without holding CO₂PL certification exhibit no statistically significant change in peer-normalised technical patent quality (ATTC = 0.160). These results imply that, in this subsample, there is no reliable evidence of an innovation response to winning a CO₂PL-applied contract in the absence of CO₂PL certification.

Winning while ISO 14001 certified only. Even though the ISO treatment sometimes shows larger point estimates, ISO-certified firms exhibit statistically insignificant effects (ATTC = 0.216). ISO certification provides general environmental-management routines but does not embed the procurement-linked carbon-performance incentives specific to the Ladder.

Winning while holding both CO₂PL and ISO. Firms holding both CO₂PL and ISO 14001 certifications exhibit the strongest and statistically significant effect (ATTC = 0.317). This pattern is consistent with the possibility that broader EMS adoption may be associated with a more pronounced innovation response when firms are exposed to CO₂PL-applied procurement.

Empirical evidence on multi-certification systems is limited. Some studies suggest that combining quality, environmental, and safety management standards can strengthen internal coordination and organisational routines, thereby creating conditions that may support innovation-related activities (Wiengarten et al., 2017). These findings, however, primarily concern perceived organisational performance rather than direct technological outcomes. In this light, our results should be viewed as suggestive. The data indicate that firms with multiple EMS certifications display stronger responses, but further research is needed to identify the mechanisms behind this pattern.

Across the comparison groups and within the five-year pre- and post-treatment window, statistically significant innovation responses are observed only for firms that either already hold CO₂PL certification at the time they win a CO₂PL-applied tender or obtain certification within the prescribed post-award compliance window. The largest point estimate arises for firms holding both CO₂PL and ISO 14001 certification, although the magnitude is similar to that for firms holding CO₂PL alone. This pattern underscores the central role of timely certification in shaping firms' innovation trajectories in the context of CO₂PL-based procurement.

Table 4: Average treatment effects on technical quality for CO₂PL-related tender wins

	PL win + PL cert.	PL win w/o cert.	PL win + ISO	PL win + both
Simple ATT	0.426 (0.26)	0.424 (0.38)	0.441 (0.34)	0.399 (0.28)
ATTC (-5, +5)	0.282* (0.14)	0.160 (0.18)	0.216 (0.15)	0.317* (0.15)
Firm fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
<i>N</i>	10,793	10,816	10,819	10,813

Notes: Standard errors clustered at the firm level are shown in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable is peer-normalised technical patent quality score. Columns correspond to four treatment definitions: (1) first CO₂PL-applying tender win while CO₂PL certified; (2) first CO₂PL-applying tender win without CO₂PL certification; (3) first CO₂PL-applying tender win while ISO 14001 certified only; (4) first CO₂PL-applying tender win while holding both CO₂PL and ISO 14001 certification. Event-study averages computed over the window [-5, +5].

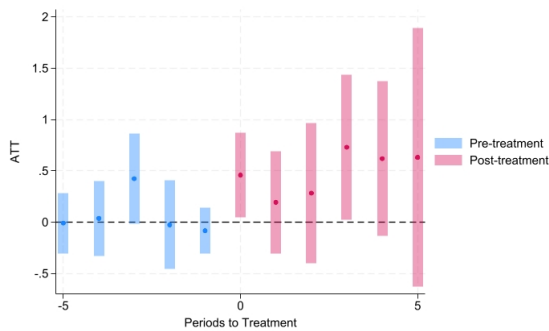


Figure 1: CO₂PL-applied win + CO₂PL

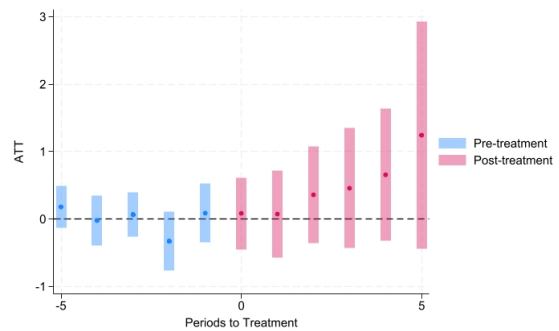


Figure 2: CO₂PL-applied win without CO₂PL

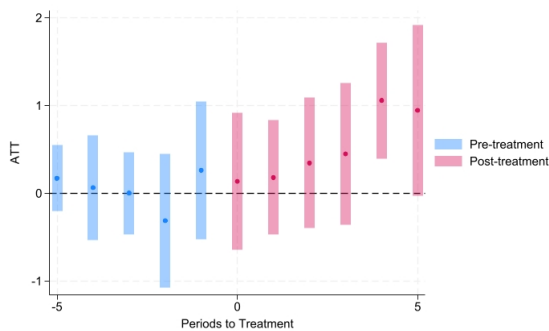


Figure 3: CO₂PL-applied win + ISO 14001

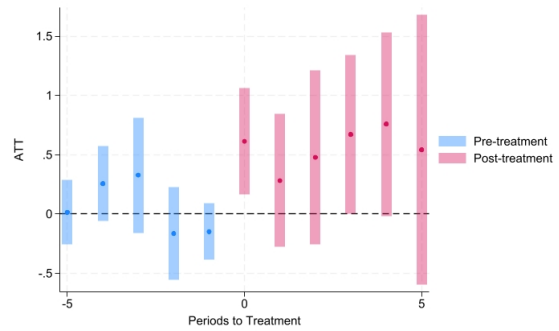


Figure 4: CO₂PL-applied win + both certifications

Figure 5: Dynamic treatment effects of first-time winning a CO₂PL-applied contract, by certification status

4.3 Heterogeneity by Contract Value in CO₂PL-Applied Procurement

The scale of procurement varies substantially across works, services, and supply contracts. A single price threshold is therefore inappropriate for identifying “small” or “large” tenders in CO₂PL-applied procurement. To address this, we classify CO₂PL-applied contracts using procurement-type-specific value bands¹:

- **Small:** works <€3 million; services <€1 million; supplies <€0.5 million,
- **Medium:** works €3–15 million; services €1–5 million; supplies €0.5–2 million,
- **Large:** works ≥€15 million; services ≥€5 million; supplies ≥€2 million.

For each group, we construct a staggered treatment indicator capturing the first year in which a firm wins a CO₂PL-applied contract while already holding CO₂PL certification.

The results reveal heterogeneity across contract-value categories. As reported in Table 5, small CO₂PL-applied contracts produce negligible effects on peer-normalised technical patent quality score (ATT \approx -0.03; ATTC \approx 0.01), and the corresponding event-study profile in Figure 13 shows no systematic post-treatment pattern. Medium-sized contracts exhibit somewhat larger point estimates (ATT \approx 0.51; ATTC \approx 0.30), but the event-study evidence in Figure 13 indicates violations of the parallel-trends assumption, preventing a causal interpretation.

In contrast, the large-contract group exhibits clear and robust treatment effects. As shown in Table 5, the simple ATT is approximately 1.03 ($p = 0.021$), and the event-study averages are sizable across both windows: ATTC = 0.51 ($p = 0.059$) for the standard $[-5, +5]$ window and ATTC = 0.56 ($p = 0.048$) when extending the post-treatment window to $[-5, +6]$. The dynamic estimates presented in Figure 13 display a pronounced post-treatment rise.

These results indicate that innovation responses to CO₂PL-applied procurement are concentrated exclusively among firms receiving large contracts. Smaller and medium-sized CO₂PL awards might not generate sufficient scale, complexity, or performance requirements to induce measurable improvements in technological capability. Only large tenders appear to generate conditions conducive to measurable technology-oriented innovation responses.

¹Contract-value groups are defined using procurement-type-specific percentiles of CO₂PL-applied contract prices. Works, services, and supplies exhibit highly skewed distributions (medians: €9.9M, €3.4M, and €1.25M, respectively), so thresholds are set relative to each type's empirical range: works <€3M / €3-15M / ≥€15M; services <€1M / €1-5M / ≥€5M; supplies <€0.5M / €0.5-2M / ≥€2M.

Table 5: Average treatment effects by size of CO₂PL-applied contract

	Small contracts	Medium contracts	Large contracts
Simple ATT	-0.027 (0.47)	0.515 (0.36)	1.031* (0.45)
ATTC (-5, +5)	0.011 (0.24)	0.302 (0.17)	0.506 (0.27)
Firm fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
<i>N</i>	10,824	10,815	10,700

Notes: Standard errors clustered at the firm level are shown in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable is peer-normalised technical patent quality score. Contract-value groups defined as: (1) *Small*: < 1 million; (2) *Medium*: €1–5 million; (3) *Large*: > 5 million. Estimates correspond to the first CO₂PL-applied tender wins among firms already holding CO₂PL certification. Event-study averages computed over the window $[-5, +5]$. A robustness specification using $[-5, +6]$ yields a statistically significant ATTC of 0.56 ($p = 0.048$) for large contracts.

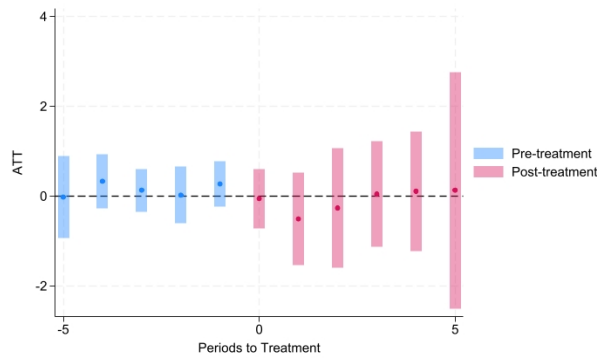


Figure 6: Small contracts

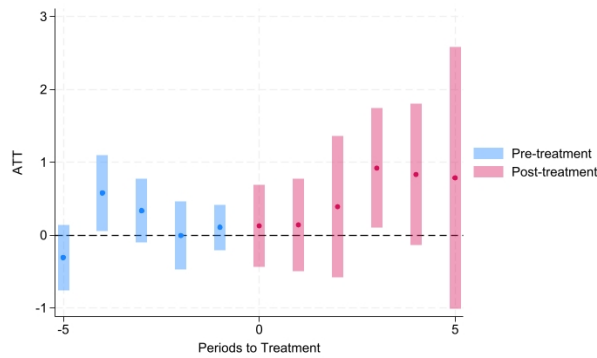


Figure 7: Medium contracts

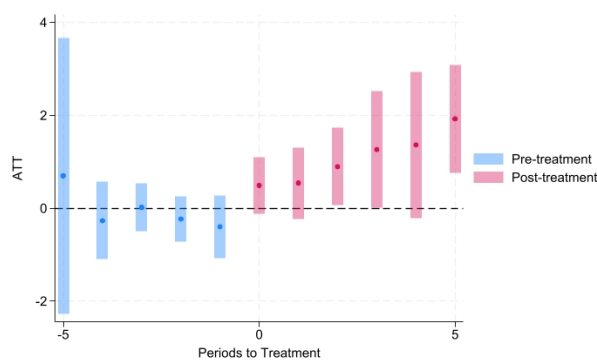


Figure 8: Large contracts

Figure 9: Dynamic treatment effects of CO₂PL-applied tender wins on peer-normalised technical patent quality score, by contract value

4.4 Heterogeneity by Sectoral R&D Intensity

To examine whether firms' technological environments moderate innovation responses to CO₂PL-applied procurement, we draw on the OECD sectoral R&D intensity taxonomy (Galindo-Rueda and

Verger, 2016). The OECD classification assigns industries to five R&D-intensity groups. Given the distribution of firms in our dataset, we aggregate these into three analytically meaningful categories: (i) low R&D intensity (OECD level 1), (ii) medium R&D intensity (levels 2–3), and (iii) high R&D intensity (levels 4–5). For each category, we construct a staggered treatment indicator capturing the first year in which a firm wins a CO₂PL-applied tender while already holding CO₂PL certification.

We found heterogeneity across R&D-intensity groups. Firms operating in *low R&D-intensity industries* exhibit no detectable innovation response: the ATT and ATTC estimates are small and statistically insignificant, and dynamic coefficients show no systematic post-treatment pattern. For *medium R&D-intensity industries*, point estimates are somewhat larger but remain imprecisely estimated; the event-study diagnostics indicate a violation of parallel trends at $t = -1$, limiting interpretation.

In contrast, firms in *high R&D-intensity industries* exhibit large and statistically significant improvements in technical patent quality following a CO₂PL-applied tender win. The simple ATT estimate is 1.21 ($p = 0.007$), and the ATTC over the $[-5, +5]$ window is 0.72 ($p = 0.022$). Dynamic estimates indicate sustained post-treatment gains beginning around $t + 2$, consistent with a gradual innovation response.

These patterns suggest that the effectiveness of CO₂PL-applied procurement in stimulating technological upgrading depends critically on firms' absorptive capacity and baseline R&D intensity. High-R&D sectors appear uniquely positioned to translate procurement-driven CO₂ requirements into measurable improvements in innovation outcomes.

Table 6: Average treatment effects by sectoral R&D intensity

	Low R&D	Medium R&D	High R&D
Simple ATT	0.215	0.513	1.213**
	(1.01)	(0.77)	(2.71)
ATTC (-5, +5)	0.103	0.438	0.724*
	(0.95)	(1.34)	(2.30)
Firm fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
<i>N</i>	10,823	10,660	10,855

Notes: Standard errors clustered at the firm level are shown in parentheses. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. The dependent variable is peer-normalised technical patent quality score. R&D-intensity groups follow OECD taxonomy (Galindo-Rueda and Verger, 2016), aggregated into low (1), medium (2–3), and high (4–5). Treatment corresponds to the first CO₂PL-applied tender win for firms already holding CO₂PL certification. Event-study averages computed over the window $[-5, +5]$.

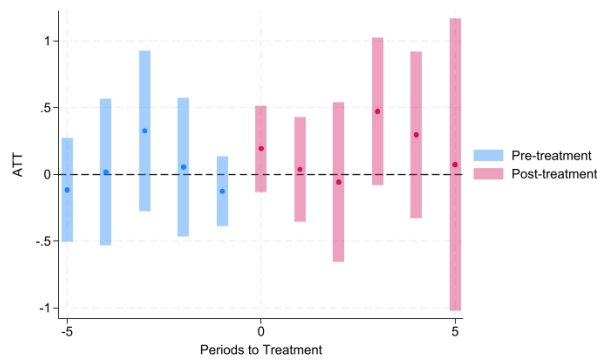


Figure 10: Low R&D intensity sector

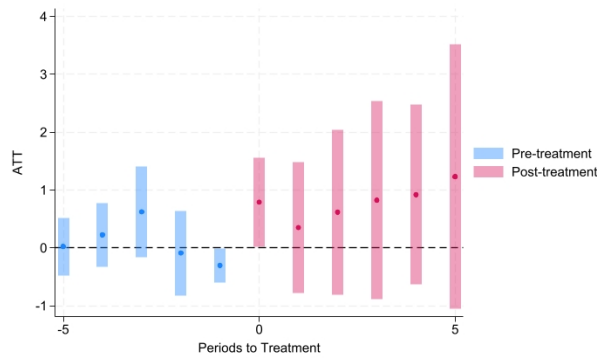


Figure 11: Medium R&D intensity sector

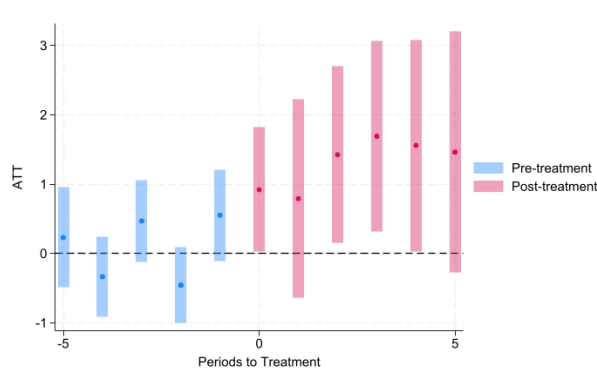


Figure 12: High R&D intensity sector

Figure 13: Dynamic treatment effects of CO₂PL-applied tender wins on peer-normalised technical patent quality score, by R&D intensity of sector

4.5 Additional Heterogeneity Analyses

In addition to heterogeneity by contract value and sectoral R&D intensity, we conducted two supplementary exercises: (i) heterogeneity by firm size, and (ii) heterogeneity by procurement type

(works, services, and supplies). These analyses are included for completeness but are not central to the interpretation of our main results. First, classifying firms into SMEs, large firms, and very large firms using Orbis size groups yields no statistically significant treatment effects in any category (Appendix 1.1). Both the ATT and ATTC estimates remain small and imprecise, and the event-study coefficients show no systematic post-treatment response, suggesting that firm size does not meaningfully mediate the innovation effects of CO₂PL-applied procurement.

Second, disaggregating tenders by procurement type reveals a pattern in which work-related procurement may deliver larger innovation responses than other categories. Works contracts alone produce marginally significant improvements in technical patent quality (ATT = 0.64, $p = 0.087$; ATTC = 0.40, $p = 0.068$), and the pooled works + services category exhibits a borderline significant ATTC of 0.28 ($p = 0.082$). The strongest evidence arises from the works + supplies grouping, which generates statistically significant and robust effects across both ATT (0.63, $p = 0.048$) and ATTC (0.39, $p = 0.026$). Taken together, these patterns tentatively suggest that works-related procurement may provide conditions more conducive to technology-oriented innovation. However, because these results are not uniformly robust across all sectoral specifications, they should be interpreted as indicative rather than definitive. Full estimates are reported in Appendix 1.2.

5 Robustness Checks and Alternative Specifications

This section assesses the robustness of the main findings using a placebo test, alternative outcome measures, and a population-wide comparison. Across all exercises, the core result remains: measurable innovation improvements arise only when firms win a CO₂PL-applied public contract while already holding CO₂PL certification. No other procurement channel or innovation measure produces comparable evidence.

5.1 Placebo Test: Green Public Procurement Without CO₂PL Application

We first examine whether general green public procurement (GPP), identified through environmental terminology in tender descriptions, produces innovation effects comparable to CO₂PL-applied procurement. Unlike the Ladder, GPP tenders do not impose structured carbon-management requirements.

To ensure comparability with our main analyses, we estimate four treatment definitions based on firms' certification status at the time of their first GPP-labelled tender win. Across most specifications, both the simple ATT and the aggregated ATTC are small in magnitude and statistically

indistinguishable from zero (Appendix 1.3).

One exception is the case in which a firm wins a GPP-labelled contract while already CO₂PL certified: the ATTC estimate of 0.23 ($p = 0.054$) is close to conventional significance thresholds. This pattern may tentatively suggest that certified firms can leverage their carbon-management capabilities when responding to environmentally framed tenders, even when CO₂PL is not explicitly applied. However, because the corresponding ATT estimate remains imprecise and the result is not replicated across other certification categories, this finding should be interpreted cautiously.

Overall, the evidence indicates that participation in generic GPP does not meaningfully drive the innovation effects documented for CO₂PL-applied procurement. The structured and performance-based design of the Ladder appears to be the mechanism most directly associated with the observed improvements in technical patent quality.

5.2 Alternative Innovation Metrics

To verify that the main results are not driven by the choice of innovation indicator, we re-estimate the event-study specification using three alternative outcomes: (i) annual patent counts, (ii) peer-normalised IP quality scores, and (iii) peer-normalised market attractiveness scores.

Across all measures, point estimates are imprecisely estimated and statistically insignificant (Appendix 1.4). Patent counts show no systematic post-treatment change, and both IP quality and market attractiveness fluctuate around zero without discernible trends. The absence of effects across these wider innovation metrics suggests that CO₂PL-applied procurement does not increase patenting volume or commercial IP valuation; instead, the effects are concentrated in technical patent quality, which captures the technological depth of firms' inventive output.

5.3 Population-Level Test: Full Dutch Firm Sample

Finally, we assess whether the innovation effects generalise to the broader population of Dutch firms with IP coverage in Orbis (approximately 200,000 firms per year). Up to this point, our analysis has been confined to Dutch firms that have either won at least one public procurement tender or have ever held CO₂PL certification, as these firms constitute the relevant pool of potential participants in public procurement and CO₂PL-related processes. The broader sample includes many firms that never participate in public procurement and therefore do not face CO₂PL-related incentives. For this reason, population-wide estimates do not constitute a causal benchmark but serve as a diagnostic check.

Across the four treatment definitions—CO₂PL-certified, uncertified, ISO-only, and jointly CO₂PL- and ISO-certified—most ATT and ATTC estimates are small and statistically insignificant (Appendix 1.5). The exception is the group of firms holding both CO₂PL and ISO 14001 certification at the time of their first tender win involving CO₂PL criteria. For this group, the aggregated event-study effect is statistically significant (ATTC = 0.33, $p = 0.032$), indicating that firms with dual environmental management systems may be better positioned to convert procurement-related requirements into improvements in technical patent quality.

However, because no comparable effects appear in the other certification categories, and because most firms in the population sample do not supply the public sector, this result should be interpreted cautiously. Overall, the evidence indicates that the innovation response to CO₂PL-applied procurement is concentrated among procurement-active firms rather than the broader firm population.

Overall, the robustness checks reinforce the core conclusion: Innovation responses arise specifically from CO₂PL-applied public procurement and depend on firms' existing CO₂PL certification status. General green procurement, alternative IP measures, and population-wide comparisons do not produce comparable effects.

6 Discussion and Conclusion

This report provides the comprehensive empirical assessment of whether the application of CO₂PL in Dutch public procurement promotes measurable improvements in firms' innovation quality. By combining detailed procurement records with firm-level certification data and patent-based innovation indicators, and by exploiting the quasi-random timing of firms' first exposure to CO₂PL-applied tenders, the analysis evaluates how procurement-based carbon-management incentives translate into changes in technological performance.

Three main findings emerge.

First, CO₂PL-applied procurement is associated with statistically significant improvements in technical patent quality, but only among firms that are certified under the Ladder at the time of their first CO₂PL-applied contract or that obtain certification within two years of contract award. Event-study estimates reveal no evidence of pre-treatment trends and indicate sustained post-treatment gains in peer-normalised technical quality, supporting a causal interpretation. In contrast, firms winning CO₂PL-applied tenders without certification, or holding only ISO 14001 certification, do not exhibit

comparable responses. These results suggest that innovation effects are driven by the structured carbon-management routines embedded in CO₂PL certification rather than by procurement participation alone or by generic environmental management systems.

Second, innovation responses are highly heterogeneous. Substantial gains are concentrated in two settings. First, only firms winning large, high-value CO₂PL-applied contracts exhibit robust improvements in technical patent quality, whereas small and medium-sized contracts generate no measurable effects. This highlights the importance of contract scale in inducing meaningful technological adjustments. Second, firms operating in R&D-intensive sectors experience the strongest post-treatment gains, while firms in low- and medium-intensity sectors show no systematic improvement. Supplementary analyses indicate that firm size plays only a limited role and that differences across procurement categories (works, services, supplies) are comparatively weak. Together, these results imply that innovation responses are primarily shaped by project scale and sectoral technological capacity rather than by organisational size or administrative procurement classifications.

Third, the observed innovation effects appear specific to the CO₂PL framework. Placebo tests using GPP-labelled tenders yield no consistent evidence of innovation gains, and alternative innovation indicators—such as patent counts, overall IP quality, and market attractiveness—show no systematic responses. A population-wide analysis of Dutch IP-active firms further indicates that effects do not generalise beyond procurement participants, with only modest responses among firms holding both CO₂PL and ISO certifications. These findings reinforce the view that the Ladder's structured, performance-based design constitutes the central mechanism underpinning the observed innovation improvements.

Taken together, the evidence indicates that CO₂PL-based procurement can stimulate technology-oriented innovation, but in a selective rather than universal manner. Innovation effects arise only when procurement explicitly applies CO₂PL criteria, firms comply with certification requirements, contract values are sufficiently large to justify substantial technical investment, and firms operate in technologically intensive sectors.

Several limitations should be acknowledged. First, the analysis relies on firm-level certification data and does not incorporate project declarations, which limits insight into contract-level implementation. Second, innovation is measured using patent-based indicators, which may not fully capture process or organisational innovation. Third, this study has not yet undergone formal peer review and should therefore be interpreted as providing preliminary evidence. Accordingly, the results should be viewed as indicating a robust association between CO₂PL-linked procurement exposure and innovation among certified firms, rather than as definitive proof of universal policy effects.

From a policy perspective, the findings suggest that procurement strategies that create credible and timely incentives for CO₂PL certification are more likely to generate innovation benefits. Strengthening the integration of CO₂PL criteria in high-value tenders, reinforcing post-award compliance mechanisms, and complementing procurement incentives with targeted innovation-support instruments in lower-R&D sectors may further enhance the Ladder's contribution to the Dutch sustainability transition.

Bibliography

- Andries, P. and Stephan, U. (2019). Environmental innovation and firm performance: How firm size and motives matter. *Sustainability*, 11(13):3585.
- Appelt, S. and Galindo-Rueda, F. (2016). Measuring the link between public procurement and innovation.
- Bastianin, A., Castelnovo, P., Florio, M., and Giunta, A. (2022). Big science and innovation: gestation lag from procurement to patents for cern suppliers. *The Journal of Technology Transfer*, 47(2):531–555.
- Blind, K., Pohlisch, J., and Rainville, A. (2020). Innovation and standardization as drivers of companies' success in public procurement: an empirical analysis. *The Journal of Technology Transfer*, 45(3):664–693.
- Boot Advocaten (2015). Guideline for tenders for energy performance contracts. Technical report, Netherlands Enterprise Agency (RVO), The Hague, Netherlands. English edition of the Leidraad Prestatiecontracten.
- Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.
- Camilleri, M. A. (2022). The rationale for iso 14001 certification: A systematic review and a cost-benefit analysis. *Corporate Social Responsibility and Environmental Management*, 29(4):1067–1083.
- Chiappinelli, O., Dalò, A., Giuffrida, L. M., and Titl, V. (2025). The greener, the better? evidence from government contractors. Technical report, CESifo Working Paper.
- Crespi, F. and Guarascio, D. (2019). The demand-pull effect of public procurement on innovation and industrial renewal. *Industrial and Corporate Change*, 28(4):793–815.
- Elegbede, O., Schaefer, E., Gutierrez, F., and Metri, S. (2022). Foreign direct investment and intellectual property innovation in the united states: A profile of global and u.s. innovators. Technical report, SelectUSA, U.S. Department of Commerce.
- European Commission (2016). Buying green! a handbook on green public procurement (3rd ed.). https://sustainable-procurement.org/fileadmin/user_upload/layout/Documents/Buying-Green-Handbook-3rd-Edition.pdf. Accessed Month Day, Year.
- European Commission (2025). Sustainability and circularity as a starting point for ict procurement. <https://green-forum.ec.europa.eu/green-business/green-public-procurement/good-practice-library/>

sustainability-and-circularity-starting-point-ict-procurement_en. Good Practice Library – Green Business, retrieved 28 Nov 2025.

Galindo-Rueda, F. and Verger, F. (2016). Oecd taxonomy of economic activities based on r&d intensity.

Hizarci-Payne, A. K., İpek, İ., and Kurt Gümüş, G. (2021). How environmental innovation influences firm performance: A meta-analytic review. *Business Strategy and the Environment*, 30(2):1174–1190.

Rennings, K. and Rammer, C. (2011). The impact of regulation-driven environmental innovation on innovation success and firm performance. *Industry and innovation*, 18(03):255–283.

SKAO (2024). *CO₂ Performance Ladder: Handbook 3.1 (English Version)*. Stichting Klimaatvriendelijk Aanbesteden & Ondernemen (SKAO), Utrecht, The Netherlands. Version 3.1, published December 2024.

Uyarra, E., Edler, J., Garcia-Estevez, J., Georghiou, L., and Yeow, J. (2014). Barriers to innovation through public procurement: A supplier perspective. *Technovation*, 34(10):631–645.

Uyarra, E. and Flanagan, K. (2010). Understanding the innovation impacts of public procurement. *European planning studies*, 18(1):123–143.

Wagner, M. (2007). On the relationship between environmental management, environmental innovation and patenting: Evidence from german manufacturing firms. *Research policy*, 36(10):1587–1602.

Wiengarten, F., Humphreys, P., Onofrei, G., and Fynes, B. (2017). The adoption of multiple certification standards: perceived performance implications of quality, environmental and health & safety certifications. *Production Planning & Control*, 28(2):131–141.

Ziegler, A. and Nogareda, J. S. (2009). Environmental management systems and technological environmental innovations: Exploring the causal relationship. *Research Policy*, 38(5):885–893.

1 Appendix

1.1 Heterogeneity by Firm Size

Table 7: Average treatment effects by firm size

	Small/Medium firms	Large firms	Very large firms
Simple ATT	0.642 (0.45)	0.195 (0.43)	0.279 (0.41)
ATTC (-5, +5)	0.292 (0.23)	0.289 (0.25)	0.215 (0.22)
Firm fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
<i>N</i>	10,813	10,718	10,718

Notes: Standard errors clustered at the firm level are shown in parentheses. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. The dependent variable is the peer-normalised technical patent quality score. Firm-size categories based on orbis' classification: small/medium, large, and very large. Treatment corresponds to the first CO₂PL-applied tender win when the firm holds CO₂PL certification. Event-study averages computed over $[-5, +5]$. None of the estimates are statistically significant, indicating no meaningful heterogeneity by firm size.

1.2 Heterogeneity by Tender Type

Table 8: Average treatment effects across sector-specific CO₂PL-applied tender definitions

	Works	Serv.	Supp.	Works+Serv.	Serv.+Supp.	Works+Supp.
Simple ATT	0.641	0.278	1.018	0.410	0.321	0.625**
	(0.36)	(0.46)	(1.20)	(0.28)	(0.39)	(0.32)
ATTC (-5, +5)	0.397	0.223	0.551	0.275	0.235	0.390**
	(0.22)	(0.23)	(0.56)	(0.16)	(0.19)	(0.18)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	10,807	10,813	10,835	10,791	10,818	10,805

Notes: Standard errors clustered at the firm level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is peer-normalised technical patent quality score. Estimates correspond to first CO₂PL-applied tender wins while already CO₂PL certified. Event-study averages computed over the window $[-5, +5]$.

1.3 Robustness: Innovation Effects of GPP-Labelled Tender Wins

Table 9: Average treatment effects on technical quality for GPP-related tender wins

	GPP win + PL cert.	GPP win w/o PL cert.	GPP win + ISO	GPP win + both
Simple ATT	0.240 (0.42)	0.244 (0.26)	0.417 (0.26)	0.284 (0.64)
ATTC (-5, +5)	0.227 (0.12)	0.103 (0.11)	0.139 (0.11)	0.286 (0.18)
Firm fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
<i>N</i>	10,754	10,765	10,810	10,796

Notes: Standard errors clustered at the firm level are shown in parentheses. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. The dependent variable is peer-normalised technical patent quality score. Columns correspond to four treatment definitions: (1) first GPP-labelled tender win while CO₂PL certified; (2) first GPP-labelled tender win without CO₂PL certification; (3) first GPP-labelled tender win while ISO 14001 certified only; (4) first GPP-labelled tender win while holding both CO₂PL and ISO 14001 certification. Event-study averages computed over the window $[-5, +5]$.

1.4 Robustness: Alternative Innovation Outcome Measures

Table 10: CO₂PL-applied tender wins and alternative innovation indicators

	Yearly patent count	IP quality	Market attractiveness
Simple ATT	8.682 (8.368)	0.250 (0.424)	-0.531 (3.012)
ATTC (-5,+5)	3.127 (3.563)	0.287 (0.215)	0.777 (1.536)
Firm fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
<i>N</i>	10,843	10,793	10,793

Notes: Standard errors clustered at the firm level are shown in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Event-study averages computed over window $[-5, +5]$. Estimates correspond to first CO₂PL-applied tender wins while already CO₂PL certified. The dependent variable in column (1) is the yearly patent count. Column (2) uses the IP quality score, normalised relative to firms' peer groups. Column (3) uses the peer-group-normalised market attractiveness score.

1.5 CO₂PL-Related Tender Wins in the Full Dutch Firm Sample

Table 11: Average treatment effects on technical quality for CO₂PL-related tender wins (all Dutch firms)

	PL win + PL cert.	PL win w/o cert.	PL win + ISO	PL win + both
Simple ATT	0.438 (0.33)	0.483 (0.36)	0.548 (0.43)	0.433 (0.27)
ATTC (-5, +5)	0.283 (0.29)	0.187 (0.16)	0.256 (0.20)	0.328* (0.15)
Firm fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
<i>N</i>	201,173	200,905	200,882	201,087

Notes: Standard errors clustered at the firm level are shown in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable is peer-normalised technical patent quality score. Columns compare alternative treatment definitions: (1) first CO₂PL-applying tender win while CO₂PL certified; (2) first CO₂PL-applying tender win without CO₂PL certification; (3) first CO₂PL-applying tender win while ISO 14001 certified only; (4) first CO₂PL-applying tender win while holding *both* CO₂PL and ISO 14001 certification. Event-study averages computed over the window [-5, +5].

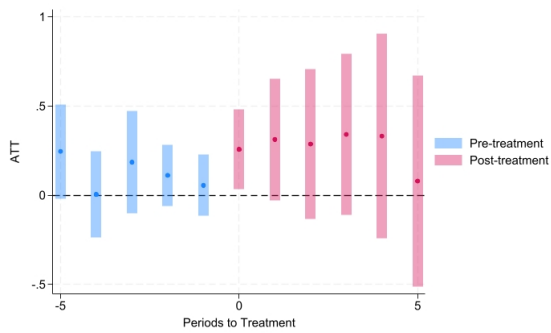


Figure 14: CO₂PL-applied win + CO₂PL (all IP firms)

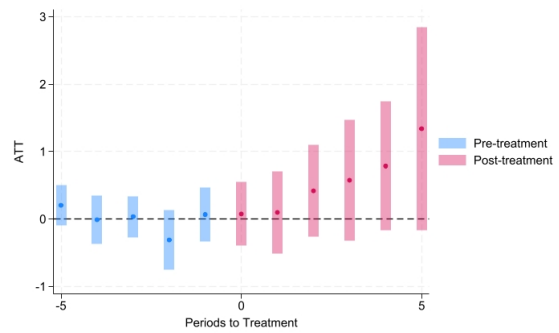


Figure 15: CO₂PL-applied win without CO₂PL (all IP firms)

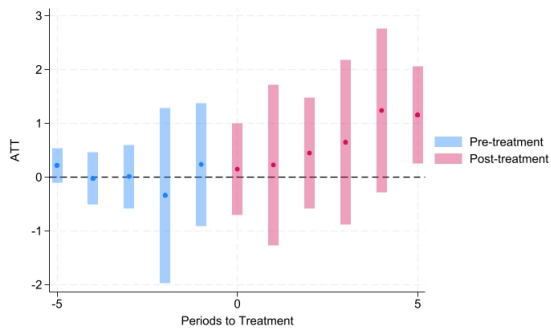


Figure 16: CO₂PL-applied win + ISO 14001 only (all IP firms)

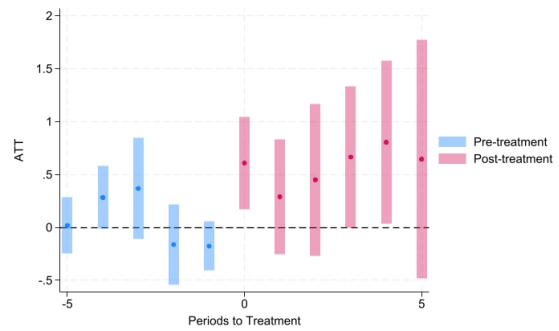


Figure 17: CO₂PL-applied win + both certifications (all IP firms)

Figure 18: Dynamic treatment effects of winning CO₂PL-applied contracts under alternative certification states (full IP firm population)