# The Impact of Cross-Border E-Commerce Pilot Zones on Regional Economic Growth: Evidence from the Yangtze and Pearl River Deltas, 2010-2023

**Qiong Song*., Chukiat Chaiboonsri., Anuphak Saosaovaphak**

**Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand**

***Corresponding Author**

**DOI: https://doi.org/10.47772/IJRISS.2026.100800490**

**Received: 22 August 2026; Accepted: 27 August 2026; Published: 09 September 2026**

## ABSTRACT

This study evaluates whether approval of China's Cross-Border E-Commerce (CBEC) comprehensive pilot zones changes city-level economic growth. The analysis uses a balanced panel of 39 cities in the Yangtze River Delta and Pearl River Delta from 2010 to 2023, comprising 546 city-year observations, 36 treated cities and three cities never treated by 2023. Because approval occurred in several cohorts, the main specification applies the Callaway-Sant'Anna group-time average treatment effect estimator with not-yet-treated controls and city-clustered standard errors. A joint test of event-time coefficients from five to two years before approval does not reject parallel trends (Wald = 6.540, p = 0.162). The estimated average treatment effect on treated cities is -0.0912 log points (SE = 0.0333, p = 0.006; 95% CI [-0.1565, -0.0259]), equivalent to approximately an 8.72% reduction in nominal GDP per capita. Effects are negative and significant in the approval year and the following year, providing reduced-form support for the revised short-run cost hypothesis: compliance, technological adaptation and resource-reallocation costs may temporarily outweigh initial gains. This interpretation is theoretical rather than a directly observed mediator. The estimate remains negative when Shanghai and Shenzhen are excluded, when the sample begins in 2014, and when 2020 is omitted, but it is imprecise when only the three never-treated cities are used as controls; a 500-draw cohort-randomisation placebo gives an empirical p-value of 0.094. On a common 2010-2021 horizon, both regional estimates are negative, while their direct difference is not significant (p = 0.161). An exploratory logistics analysis finds that policy exposure raises per-capita express-delivery revenue, but the activity-to-growth path and bootstrapped indirect effect are insignificant. The evidence therefore supports a conditional short- to medium-run negative effect and calls for longer evaluation horizons, transitional support and direct measurement of firm and worker adjustment.

**Keywords:** Cross-border e-commerce, Comprehensive pilot zones, Regional economic growth, Staggered difference-in-differences, Adjustment costs

## INTRODUCTION

Cross-border e-commerce (CBEC) combines digital market matching with physical trade fulfilment. Online platforms can reduce search and information costs, yet transactions still require payment settlement, customs processing, warehousing and delivery. The economic effect of CBEC therefore depends on the interaction between digital access and local institutions, infrastructure and firm capabilities (Freund & Weinhold, 2004; Gomez-Herrera et al., 2014; Hummels & Schaur, 2013; Terzi, 2011). China offers a particularly useful setting because the State Council introduced CBEC comprehensive pilot zones in successive city cohorts from 2015 onward. The policy package bundles customs, taxation, foreign-exchange and logistics arrangements, so its city-level effect is a reduced-form policy effect rather than the effect of one isolated instrument.

The scale of the sector makes the question economically important. The China Academy of Information and Communications Technology (CAICT, 2024) reports CBEC imports and exports of 2.38 trillion yuan in 2023, equal to 5.7% of national foreign trade. Aggregate expansion, however, does not show whether treated cities grew faster than they would have without designation. A pilot zone may improve market access and digital capability, but implementation can also require new compliance routines, information systems, staff training and resource reallocation. Such costs may be concentrated around approval and may coexist with benefits that emerge only later. This timing is consistent with a Schumpeterian adjustment view in which entry, exit and reallocation temporarily disrupt existing production before any productivity gains materialise (Aghion & Howitt, 1992; Schumpeter, 1942).

Existing empirical results are mixed. Some studies report positive effects on city growth, resident income, exports or industrial upgrading (Cai et al., 2025; Li et al., 2023; Zhong et al., 2022), whereas other work emphasises export creation and reallocation effects (Cheng et al., 2025). Comparisons are complicated by staggered treatment timing. Conventional two-way fixed-effects difference-in-differences (TWFE) can combine cohort effects with negative weights when treatment effects are heterogeneous (Baker et al., 2022; Goodman-Bacon, 2021). The interpretation of a single coefficient can also depend on whether already-treated cities are used as controls.

This study addresses these issues with a balanced panel of 39 cities from 2010 to 2023: 30 in the Yangtze River Delta (YRD) and nine in the Pearl River Delta (PRD). Thirty-six cities entered the policy in the 2015, 2016, 2018, 2019, 2020 or 2022 cohorts, while three remained untreated. The main estimator is the Callaway-Sant'Anna (2021) group-time ATT with not-yet-treated controls and city-clustered standard errors. The design reports a joint pre-trend test, event-time support, alternative comparison groups, a randomised cohort-assignment placebo and a direct regional contrast. Logistics is studied only as an exploratory activity channel because the available measure, per-capita express-delivery revenue, does not observe clearance speed, reliability, route coverage or unit cost.

The study makes three contributions. First, it provides a cohort-aware estimate of the aggregate city-level effect and explicitly separates treatment-effect heterogeneity from heteroskedasticity-robust inference. Second, it documents a negative near-term reduced-form pattern while reporting the comparison-group and event-time limitations that qualify its interpretation. Third, it narrows mechanism claims: the revised H2 is evaluated using early event-time effects, while logistics mediation is treated as supplementary and is not claimed when the indirect effect is statistically indistinguishable from zero. The remainder of the paper reviews the relevant theory and evidence, describes the research design, presents results and discusses implications and limitations.

## LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

### CBEC, market access and regional growth

Digital trade can lower search costs, make small firms more visible to foreign buyers and shorten the distance between producers and consumers. Internet use has been associated with international trade expansion, but the effect depends on complementary transport and institutional systems (Freund & Weinhold, 2004; Gomez-Herrera et al., 2014). At the regional level, agglomeration theories predict that shared suppliers, labour pools and knowledge spillovers can reinforce activity in dense places (Marshall, 1890; Duranton & Puga, 2004; Rosenthal & Strange, 2004). New economic geography similarly emphasises increasing returns and market-access feedbacks (Krugman, 1991; Fujita et al., 1999). These perspectives support a positive policy effect when a pilot zone coordinates digital infrastructure, customs procedures and logistics.

Chinese evidence is broadly consistent with several beneficial margins. Zhong et al. (2022) report positive city-level growth effects for pilot zones, and Cai et al. (2025) find higher resident income. Other studies identify industrial upgrading, financing relief, export creation or entrepreneurial responses as possible channels (Li et al., 2023; Wang et al., 2025). Yet a comprehensive pilot zone is a bundled place-based intervention. Export creation may be accompanied by reallocation from incumbent firms or neighbouring cities (Cheng et al., 2025), and logistics turnover may rise without a measurable improvement in efficiency. A city-level GDP outcome is therefore a net aggregate outcome, not a direct measure of any individual mechanism.

### Adjustment costs and creative destruction

The growth literature also permits a negative short-run effect. New systems require firms and public agencies to connect customs, tax, payment and inventory processes; workers must acquire new skills; and capital may move from established activities toward uncertain digital ventures. The resulting adjustment can reduce measured output during the transition. Creative destruction theory describes growth as a process in which new technologies and firms displace older arrangements, with the timing of costs and gains depending on reallocation frictions (Aghion & Howitt, 1992; Melitz, 2003; Schumpeter, 1942). Evidence on trade adjustment similarly shows that local gains can be accompanied by worker and sectoral disruption (Autor et al., 2014). This mechanism is plausible for CBEC policy, but the city-level data do not directly measure compliance spending, firm exits, worker transitions or future productivity.

**Identification in staggered policy settings**

The timing of pilot-zone approval varies by city. When effects differ across cohorts or event time, TWFE may use earlier-treated units as controls for later-treated units and produce a weighted average that is difficult to interpret (Baker et al., 2022; de Chaisemartin & D'Haultfœuille, 2020; Goodman-Bacon, 2021). Callaway and Sant'Anna (2021) instead estimate group-time effects using valid untreated or not-yet-treated comparisons and then aggregate them. Sun and Abraham (2021) provide a related event-study approach that avoids contamination from heterogeneous treatment effects. These developments motivate the estimator and diagnostics used here. These developments motivate the estimator and diagnostics used here. For a comprehensive review of the evolving DiD literature, see Baker, Callaway, Cunningham, Goodman-Bacon, and Sant'Anna (2026).

**Hypotheses**

The first hypothesis is deliberately two-sided because the policy bundles several instruments and prior evidence is mixed. H1: CBEC comprehensive pilot-zone approval has a statistically significant effect on log nominal GDP per capita in treated cities relative to eligible comparison cities; the sign is determined empirically.

The revised short-run hypothesis gives a directional prediction. H2: In the short run, CBEC pilot-zone approval has a negative effect on city-level economic growth because policy-related compliance, technological adaptation and resource-reallocation costs may temporarily outweigh the initial benefits. The logistics-activity path is a separate supplementary exploration and is not used to test H2; H2 is evaluated using the early post-approval event-time ATT estimates. H2 is a reduced-form prediction about early post-approval outcomes. It does not claim that the cost construct is observed as a separately identified mediator.

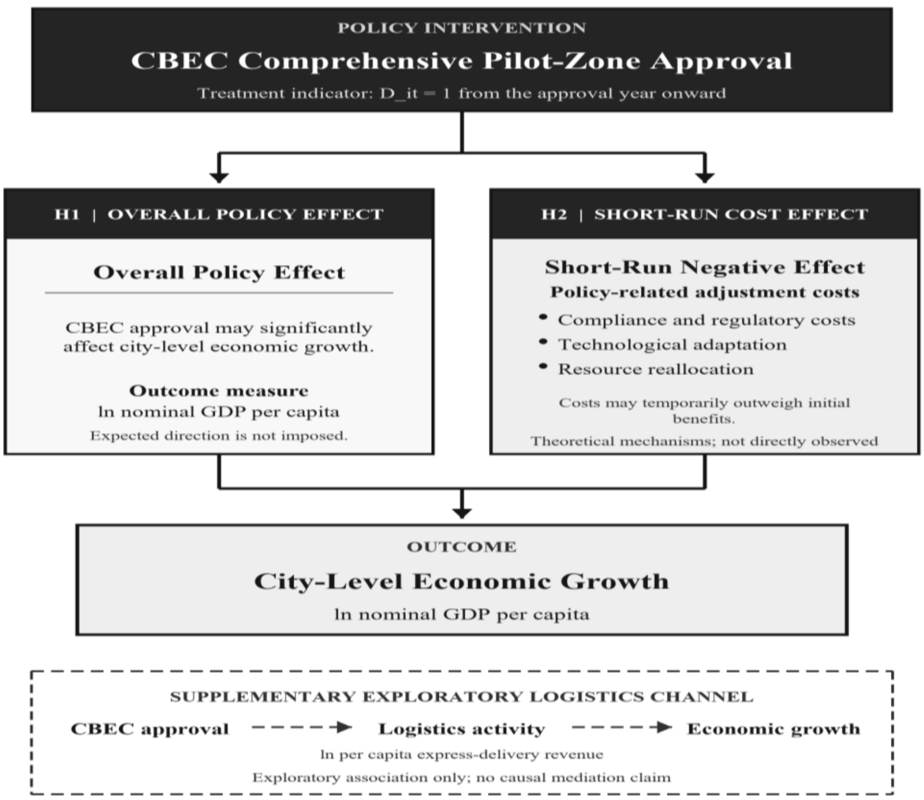


Figure 1: Conceptual framework linking pilot-zone approval, adjustment costs and outcomes

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

*Note: Solid arrows denote estimated causal paths. Dashed arrows denote theoretical channels not directly observed in city-level data. Dotted arrows denote supplementary exploratory analyses that do not establish causal mediation. H1 is evaluated using the main group-time ATT; H2 is evaluated using the approval-year and following-year event-time ATT estimates. Regional heterogeneity is assessed via a direct bootstrap test of the YRD-minus-PRD difference.*

# METHODOLOGY

### Study setting, sample and data

The main sample is a balanced panel of 39 prefecture-level cities observed annually from 2010 to 2023. It contains 30 YRD cities and nine PRD cities. Policy cohorts are defined by official approval year: 2015, 2016, 2018, 2019, 2020 and 2022. Three YRD cities were never treated by the end of 2023. The outcome is the natural logarithm of nominal GDP per capita in current yuan. We use nominal GDP per capita for three reasons. First, a consistent city-level GDP deflator is not available for all cities across the 2010–2023 panel. Second, the DID estimator differences out common time shocks, including common price trends, so to the extent that price changes are similar across cities, they do not bias the estimated treatment effect. Third, the policy effect on nominal income is itself policy-relevant, particularly for local fiscal capacity and tax revenues. Nevertheless, the use of nominal GDP means that the estimated effect may partly reflect price movements; this limitation is addressed in the robustness checks and discussed in the limitations section. The exploratory channel sample covers 25 cities from 2019 to 2024 and 150 observations.

City GDP and population information are assembled from official statistical sources and yearbooks; treatment timing comes from official pilot-zone approval records; express-delivery revenue is drawn from city postal statistics. The panel is audited for duplicate city-year keys, missing outcomes, cohort consistency and treatment timing. Because a complete, consistently measured set of time-varying controls is unavailable for all cities and years, identification relies on the stated parallel-trends and no-anticipation assumptions rather than on a saturated covariate model.

A key limitation of the design is that only three cities remain untreated through 2023, and all are located in the YRD. The main results are therefore identified primarily through the not-yet-treated comparison, which relies on the assumption that later-treated cities provide a valid counterfactual for earlier-treated ones before their own treatment. The statistical imprecision of the never-treated-only estimate (reported in the robustness section) indicates that inferences about the average treatment effect are conditional on the choice of comparison group.

The main specification is unconditional on time-varying covariates. This choice reflects data constraints rather than a conceptual preference: a complete and consistently measured set of time-varying controls is not available for all 39 cities over the 14-year period. To the extent that time-varying confounders exist, the estimates may be biased. The pre-trend test provides some support for the identifying assumption, but it cannot rule out all forms of unobserved time-varying selection.

Table 1: Sample, data sources and variable definitions

| Element | Definition / source | Role |
|---|---|---|
| Unit of analysis | City-year; 39 cities, 2010-2023 | Main panel |
| Outcome | ln(nominal GDP per capita, yuan) | Economic growth |
| Treatment | 1 from official CBEC pilot-zone approval year | Policy exposure |
| Channel proxy | ln(per-capita express-delivery revenue) | Exploratory activity |
| Treatment cohorts | 2015, 2016, 2018, 2019, 2020, 2022 | Staggered timing |
| Comparison | Not-yet-treated cities; never-treated sensitivity | Identification |

*Note: The logistics proxy measures activity intensity and should not be interpreted as technical efficiency.*

**Group-time ATT estimator**

Let $Y_{it}$ (1) and $Y_{it}$ (0) denote potential outcomes for city i in year t with and without treatment. Let $G_i$ be the first approval year, and let $D_{it}$ = 1[t >= $G_i$] for treated cities. The group-time average treatment effect for cohort g at time t is:

$$ATT(g,t) = E[Y_{it}(g) - Y_{it}(0) \mid G_i = g] \tag{1}$$

Following Callaway and Sant'Anna (2021), each ATT(g,t) is estimated using not-yet-treated cities as controls at time t and an outcome-regression adjustment. The aggregate ATT is a weighted average over identified post-treatment group-time effects:

$$ATT = \sum_{g\in\mathcal{G}}\sum_{t\geq g} w(g,t)\cdot ATT(g,t),\ where \sum_{g\in\mathcal{G}}\sum_{t\geq g} w(g,t) = 1 \tag{2}$$

The main standard errors are clustered by city. The design avoids using already-treated cities as counterfactuals for later cohorts and makes the composition of each event-time estimate explicit.

**Dynamic effects and pre-trend test**

Event time is k = t - $G_i$. The dynamic aggregation is reported for leads and lags supported by the panel, with k = -1 as the omitted reference period. The pre-treatment diagnostic tests whether the coefficients from k = -5 to k = -2 are jointly zero:

$$Y_{it} = \alpha_i + \lambda_t + \sum_{k\neq -1} \beta_k \cdot 1[t - G_i = k] + \varepsilon_{it} \tag{3}$$

$$H_0 = \beta_{-5} = \beta_{-4} = \beta_{-3} = \beta_{-2} = 0 \tag{4}$$

The event-time plot is interpreted with support counts. Longer lags can be estimated from fewer cohorts and treated cities, so they are not treated as a monotonic trajectory or as evidence of a long-run equilibrium effect.

**Benchmarks, robustness and regional contrast**

A conventional TWFE benchmark is estimated for comparison:

$$Y_{it} = \alpha_i + \lambda_t + \delta D_{it} + \varepsilon_{it} \tag{5}$$

The robustness design repeats the main estimator with never-treated controls only, excludes Shanghai and Shenzhen, starts the sample in 2014, omits calendar year 2020, and reassigns cohort labels in 500 placebo draws. Regional estimates use a common 2010-2021 horizon. The direct regional contrast is:

$$\Delta_{YRD-PRD} = ATT_{YRD} - ATT_{PRD} \tag{6}$$

A statistically insignificant regional contrast is not evidence that the regions are identical; it indicates that the data do not establish a difference under the common-horizon design.

**Exploratory logistics channel**

The channel analysis is intentionally supplementary. It is not used to test H2, which is evaluated using the event-study estimates in the Results section. This analysis is instead designed to explore whether a logistics-activity proxy shows patterns consistent with an indirect policy channel. Following the general causal-mediation framework of Imai et al. (2010), let $M_{it}$ denote log per-capita express-delivery revenue. The policy-to-activity and activity-to-growth paths are estimated on the common channel sample:

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

$$M_{it} = \alpha_i + \lambda_t + aD_{it} + u_{it} \tag{7}$$

$$Y_{it} = \alpha_i + \lambda_t + c'D_{it} + bM_{it} + v_{it} \tag{8}$$

$$Indirect\ Effect = a \times b \tag{9}$$

The indirect effect is evaluated with a bootstrap confidence interval and is not interpreted as causal mediation when the mediator is a scale proxy, the sample is small, or the mediator-outcome path is statistically weak.

## RESULTS

### Descriptive evidence and cohort composition

Table 2 summarizes the main variables and treatment composition. Average nominal GDP per capita is 91,587 yuan, while mean log GDP per capita is 11.312. The policy indicator equals one in 32.1% of city-year observations. The largest cohort is 2020, with 17 cities; the earliest cohort contains one city. This uneven composition is one reason to report cohort support rather than rely on a single post-treatment average.

Table 2: Descriptive statistics

| Panel A: Variable | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| GDP per capita (yuan) | 546 | 91,587.1 | 41,705.1 | 17,400 | 206,278 |
| Log GDP per capita | 546 | 11.312 | 0.496 | 9.764 | 12.237 |
| Policy exposure | 546 | 0.321 | 0.467 | 0 | 1 |
| Log express revenue (channel) | 150 | 6.623 | 0.557 | 5.337 | 7.852 |

*Note: The 36 treated cities are distributed across 2015, 2016, 2018, 2019, 2020 and 2022 cohorts; three YRD cities are never treated. Panel A uses the 546-observation main panel except for the logistics proxy, which uses 150 observations.*

### Main estimate and dynamic pattern

The pre-treatment event-time coefficients are small relative to their confidence intervals. The joint Wald test for $k = -5$ through $k = -2$ is 6.540 with $p = 0.162$, so the specified parallel-trends null is not rejected. This result supports the identifying window but does not prove that all unobserved time-varying differences are absent.

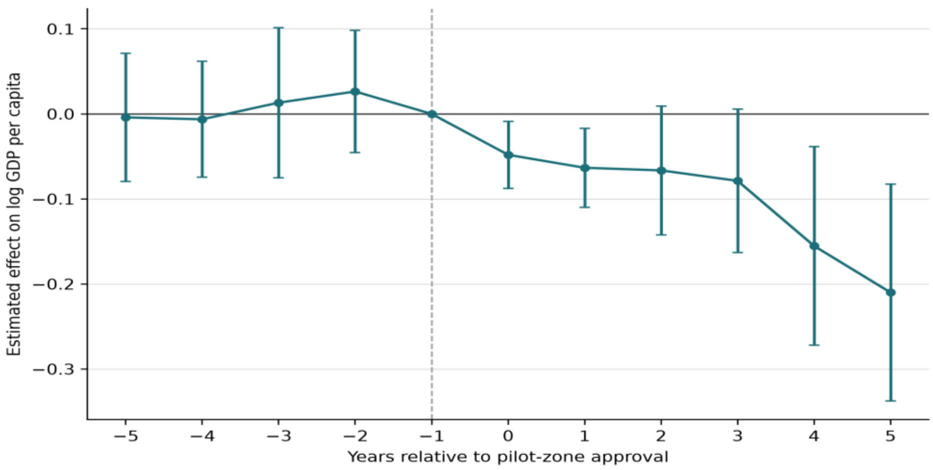


Figure 2: Group-Time Event-Study Estimates with 95% Confidence Intervals

*Note: Event time -1 is the reference period. Later event times rely on fewer cohorts and treated cities; coefficient magnitude must therefore be interpreted together with support.*

Figure 2 displays the event-time estimates. The approval-year effect is -0.0481 (p = 0.0167) and the one-year effect is -0.0634 (p = 0.0075). The estimates for k = 2 and k = 3 are negative but less precise. Effects at k = 4 and k = 5 are more negative, yet they are supported by four and three contributing cohorts respectively. The visual pattern is therefore consistent with early adjustment costs, but it does not establish a continuously worsening path.

Table 3: Main ATT and selected event-time estimates

| Estimate | ATT | SE | p-value | 95% CI | Support |
|---|---|---|---|---|---|
| Main group-time ATT | -0.0912 | 0.0333 | 0.006 | [-0.1565, -0.0259] | 36 treated cities |
| Event time k = 0 | -0.0481 | 0.0201 | 0.0167 | [-0.0875, -0.0087] | 6 cohorts /36 cities |
| Event time k = 1 | -0.0634 | 0.0237 | 0.0075 | [-0.1098, -0.0169] | 6 cohorts /36 cities |
| Event time k = 2 | -0.0665 | 0.0386 | 0.0855 | [-0.1422, 0.0093] | 5 cohorts /31 cities |
| Event time k = 3 | -0.0787 | 0.0431 | 0.0677 | [-0.1631, 0.0057] | 5 cohorts /31 cities |

*Note: The main ATT is -0.0912 log points. exp(-0.0912)-1 = -8.72%. Event-time estimates use k = -1 as the omitted period.*

The main ATT is -0.0912 log points (SE = 0.0333, p = 0.0062; 95% CI [-0.1565, -0.0259]), corresponding to approximately -8.72% in the nominal GDP-per-capita measure. H1 is supported because the effect is statistically significant, with a negative sign determined by the data. H2 is also supported under the prespecified short-run rule because both k=0 and k=1 are negative and significant. The result should be described as reduced-form support: the panel shows the early outcome pattern, not the cost components themselves. This interpretation is consistent with a short-run adjustment-cost story, but the aggregate city-level data do not directly observe compliance expenditure, technology adaptation, or resource reallocation. The H2 result should be read as reduced-form support rather than evidence of a separately identified mechanism.

**Robustness and placebo evidence**

The main estimate is stable across several sample and calendar-year checks. Excluding Shanghai and Shenzhen gives -0.0904 (p = 0.0060); starting in 2014 gives -0.0912 (p = 0.0039); and omitting 2020 gives -0.0815 (p = 0.0131). The TWFE benchmark is smaller and insignificant (-0.0382, p = 0.3311), illustrating why the group-time estimator is used for the primary inference. The estimate falls to -0.0599 and becomes imprecise when only three never-treated cities are controls. All three are in the YRD, so this sensitivity test has limited geographic breadth.

Table 4: Robustness and placebo checks

| Specification | Estimate | SE | p-value | Interpretation |
|---|---|---|---|---|
| Main: not-yet-treated controls | -0.0912 | 0.0333 | 0.0062 | Significant |
| Never-treated controls only | -0.0599 | 0.0406 | 0.1403 | Imprecise |
| Exclude Shanghai and Shenzhen | -0.0904 | 0.0329 | 0.0060 | Stable |
| Start sample in 2014 | -0.0912 | 0.0316 | 0.0039 | Stable |

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

| | | | | |
|---|---|---|---|---|
| Omit calendar year 2020 | -0.0815 | 0.0328 | 0.0131 | Stable |
| TWFE benchmark | -0.0382 | 0.0392 | 0.3311 | Benchmark only |
| 500 cohort randomisations | Observed -0.0912 | SD 0.0521 | 0.0938 | Suggestive at 10% |

In the cohort-randomisation placebo, the observed ATT lies near the lower tail of the 500 randomised estimates. The empirical two-sided p-value is 0.0938. This is suggestive at the 10% level but does not meet a 5% randomisation-inference threshold. Taken together, the checks support a conditional negative effect while showing that inference is sensitive to the comparison group and assignment design.

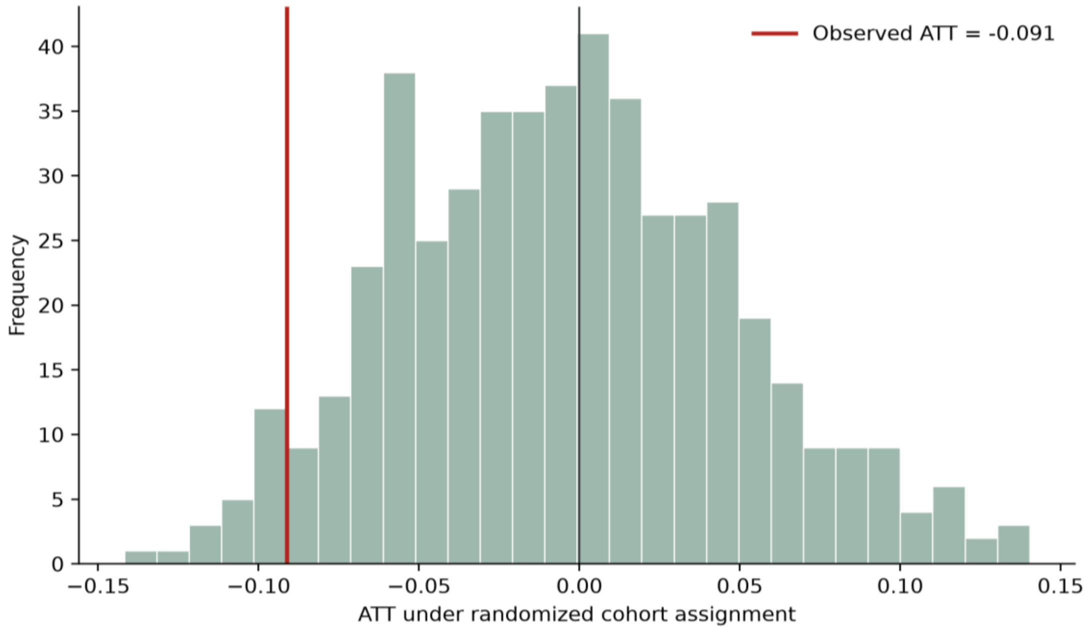


Figure 3: Distribution of randomised treatment-cohort placebo estimates

**Regional comparison**

Regional estimates are computed on the common 2010-2021 horizon to avoid comparing different post-treatment support. The YRD estimate is -0.0623 ($p = 0.0450$) and the PRD estimate is -0.0971 ($p = 0.0002$). The direct YRD-minus-PRD contrast is 0.0348 with a bootstrap p-value of 0.1607 and 95% CI [-0.0123, 0.0851]. Therefore, both regional estimates are negative, but the data do not establish a statistically significant regional difference.

**Exploratory logistics activity channel**

In the 25-city channel sample, pilot-zone exposure increases log per-capita express-delivery revenue (path a = 0.0947, $p < 0.001$). However, logistics activity is not significantly associated with log GDP per capita after city and year effects are included (path b = 0.0198, $p = 0.822$). The indirect effect is 0.00187, with bootstrap SE = 0.00724, $p = 0.684$ and 95% CI [-0.01234, 0.01720]. The total policy effect in this smaller sample is -0.0430 ($p = 0.0304$) and the direct effect is -0.0449 ($p = 0.0337$). These results do not support causal mediation through the available proxy. They do not show that logistics is economically irrelevant; they show that revenue intensity is not a sufficient measure of efficiency and does not identify a significant mediator-outcome path.

Table 5: Regional Contrast and Exploratory Logistics-Channel Results

| **Component** | **Estimate** | **Std. Err.** | **p-value** | **95% CI** |
|---|---|---|---|---|
| YRD ATT, common horizon | -0.0623 | 0.0311 | 0.0450 | [-0.1232, -0.0014] |

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

| | | | | |
|---|---|---|---|---|
| PRD ATT, common horizon | -0.0971 | 0.0264 | 0.0002 | [-0.1489, -0.0453] |
| YRD – PRD difference | 0.0348 | 0.0248 | 0.1607 | [-0.0123, 0.0851] |
| Policy → logistics activity (a) | 0.0947 | 0.0185 | < 0.001 | [0.0584, 0.1310] |
| Logistics activity → GDP per capita (b) | 0.0198 | 0.0874 | 0.8216 | [-0.1515, 0.1911] |
| Indirect effect (a × b) | 0.00187 | 0.00724 | 0.684 | [-0.01234, 0.01720] |

## DISCUSSION

The central finding is a negative short- to medium-run association between pilot-zone approval and nominal GDP per capita under the main group-time design. Several interpretations are possible. First, designation may require firms and local agencies to incur front-loaded compliance and technology-adaptation costs. Second, the aggregate GDP measure may capture losses in traditional activities alongside gains in CBEC-related activity. Third, policy may reallocate activity across firms or neighbouring cities rather than raise the local aggregate immediately. These interpretations are consistent with creative destruction and trade-adjustment perspectives, but the present design cannot distinguish among them because it does not observe firm exits, worker transitions, compliance expenditure or productivity.

The result is also a reminder that timing and comparison groups matter. The early event-time coefficients are negative and significant, while later coefficients are less precisely supported. The never-treated-only sensitivity removes significance, and the placebo result is suggestive rather than decisive. The findings therefore should not be generalized as proof that pilot zones permanently reduce growth. They indicate that the short-run outcome is not automatically positive and that a longer horizon and more direct adjustment measures are needed to assess whether later benefits offset initial costs.

The regional analysis does not support a simple YRD-versus-PRD ranking. Both estimates are negative on a common horizon, but their difference is not statistically significant. This suggests that local institutional capacity, industrial composition, market access and implementation quality may be more informative than broad regional labels. Similarly, higher express-delivery revenue does not establish improved logistics efficiency. Policy evaluation should distinguish activity volume from time, reliability, unit cost and network coverage.

## CONCLUSION AND POLICY IMPLICATIONS

This paper evaluates China's CBEC comprehensive pilot-zone programme with a balanced panel of 39 cities over 2010-2023. Using the Callaway-Sant'Anna group-time ATT estimator and not-yet-treated controls, the estimated effect on nominal GDP per capita is -0.0912 log points, approximately -8.72%, and statistically significant at conventional levels. The pre-treatment joint test does not reject the specified parallel-trends window. Negative effects are concentrated in the approval year and the following year, which supports the revised H2 as a short-run reduced-form prediction. The proposed compliance and adaptation costs remain an interpretation rather than a directly observed mediator.

The evidence has four policy implications. First, evaluations should pre-specify longer horizons and report the number of cohorts supporting each event time. A negative first-year effect should not be treated as a final assessment of a policy intended to change trade infrastructure and firm capabilities. Second, governments can reduce transition risk with technical assistance, interoperable customs and tax systems, training, and temporary support for smaller firms facing compliance costs. Third, policy design should monitor adjustment outcomes directly, including firm entry and exit, employment, wages, customs time, delivery reliability and unit logistics costs. Fourth, regional implementation should be based on local capacity and industrial structure rather than an assumed advantage of either delta.

For firms and logistics providers, the findings imply that pilot-zone approval is not a guarantee of immediate demand or productivity. Investments should be staged against measurable compliance and service milestones, and revenue growth should not be used as a substitute for efficiency indicators. For researchers, the results motivate broader city coverage, additional post-2023 observations and firm-level data that can separate creation, reallocation and displacement.

**Limitations And Future Research**

The sample covers only 39 cities in two economically advanced coastal regions (YRD and PRD). External validity to central, western, or northeastern China is therefore uncertain. The results should be interpreted as conditional on the institutional and geographic context of these two regions, and should not be generalized automatically to less-connected or less-developed cities.

The estimates rely on unconditional parallel trends because a complete set of time-varying controls is unavailable for all 39 cities. Only three cities are never treated, and all are in the YRD; the loss of significance in the never-treated sensitivity shows comparison-group sensitivity. Long event-time estimates also rely on fewer cohorts. The logistics analysis has only 25 cities and 150 observations, and its revenue proxy mixes activity scale with possible price and volume changes. In addition, the outcome variable is nominal GDP per capita. To the extent that city-level price trends differ systematically by treatment timing, the estimated effect may partly capture relative price changes rather than real output changes. A consistent city-level deflator was not available for the full panel, so we are unable to isolate real output effects directly. A robustness check using real GDP per capita was not feasible for the same reason. This remains an important direction for future research as more refined price data become available. This limitation should be kept in mind when interpreting the magnitude of the estimated ATT. Finally, nominal GDP per capita does not capture distribution, employment quality, firm innovation or welfare, and the design does not model spatial spillovers or the interaction between COVID-19 and trade-oriented cities.

Future research should extend the observation period, include cities outside the two coastal deltas, and develop direct measures of customs performance and logistics efficiency. Recent work by Luo, Kamarudin, Soh, and Shan (2026), for example, provides province-level evidence on the complementarity between logistics efficiency and AI capability in shaping CBEC outcomes, offering a useful direction for further investigation. Researchers should also link policy cohorts to firm registries, payroll records, customs transactions and innovation data. Spatial econometric designs and mediation methods based on explicit mediator-outcome assumptions could help determine whether the short-run aggregate decline reflects firm exit, worker reallocation, temporary implementation costs or displacement to neighbouring cities.

**Data Availability**

The study uses secondary city-level data assembled from official statistical yearbooks, pilot-zone approval records and postal statistics. The constructed dataset and associated analysis code are available from the author upon reasonable request, subject to the terms imposed by the original data providers and relevant institutional requirements.

**Conflict Of Interest**

The author declares no conflicts of interest.

**Funding**

The author received no external funding for this research.

**Ethical Approval**

Ethical approval was not required because this study uses secondary aggregate city-level data and does not involve human participants, personal data or animals.

## ACKNOWLEDGEMENTS

The author thanks Assoc. Prof. Dr. Chukiat Chaiboonsri and Asst. Prof. Dr. Anuphak Saosaovaphak for their guidance and constructive comments, and the Faculty of Economics at Chiang Mai University for its academic support.

## REFERENCES

1. Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. Econometrica, 60(2), 323-351. https://doi.org/10.2307/2951599
2. Aghion, P., Blundell, R., Griffith, R., Howitt, P., & Prantl, S. (2009). The effects of entry on incumbent innovation and productivity. Review of Economics and Statistics, 91(1), 20-32. https://doi.org/10.1162/rest.91.1.20
3. Autor, D. H., Dorn, D., Hanson, G. H., & Song, J. (2014). Trade adjustment: Worker-level evidence. Quarterly Journal of Economics, 129(4), 1799-1860. https://doi.org/10.1093/qje/qju026
4. Baker, A. C., Larcker, D. F., & Wang, C. C. Y. (2022). How much should we trust staggered difference-in-differences estimates? Journal of Financial Economics, 144(2), 370-395. https://doi.org/10.1016/j.jfineco.2022.01.004
5. Baker, A., Callaway, B., Cunningham, S., Goodman-Bacon, A., & Sant'Anna, P. H. C. (2026). Difference-in-differences designs: A practitioner‘s guide. Journal of Economic Literature, 64(2), 498-557. https://doi.org/10.1257/jel.20251650
6. Cai, L., Du, Q., Bao, L., Boateng, A., & Li, Z. (2025). Does cross-border e-commerce development promote the improvement of residents' income level? Evidence from China. Research in International Business and Finance, 77, 102907. https://doi.org/10.1016/j.ribaf.2025.102907
7. Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200-230. https://doi.org/10.1016/j.jeconom.2020.12.001
8. Cheng, S., Song, Y., & Tu, Y. (2025). Place-based policies, creation and reallocation effects on city exports: Insights from China's cross-border e-commerce comprehensive pilot zones. China Economic Review, 93, 102479. https://doi.org/10.1016/j.chieco.2025.102479
9. China Academy of Information and Communications Technology. (2024). Report on China's digital economy development. https://www.caict.ac.cn/kxyj/qwfb/bps/
10. de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review, 110(9), 2964-2996. https://doi.org/10.1257/aer.20181169
11. Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In Handbook of Regional and Urban Economics (Vol. 4, pp. 2063-2117). Elsevier. https://doi.org/10.1016/S1574-0080(04)80005-1
12. Freund, C., & Weinhold, D. (2004). The effect of the Internet on international trade. Journal of International Economics, 62(1), 171-189. https://doi.org/10.1016/S0022-1996(03)00059-X
13. Fujita, M., Krugman, P., & Venables, A. J. (1999). The spatial economy: Cities, regions, and international trade. MIT Press. https://doi.org/10.7551/mitpress/6389.001.0001
14. Gomez-Herrera, E., Martens, B., & Turlea, G. (2014). The drivers and impediments for cross-border e-commerce in the EU. Information Economics and Policy, 28, 83-96. https://doi.org/10.1016/j.infoecopol.2014.05.002
15. Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2), 254-277. https://doi.org/10.1016/j.jeconom.2021.03.014
16. Hummels, D. L., & Schaur, G. (2013). Time as a trade barrier. American Economic Review, 103(7), 2935-2959. https://doi.org/10.1257/aer.103.7.2935
17. Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. Psychological Methods, 15(4), 309-334. https://doi.org/10.1037/a0020761
18. Krugman, P. (1991). Increasing returns and economic geography. Journal of Political Economy, 99(3), 483-499. https://doi.org/10.1086/261763
19. Li, C., Zhang, J., & Liu, B. (2023). CBEC pilot zone policy and industrial structure upgrading: An empirical study based on prefecture-level city panel data. Finance and Trade Economics, 44(2), 115-

129. https://oversea.cnki.net/kns8s/defaultresult/index?kw=CBEC+pilot+zone+industrial+structure+upgrading+Li+Zhang+Liu&language=EN

20. Luo, H., Kamarudin, F., Soh, W., & Shan, Z. (2026). Fulfilment efficiency, AI capability, and cross-border e-commerce development in China: Complementarities, regional heterogeneity, and resource-saving potential. Sustainability, 18(3), 1202. https://doi.org/10.3390/su18031202
21. Marshall, A. (1890). Principles of economics. Macmillan. https://www.econlib.org/library/Marshall/marP.html
22. Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica, 71(6), 1695-1725. https://doi.org/10.1111/1468-0262.00467
23. Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In Handbook of Regional and Urban Economics (Vol. 4, pp. 2119-2171). Elsevier. https://doi.org/10.1016/S1574-0080(04)80006-3
24. Schumpeter, J. A. (1942). Capitalism, socialism and democracy. Harper & Brothers. https://www.routledge.com/Capitalism-Socialism-and-Democracy/Schumpeter/p/book/9780415567893
25. Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics, 225(2), 175-199. https://doi.org/10.1016/j.jeconom.2020.09.006
26. Terzi, N. (2011). The impact of e-commerce on international trade and employment. Procedia - Social and Behavioral Sciences, 24, 745-753. https://doi.org/10.1016/j.sbspro.2011.09.010
27. Wang, W., Sun, M., & Zhou, D. (2025). The impact of cross-border e-commerce comprehensive pilot zone on corporate financial constraints in China. Humanities and Social Sciences Communications, 12, Article 1107. https://doi.org/10.1057/s41599-025-05409-3
28. Zhong, M., Wang, Z., & Ge, X. (2022). Does cross-border e-commerce promote economic growth? Empirical research on China's pilot zones. Sustainability, 14(17), 11032. https://doi.org/10.3390/su141711032