Case study - Dynamic effects of support programs

Statistics Canada

Economic Social Analysis and Modelling Division

February 2026

Delivering insight through data for a better Canada

Framework & key message

Evaluate programs that serve businesses

- **Goal**: Measure quasi-causal effects on sales, employment, and longevity of businesses supported by federal agencies (BDC, FCC, EDC)
- **Toolkit**: DiD, event study (ES), survival analysis (SA) + linked data from Statistics Canada (BR, CEEDD, NALMF, DSD).
- **Key message**: EE validates DiD assumptions and highlights impact dynamics (timing, persistence, anticipation).
- **Illustrative case**: application to BDC (sales and employment growth) — demonstration of feasibility and value for decision-makers.

Evaluation context

Why these methods in the public sector?

- **Real constraints**: no randomization, fragmented data, staggered deployments, often post-implementation evaluations.
- **Solution**: secure linkages and long time series→ rich panels, standardized indicators, traceability
- **Clarify the causal**: distinguish between confounding (age, size, market), mediators (e.g. marketing investment after the loan) and moderators (industry).
- **Positioning**: RCT/RDD/IV often impractical → quasi-experimental (DiD/EE/AS) = rigor/feasibility balance.

DiD (basis of inference)

Difference-in-Differences (DiD)

- **Idea**: compare pre→post Δ between treated and comparable untreated → ATT. Fixed Difference Control + Common Shocks.
- **Key Assumption**: Parallel trends (plausibility supported by pre-treatment series, matching/IPW).
- **“Staggered adoption":** TWFE models can mix cohorts (negative weights). Favor ATT(g,t) and proper aggregations (e.g., Callaway-Sant'Anna).

Event Study: The Complement to DiD

Event Study (EE)

- **What it does**: estimates the dynamic effects around t=0 (first exposure): anticipated effects (anticipation), delayed effects (persistence).
- **What it's used for**:
  - Diagnose parallelism (pre-trends ~ 0),
  - Detect anticipation (leads ≠ 0),
  - Explain the timing (immediate or delayed effect, J-curve, loss of momentum).

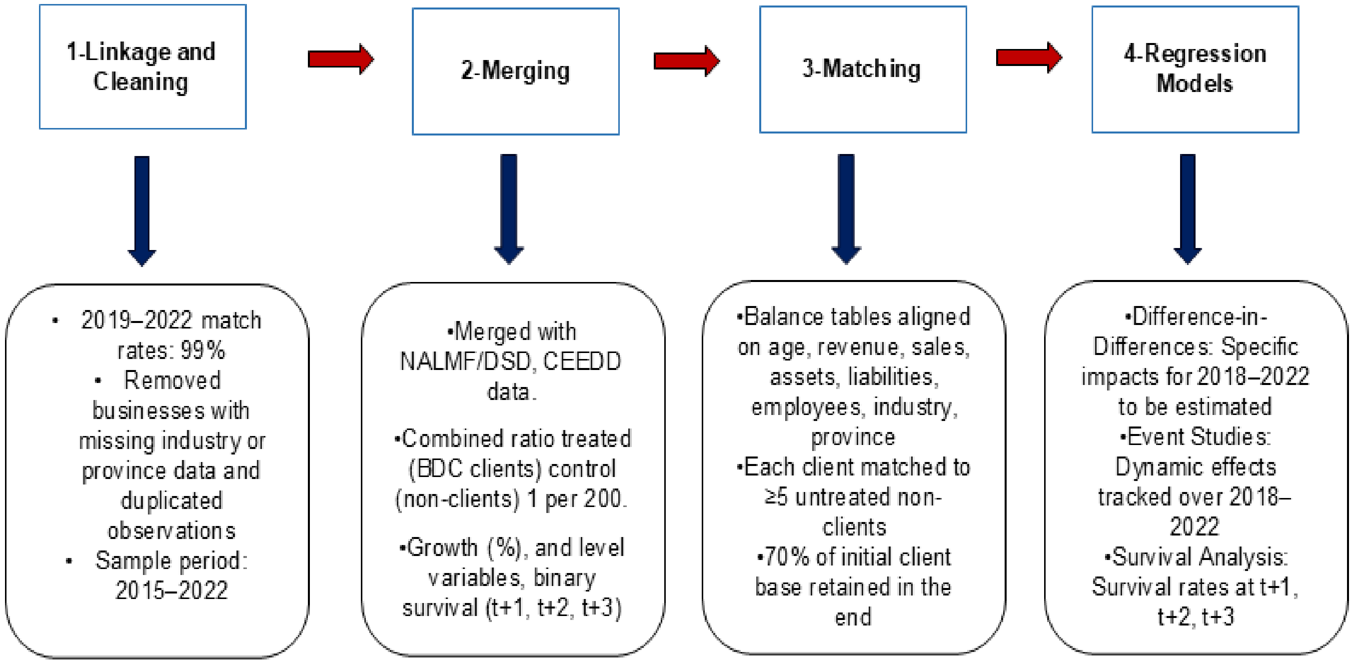
- **Best practices**: windows with sufficient support, grouping, valid controls (never/not yet processed), enterprise-level SE cluster.

Application to Agency Programs

- **Data**: customers → BR → CEEDD merger (NALMF, DSD); construction dependent variables, time of processing, control variables.
- **Counterfactual**: matching/IPW for comparability; DiD + EE estimation; complement with Survival (t+1/t+2/t+3, p.p. gain).
- **BDC cases (e.g.):** significant gains in sales/employment; EE shows flat pre-trends and a post-t rise consistent with DiD. (See Measuring BDC s impact on its clients Analysis)
- To remember:
    - DiD summarizes the average effect,
    - ES brings credibility + timing,
    - Survival informs resilience

Methodological framework: overview

- **Objective: To isolate quasi-causal effects of federal programs from external factors (firm characteristics, sectoral trends, macro shocks).**
- 4-step framework:
  - Coupling of customer data to the Business Register (BR) + cleaning,
  - Merger with business bases (CEEDD: NALMF, DSD),
  - Construction of the counterfactual (matching/weighting),
  - Econometric models: DiD, Event study, Survival.



Step 2: Merge & build the analytical panel

- **Outputs**: longitudinal series (sales, $CA exports, employment) for treated and controls
- **Key Databases (CEEDD):**
  - NALMF = main base for incorporated companies with employees; does not cover unincorporated women without employees,
  - DSD = demographic dimensions (gender, age, immigration status, disability; Indigenous homeowners and racialized groups),
  - Supplements if needed: T1FD / T2S50 for self-employed workers/specific structures.

- **Constructed variables**: treatment indicators (timing/intensity), control variables (size, industry, pre-treatment trends), dependent variables (employment/sales growth, productivity, survival). Use aggregations/delays to limit endogeneity.

Steps 3 & 4: Counterfactual & Templates

Step 3 — Matching (construction of the control group)

- **Principle**: to match similar non-treaties (industry, size, income, location, employment) to treated, to improve comparability.
- **Methods**: PSM, nearest neighbors, caliber, kernel matching (KM) — weights all untreated cases according to their similarity, the choice of bandwidth is crucial.
- **Caveats**: Matching does not guarantee parallel trends, may reduce sample size, and may be sensitive to the required algorithm → diagnostics and sensitivity testing.

Steps 3 & 4: Counterfactual & Templates

Step 4 — Models: Event Studies

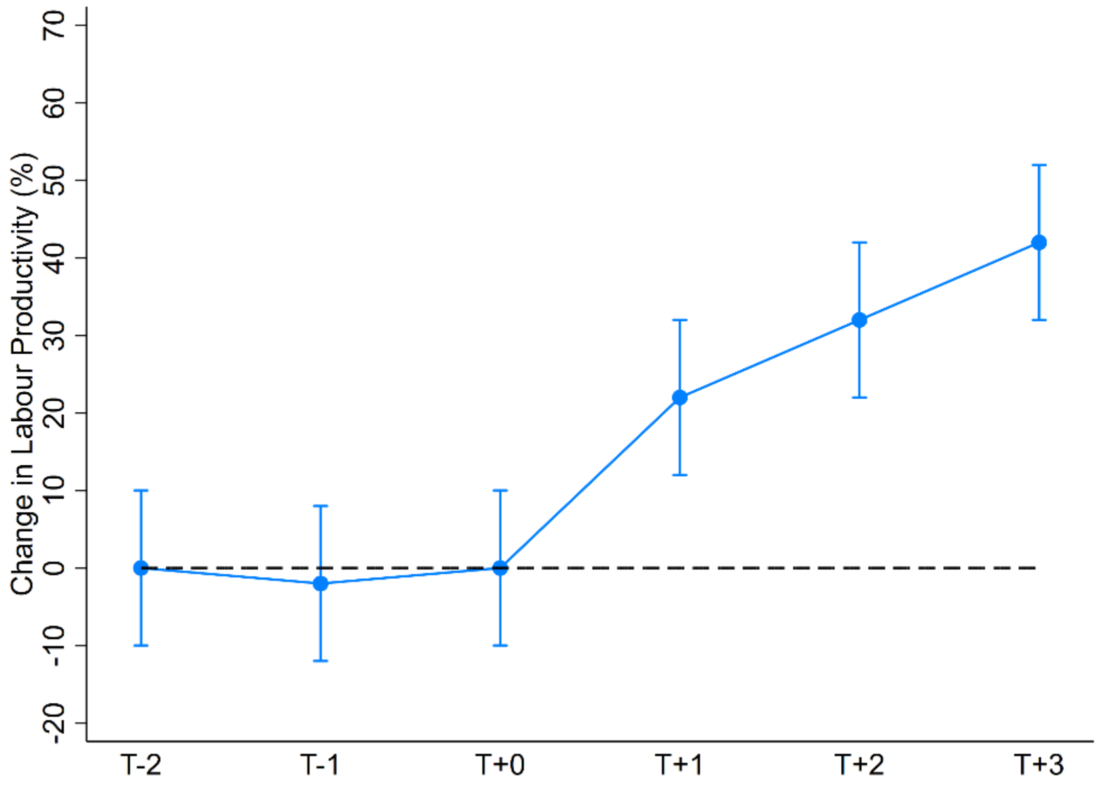
- **Aim**: To estimate, for each k, the incremental impact of BDC services on income/employment (or productivity) compared to the control group in relative time t+k
- **Counterfactual (IPWRA):** Selection of the control by inverse probability weight (IPW) and regression adjustment, using the means of the 3 pre-treatment years (age, assets, liabilities, employees, revenues, sales of goods and services), to mitigate anticipation and annual volatility.
- **Regions and sectors**: included to control for regional and sectoral heterogeneity.

Steps 3 & 4: Counterfactual & Templates

- **Specification (relative time):**
  $y_{it} = \alpha + \sum_{k \neq -1} \beta_k \, \mathbf{1}\{\tau_{it} = k\} + \gamma_i + \delta_t + X'_{it}\theta + \varepsilon_{it},$
  where k = −1 as the reference period; SE grouped at the enterprise level; controls/weights from IPW matching.
- **Cohorts by year of first service**: for a 2018 cohort, ATT_2018... ATT_2022 (t+0→t+4) depending on availability; Later cohorts have shorter windows (e.g., 2019: t+0→t+3, etc.)
- **Alignment and aggregation:** all cohorts are re-coded in relative time (t+0... t+4/5), and then aggregated to estimate common ATT_{t+k} (pooling)—allowing for a single trajectory of average impact over time.

Steps 3 & 4: Counterfactual & Templates

- No improvement at (t+0), but significant and increasing gains appear in the years that follow.
- These effects are ignored by descriptive statistics, which usually report only one average result.
- **Event studies** display results for each time period, helping analysts determine whether the effects are fading, stabilizing, or continuing to increase.

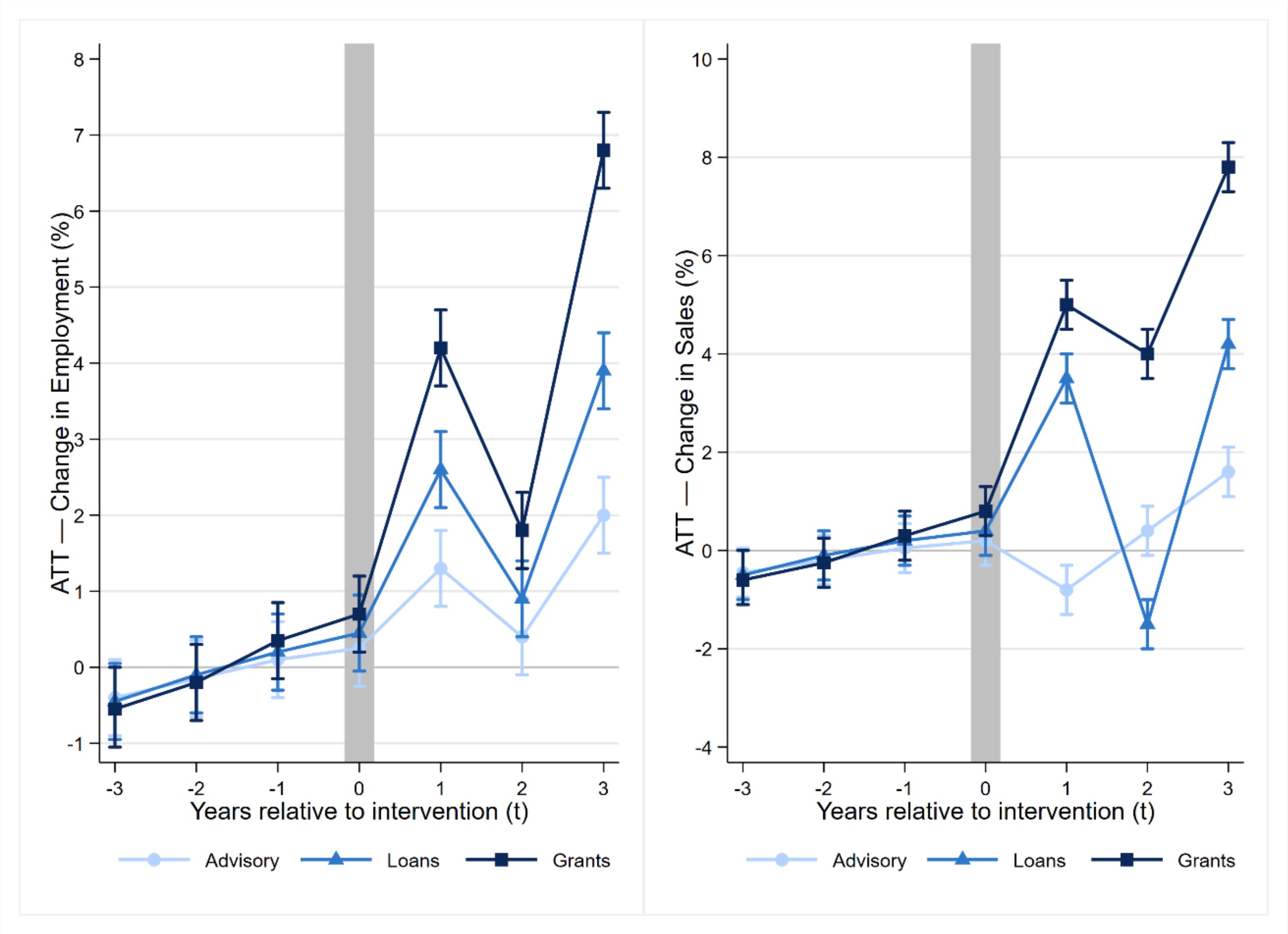


Case Study

- **Typical program (SMEs, Canada):** financial support (grants/bursaries, loans/guarantees), advice/training. Question: does this support increase employment and sales?
- **Observation period**: e.g. 2018 cohorts followed 2015→2021 to compare before/during/after the intervention, considering timing of first service (t=0), age, size, region, and area.
- **Database**: linkage of the list of supported enterprises to the StatCan enterprise database and then to the CEEDD files (e.g., NALMF, DSD) to construct a rich longitudinal panel.
- **Comparison of treated and controls**: supported firms are compared with similar non-supported firms to isolate the effect of the program from external factors (sector cycle, macro shocks).

Event Study Results (Employment & Sales)

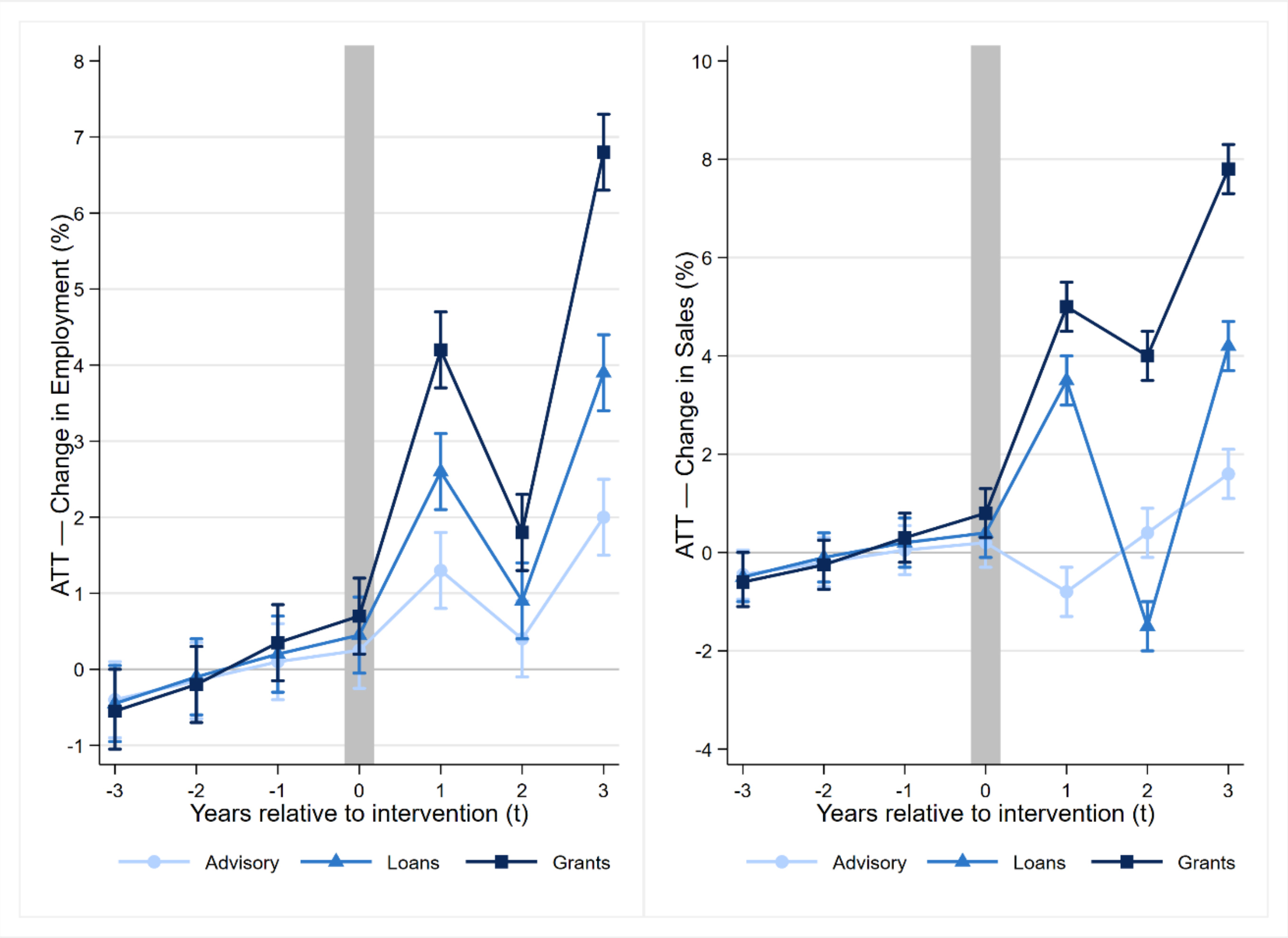
- **Pre-treatment diagnosis**: t−3 coefficients... T+0 close to zero → plausible parallel pre-trends between customers and controls.
- **After t=0 (emploi):**
  - **Subsidies**: the most marked effect; T+1 increase, pandemic year trough, Q+3 rebound (~> 6 p.p.).
  - **Loans**: moderate profile; gain in t+1, slowdown in t+2, recovery in t+3.
  - **Advisory**: modest (≤ ~2 p.p. over the period).



Event Study Results (Employment & Sales)

- **After t=0 (sales):** similar trend; for loans, a more pronounced decline around the pandemic and then catching up; subsidies remain dominant

- EE shows when effects appear, how they respond to shocks (pandemic) and for whom they are strongest (differences by type of service).



Event study and DiD: implications

- **DiD** → "average effect size" at key steps (t+1, t+2, t+3); Easy to communicate. Limit: Restricted dynamic effect.
- **Event study**→ "impact dynamics": Validates hypotheses (pre-trends, anticipation)
  - Discovers timing (take-off, persistence/recovery),
  - Contextualizes shocks (e.g., pandemic) and guides the optimization of the mix of services (grants vs. loans vs. advice).
  - Comparisons between services: caution – each ATT reflects a different matched sample (selection), so impact profiles and windows are mostly compared rather than an "absolute ranking".

Limitations and Extension

- **TWFE — key non-testable hypothesis**: parallel trends are necessary but not observable for the counterfactual; Unobserved confounders, policy anticipation or differentiated shocks bias β. Even with similar pre-trends, there is no guarantee that they will be maintained post-treatment

- **Staggered adoption**: With a standard TWFE, negative weights and contamination from previously treated units can produce misleading estimates.
  - **Limitations**: may be less efficient (larger samples required), small cohorts, → omitted/imprecise effects; is still based on non-testable parallel (conditional) trends, even if pre-trend tests are more readable.

Conclusion

- **Why evaluate**? Organizations want to strengthen accountability, optimize resource allocation and improve program results.
- To offer a **combined framework accessible to agencies**:
  - DiD → summarizes the average effect at key horizons,
  - Event Study → validates the hypotheses (pre-trends/anticipation) and reveals the dynamics of the effects,
  - Survival→ links performance gains to business sustainability.
- **Value-added**: These approaches complement qualitative methods (interviews, case studies) to better isolate program impacts.