

# Big Wins, Small Net Gains: Direct and Spillover Effects of First Industry Entries in Puerto Rico

Jorge A. Arroyo  
*Independent Researcher*  
arroyo.jorgeantonio@gmail.com

November 2025

## Abstract

I study how first sizable industry entries reshape local and neighboring labor markets in Puerto Rico. Using over a decade of quarterly municipality–industry data (2014Q1–2025Q1), I identify “first sizable entries” as large, persistent jumps in establishments, covered employment, and wage bill, and treat these as shocks to local industry presence at the municipio–industry level.

Methodologically, I combine staggered-adoption difference-in-differences estimators that are robust to heterogeneous treatment timing with an imputation-based event-study approach, and I use a doubly robust difference-in-differences framework that explicitly allows for interference through pre-specified exposure mappings on a contiguity graph. The estimates show large and persistent direct gains in covered employment and wage bill in the treated municipality–industry cells over 0–16 quarters. Same-industry neighbors experience sizable short-run gains that reverse over the medium run, while within-municipality cross-industry and neighbor all-industries spillovers are small and imprecisely estimated. Once these spillovers are taken into account and spatially robust inference and sensitivity checks are applied, the net regional 0–16 quarter effect on covered employment is positive but modest in magnitude and estimated with considerable uncertainty. The results imply that first sizable entries generate substantial local gains where they occur, but much smaller and less precisely measured net employment gains for the broader regional economy, highlighting the importance of accounting for spatial spillovers when evaluating place-based policies.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Motivation . . . . .	5
1.2	Contributions . . . . .	6
1.3	Research Questions . . . . .	7
1.4	Structure of the Paper . . . . .	8
<b>2</b>	<b>Conceptual Framework &amp; Related Literature</b>	<b>8</b>
2.1	Overview and Objectives . . . . .	8
2.2	Position in the Econometric Literature . . . . .	9
2.3	Modern Staggered DiD (General Framework) . . . . .	9
2.4	Direct Effects Estimators (CS & BJS) . . . . .	10
2.5	Interference-Aware DiD & Exposure Mappings . . . . .	10
2.6	Spatially Robust Inference (Conley, C-SCPC, $\pm$ TMO) . . . . .	11
<b>3</b>	<b>Data</b>	<b>12</b>
3.1	Raw Ingestion . . . . .	12
3.2	Quarter Index . . . . .	12
3.3	Industry Harmonization . . . . .	13
3.4	Geographic Layer (Municipios) . . . . .	13
3.5	Panel Construction . . . . .	14
3.6	Heterogeneity Strata . . . . .	15
3.7	Auxiliary Data for TMO Analysis . . . . .	16
3.8	Descriptive Statistics and Sample Composition . . . . .	16
<b>4</b>	<b>Treatment &amp; Exposure Construction</b>	<b>19</b>
4.1	Direct Treatment Definition (Large Entry Events) . . . . .	19
4.2	Exposure Mappings (General Framework) . . . . .	19
4.3	Specific Exposure Channels & Contrasts . . . . .	20
4.4	W Matrix (Spatial Weights) . . . . .	21
4.5	Lagged Exposure Histories ( $L = 4$ ) . . . . .	21
4.6	Timing and Geography of First Sizable Entries . . . . .	21
<b>5</b>	<b>Estimands</b>	<b>26</b>
5.1	Direct ATT ( $ATT^{own}(\ell)$ ) . . . . .	27
5.2	Neighbor & Cross-Industry ATTs (Spillover Estimands) . . . . .	27
<b>6</b>	<b>Identification</b>	<b>30</b>
6.1	Parallel Trends Conditional on Exposure Histories . . . . .	30
6.2	Limited Anticipation . . . . .	31
6.3	Interference Channels . . . . .	31
6.4	Exposure Positivity and Overlap . . . . .	32
6.5	Sensitivity to Violations of Parallel Trends . . . . .	32
6.6	Summary of Key Identification Assumptions . . . . .	33

<b>7 Estimation Strategy</b>	<b>34</b>
7.1 Balanced vs. Unbalanced Horizons . . . . .	34
7.2 Direct Effects: Primary Specification (CS) . . . . .	34
7.3 Direct Effects: Cross-Check (BJS) . . . . .	35
7.4 Spillovers: Primary Channel (DR–DiD) . . . . .	35
7.5 Cross-Fitting with Municipality-Respecting Folds . . . . .	36
7.6 Heterogeneous Effects: Auxiliary TWFE Specification . . . . .	37
<b>8 Fixed Effects and Inference</b>	<b>37</b>
8.1 Event-Study and Cumulative Inference . . . . .	38
8.2 Clustering at the Level of Treatment Assignment . . . . .	38
8.3 Moran’s I Residual Diagnostic (“Trigger”) . . . . .	39
8.4 Primary Spatial Inference: Conditional SCPBC . . . . .	39
8.5 Conley/SHAC Robustness Checks . . . . .	40
8.6 Two-Way Clustering Sensitivity . . . . .	40
<b>9 Diagnostics and Falsification</b>	<b>41</b>
9.1 Pre-Trends and Event-Study Diagnostics . . . . .	41
9.2 HonestDiD Sensitivity Analysis . . . . .	42
9.3 Overlap and Positivity Diagnostics . . . . .	42
9.4 Spatial Dependence Diagnostics . . . . .	43
9.5 Multiple-Testing Adjustments . . . . .	43
9.6 Diagnostic Results . . . . .	44
<b>10 Results</b>	<b>47</b>
10.1 Direct Dynamic Effects (CS Event Studies) . . . . .	47
10.2 Cumulative Direct Effects by Slice . . . . .	47
10.3 Spillovers and Decomposition of Total Effects (DR–DiD) . . . . .	49
<b>11 Heterogeneity</b>	<b>54</b>
<b>12 Interpretation and Policy Implications</b>	<b>58</b>
12.1 Magnitude of Direct Effects . . . . .	58
12.2 Spillovers, Crowding Out, and Net Effects . . . . .	59
12.3 Differences Across Space and Sectors . . . . .	60
12.4 Implications for Local Development Policy . . . . .	60
<b>13 Robustness</b>	<b>61</b>
13.1 Estimator Comparison: CS vs. BJS . . . . .	61
13.2 Sensitivity to Parallel-Trends Violations . . . . .	61
13.3 Inference Methods for DR–DiD Spillover Effects . . . . .	62
13.4 Specification Robustness: Exposure-History Definitions . . . . .	62
<b>14 Conclusion</b>	<b>66</b>
<b>Appendix</b>	<b>72</b>

<b>A Spatial Weights and Neighbor Structure</b>	<b>72</b>
A.1 Neighbor Degree Distribution under $W$	72
<b>B Additional Event-Study and Exposure Plots</b>	<b>73</b>
B.1 CS vs. BJS Event-Studies for Secondary Outcomes	73
B.2 Balanced vs. Unbalanced Horizons for Covered Employment	73
B.3 Exposure-Augmented Event-Study by Spillover Channel	74
<b>C Spillover Slice Estimates and Overlap Diagnostics</b>	<b>76</b>
C.1 DR–DiD Slice Effects for Secondary Outcomes	76
C.2 Exposure Overlap Across Channels and Slices	76
<b>D Spatial Dependence and Inference Comparisons</b>	<b>82</b>
D.1 Summary of Moran's $I$ Residual Diagnostics	82
D.2 Inference Comparisons for DR–DiD Slice Effects	82
<b>E Full Heterogeneity Grid</b>	<b>84</b>
<b>F ACS Auxiliary Socioeconomic Indicators Used in the TMO Analysis</b>	<b>84</b>
<b>G Reproducibility and Computational Workflow</b>	<b>86</b>
G.1 Raw Data Ingestion and Storage	86
G.2 Spatial Data Processing and GIS Artifacts	86
G.3 Intermediate Data Artifacts and Exposure Workflow	86
G.4 Exposure Trimming and Implementation Flags	87
G.5 Balanced Event-Time Horizon Indicator	87
G.6 DR–DiD Slice Implementation Details	88
G.7 Overlap Diagnostics Artifact	88
G.8 Heterogeneity Multiple-Testing Artifact	88

# 1 Introduction

## 1.1 Motivation

Large, persistent entries of new industries are central objects of local development policy. Governments devote substantial resources to attracting and retaining firms—through tax incentives, subsidies, and complementary business support programs—with the hope that major entries will generate jobs not only in the entering industry, but also in related sectors and the broader local economy (Trinajstić et al., 2022; Slattery and Zidar, 2020). Evidence from regional firm subsidies, often targeted at manufacturing but with documented spillovers to other sectors and economically connected regions, shows that such programs can meaningfully raise local employment and reshape the composition of economic activity (Siegloch et al., 2025; Aobdia et al., 2024). Yet it remains unclear how these large entries reshape local labor markets over time, how far their effects travel across space, and whether they primarily benefit the entering industry or diffuse more broadly across the local economy.

Two empirical challenges make these questions difficult to answer. First, much of the existing work on regional shocks and place-based policies relies on conventional two-way fixed effects (TWFE) difference-in-differences (DiD) designs. A rapidly expanding econometric literature shows that TWFE estimators can perform poorly in staggered adoption settings with heterogeneous treatment effects, generating contaminated dynamic coefficients and misleading pre-trend diagnostics (Sun and Abraham, 2021; Roth et al., 2023; Baker et al., 2025). Any evaluation of large industry entries that relies on staggered adoption must therefore use DiD methods explicitly designed to handle heterogeneous effects, rather than standard TWFE event studies.

Second, standard DiD frameworks typically assume no *interference* across units: one municipality’s treatment is assumed not to affect another municipality’s outcomes. This assumption is implausible in settings where firm entry in one location can reshape employment in nearby municipios or in related industries. Recent work develops interference-aware DiD frameworks and exposure mappings that allow for spillover effects under suitable conditions and emphasize the need to separate own-treatment effects from spillovers on other units (Liu et al., 2019; Xu, 2023; Fiorini et al., 2024; Shahn et al., 2024). These developments are especially relevant when the policy motivation is explicitly about regional spillovers and cross-industry linkages.

This paper brings these insights to a new empirical setting: large, persistent industry entries in Puerto Rico, observed at the municipality–industry level using administrative employment records. I treat first sizable entry events—defined as large, durable jumps in establishments, covered employment, and wage bill—as shocks to local industry presence and

use an empirical framework that separately identifies direct effects on the treated municipality–industry cell and spillover effects operating through spatial and cross-industry linkages. The central question is whether these “big wins” for the host municipio generate equally large net gains for the broader regional economy, or whether they mainly reallocate activity across space within the same industry.

## 1.2 Contributions

This paper studies how large, persistent industry entries reshape local and neighboring labor markets in Puerto Rico. I focus on municipality–industry cells and exploit first sizable entry events—identified from administrative employment records—as shocks to local industry presence. The analysis delivers four main sets of contributions.

First, I measure the *direct* effects of a first sizable entry on the treated municipality–industry cell. Using modern staggered-adoption DiD estimators ([Callaway and Sant’Anna, 2021](#); [Borusyak et al., 2024](#)), I estimate dynamic event-study paths and cumulative 0–16 quarter effects for establishments, covered employment, and the real wage bill. These estimators avoid using already-treated units as controls and are robust to heterogeneous treatment effects, addressing the concerns raised by recent work on TWFE event studies ([Sun and Abraham, 2021](#); [Roth et al., 2023](#)). The estimates show large and persistent gains in employment and wage bills in the entering municipality–industry cell.

Second, I quantify *spillovers to neighboring municipios*. Building on interference-aware DiD ideas and exposure-mapping strategies ([Aronow et al., 2020](#); [Xu, 2023](#); [Shahn et al., 2024](#)), I pre-specify an exposure mapping over a contiguity-based spatial graph of municipios to estimate how entry in one location affects outcomes in adjacent locations. A doubly robust DiD framework with exposure histories delivers slice-level estimates that decompose post-entry changes in employment into direct and spillover components. The results show sizable short-run gains for same-industry neighbors that are offset by longer-run negative same-industry spillovers.

Third, I decompose *cross-industry channels*. Within each municipality, I distinguish (i) within-municipality cross-industry effects—entry in industry  $k$  affecting other industries  $k' \neq k$  in the same municipality—from (ii) neighbor all-industries spillovers—entry in  $k$  at municipality  $i$  affecting all industries  $k'$  (including  $k$ ) in bordering municipios. This decomposition, implemented through separate exposure mappings, allows me to assess whether gains from entry are concentrated in the entering industry, spill across industries locally, or diffuse across space more broadly. In practice, cross-industry and neighbor-all spillovers are small and imprecisely estimated relative to the same-industry neighbor channel.

Fourth, I translate dynamic effects into *policy-relevant cumulative impacts*. I aggregate event-time paths into cumulative 0–16 quarter effects with simultaneous confidence bands and compute “policy slices” (0–4, 5–8, 9–16, 0–16 quarters) that map into horizons relevant for investment incentives, workforce training, and regional coordination. I also explore heterogeneity across pre-defined strata—tradable vs. non-tradable sectors, metro vs. non-metro municipios, and high- vs. low-wage sectors—using an auxiliary fixed-effects specification. Combining the direct and spillover components, I provide simple summary measures of cumulative gains that are suited for budgeting and return-on-investment discussions.

Previewing the main findings, first sizable entries generate very large and persistent direct gains in covered employment and wage bills in the treated municipality–industry cell, but these “big wins” are accompanied by non-trivial negative same-industry spillovers across space. Over a four-year horizon, net regional gains in covered employment are positive but modest and imprecisely estimated, implying that plant-level job counts can substantially overstate the net employment gains relevant for island-wide policy evaluation.

### 1.3 Research Questions

The analysis is organized around the following pre-specified questions:

**Q1: Direct effects.** For a municipality  $i$  and industry  $k$ , what is the dynamic percentage change in establishments, covered employment, and wage bill over 0–16 quarters after the first sizable entry event?

**Q2: Spillovers to same-industry neighbors.** How do outcomes in neighboring municipios change when adjacent municipios experience a large entry in the same industry?

**Q3: Cross-industry channels.**

**Q3a: Within-municipality cross-industry spillovers.** For an entry in industry  $k$  at municipality  $i$ , what are the dynamic percentage changes in outcomes for other industries  $k' \neq k$  within the same municipality?

**Q3b: Neighbor all-industries spillovers.** For an entry in industry  $k$  at municipality  $i$ , what are the dynamic percentage changes in outcomes for all industries  $k'$  (including  $k$ ) in bordering municipios?

**Q4: Heterogeneity.** How do direct and spillover effects vary across pre-defined strata—tradable vs. non-tradable industries, metro vs. non-metro municipios, and high- vs. low-wage sectors defined by pre-period median real wages (2014Q1–2019Q3)?

**Q5: Cumulative value.** What are the cumulative 0–16 quarter gains (direct + spillover + cross-industry components where relevant), and how large are these effects in magnitude and uncertainty from a budgeting or return-on-investment perspective?

## 1.4 Structure of the Paper

The remainder of the paper proceeds as follows. Section 2 reviews the conceptual framework and related econometric literature on staggered DiD, interference-aware DiD, and spatially robust inference. Section 3 describes the data, variables, spatial aggregation, and construction of heterogeneity strata. Section 4 formalizes the treatment assignment and exposure mapping procedures. Section 5 defines the estimands of interest, distinguishing direct effects from spillover channels. Section 6 sets out the identifying assumptions, including conditional parallel trends under interference and exposure-positivity conditions. Section 7 details the estimation procedures for direct and spillover effects and the auxiliary heterogeneity specification. Section 8 presents the cluster- and spatially robust inference procedures. Section 9 reports pre-trend tests, overlap and exposure-positivity diagnostics, spatial autocorrelation checks, and multiple-testing adjustments. Section 10 presents the main results, including dynamic and cumulative effects for all exposure channels. Section 11 explores heterogeneous effects across industry types and regions. Section 12 discusses the economic interpretation and policy implications of the findings. Section 13 reports sensitivity analyses for estimator choice, exposure-history definitions, and inference methods. Section 14 concludes.

# 2 Conceptual Framework & Related Literature

## 2.1 Overview and Objectives

The empirical setting combines administrative employment and wage records for Puerto Rico’s 78 municipios with a spatial representation of adjacency between municipios. The unit of analysis is a municipality–industry cell  $(i, k)$  observed quarterly. I define a “first sizable entry” as a large, persistent increase in establishments, employment, or wage bill for industry  $k$  in municipality  $i$  (Section 4.1) and interpret this as a shock to local industry presence.

The main objects of interest are dynamic and cumulative treatment effects along four dimensions. First, I study the *direct* impact of a first sizable entry on the treated municipality–industry cell: how establishments, covered employment, and wage bill in  $(i, k)$  evolve over the 0–16 quarters following entry (Q1). Second, I measure *same-industry spillovers* to neighboring municipios: whether entry in industry  $k$  in municipality  $i$  raises outcomes in the same industry  $k$  in adjacent municipios (Q2). Third, I quantify *cross-industry channels*: (i)

within-municipality cross-industry effects—entry in  $k$  affecting  $k' \neq k$  at the same  $i$ —and (ii) neighbor all-industries spillovers—entry in  $k$  affecting all industries  $k'$  in bordering municipios (Q3a, Q3b). Finally, I examine heterogeneity across pre-defined strata (Q4) and aggregate these dynamic paths into *cumulative 0–16 quarter effects* and related budgeting metrics (Q5).

Answering these questions requires methods that accommodate staggered adoption, heterogeneous treatment effects, interference across space and industries, and spatially correlated shocks. The remainder of this section situates the estimators and inference procedures used in the paper within three strands of the recent econometric literature: modern staggered DiD, interference-aware DiD and exposure mappings, and spatially robust inference.

## 2.2 Position in the Econometric Literature

My methodology is situated at the intersection of three recent advances in the econometrics literature: (1) heterogeneity-robust estimators for staggered difference-in-differences (DiD) designs; (2) DiD models that explicitly account for interference and spillovers; and (3) finite-sample-robust inference methods for spatial and panel data. This section provides a conceptual overview of these strands and how they inform the design of my empirical framework; formal estimands, identification assumptions, and implementation details are developed in Sections 5–8.

## 2.3 Modern Staggered DiD (General Framework)

The canonical two-way fixed effects (TWFE) regression, long the standard for DiD analysis, has been shown to be problematic in staggered adoption settings with heterogeneous treatment effects. When treatment timing varies, the TWFE estimator identifies a weighted average of individual treatment effects where some weights can be negative, because already-treated units are used as controls in “forbidden comparisons” (Roth et al., 2023; Baker et al., 2025). This can lead to biased estimates or even incorrectly signed coefficients.

Simulation studies confirm that while traditional TWFE performs well under homogeneous effects, its performance “diminish[es] notably” and exhibits “notably high bias” in the presence of dynamic and heterogeneous effects (Wang et al., 2024). Building on these insights, I adopt “modern” DiD estimators that are designed for staggered adoption with heterogeneous treatment effects and that avoid using already-treated units as controls.

## 2.4 Direct Effects Estimators (CS & BJS)

For my primary analysis of direct effects, I use the [Callaway and Sant'Anna \(2021\)](#) estimator. This approach defines the **group-time average treatment effect**,  $\text{ATT}(g, t)$ —the average effect for cohort  $g$  at time  $t$ —as the central building block for analysis. By structuring the design around cohort-specific comparisons to clean control groups, the CS framework avoids the contamination issues that arise in conventional TWFE event studies.

As a key cross-check and robustness exercise, I also employ the **imputation-based estimator** of [Borusyak et al. \(2024\)](#). This “BJS” method fits a TWFE model on “clean” (never- or not-yet-treated) observations and uses the estimated coefficients to impute counterfactual outcomes for treated observations ([Borusyak et al., 2024](#)). It is known to be an efficient estimator because it leverages all pre-treatment periods for imputation ([Borusyak et al., 2024](#)).

The use of these robust estimators is motivated by the diagnostic findings of [Sun and Abraham \(2021\)](#), who demonstrate that coefficients in a traditional event study (TWFE with leads and lags) are “contaminated by effects from other periods.” This contamination means that “apparent pre-trends can arise solely from treatment effects heterogeneity,” rendering standard pre-trend tests invalid and motivating estimators like CS and BJS that rely only on clean control groups ([Sun and Abraham, 2021](#)). This identification logic is related to the **Local Projections (LP-DiD)** framework, which [Dube et al. \(2025\)](#) show can replicate the CS and BJS estimators through a “clean control” condition ([Dube et al., 2025](#)). Practical implementation choices for CS and BJS in my setting are discussed in Section 7.

## 2.5 Interference-Aware DiD & Exposure Mappings

A core contribution of my paper is to measure spillovers. I therefore build on the literature that relaxes the Stable Unit Treatment Value Assumption (SUTVA). The foundational concept is the **exposure mapping**, which formalizes interference by defining a unit’s outcome as a function of the treatment status of its neighbors ([Aronow et al., 2020](#)).

Following [Aronow et al. \(2020\)](#) and [Leung \(2024\)](#), I define spillover effects as **contrasts across exposures**, for example, the effect of having one treated neighbor versus zero ([Leung, 2024](#)). My paper specifies three distinct exposure mappings to capture different spillover channels: (i) same-industry neighbor exposure; (ii) within-municipality cross-industry exposure; and (iii) all-industries neighbor exposure, where each source industry  $k$  generates exposures affecting all target industries  $k'$ —including  $k$  itself—in neighboring municipios. For the cross-industry and neighbor-all channels, I pool exposures by summing contributions across source industries. These mappings are formally defined in Section 4 and linked to

estimands in Section 5.

Xu (2023) formally establish identification of the **Direct ATT (DATT)** and **Spillover ATT (SATT)** in a DiD framework under a “modified parallel trends” assumption, namely **parallel trends conditional on exposure history**, which I adopt (Xu, 2023). This identification strategy is also formalized in the **Structural Nested Mean Model (SNMM)** literature. Shahn et al. (2024) connect DiD to SNMMs using the same exposure mapping framework and identify effects under an equivalent “network conditional parallel trends” assumption based on exposure history.

My approach relates to other recent work on interference in DiD. Fiorini et al. (2024) offer a “simple approach” that identifies the direct effect (their “ATT without interference”) by assuming spillovers cease once a unit is treated; in contrast, I allow spillovers and estimate direct and indirect effects jointly. Peng et al. (2025) motivate neighborhood-structured interference via their Differences-in-Neighbors estimator under a formal Neighborhood Interference assumption; I use their framework as conceptual background but do not employ DN for estimation. Estimation details for the interference-aware DiD components, including how I operationalize exposure histories and parallel-trends assumptions, are provided in Sections 6 and 7. Descriptive exposure-augmented event studies used as diagnostics are discussed in Section 9.

## 2.6 Spatially Robust Inference (Conley, C-SCPC, $\pm$ TMO)

To conduct inference, I must account for spatial correlation in my panel. The canonical approach is the Conley (1999) estimator, a nonparametric method that is a spatial analogue of time-series HAC estimators (Conley and Molinari, 2005). However, these and related methods can suffer from size distortions (i.e., poor finite-sample performance) in non-stationary settings such as panel data and DiD (Müller and Watson, 2023).

My primary inference strategy, therefore, is the **C-SCPC (Conditional Spatial Correlation Principal Components)** method proposed by Müller and Watson (2023). This is a robustified version of the SCPC method specifically developed to have good size properties in empirically relevant panel settings. As an additional sensitivity check, I also employ the **Thresholding Multiple Outcomes (TMO)** procedure from Della Vigna et al. (2025), which is designed to adjust standard errors for cross-sectional dependence. The application of these spatially robust covariance estimators to my DiD estimators is detailed in Section 8.

## 3 Data

In this section, I describe the construction of the dataset used in the analysis, including the ingestion of raw employment files, the construction of a consistent quarterly index, the harmonization of industry classifications, the creation of the municipios geometry layer for spatial analysis, and the aggregation to a balanced municipality–industry–quarter panel. These steps yield the core panel used for treatment assignment, exposure mapping, and estimation.

### 3.1 Raw Ingestion

The main data source is the Puerto Rico Department of Labor and Human Resources' *Composición Industrial por Municipio* portal ([Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas, 2025](#)). The *Composición Industrial por Municipio* tables are produced by the Programa Censo Trimestral de Empleo y Salarios (Puerto Rico's implementation of the U.S. Quarterly Census of Employment and Wages, QCEW; [Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas del Trabajo, nd](#)). They compile administrative records from the unemployment insurance system and quarterly employer tax reports, covering nearly all wage-and-salary employment in Puerto Rico.

Each quarterly release reports, for every municipality and industry (classified under the North American Industry Classification System, NAICS), the number of establishments, total employment, total wages, and average wages. The dataset excludes self-employed and unpaid workers, but includes full- and part-time payroll employees, corporate officers, and federal personnel assigned to Puerto Rico. Establishments are assigned to municipios by physical location, with special codes identifying multi-municipality aggregates and unclassified sites. Detailed variable definitions, NAICS versioning, and coverage notes follow the official quarterly publication ([Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas del Trabajo, 2025a](#)) and the program's methodological report ([Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas del Trabajo, nd](#)). All files are processed using a consistent schema, cleaned, and stacked into a single quarterly dataset; implementation details are documented in Section G.

### 3.2 Quarter Index

I construct a complete quarterly time index spanning from 2014Q1 to 2025Q1. The index contains the calendar year, quarter number, a label (e.g., “2018Q3”), the start date of the

quarter, and a sequential integer variable  $t$ . These identifiers provide a harmonized timeline used for event-time alignment and temporal joins across sources.

### 3.3 Industry Harmonization

Industry codes appear in different formats and levels of aggregation across files. I harmonize all NAICS labels by normalizing the underlying text and standardizing numeric strings, and then merge the cleaned records against a manually curated NAICS mapping table. This table specifies both exact codes and two-digit ranges, with range entries expanded to cover the full set of corresponding sectors. When English titles are present for a mapped range, they are propagated to all sectors covered by that range. The resulting variable `naics_harmonized` combines observed and mapped identifiers into a single consistent classification used throughout the analysis. Further details on the mapping procedure and its implementation are provided in Section G.

### 3.4 Geographic Layer (Municipios)

I build a municipios geometry layer for Puerto Rico by filtering TIGER/Line county shapefiles to the state code `STATEFP=72`, ensuring one observation per municipality. I standardize identifiers (`STATEFP`, `COUNTYFP`, `GEOID`) and rename the municipality name field to `NAME`, trimming whitespace while preserving official casing. I de-duplicate on `GEOID`, sort by `GEOID` to stabilize joins, and assert exactly 78 municipios post-processing.

The canonical layer is saved as a GeoPackage (layer name `municipios`), with a companion table recording attribute metadata and the coordinate reference system used. To compute area and perimeter diagnostics, and to support distance-based spatial weights, I project geometries to the Puerto Rico StatePlane coordinate system (EPSG:32161).<sup>1</sup> The resulting geometry layer forms the spatial foundation for the contiguity and distance-band weights used in the exposure mappings described in Section 4. Geometry validation and repair procedures follow standard OGC and GEOS practices and are documented in Section G.<sup>2</sup>

I adopt a coastline-respecting contiguity representation and compute centroid-based distance metrics in the local StatePlane system to enable robustness checks based on distance-band weights, consistent with diffusion on a bounded island domain (Kikuchi, 2025).

---

<sup>1</sup>EPSG Geodetic Parameter Registry (2025) defines EPSG:32161—NAD83 / Puerto Rico & Virgin Islands, units in meters. This projection minimizes distortion for distance-based spatial weights.

<sup>2</sup>Administrative boundary semantics follow U.S. Census Bureau (2025), which defines `STATEFP`, `COUNTYFP`, and `GEOID` for county-level shapefiles.

### 3.5 Panel Construction

I aggregate the harmonized industry data to the municipality–industry–quarter level, denoted (GEOID, naics\_code, quarter), where GEOID follows the 5-digit FIPS format (72xxx). Before aggregation, I apply exclusion filters to remove multi-municipality aggregates and unclassified observations. Specifically, I drop municipality codes 995 and 999 (which represent “Multi Municipio” and “No Codificado” records) and municipality names matching these labels, and exclude observations with NAICS codes 99 or 10 (unclassified establishments). Municipality codes are standardized by extracting the last three digits from the original code field and prepending “72” to construct valid 5-digit FIPS identifiers, ensuring consistency with the canonical geometry layer.

For each (GEOID, naics\_code, quarter) triplet, I compute four core outcome variables by summing over all reporting units: `establishments` (count of establishments), `covered_emp` (total covered employment), `total_wages` (aggregate payroll), and `avg_wage_level` (defined as total wages divided by covered employment when the latter is positive, and missing otherwise).

To construct real wage measures, I load the monthly CPI series (*Concatenación*) from the Puerto Rico Department of Labor and Human Resources ([Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas, 2024](#)). The CPI corresponds to the official *Índice de Precios al Consumidor en Puerto Rico* ([Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas del Trabajo, 2025b](#)), prepared by the Negociado de Estadísticas del Trabajo as part of its legal mandate under the Department’s Organic Law (Ley Núm. 15 de 1931). The index measures average changes in retail prices of goods and services consumed by all families in Puerto Rico, using a fixed-weight aggregation method with December 2006 = 100 as the base period and geometric means for price averaging since the 2010 revision. The consumption basket is derived from the *Estudio de Ingresos y Gastos del Consumidor Urbano de Puerto Rico (1999–2003)*, grouped into eight main categories: Food and Beverages, Housing, Apparel, Transportation, Medical Care, Recreation, Education and Communication, and Other Goods and Services.

I compute quarterly means by averaging the three months in each quarter and rebase the index so that the 2020 average equals 100. Nominal wages are then deflated to obtain `total_wages_real_2020` and `avg_wage_level_real_2020`. All outcome variables (establishments, employment, nominal wages, and real wages) are log-transformed where positive, yielding `log_establishments`, `log_covered_emp`, `log_total_wages`, and `log_total_wages_real_2020`. To ensure a balanced time structure, I construct a complete skeleton by crossing all observed (GEOID, naics\_code) pairs with the full quarterly index from Section 3.2. This guarantees that every municipality–industry pair that appears in

the data has a record for every quarter from 2014Q1 to 2025Q1, with missing observations explicitly represented as NA. A stable industry label (`industry_name`) is retained for each `naics_code` by selecting the most frequently observed English industry name, which is used in figures and tables. The resulting panel serves as the base dataset for all subsequent treatment assignment, exposure mapping, and estimation stages.

### 3.6 Heterogeneity Strata

The heterogeneity analysis relies on time-invariant strata defined at the municipality–industry level (`GEOID`, `naics_code`). I construct three such strata: (i) tradable vs. non-tradable sectors, (ii) metro vs. non-metro municipios, and (iii) high- vs. low-wage sectors. These labels are computed once from the pre-period data and then merged onto the full panel.

**Tradable vs. non-tradable sectors.** Using the harmonized NAICS codes from Section 3.3, I map each industry to its 2-digit NAICS sector and classify sectors as tradable or non-tradable. Sectors 11, 21, 31–33, 42, 48–49, 51, 52, 54, and 55 are treated as tradable, reflecting goods and services that can be supplied across regions. Sectors 22, 23, 44–45, 53, 56, 61, 62, 71, 72, 81, 92, and 99 are treated as non-tradable, reflecting local services and public administration. Any remaining 2-digit sectors not appearing in this list are conservatively assigned to the non-tradable group. The resulting tradable indicator is constant over time and across municipios for a given industry.

**Metro vs. non-metro municipios.** I classify municipios as metro or non-metro using Core Based Statistical Area delineations from OMB Bulletin No. 23-01 ([Office of Management and Budget, 2023](#)). This table contains a binary metro indicator, which I merge onto the panel by municipality identifier `GEOID`. Municipios lacking a valid classification are treated as non-metro. The metro status is thus time-invariant and shared by all industries within a municipio.

**High- vs. low-wage sectors.** Finally, I construct wage-based strata using the real average wage level `avg_wage_level_real_2020` defined in Section 3.5. I restrict attention to the pre-program period 2014Q1–2019Q3 and, for each industry  $k$ , compute the median of `avg_wage_level_real_2020` across all municipios and quarters in this window. Let  $\tilde{w}_k$  denote this industry-level pre-period median and let  $\tilde{w}$  denote the median of  $\{\tilde{w}_k\}$  across all industries with sufficient data. Industries with  $\tilde{w}_k \geq \tilde{w}$  are classified as high-wage and the remainder as low-wage. Industries with insufficient pre-period wage data are assigned

to a residual “unknown” group, which I exclude from the high vs. low comparisons in the heterogeneity analysis.

### 3.7 Auxiliary Data for TMO Analysis

To construct the auxiliary outcomes required for the Thresholding Multiple Outcomes (TMO) sensitivity analysis ([Della Vigna et al., 2025](#)), I build a separate municipal-level panel using tables from the ACS 5-year estimates for 2013–2023 ([U.S. Census Bureau, 2023](#)). These tables provide socioeconomic metrics on demographics, employment, and housing. I construct a set of 44 indicators, compute year-over-year changes, and upsample the resulting annual series to the quarterly frequency of the main analysis. This auxiliary panel is used solely to estimate the cross-location correlation matrix that enters the TMO adjustment; further implementation details are provided in [Section 8](#) and [Section G](#).

### 3.8 Descriptive Statistics and Sample Composition

[Table 1](#) summarizes the pre-program distribution of outcomes at the municipality–industry–quarter level over 2014Q1–2019Q3. Panel A reports means and standard deviations of the logarithms of covered employment, establishments, and the real wage bill, highlighting that municipality–industry cells are typically small but exhibit substantial dispersion, with a heavy upper tail of large sectors. Panel B shows that a large majority of municipality–industry cells belong to tradable sectors, most are located in metro municipios, and high- and low-wage industries are roughly evenly split, indicating that the balanced panel covers a broad cross-section of Puerto Rico’s local production structure.

[Figure 1](#) maps the canonical municipios geometry and highlights the metro versus non-metro split used in the heterogeneity analysis. Metro municipios, as defined by OMB Bulletin 23-01, cluster along the north coast and southern corridor, while non-metro municipios are concentrated in the central mountainous region and offshore islands. This same geography underlies the contiguity matrix used to construct the spatial weights in [Section 4](#).

Metro vs. Non-Metro Municipios (Figure D1)

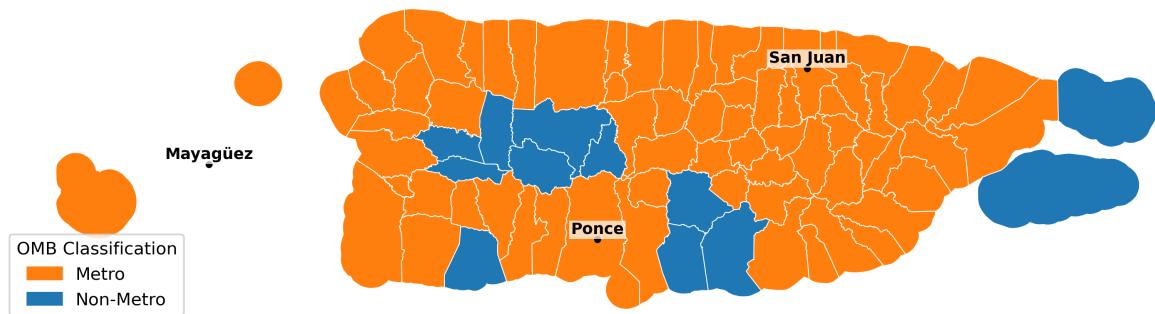


Figure 1: Metro vs. Non-Metro Municipios, OMB Classification

*Notes:* The figure shows Puerto Rico's 78 municipios using the canonical geometry layer described in Section 3.4. Metro municipios, defined using OMB Bulletin 23-01, are shaded in one color and non-metro municipios in another; San Juan, Ponce, and Mayagüez are labeled for orientation. This geography underlies the contiguity-based spatial-weights matrix and the metro/non-metro heterogeneity split used in subsequent sections.

Table 1: Descriptive Statistics and Strata Counts, Pre-Period 2014Q1–2019Q3

	Mean	SD	Observations
<b>Panel A: Municipality–Industry–Quarter Outcomes (Logs)</b>			
log(covered employment)	5.026	1.771	73,554
log(establishments)	1.861	1.523	73,554
log(real wage bill, 2020 USD)	13.581	1.935	73,554
Number of municipios	78		
Number of municipality–industry cells	3,198		
Number of quarters	23		
<b>Panel B: Municipality–Industry Cells by Strata</b>			
<i>Tradable status</i>			
Tradable sectors	2,340		73.2%
Non-tradable sectors	858		26.8%
<i>Metro status</i>			
Metro municipios	2,706		84.6%
Non-metro municipios	492		15.4%
<i>Wage stratum</i>			
High-wage industries	1,638		51.2%
Low-wage industries	1,560		48.8%
Unknown wage industries	0		0.0%

*Notes:* Panel A reports means and standard deviations of the logarithms of covered employment, establishments, and the real wage bill across all municipality–industry–quarter observations in 2014Q1–2019Q3. The real wage bill is measured in 2020 dollars. The pre-program sample comprises 78 municipios, 3,198 municipality–industry cells, and 23 quarters, for a total of 73,554 observations. Panel B reports the number and share of municipality–industry cells in the balanced pre-period panel by tradability, metro status, and wage stratum; shares are relative to the 3,198 municipality–industry cells. In the current sample all industries have sufficient pre-period wage data, so the “Unknown wage industries” category contains zero cells and is included for completeness.

## 4 Treatment & Exposure Construction

### 4.1 Direct Treatment Definition (Large Entry Events)

Before defining exposure mappings, I first define the direct treatment event  $G_{ik}$ . The unit of analysis is a municipality–industry pair  $(i, k)$ , and “treatment” is a large entry event derived from establishment counts, employment, and real wage-bill data (`total_wages_real_2020`; Section 3). In the baseline, the event is the first quarter  $t$  in which  $(i, k)$  satisfies any of three establishment-based persistent trigger conditions (T1–T3).

1. **T1 (New Entry):** Zero establishments for at least the previous eight quarters ( $\ell \in [-8, -1]$ ), then  $\geq 1$  establishment in quarter  $t$  and again in  $t+1$ .
2. **T2 (Small-to-Large Growth):**  $\leq 1$  establishment in  $t-1$ , then  $\geq 2$  in  $t$  and again in  $t+1$ .
3. **T3 (Top-Decile Establishment Jump):** The quarter-over-quarter change in establishments ( $\Delta \text{est}_t$ ) is at least the 90th percentile of such changes within the same NAICS during 2014Q1–2019Q3, with a minimum absolute increase of one. Persistence requires a non-decrease in the next quarter:  $\text{est}_{t+1} \geq \text{est}_t$ .

When multiple rules are met in the same quarter, I apply a fixed priority order T1 > T2 > T3. The cohort identifier  $G_{ik}$  is defined as the earliest quarter that satisfies any of T1–T3. The persistence requirements distinguish durable entries from temporary fluctuations. I interpret these large, persistent jumps as first sizable entries that proxy for major expansions of industry presence in a municipality. The cohort variable  $G_{ik}$  underpins the staggered-adoption design used throughout the paper; formal estimands based on  $G_{ik}$  are defined in Section 5, and event-time windows are discussed in Section 7. These entry events provide the variation used to study the direct effects in Q1 and the spillover channels in Q2–Q3b.

### 4.2 Exposure Mappings (General Framework)

The spillover analysis is based on explicit exposure mappings that translate neighbors’ treatment into exposure variables for each unit. I build the channels (Expo, XExpo, and NExpo) using an adjacency matrix ( $W$ ) as the basis for a neighbor-based framework. Following Aronow et al. (2020), an *exposure mapping* is a function that maps a unit’s network connections (from an adjacency matrix) and the full treatment vector to a specific exposure value.

The exposure histories constructed from these mappings play a central role in separating direct and spillover effects. Xu (2023) shows that, in a DiD framework with interference,

identification of direct and spillover average treatment effects can be achieved under a parallel trends assumption that holds conditional on exposure histories. Similarly, [Shahn et al. \(2024\)](#) formalize how spillover effects can depend on both past and concurrent exposure, motivating my use of exposure mappings that track exposure over time.

The three channels used in this paper are necessarily a simplification of the true, unknown interference structure. [Sävje \(2023\)](#) argues that such simplification is both justified and necessary, and recommends separating the role of an exposure mapping (defining an interpretable estimand) from the impossible task of fully capturing the underlying causal structure. This perspective justifies defining Expo, XExpo, and NExpo as interpretable channels of interest that capture economically relevant forms of interference.

### 4.3 Specific Exposure Channels & Contrasts

To separately identify different spillover effects, I rely on the concept of *exposure contrasts*. [Leung \(2024\)](#) defines such contrasts as nonparametric comparisons of expected outcomes under different realizations of an exposure mapping. In my setting, this framework supports contrasts based on, for example, having high versus low exposure to treated neighbors under a given mapping.

I specify three distinct exposure channels:

- **Same-industry neighbor exposure (Expo):** exposure of  $(i, k)$  to treatment in the same industry  $k$  in neighboring municipios.
- **Within-municipality cross-industry exposure (XExpo):** exposure of  $(i, k)$  to treatment in other industries  $k' \neq k$  within the same municipio  $i$ .
- **Neighbor all-industries exposure (NExpo):** exposure of  $(i, k)$  to treatment in all industries  $k'$  (including  $k$ ) in neighboring municipios.

These mappings are designed to distinguish distinct economic channels. By defining separate exposure histories for same-industry versus cross-industry neighbors, I allow for the possibility that competitive effects within the same industry ([Slattery and Zidar, 2020](#)) and agglomeration or demand spillovers across industries ([Siegloch et al., 2025](#)) operate simultaneously but with different signs and magnitudes. Formal estimands based on exposure contrasts under each channel are introduced in Section 5.

## 4.4 W Matrix (Spatial Weights)

The spatial weight matrix  $W$  is constructed from the municipios geometry layer for Puerto Rico (Section 3.4). I use a primary contiguity matrix that begins with Queen contiguity (polygon-based adjacency) and supplements it with  $k = 3$  nearest-neighbor (KNN) links for municipios that are geographically isolated under pure contiguity. These additional edges are added only for isolates, following a “Queen + KNN plug-in” rule.

After adding these links, the neighbor graph is made symmetric (if  $i$  is a neighbor of  $j$ , then  $j$  is a neighbor of  $i$ ), and the resulting weight matrix is then row-standardized. Row-standardization implies that each row of  $W$  sums to one, so exposure variables formed as weighted averages of neighbors’ treatment can be interpreted as mean treatment among a unit’s neighbors. The underlying neighbor relation is undirected, even though the row-standardized matrix need not be numerically symmetric. These weights serve as the basis for the same-industry and all-industries neighbor exposure measures described above.

## 4.5 Lagged Exposure Histories ( $L = 4$ )

In addition to contemporaneous exposure, I summarize each channel’s exposure over time using lagged histories. For each exposure channel (direct own treatment, within-municipality cross-industry exposure, and neighbor-all exposure), I construct two history summaries:

- an “*any-since-adoption*” exposure share, which records the fraction of post-entry periods up to time  $t$  in which the unit (or its neighbors, under the relevant channel) has been exposed; and
- a “*last four post-treatment quarters*” share, which records the average exposure over the most recent  $L = 4$  post-entry quarters.

These summaries are defined analogously across channels, with the underlying exposure at each date determined by the relevant mapping (Expo, XExpo, or NExpo). The resulting history variables are used both to define estimands that depend on exposure duration and to condition on exposure history in the identification and estimation procedures. Their role in the identifying assumptions and doubly-robust estimation strategy is detailed in Sections 6 and 7.

## 4.6 Timing and Geography of First Sizable Entries

The treatment rule in Section 4.1 yields 1,871 municipality–industry cells with a first sizable entry over 2014Q2–2024Q4. Figure 2 plots the distribution of cohort dates  $G_{ik}$  by calendar

quarter. Entries are heavily front-loaded: the bulk of first sizable entries occur between 2014 and 2016, after which the number of new treated municipality–industry cells declines steadily and falls to relatively low levels by the early 2020s. There are no first sizable entries in the initial quarter 2014Q1 or after 2024Q4, consistent with the maturing of the industrial structure over the sample window.

Figure 3 maps the cumulative number of first sizable entries by municipio. Every municipio experiences at least ten entry events, but the intensity of entry is unevenly distributed across space: municipios receive between 10 and 40 first sizable entries, with a median of 24 and an interquartile range of 20 to 27 entries. The map bins municipios into five quintile groups based on their entry counts (10–18, 19–22, 23–26, 27–28, and 29–40 entries), highlighting the concentration of entering industries in coastal and metropolitan municipios and the relatively lower intensity in many interior municipios. This geographic heterogeneity motivates the explicit modeling of spatial spillovers through the exposure mappings introduced above and the spillover estimands and identification strategy developed in Sections 5–6.

Tables 2 and 3 decompose treated municipality–industry cells by trigger type. About three quarters of first sizable entries arise from large, top-decile establishment jumps (T3), while the remainder are split between “new” entries (T1) from previously empty cells and “small-to-large” expansions (T2) that move from at most one establishment to at least two. Roughly one third of treated cells under each trigger contribute to the balanced  $[-8, +16]$  event-time window used in the main event-study analysis, and the median pre-period employment levels differ sharply across trigger types. In particular, pre-period median covered employment is about five workers for T1 cells, eighteen workers for T2 cells, and 151 workers for T3 cells, indicating that top-decile establishment jumps tend to occur in substantially larger municipality–industry cells. These diagnostics help assess whether the composition of treated cells differs systematically across trigger types and whether the balanced-window sample is skewed toward particular kinds of entries.

Figure 4 provides a simple “validity check” on the treatment definition by plotting average establishment counts around the entry event for each trigger type, using only municipality–industry cells that belong to the balanced window. For T1 (new entries), average establishments are essentially zero in the eight quarters before  $t = 0$  and jump to roughly four establishments at the event quarter, remaining persistently positive thereafter. T2 (small-to-large growth) cells move from about one establishment pre-event to around two post-event. For T3 (top-decile jumps), average establishments are flat at roughly 26 in the pre-period and then jump sharply to nearly 80 at  $t = 0$ , staying elevated for several quarters. These patterns are consistent with the intended interpretation of the triggers as large, persistent changes in local industry presence rather than transitory noise.

Figure T1: Timing of First Sizable Entries (2014–2025)

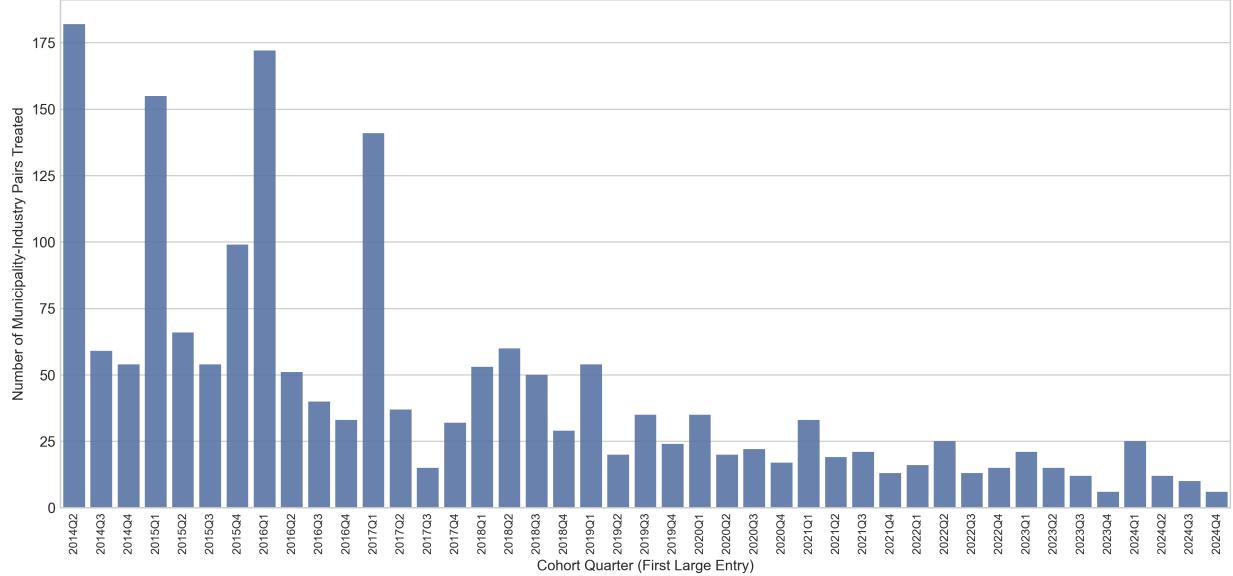


Figure 2: Number of First Sizable Entry Events by Cohort Quarter, 2014–2024

*Notes:* The figure plots the number of municipality–industry cells  $(i, k)$  whose first sizable entry date  $G_{ik}$  falls in each calendar quarter from 2014Q2 to 2024Q4. Treatment events are defined using the establishment-based trigger rules in Section 4.1. The shaded region marks the pre-program period 2014Q1–2019Q3 used for defining exposure histories and pre-period diagnostics; there are no first sizable entries in 2014Q1.

Figure T2: Intensity of First Sizable Entries by Municipio

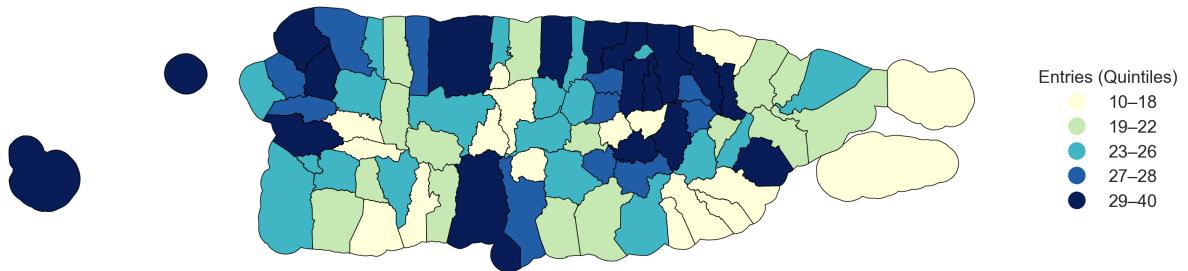


Figure 3: Intensity of First Sizable Entries by Municipio

*Notes:* The figure maps Puerto Rico's 78 municipios and shades each according to the total number of industries  $k$  with a first sizable entry  $G_{ik}$  between 2014Q2 and 2024Q4. Shading corresponds to quintiles of the municipio-level entry count with bins 10–18, 19–22, 23–26, 27–28, and 29–40 entries; darker shades indicate municipios with more entering industries. The underlying geometry and contiguity structure match the spatial layer described in Section 3.4 and the spatial-weights matrix in Section 4.4.

Table 2: Composition of First Sizable Entry Triggers: Counts and Shares

Trigger type	Treated cells	Share of treated (%)
T1: New entry	240	12.8
T2: Small-to-large growth	210	11.2
T3: Top-decile establishment jump	1,421	76.0
Total	1,871	100.0

*Notes:* This table reports the number of treated municipality–industry cells and their share of all treated cells by trigger type. T1 events are new entries from at least eight quarters of zero establishments. T2 events move from at most one establishment to at least two with persistence. T3 events are top-decile establishment jumps with persistence, as defined in Section 4.1.

Table 3: Composition of First Sizable Entry Triggers: Balanced Window and Pre-Period Employment

Trigger type	Share in balanced window (%)	Median pre-period log(cov. emp)	Approx. pre-period cov. emp.
T1: New entry	37.3	1.61	5
T2: Small-to-large growth	35.8	2.89	18
T3: Top-decile establishment jump	33.4	5.02	151
All triggers	34.2	—	—

*Notes:* “Share in balanced window” is the fraction of treated municipality–industry cells under each trigger that have a complete  $[-8, +16]$  event-time window. Median pre-period  $\log(\text{covered employment})$  is computed over 2014Q1–2019Q3 for each trigger type. The last column shows approximate employment levels obtained by exponentiating the medians and rounding to the nearest worker.

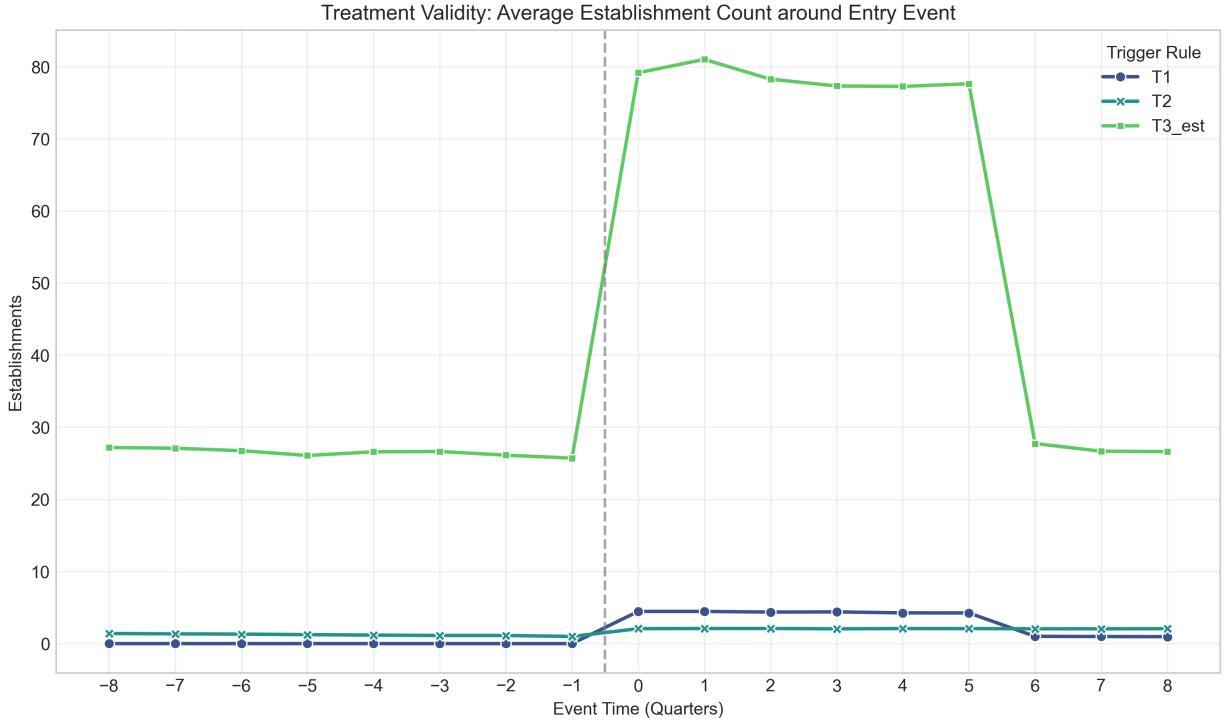


Figure 4: Average Establishment Count Around First Sizable Entry, by Trigger Type

*Notes:* The figure plots mean establishment counts by event time  $\ell \in [-8, 8]$  for municipality-industry cells in the balanced window, separately by trigger type (T1, T2, T3). The dashed vertical line between  $\ell = -1$  and  $\ell = 0$  marks the event quarter. T1 (new entries) and T2 (small-to-large growth) show discrete increases in establishments at  $\ell = 0$  from very low pre-period levels, while T3 (top-decile jumps) exhibits a large, persistent increase from roughly 26 establishments pre-event to around 80 post-event, consistent with the interpretation of these triggers as sizable, durable entry events.

## 5 Estimands

I next formalize the treatment effect parameters that correspond to the research questions in Section 1. For each municipality–industry cell  $(i, k)$ , let  $G_{ik}$  denote the first quarter of a sizable entry (Section 4.1) and let  $\ell = t - G_{ik}$  be time since entry. The primary direct-effect target is the dynamic path  $\text{ATT}^{\text{own}}(\ell)$ , which measures the average effect of a first sizable entry on the three main log outcomes—`log_establishments`, `log_covered_emp`, and the real wage bill `log_total_wages_real_2020` ( $2020 = 100$ )—in the treated municipality–industry cell at horizon  $\ell$  (Q1). I also define slice-level direct and spillover effects that summarize the average impact over post-treatment windows such as 0–4, 5–8, 9–16, and 0–16 quarters, corresponding to the “policy slices” used to summarize cumulative gains (Q5).

Following [Callaway and Sant'Anna \(2021\)](#), I use the *group-time average treatment effect*, denoted  $\text{ATT}(g, t)$ , as the basic building block and aggregate these effects across cohorts sharing the same relative event time  $\ell = t - g$  to obtain  $\text{ATT}^{\text{own}}(\ell)$ . I then extend this framework to spillover channels by defining exposure-based slice estimands that hold own treatment fixed while varying exposure to neighbors’ entries along three pre-specified channels: same-industry neighbors, within-municipality cross-industry, and neighbor all-industries (Q2–Q3b). Unless stated otherwise, these dynamic and slice-level parameters are defined with unit weights over municipality–industry cells: the aggregation weights place equal mass on each treated municipality–industry cell, rather than, for example, employment weights, so the target population is the average treated municipality–industry cell rather than the average worker.

To explicitly parameterize treatment effects by horizon, I rely on [Borusyak et al. \(2024\)](#), who define the *average treatment effect  $h$  periods since treatment* as a common target in event-study analyses. Their formulation clarifies that each dynamic effect can be viewed as a weighted average of underlying potential-outcome contrasts, with the weights depending on the timing structure of treatment adoption. This representation complements the aggregation logic in [Callaway and Sant'Anna \(2021\)](#) and emphasizes that  $\text{ATT}^{\text{own}}(\ell)$  can be interpreted as an ATT-by-horizon object.

The concept of cohort-time treatment effects is developed by [Sun and Abraham \(2021\)](#), who define cohort-specific average treatment effects on the treated and show that conventional two-way fixed effects (TWFE) event-study estimators confound these effects across groups when treatment effects are heterogeneous. Their framework motivates the use of heterogeneity-robust event-time estimands such as  $\text{ATT}^{\text{own}}(\ell)$ , which aggregate cohort-time effects using interpretable and convex weighting schemes.

## 5.1 Direct ATT ( $\text{ATT}^{\text{own}}(\ell)$ )

In the context of Q1, for a given outcome  $Y$  (log covered employment, log establishments, or log real total wages),  $\text{ATT}^{\text{own}}(\ell)$  can be read as the average percentage effect on  $Y$  in a municipality–industry cell that experienced a first sizable entry  $\ell$  quarters ago, relative to what would have happened in the absence of such an entry. The cumulative quantity

$$\text{ATT}_{\text{cumu}}^{\text{own}}(0, 16) = \sum_{\ell=0}^{16} \text{ATT}^{\text{own}}(\ell)$$

summarizes the total impact over the first 16 post-entry quarters and is the main direct-effect object used in the budgeting and return-on-investment discussion in Q4.

For notational simplicity, I suppress the industry index  $k$  and write  $i$  for the municipality–industry cell  $(i, k)$ .

Formally, let  $Y_{i,t}(1)$  and  $Y_{i,t}(0)$  denote the potential outcomes for unit  $i$  at time  $t$  with and without a first sizable entry at  $G_i = g$ . The primary estimand of interest is the *direct* average treatment effect on the treated (ATT) at event-time horizon  $\ell$ , denoted  $\text{ATT}^{\text{own}}(\ell)$ . Following [Callaway and Sant'Anna \(2021\)](#), I define this as

$$\text{ATT}^{\text{own}}(\ell) \equiv \mathbb{E}[Y_{i,g+\ell}(1) - Y_{i,g+\ell}(0) \mid G_i = g],$$

with aggregation across cohorts sharing the same  $\ell$  implemented via the group-time ATT structure. Each  $\text{ATT}(g, t)$  measures the average effect for units first treated in period  $g$  at time  $t$ , and aggregating across cohorts with the same event-time  $\ell$  yields the dynamic event-study effect  $\text{ATT}^{\text{own}}(\ell)$ .

This cohort-time perspective parallels the ATT-by-horizon notation of [Borusyak et al. \(2024\)](#), in which dynamic effects are indexed directly by the number of periods since treatment. In both cases, the estimand at horizon  $\ell$  is a weighted average of potential-outcome contrasts for cohorts that have been treated for  $\ell$  periods, with weights determined by the distribution of adoption dates. The cumulative  $\text{ATT}_{\text{cumu}}^{\text{own}}(0, 16)$  can be viewed as a summary of these horizon-specific effects over the post-entry window of interest. Practical implementation choices, including reference periods and limited-anticipation windows, are described in Sections 6 and 7.

## 5.2 Neighbor & Cross-Industry ATTs (Spillover Estimands)

The spillover estimands answer Q2–Q3b. For a given slice  $[a, b]$ ,  $\text{DATT}[a, b]$  captures the average direct effect of own entry over horizons  $\ell \in [a, b]$  on the treated municipality–industry

cell, while  $SATT_{\text{same}}[a, b]$ ,  $SATT_{\text{cross}}[a, b]$ , and  $SATT_{\text{all}}[a, b]$  capture the average effects of increasing exposure along the same-industry neighbor, within-municipality cross-industry, and neighbor all-industries channels, respectively, holding own treatment fixed. Thus:

- $SATT_{\text{same}}[a, b]$  corresponds to Q2, measuring how outcomes in the same industry respond to neighboring municipios' entries over the slice;
- $SATT_{\text{cross}}[a, b]$  corresponds to Q3a, capturing cross-industry spillovers within the same municipality;
- $SATT_{\text{all}}[a, b]$  corresponds to Q3b, capturing spillovers to all industries in bordering municipios.

The aggregate spillover  $SATT[a, b]$  and total effect  $TATT[a, b]$  then decompose the post-entry impact into direct and spillover components that feed into the cumulative-value calculations in Q4.

I adopt an *exposure-mapping* approach in which neighbors' assignments are summarized by low-dimensional exposure variables and their histories (Section 4), and I define spillover estimands as exposure *contrasts* that hold a unit's own treatment fixed. Conceptually, these correspond to the “indirect” or “spillover” effects under partial interference (Liu et al., 2019), implemented here in a staggered DiD setting with exposure mappings that vary across space, time, and industries.

**Slices and notation.** Let the event-time slices be  $[a, b] \in \{[0, 4], [5, 8], [9, 16], [0, 16]\}$  and denote the slice length by  $L_s = b - a + 1$ . For each slice, I define an own-treatment indicator

$$D_{\text{slice}} = \mathbf{1}\{t - g \in [a, b], A_{i,t} = 1\},$$

where  $A_{i,t}$  is the indicator of own treatment for unit  $i$  at time  $t$ . For the spillover channels, I summarize exposure over the slice by *summing* the underlying per-period exposure processes. For example, with  $E_{i,t}^{\text{same}}(\ell)$  denoting the same-industry neighbor exposure at event time  $\ell$ ,

$$S_{\text{same,slice}} = \sum_{\ell \in [a, b]} E_{i,t}^{\text{same}}(\ell),$$

and analogous definitions apply to  $S_{\text{cross,slice}}$  and  $S_{\text{all,slice}}$ , pooling across source industries where appropriate (consistent with the exposure construction in Section 4). Because each  $E_{i,t}(\ell)$  is a share in  $[0, 1]$ , the slice sums  $S_{\cdot, \text{slice}}$  lie in  $[0, L_s]$ . A one-unit increase in, for example,  $S_{\text{same,slice}}$  corresponds to an additional quarter of full exposure; moving from zero exposure to 100% exposure in every quarter of the slice corresponds to an  $L_s$ -unit increase, so the effect of

that full-slice contrast is  $L_s$  times the corresponding slice coefficient. For overlap diagnostics and the  $\epsilon$ -trimming rule, I also work with slice-average shares  $\bar{S}_{\cdot,\text{slice}} = S_{\cdot,\text{slice}}/L_s \in [0, 1]$  and implement trimming in terms of these averages. In estimation, the slice sums  $S_{\cdot,\text{slice}}$  enter the doubly robust DiD regressions as the main spillover regressors, while the slice averages  $\bar{S}_{\cdot,\text{slice}}$  are used solely for overlap diagnostics and the  $\epsilon$ -trimming rule.

**Slice-level spillover estimands.** I define direct and channel-specific spillover targets as average *partial* effects within a slice, holding the other channels fixed. Let  $Y_{i,t}(d, s_{\text{same}}, s_{\text{cross}}, s_{\text{nall}})$  denote the potential outcome for unit  $i$  at time  $t$  under own-treatment status  $d$  and exposure slices  $(s_{\text{same}}, s_{\text{cross}}, s_{\text{nall}})$ . Then

$$\text{DATT}[a, b] \equiv \mathbb{E} \left[ Y_{i,t}(1, S_{\text{same},\text{slice}}, S_{\text{cross},\text{slice}}, S_{\text{nall},\text{slice}}) - Y_{i,t}(0, S_{\text{same},\text{slice}}, S_{\text{cross},\text{slice}}, S_{\text{nall},\text{slice}}) \mid t - g \in [a, b] \right],$$

$$\text{SATT}_{\text{same}}[a, b] \equiv \mathbb{E} \left[ Y_{i,t}(0, S_{\text{same},\text{slice}}+1, S_{\text{cross},\text{slice}}, S_{\text{nall},\text{slice}}) - Y_{i,t}(0, S_{\text{same},\text{slice}}, S_{\text{cross},\text{slice}}, S_{\text{nall},\text{slice}}) \mid t - g \in [a, b] \right],$$

$$\text{SATT}_{\text{cross}}[a, b] \equiv \mathbb{E} \left[ Y_{i,t}(0, S_{\text{same},\text{slice}}, S_{\text{cross},\text{slice}}+1, S_{\text{nall},\text{slice}}) - Y_{i,t}(0, S_{\text{same},\text{slice}}, S_{\text{cross},\text{slice}}, S_{\text{nall},\text{slice}}) \mid t - g \in [a, b] \right],$$

$$\text{SATT}_{\text{nall}}[a, b] \equiv \mathbb{E} \left[ Y_{i,t}(0, S_{\text{same},\text{slice}}, S_{\text{cross},\text{slice}}, S_{\text{nall},\text{slice}}+1) - Y_{i,t}(0, S_{\text{same},\text{slice}}, S_{\text{cross},\text{slice}}, S_{\text{nall},\text{slice}}) \mid t - g \in [a, b] \right].$$

Because the slice exposures  $S_{\cdot,\text{slice}}$  are defined as sums of per-period exposure shares over  $L_s$  quarters, a one-unit change in  $S_{\cdot,\text{slice}}$  should be interpreted as the effect of an additional quarter of full exposure, while the effect of moving from zero exposure to full exposure in all  $L_s$  quarters of the slice is  $L_s$  times the corresponding slice-level coefficient. I summarize total spillovers by

$$\text{SATT}[a, b] = \text{SATT}_{\text{same}}[a, b] + \text{SATT}_{\text{cross}}[a, b] + \text{SATT}_{\text{nall}}[a, b],$$

and the total effect by

$$\text{TATT}[a, b] = \text{DATT}[a, b] + \text{SATT}[a, b].$$

These targets fit within the exposure-contrast perspective of Aronow et al. (2020) and Leung (2024), where spillover effects are defined as differences in mean outcomes across exposure levels, holding own treatment fixed. In my application, the exposure mappings and histories that enter these definitions are described in Section 4, and the assumptions under which the slice-level DATT and SATT objects admit a causal interpretation are discussed in Section 6. Estimation strategies that deliver consistent estimators of these parameters, including the doubly robust DiD procedures with exposure histories, are presented in Section 7.

## 6 Identification

### 6.1 Parallel Trends Conditional on Exposure Histories

I follow Xu (2023) in adopting a *parallel trends under interference, conditional on exposure histories* assumption as the core identifying restriction. Intuitively, conditional on past exposure histories, potential outcome trends would have evolved in parallel across exposure groups in the absence of treatment. This corresponds to Assumptions 3 and 3' in Xu (2023), which generalize the classical difference-in-differences (DiD) assumption to settings with interference.

To formalize this idea, I use the terminology of Shahn et al. (2024), who define a *conditional parallel trends indexed by exposure histories* assumption and connect it to doubly robust DiD estimators. In the special case of no interference, this formulation nests the standard DiD setting of Callaway and Sant'Anna (2021), where conditioning is on pre-treatment covariates rather than on exposure paths.

Under this condition, the direct and spillover estimands defined in Section 5 admit a causal interpretation. When the exposure mapping is misspecified, I follow Xu (2023) and Sävje (2023) in interpreting the direct path as an *expected direct ATT (EDATT)* and the spillover targets as *expected exposure contrasts* with respect to the chosen mappings. In practice, the exposure histories that index this assumption enter the estimator via a small set of exposure-history summaries (for example, “any-since-adoption” and “last four post-treatment quarters” exposure shares) that are included as controls in the nuisance functions, rather than via the full exposure path. Estimation procedures that implement this assumption using doubly robust DiD and cross-fitting are described in Section 7.

## 6.2 Limited Anticipation

I assume a limited anticipation window of  $\delta$  quarters (baseline:  $\delta = 2$ , robustness:  $\delta = 1$ ). The economic content of this restriction is that agents may adjust in a small number of pre-treatment quarters, but not arbitrarily far in advance of a first sizable entry. Following Callaway and Sant'Anna (2021), I operationalize this assumption by using a cohort-specific pre-treatment period as the reference outcome, separated from the treatment date by a buffer of  $\delta$  quarters.

For a unit first treated in quarter  $g$ , event-time contrasts at horizon  $\ell$  compare outcomes  $\ell$  periods after adoption to outcomes in this cohort-specific reference period. Anticipatory leads  $\ell \in \{-\delta, \dots, -1\}$  are treated as diagnostics for pre-trends rather than as identifying moments for the dynamic effect path. Implementation details on the choice of reference period and the handling of anticipatory leads in estimation and inference are reported in Sections 7 and 8.

## 6.3 Interference Channels

Following the exposure-mapping framework of Aronow et al. (2020), I assume that interference operates through three pre-specified channels and that no other spillover paths are present. Each unit's exposure depends on its own treatment and on the treatment assignments of its neighbors within a specified network. I adopt the *partial interference* assumption of Liu et al. (2019), which confines spillovers to predefined groups or networks rather than allowing unrestricted dependence.

In the primary specification, neighborhood interference is implemented at a single hop ( $h = 1$ ): municipios are affected by the treatment status of their first-order neighbors in the contiguity graph and not by more distant locations. Multi-hop exposure measures are computed for diagnostics but are not used to define the primary estimands. The three channels are:

1. same-industry neighbor exposure (Expo), capturing exposure of  $(i, k)$  to entries in the same industry  $k$  in neighboring municipios;
2. within-municipality cross-industry exposure (XExpo), capturing exposure of  $(i, k)$  to entries in other industries  $k' \neq k$  within the same municipio  $i$ ;
3. neighbor all-industries exposure (NExpo), capturing exposure of  $(i, k)$  to entries in all industries  $k'$  (including  $k$ ) in neighboring municipios.

These channels are designed as interpretable exposure contrasts in the sense of Leung (2024) and correspond to the economic mechanisms discussed in Section 2. They underlie

the spillover estimands introduced in Section 5 and are constructed from the spatial weights matrix  $W$  described in Section 4.4.

## 6.4 Exposure Positivity and Overlap

Identification additionally requires *exposure positivity*. Intuitively, for each exposure contrast of interest, there must be sufficient variation in exposure levels among otherwise comparable units so that counterfactual outcomes are well defined and estimable. I follow Xu (2023) and Shahn et al. (2024) in treating overlap in exposure histories as necessary for DiD with interference.

In practice, I enforce exposure positivity by trimming observations in regions of very low or very high exposure where empirical support is weak. Specifically, I bound the slice-average exposure shares away from zero and one, retaining only units whose exposure lies in an interior interval  $[\epsilon, 1 - \epsilon]$  for each channel and slice, with a small  $\epsilon > 0$ . This trimming rule is motivated by the overlap conditions in Crump et al. (2009) and the discussion of exposure-positivity in Aronow et al. (2020), Sävje (2023), and Xu (2023). The exact choice of  $\epsilon$  and the implementation of trimming and minimum-sample thresholds are described in Sections 7 and 9.

## 6.5 Sensitivity to Violations of Parallel Trends

The assumptions above provide a baseline causal interpretation for the direct and spillover estimands, but may be violated in practice. To assess robustness to deviations from parallel trends, I use the *Honest DiD* framework of Rambachan and Roth (2023), which constructs robust confidence sets under explicit restrictions on how much the path of treatment effects can deviate from parallel trends in pre-treatment periods.

I focus on the smoothness restriction  $\mathcal{SD}(M)$ , which bounds the curvature (second differences) of the effect path and is well suited to settings with potentially accelerating dynamic responses. This choice aligns with recommendations in recent DiD syntheses (Roth et al., 2023) and applied evaluations (Wang et al., 2024). Following the suggestion in Xu (2023), I interpret these sensitivity analyses in the presence of interference as robustness checks for both the parallel trends and exposure-mapping assumptions. The implementation of Honest DiD, including the grid of  $M$  values and the visualization of robust bounds, is described in Section 9 and revisited in the robustness analysis in Section 13.

## 6.6 Summary of Key Identification Assumptions

For clarity, I summarize the main assumptions underlying the identification of direct and spillover effects:

- A1: Direct-effects parallel trends (no interference).** In the absence of any first sizable entries (for all municipality–industry cells), municipality–industry cells that eventually experience a first sizable entry and those that do not would have followed parallel trends within industry. This corresponds to the standard difference-in-differences assumption for the direct path in the non-interference benchmark and underlies the own-treatment ATT estimands discussed in Sections 5 and 7. Assumption **A2** below generalizes this benchmark to allow for interference via exposure histories.
- A2: Parallel trends conditional on exposure histories (with interference).** In the presence of interference, potential outcome trends are assumed to evolve in parallel across exposure groups *conditional on past exposure histories and the observed controls used in the nuisance functions*, as in Xu (2023) and Shahn et al. (2024). This assumption justifies the use of doubly robust DiD with exposure-history summaries (for example, “any-since-adoption” or “last four post-treatment quarters” exposure shares) and fixed effects entering the nuisance functions for the spillover estimands.
- A3: Limited anticipation.** Agents may adjust outcomes in the last few pre-treatment quarters, but there is no systematic anticipation earlier than the limited-anticipation window ( $\delta$  quarters). Identification is implemented by differencing outcomes relative to the cohort-specific reference period  $g - \delta - 1$  and excluding anticipatory leads  $\ell \in \{-\delta, \dots, -1\}$  from the uniform-band support set, as described in Sections 7 and 8.
- A4: Restricted interference structure.** Spillovers operate only through the three pre-specified exposure channels—same-industry neighbors, within-municipality cross-industry, and neighbor all-industries—defined on a queen-contiguity plus KNN graph with a single-hop neighborhood ( $h = 1$ ). I do not model higher-order or long-range network effects beyond this structure (Section 4.4).
- A5: Exposure positivity and overlap.** Within each event-time slice and channel, there is sufficient overlap in exposure shares across units to support the estimation of exposure contrasts. In practice, this is enforced by two-sided  $\epsilon$ -trimming of slice-average exposure shares (default  $\epsilon = 0.02$ ), the balanced-horizon restriction, and minimum post-trim sample thresholds, as detailed in Sections 7, 9, and G.3.

**A6: Common shocks controlled by fixed effects.** Time-varying shocks that affect municipality–industry cells within an industry or across Puerto Rico are assumed to be absorbed by the included unit and time fixed effects (and, where relevant, industry-by-time fixed effects) and by the observed exposure histories. Identification rules out additional unobserved shocks that would systematically shift trends for treated or highly exposed units relative to their comparison units, beyond what is captured by these controls.

## 7 Estimation Strategy

Across all specifications, I treat `log_covered_emp` as the primary outcome. Complementary outcomes include `log_establishments` and the log real wage bill, `log_total_wages_real_2020`. Nominal wage measures, including `log_total_wages`, are used only for internal checks and are not reported as main outcomes.

### 7.1 Balanced vs. Unbalanced Horizons

I distinguish between balanced and unbalanced horizons in event-time aggregations following [Callaway and Sant'Anna \(2021\)](#). A *balanced* horizon restricts estimation to event times that are identified for all treated cohorts, avoiding composition effects from treatment-timing heterogeneity. [Borusyak et al. \(2024\)](#) define balanced aggregation as one where each horizon is estimated from the same set of units, emphasizing efficiency and unbiasedness. In contrast, unbalanced aggregations allow varying support across event times, which can distort dynamic patterns. As shown by [Dube et al. \(2025\)](#), restricting to balanced horizons mitigates composition biases common in naive event studies. I therefore treat a balanced event-time horizon as primary and use unbalanced overlays for comparison, following the recommendation in [Baker et al. \(2025\)](#) to guard against “imbalance in event time” when aggregating treatment effects.

In practice, balanced horizons are implemented by requiring that treated municipality–industry cells have sufficient pre- and post-treatment support to contribute to the full event-time window, while retaining all available control observations. Details of this implementation are documented in Section [G](#).

### 7.2 Direct Effects: Primary Specification (CS)

For the primary analysis of direct effects, I estimate group-time average treatment effects  $\text{ATT}(g, t)$  using the estimator of [Callaway and Sant'Anna \(2021\)](#). This approach constructs

cohort-specific treatment effects relative to a comparison set and then aggregates them to event time. In my specification, the comparison set consists of units that are untreated at time  $t$ , combining not-yet-treated and never-treated municipality–industry cells within the same industry (an NY design). I use the unconditional version of the estimator (without additional covariates) and impose a limited-anticipation window by differencing outcomes relative to a cohort-specific pre-event baseline quarter  $g - (\delta + 1)$ , as described in Section 6.

Dynamic event-time effects  $\text{ATT}^{\text{own}}(\ell)$  are obtained by aggregating the group–time effects across cohorts with the same relative event time  $\ell = t - g$ . Balanced horizons are emphasized as the primary series, while unbalanced event-time paths are reported as overlays for comparison and composition diagnostics. Uncertainty for these estimates is obtained using uniform (simultaneous) confidence bands based on multiplier bootstrap procedures; the inference details are provided in Section 8.

### 7.3 Direct Effects: Cross-Check (BJS)

As a complementary check, I implement the imputation-based estimator of [Borusyak et al. \(2024\)](#). This method fits a two-way fixed-effects model (unit and time effects) to untreated observations only (within industry), uses the fitted model to impute untreated potential outcomes  $\widehat{Y}_{it}(0)$  for all cells, and defines estimated treatment effects as  $Y_{it} - \widehat{Y}_{it}(0)$  on treated cells. Dynamic event-time effects  $\text{ATT}^{\text{own}}(\ell)$  are then constructed by averaging these imputed treatment effects across treated municipality–industry cells at event time  $\ell$ , with equal weight on each treated unit.

Balanced horizons are again used as the primary window, with unbalanced paths reported as a robustness overlay. Standard errors and simultaneous confidence bands for the BJS-based event-time path are obtained using GEOID-cluster bootstrap procedures; these inference methods are described in Section 8. Cumulative effects over  $\ell \in [0, 16]$  are computed by summing the dynamic coefficients and using the same bootstrap draws to obtain uncertainty for the cumulative statistics.

### 7.4 Spillovers: Primary Channel (DR–DiD)

To estimate spillover effects, I use a doubly robust (DR) DiD estimator with cross-fitting, following [Xu \(2023\)](#) and [Shahn et al. \(2024\)](#). The analysis is organized in terms of event-time slices  $[a, b] \in \{[0, 4], [5, 8], [9, 16], [0, 16]\}$ , corresponding to the policy windows described in Section 5. For each slice, I construct:

- an own-treatment indicator  $D_{\text{slice}}$  that captures whether the unit is treated in the slice, and

- three exposure-slice variables  $S_{\text{same,slice}}$ ,  $S_{\text{cross,slice}}$ , and  $S_{\text{null,slice}}$ , obtained by summing the per- $\ell$  exposure shares for the same-industry neighbor, within-municipality cross-industry, and neighbor all-industries channels over  $\ell \in [a, b]$ .

These regressors are constructed from the share-based exposure measures defined in Section 4. Because each per- $\ell$  exposure share lies in  $[0, 1]$ , the slice sums  $S_{\cdot, \text{slice}}$  lie in  $[0, L_s]$ , where  $L_s = b - a + 1$  is the slice length. A one-unit change in  $S_{\cdot, \text{slice}}$  corresponds to an additional quarter of full exposure, while moving from zero exposure to 100% exposure in every quarter of the slice corresponds to an  $L_s$ -unit change. For overlap diagnostics and the  $\epsilon$ -trimming rule, I also work with slice-average exposure shares  $\bar{S}_{\cdot, \text{slice}} = S_{\cdot, \text{slice}} / L_s \in [0, 1]$ . For each outcome and slice, I estimate a linear model in which both the outcome and the regressors are first residualized using cross-fitted nuisance models, and then regressed in a single OLS step with GEOID-clustered inference. The coefficients on  $D_{\text{slice}}$ ,  $S_{\text{same,slice}}$ ,  $S_{\text{cross,slice}}$ , and  $S_{\text{null,slice}}$  provide estimates of DATT $[a, b]$  and the channel-specific SATT parameters described in Section 5.

Balanced horizons are used as the primary specification: treated or exposed units must have sufficient support over the full post-treatment window, while control observations are retained. Before estimation, I enforce the overlap conditions described in Section 6 by trimming observations with extreme exposure shares and by imposing a minimum post-trim sample size per slice. Unbalanced versions of the slice estimates are reported for comparison and diagnostics.

## 7.5 Cross-Fitting with Municipality-Respecting Folds

Cross-fitting is used to obtain doubly robust estimators while respecting the clustering structure of the data. I implement cross-fitting with group-based folds defined at the municipality level (GroupKFold by GEOID), so that all observations from a given municipality are assigned to the same fold. This strategy is supported by [Cao \(2024\)](#), who show that cluster-respecting cross-validation is important for dependent data and that buffer zones are not required when using asymptotically linear learners.

Consistent with this guidance, I employ simple parametric learners as nuisance models. The outcome regression is estimated via ridge regression, the own-treatment propensity via logistic regression, and the exposure-slice nuisance functions via ridge regression. Exposure-history controls, constructed as described in Section 4, enter the nuisance models to implement the “parallel trends conditional on exposure histories” assumption. Positivity and availability flags are enforced as sample filters but are not used as features in the nuisance models. After cross-fitted residualization of the outcome and regressors, I apply a two-way within

transformation at the municipality–industry and quarter levels and estimate the final slice regression with GEOID-clustered standard errors.

## 7.6 Heterogeneous Effects: Auxiliary TWFE Specification

The heterogeneity analysis is based on an auxiliary two-way fixed-effects specification that uses the strata defined in Section 3.6 but does not re-estimate the full DR–DiD model for each subgroup. For each outcome  $Y_{ikt}$  and event-time slice  $[a, b] \in \{[0, 4], [5, 8], [9, 16], [0, 16]\}$ , I construct the same direct-treatment and same-industry spillover regressors as in the main DR–DiD specification, namely  $D_{\text{slice}}$  and a same-industry exposure slice  $S_{\text{slice}}$  based on the same-industry neighbor shares.

I then estimate a linear model with municipality–industry and quarter fixed effects using within-demeaning. For each stratum variable  $Z \in \{\text{tradable}, \text{metro}, \text{wage\_stratum}\}$ , I interact both  $D_{\text{slice}}$  and  $S_{\text{slice}}$  with indicator functions for the strata in  $Z$  and read off stratum-specific direct and spillover coefficients from these interactions. The heterogeneity analysis focuses on the direct and same-industry neighbor spillover channels; the within-municipality cross-industry and neighbor all-industries spillovers are not re-estimated by stratum.

The estimation sample for the heterogeneity specification is restricted to the balanced event-time window so that each slice is estimated on a common set of units. There is no additional trimming or cross-fitting in this auxiliary model. When available, exposure-history controls enter additively to support the same “parallel trends conditional on exposure histories” restriction, but these controls are not interacted with the strata. Standard errors are computed using cluster-robust inference at the municipality level, and I form stratum-specific DATT, SATT, and TATT estimates for each outcome–slice–stratum combination, as well as aggregate summaries that average across strata using the distribution of municipality–industry units as weights. Adjustments for multiple testing across heterogeneous effects are reported in the diagnostics and robustness sections.

## 8 Fixed Effects and Inference

All regression specifications include municipality–industry and quarter fixed effects, as described in Section 7. Inference combines cluster-robust standard errors at the level of treatment assignment with event-study bootstrap procedures and, when warranted, spatially robust covariance estimators.

## 8.1 Event-Study and Cumulative Inference

For the direct-effect event studies, I construct uniform (simultaneous) confidence bands for the dynamic path  $\text{ATT}^{\text{own}}(\ell)$  under both the Callaway–Sant’Anna (CS) and Borusyak–Jaravel–Spiess (BJS) estimators.

For the CS estimator, I follow [Callaway and Sant’Anna \(2021\)](#) and use a Rademacher multiplier bootstrap applied to the influence-function-based aggregation over group–time cells. Each bootstrap draw perturbs the group–time contributions with independent Rademacher multipliers drawn at the municipality level, recomputes the aggregated  $\text{ATT}^{\text{own}}(\ell)$ , and forms a max- $|t|$  statistic over post-treatment event times  $\ell \geq 0$ , excluding the omitted reference bin  $\ell = -1$  and the anticipatory leads  $\ell \in \{-\delta, \dots, -1\}$  implied by the limited-anticipation window. The  $(1 - \alpha)$  quantile of this max- $|t|$  distribution is used as a critical value to construct uniform bands for all reported event times. As in [Freyaldenhoven et al. \(2019\)](#), these simultaneous bands are used for visual inference on the dynamic path.

For the BJS estimator, I obtain uncertainty from a GEOID-cluster bootstrap. Each bootstrap sample resamples municipalities with replacement, re-estimates the untreated fixed-effects model, re-imputes counterfactual outcomes, and recomputes  $\text{ATT}^{\text{own}}(\ell)$ . I then form a max- $|t|$  statistic over  $\ell \geq 0$  (excluding the reference bin  $\ell = -1$ ) and use its  $(1 - \alpha)$  quantile to construct uniform bands for the BJS-based event-time profile.

Cumulative effects over  $\ell \in [0, 16]$  are computed by summing the dynamic coefficients,

$$\widehat{\text{ATT}}_{\text{cumu}}^{\text{own}}(0, 16) = \sum_{\ell=0}^{16} \widehat{\text{ATT}}^{\text{own}}(\ell),$$

and their uncertainty is obtained from the *same* bootstrap draws used for the dynamic bands: for each draw, I sum the corresponding event-time coefficients, and I report two-sided percentile intervals from the bootstrap distribution of the cumulative effect. This procedure is applied to both CS- and BJS-based event-study estimates.

For the slice-level DR–DiD spillover estimands (Section 7.4), I use GEOID-cluster-robust standard errors as a conventional benchmark, but whenever the Moran’s- $I$  diagnostics indicate spatial dependence I base inference and the main tables on C–SCPC spatially robust standard errors, reporting cluster-robust SEs only as a secondary comparison.

## 8.2 Clustering at the Level of Treatment Assignment

I report standard errors clustered at the municipality level—the level of independent treatment assignment—following design-based reasoning in [Roth et al. \(2023\)](#). They recommend clustering at the level where treatment is independently assigned, noting that this guidance

is particularly important when the number of clusters is not large. When few treated clusters exist, [Roth et al. \(2023\)](#) further discuss alternative inference procedures, such as permutation or bootstrap methods, highlighting their relevance for small-sample settings. This clustering choice serves as the baseline for uncertainty quantification for the direct and spillover effects in Q1–Q4, as well as the cumulative-value objects in Q5, and it remains the primary method in specifications where residual spatial dependence is not detected; when Moran’s- $I$  diagnostics indicate spatial autocorrelation, the main reported inference instead relies on the C-SCPC procedure described below.

### 8.3 Moran’s I Residual Diagnostic (“Trigger”)

To determine whether spatially robust inference is warranted, I apply a residual spatial autocorrelation diagnostic based on the global Moran’s  $I$  statistic. As shown by [Chen \(2016\)](#), the Durbin–Watson test is inappropriate for spatial cross-sectional data because its value depends on the ordering of observations, whereas Moran’s  $I$  provides a permutation-invariant measure of spatial autocorrelation. Following this insight, I define the “trigger” for spatial inference as a significant Moran’s  $I$  result ( $p < 0.05$ ) using the spatial-weights matrix  $W$  (queen contiguity with a KNN plug-in for isolates, as in Section 4.4) and a permutation reference distribution.

For multivariate models (e.g., stacked residuals across multiple outcomes), I use the multivariate extension of Moran’s  $I$  proposed by [Yamada \(2024\)](#). Their formulation defines both global and local multivariate Moran’s  $I$ , which summarize spatial autocorrelation across multiple residual series as single statistics, allowing me to implement a unified test for spatial dependence across outcomes.

### 8.4 Primary Spatial Inference: Conditional SCPC

When the Moran’s- $I$  trigger indicates spatial autocorrelation, I conduct inference using Conditional Spatial Correlation Projection Control (C-SCPC), following [Müller and Watson \(2023\)](#). C-SCPC controls size conditional on regressors and spatial locations and is valid in panel or difference-in-differences (DiD) designs with fixed effects. It extends the spatial-correlation-robust method developed by [Müller and Watson \(2022\)](#), who establish finite-sample size control and asymptotic validity under weak spatial dependence. I follow their guidance in selecting the number of principal components  $q$  according to the expected-length rule and calibrate the worst-case correlation bound using plausible average spatial correlation. Together, these studies justify the use of C-SCPC as the primary spatial inference procedure once spatial dependence is detected, and in practice all specifications for which the Moran’s-

$I$  trigger is activated are reported in the main tables with C-SCPC standard errors and associated confidence intervals; conventional cluster-robust and SHAC standard errors are relegated to robustness tables and the appendix.

## 8.5 Conley/SHAC Robustness Checks

As an appendix robustness check, I compute Conley/Spatial HAC (SHAC) standard errors using distances between municipality centroids measured in the Puerto Rico StatePlane coordinate system (EPSG:32161) and converted to kilometers. The primary specification uses a 75 km cutoff radius for the spatial Bartlett kernel, calibrated to the geographic scale of Puerto Rico. This radius is large enough to capture economically relevant cross-municipality dependence on an island of roughly 170 km by 60 km while remaining small enough to avoid treating all locations as effectively equidistant, in line with common practice in Conley-type applications.

[Conley and Molinari \(2005\)](#) show that the nonparametric SHAC estimator remains consistent and practically robust to bounded errors in spatial locations or distances. However, recent work by [Conley et al. \(2023\)](#) highlights that these asymptotic estimators can suffer from size distortions in finite samples, proposing a spatial dependent wild bootstrap as a refinement. While I rely on C-SCPC as the primary finite-sample-robust procedure, I report standard asymptotic SHAC estimates here to facilitate comparison with the existing applied literature. I implement this estimator following the methodology outlined in the `conleyreg` package ([Düben, 2025](#)), specifying a Bartlett kernel and lag cutoffs appropriate for the panel structure.

In addition, [Xu \(2023\)](#) extends SHAC inference to difference-in-differences settings with interference, demonstrating that kernel-weighted spatial HAC estimators remain consistent under arbitrary spatial correlation. Finally, to evaluate the sensitivity of spatial SE adjustments, I include results adjusted by the Thresholding Multiple Outcomes (TMO) method of [Della Vigna et al. \(2025\)](#). This approach uses auxiliary outcomes to identify cross-location dependence and typically increases standard errors, providing a meaningful “ $\pm$ TMO” sensitivity band alongside the SHAC and C-SCPC estimates.

## 8.6 Two-Way Clustering Sensitivity

As a further robustness check, I estimate two-way cluster-robust standard errors. The canonical approach of [Cameron et al. \(2011\)](#) constructs two-way clustered variances as the sum of one-way cluster variances minus their intersection term, and it applies readily to non-nested structures such as municipality  $\times$  time. To address the potential for serial correlation

in time effects, I also compute the variant proposed by [Chiang et al. \(2023\)](#), which augments the two-way covariance estimator with kernel-weighted time autocovariances and selects the lag truncation using the Andrews rule. [Chiang and Sasaki \(2023\)](#) provide the theoretical justification for two-way cluster-robust standard errors under mild conditions and confirm their accuracy in typical applied settings. I treat these two-way clustered estimates as a sensitivity check relative to the baseline GEOID-clustered and spatially robust procedures described above.

## 9 Diagnostics and Falsification

I implement a series of diagnostic and falsification checks designed to assess identification credibility and robustness across several dimensions: pre-trends, sensitivity to violations of parallel trends, overlap and exposure positivity, spatial dependence, and multiple-testing adjustments.

### 9.1 Pre-Trends and Event-Study Diagnostics

Following [Sun and Abraham \(2021\)](#), I estimate the “interaction-weighted” (IW) event-study model to diagnose potential violations of parallel trends. This approach corrects the well-known contamination of naive two-way fixed effects (TWFE) estimates when treatment effects are heterogeneous across cohorts. The IW estimator isolates cohort-specific dynamics and recovers unbiased lead coefficients that serve as valid pre-trend diagnostics. As [Roth et al. \(2023\)](#) document, naive TWFE specifications may produce misleading lead coefficients due to negative weights and forbidden comparisons, motivating the use of the Sun–Abraham estimator.

I also draw on the complementary pre-trend test proposed by [Borusyak et al. \(2024\)](#), who construct uncontaminated “untreated-only” regressions that directly test for spurious pre-trends using the never-treated group. In applied work, this logic aligns with the goal of the IW estimator: to ensure that pre-treatment dynamics are estimated from valid comparisons. The convention of examining lead coefficients ( $\ell < -1$ ) as pre-trend checks is standard in policy evaluation ([Wang et al., 2024](#)), but these must be interpreted using estimators that respect treatment heterogeneity. When significant pre-trends are detected, I follow the recommendation of [Freyaldenhoven et al. \(2019\)](#), who show that pre-treatment leads can be re-purposed as instrumental variables to correct for endogeneity in treatment timing.

## 9.2 HonestDiD Sensitivity Analysis

To evaluate robustness to deviations from parallel trends, I apply the “HonestDiD” sensitivity approach of [Rambachan and Roth \(2023\)](#), implementing the smoothness (SD) restriction that constrains second differences of the event-study path ( $|\Delta^2\theta_\ell| \leq M$ ) over a grid of curvature budgets  $M$ . This framework constructs bounds on the average treatment effect on the treated (ATT) that remain valid under bounded violations of parallel trends, calibrated to observed pre-trends. As summarized by [Roth et al. \(2023\)](#), HonestDiD-style sensitivity analysis has become a best practice for credible difference-in-differences inference. I also incorporate the recommendation of [Wang et al. \(2024\)](#), who emphasize sensitivity analysis as an essential robustness check in empirical DiD research. Extending these ideas to settings with interference, [Xu \(2023\)](#) explicitly suggest adapting Rambachan–Roth-style sensitivity analyses when the exposure mapping may be misspecified, providing formal justification for implementing such a check here.

## 9.3 Overlap and Positivity Diagnostics

For settings with interference, the validity of causal identification relies critically on the overlap, or positivity, of exposure histories. [Xu \(2023\)](#) formalize the exposure-positivity condition for difference-in-differences with interference, requiring that each possible exposure path has nonzero probability. This condition parallels the more general support requirement described by [Aronow et al. \(2020\)](#) in the context of experimental spillovers. Similarly, [Shahn et al. \(2024\)](#) extend the classical positivity assumption to complex exposure mappings, while [Liu et al. \(2019\)](#) formalize group-level positivity under partial interference.

Building on [Sävje \(2023\)](#), I diagnose overlap using the distributions of slice-average exposure shares  $\bar{S}_{\cdot,\text{slice}}$  and apply a two-sided  $\epsilon$ -trim that retains observations only if their slice-average exposure lies in  $[\epsilon, 1 - \epsilon]$  (default  $\epsilon = 0.02$ ). This removes both near-zero and near-saturation regions to ensure sufficient overlap for all comparisons, following the rationale of [Crump et al. \(2009\)](#). The DR–DiD regressors themselves are defined in terms of the corresponding slice-sum exposures  $S_{\cdot,\text{slice}} = L_s \bar{S}_{\cdot,\text{slice}}$ , so trimming is implemented on the normalized shares while estimation uses their rescaled sums. I report coverage impacts by channel and slice, alongside histogram and ECDF diagnostics that compare exposure distributions before and after trimming and balancing, and I enforce a minimum post-trim sample-size threshold per slice in the main DR–DiD estimations.

## 9.4 Spatial Dependence Diagnostics

Residual spatial dependence is assessed using Moran's  $I$  with permutation inference. Following [Chen \(2016\)](#), I use a Moran's  $I$  test on model residuals as a diagnostic trigger for spatial autocorrelation, with the spatial-weights matrix  $W$  (queen contiguity with a KNN plug-in for isolates) defined in Section 4.4. To accommodate multivariate and stacked residuals, I extend the test following [Yamada \(2024\)](#), who generalize Moran's  $I$  to multivariate settings and formalize its dependence on the spatial-weights matrix. These diagnostics determine when the spatially robust inference procedures in Section 8 are applied.

## 9.5 Multiple-Testing Adjustments

Finally, I correct for multiple inference separately for the primary dynamic event-study paths and for the heterogeneous-effects summaries.

For the main dynamic event-study paths of the CS and BJS estimators (in particular, the full set of post-treatment coefficients for the primary outcome), I treat the post-treatment horizons as a high-stakes family and construct simultaneous (uniform) confidence bands using a studentized max- $|t|$  ("sup- $t$ ") statistic. Building on the influence-function representations and bootstrap schemes developed for these estimators ([Callaway and Sant'Anna, 2021](#); [Borusyak et al., 2024](#)) and the general theory of sup- $t$  bands in [Montiel Olea and Plagborg-Møller \(2019\)](#), as implemented in event-study applications such as [Freyaldenhoven et al. \(2019\)](#), I generate multiplier or cluster bootstrap draws of the event-study coefficients, compute in each draw the maximum absolute  $t$ -statistic over post-treatment event times, and take the  $(1 - \alpha)$  quantile of this max- $|t|$  distribution as the critical value. The resulting bands  $\widehat{\text{ATT}}(\ell) \pm c_{1-\alpha} \widehat{s_e}(\ell)$  provide joint coverage over all post-treatment horizons of the primary outcome and therefore act as a simple family-wise multiple-testing adjustment for the event-study path as a whole. For cumulative effects such as the 0–16 quarter ATT, I treat the cumulative estimand as a linear combination of the event-time coefficients and construct bootstrap confidence intervals from the same resampling scheme applied to this linear combination.

For the heterogeneous-effects analysis, which uses the auxiliary TWFE specification described in Section 7.6, I instead control the false discovery rate (FDR). I treat all heterogeneous DATT, SATT, and TATT coefficients within a given combination of outcome and event-time slice as a single testing family and compute Benjamini–Hochberg (BH) and Benjamini–Yekutieli (BY)  $q$ -values ([Benjamini and Hochberg, 1995](#); [Benjamini and Yekutieli, 2001](#)). The BH procedure is reported as the default FDR adjustment, while BY provides a robustness check under positive or arbitrary dependence across coefficients. These adjustments are implemented on the heterogeneity results table and reported alongside the corresponding

point estimates and standard errors in the main and appendix tables.

## 9.6 Diagnostic Results

Figure 5 reports the Sun–Abraham interaction-weighted event-study coefficients for log covered employment around the first sizable entry. Pre-treatment coefficients for  $\ell \leq -2$  are close to zero and display wide confidence intervals, with no clear systematic trend before treatment. This pattern is consistent with the conditional parallel trends assumption discussed in Section 6 and provides the baseline calibration for the HonestDiD sensitivity analysis.

Figure 6 examines overlap for the main spillover channel. The left panel shows the distribution of slice-average same-industry neighbor exposure shares  $\bar{S}_{\text{same},0-16}$  before trimming, while the right panel reports the corresponding empirical CDF. Most observations exhibit very low exposure shares, and the two-sided  $\epsilon$ -rule at 0.02 and 0.98 trims a large mass of near-zero exposure cells. Table 4 quantifies this trimming by event-time slice: between 60% and 82% of raw observations are removed, but all 78 municipios remain in the trimmed samples, ensuring geographic coverage while enforcing exposure positivity.

Table 5 summarizes the Moran’s  $I$  residual diagnostics. For the direct exposure-augmented event-study and for all DR–DiD spillover slices, Moran’s  $I$  is positive and statistically significant at conventional levels using the Queen+KNN spatial-weights matrix. Examining the time-series of quarter-specific local Moran statistics, between 4% and 18% of quarters show statistically significant spatial clustering across the five specifications reported here (detailed quarter-by-quarter diagnostics are in Appendix Table D1). These findings activate the spatial-dependence “trigger” described in Section 8 and imply that, for both the direct exposure-augmented event study and the DR–DiD spillover slices, the main reported inference is based on C–SCPC spatially robust standard errors, with cluster-robust and SHAC standard errors reported only as complementary robustness checks.

Figure DI1: Pre-treatment Event-Study Diagnostics  
Outcome: log\_covered\_emp

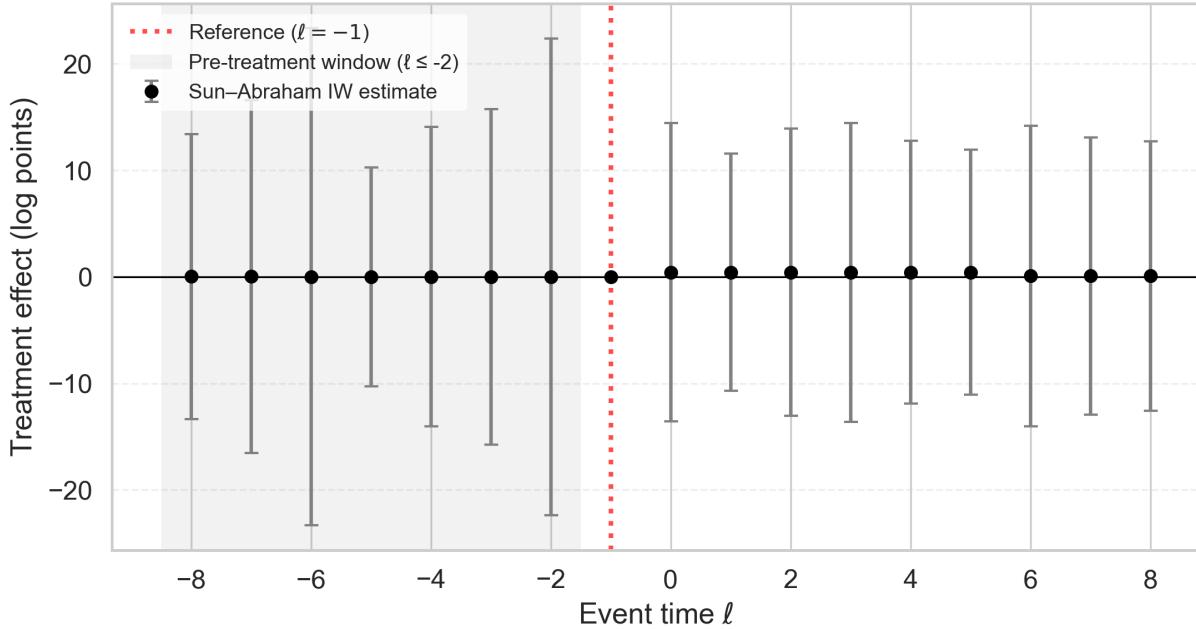


Figure 5: Pre-Treatment Event-Study Diagnostics for Direct Effects (log covered employment)

*Notes:* The figure plots Sun-Abraham interaction-weighted event-study coefficients for log\_covered\_emp from  $\ell = -8$  to  $\ell = 8$  quarters relative to the first sizable entry, omitting  $\ell = -1$  as the reference period. Black dots show point estimates and vertical bars show 95% confidence intervals. The shaded region highlights pre-treatment leads  $\ell \leq -2$ , which are used as pre-trend diagnostics rather than as identifying moments for the dynamic effect path.

Table 4: Overlap and Trimming Summary for Same-Industry Neighbor Exposure

Slice	Channel	Obs (raw)	Obs (trimmed)	% trimmed
0-4	Same-industry neighbors	27,465	8,606	68.7
5-8	Same-industry neighbors	42,002	7,633	81.8
9-16	Same-industry neighbors	52,222	10,621	79.7
0-16	Same-industry neighbors	27,351	10,899	60.2

*Notes:* This table reports overlap and trimming diagnostics for the same-industry neighbor exposure channel. “Obs (raw)” is the number of municipality–industry–quarter observations with non-missing exposure before trimming. “Obs (trimmed)” applies the two-sided  $\epsilon$ -rule at 0.02 and 0.98 and the balanced-window filter. Percent trimmed is the share removed. All 78 municipios remain represented. Additional diagnostics are in the appendix.

Figure DI2: Exposure Overlap and Trimming Regions  
(Same-industry neighbors, slice [0,16])

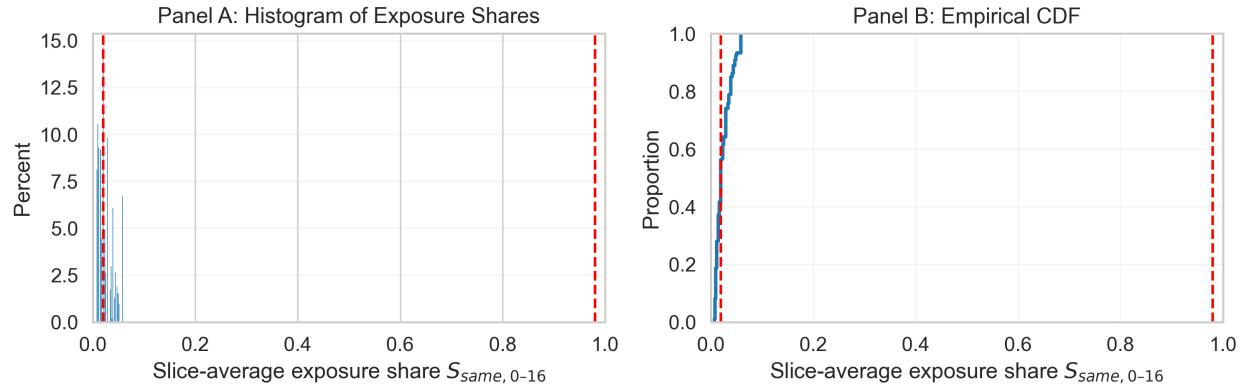


Figure 6: Exposure Overlap and Trimming for Same-Industry Neighbor Exposure, Slice [0,16]

*Notes:* Panel A shows the histogram of slice-average same-industry neighbor exposure shares  $\bar{S}_{same, 0-16}$  across municipality–industry–quarter observations before trimming. Panel B reports the corresponding empirical CDF. Vertical dashed lines mark the two-sided trimming rule at 0.02 and 0.98 used to enforce exposure positivity. For this channel and slice, the trimming rule removes about 60% of observations (see Table 4), while retaining all municipios.

Table 5: Moran's  $I$  Residual Spatial Dependence Diagnostics

Model	Slice	Moran's $I$	Permutation $p$ -value
Direct (Exposure ES dynamic)	–	0.155	0.026
DR–DiD spillover	0–16	0.286	0.002
DR–DiD spillover	0–4	0.342	0.001
DR–DiD spillover	5–8	0.272	0.001
DR–DiD spillover	9–16	0.324	0.001

*Notes:* Each row reports the global Moran's  $I$  statistic and permutation-based  $p$ -value (999 permutations) for residuals from the indicated specification and event-time slice. All diagnostics use the Queen+KNN spatial-weights matrix. All models meet the spatial-dependence threshold ( $p < 0.05$ ).

## 10 Results

This section presents the main estimates for the direct and spillover effects of first sizable entries. I begin with dynamic Callaway–Sant’Anna (CS) event-study estimates for the three primary outcomes, then summarize cumulative 0–16 quarter effects. I then turn to the doubly robust DiD (DR–DiD) slice estimands that decompose the total impact on employment into direct and spillover components across the three exposure channels.

### 10.1 Direct Dynamic Effects (CS Event Studies)

Figure 7 reports the CS event-study for `log_covered_emp`. All coefficients are normalized to the cohort-specific baseline quarter  $g - (\delta + 1)$ ; in the baseline specification with  $\delta = 2$  this corresponds to event time  $\ell = -3$ , three quarters before entry, while the last two pre-event quarters  $\ell \in \{-2, -1\}$  are treated as a limited-anticipation window used for pre-trend diagnostics. Pre-treatment leads are near zero and well inside the simultaneous confidence band, consistent with the pre-trend diagnostics in Section 9. At the time of entry, covered employment jumps by roughly 13 log points and continues to rise over the first year, stabilizing between 14 and 17 log points above the pre-entry level during the subsequent three to four years. The 95% uniform band excludes zero for all post-entry horizons in the balanced window.

Figure 8 shows analogous event studies for `log_estab` and the `log_wage_bill`. The response of establishments is large and immediate: the number of establishments in the treated municipality–industry cell increases by about 25–30 log points on impact and remains at least 25 log points above the pre-entry level through 16 quarters. The real wage bill also rises quickly, with a pattern similar to covered employment but somewhat more volatile, reflecting both changes in headcounts and in average wages. Across all three outcomes, the dynamic paths display a sharp break at entry, persistent positive effects over four years, and no evidence of systematic pre-trends.

Taken together, these dynamic estimates answer the first part of the research question on direct effects: first sizable entries are followed by large, abrupt, and persistent increases in covered employment, establishments, and wage bills in the treated municipality–industry cell, with no evidence of pre-existing trends that could account for the post-entry jump.

### 10.2 Cumulative Direct Effects by Slice

Table 6 summarizes the CS estimates as *cumulative* 0–4, 5–8, 9–16, and 0–16 quarter effects, obtained by summing the event-time coefficients  $\text{ATT}^{\text{own}}(\ell)$  over all quarters in each slice.

Figure R1: Direct Effects on Employment (CS, Balanced)

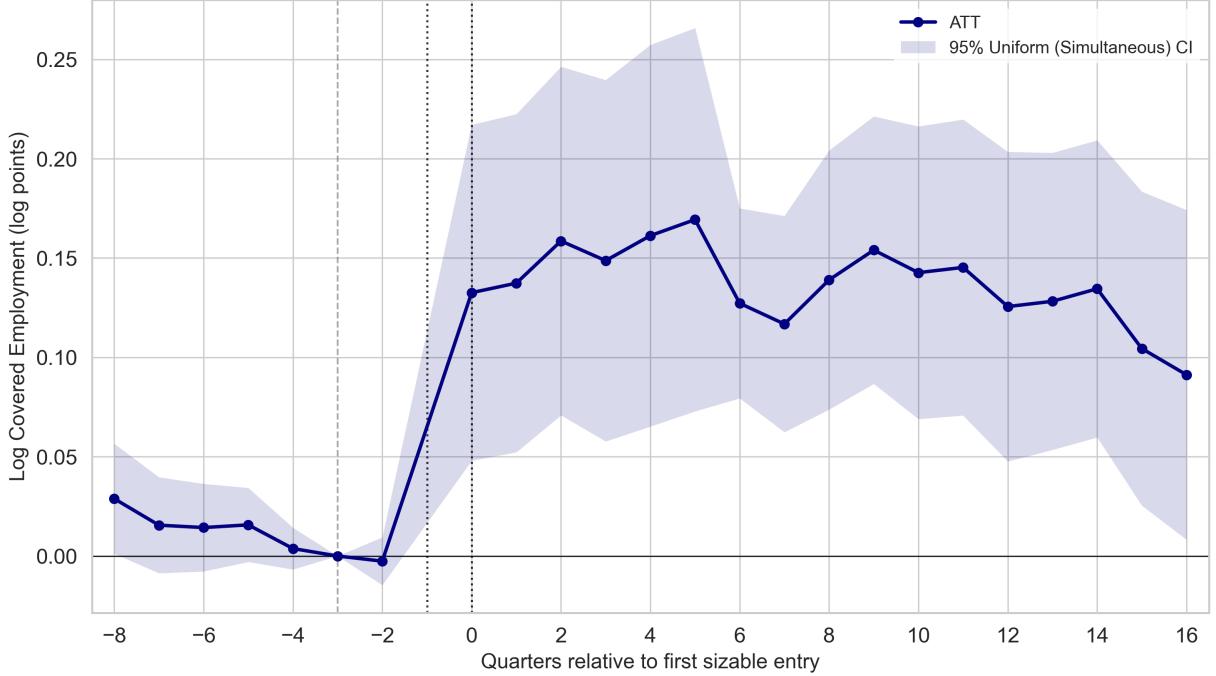


Figure 7: Direct Effects of First Sizable Entry on Covered Employment (CS, Balanced)

*Notes:* The figure plots Callaway–Sant’Anna group-time ATT estimates for `log_covered_emp`, aggregated to event time  $\ell$  using the balanced horizon. The horizontal axis reports quarters relative to the first sizable entry ( $\ell = 0$ ); the vertical axis shows the estimated treatment effect in log points. Coefficients are normalized to the cohort-specific baseline quarter  $g - (\delta + 1)$ ; in the baseline specification with  $\delta = 2$  this corresponds to event time  $\ell = -3$  (three quarters before entry). The quarters  $\ell \in \{-2, -1\}$  form a limited-anticipation window: they are used as pre-trend diagnostics but are excluded from the set of event times over which the simultaneous ( $\max_{|t|}$ ) confidence band is constructed. For visual clarity the last pre-event quarter  $\ell = -1$  is omitted from the plotted series. The solid line is the point estimate and the shaded region is the 95% simultaneous confidence band. Vertical lines at  $\ell = -3$ ,  $\ell = -1$ , and  $\ell = 0$  mark the baseline reference quarter, the last pre-event (anticipation) quarter, and the entry quarter, respectively.

For each outcome and slice, I report the sum of the dynamic ATT coefficients, together with bootstrap standard errors and percentile-based 95% confidence intervals.

For covered employment, the cumulative effect is 0.74 log points over quarters 0–4 and 0.55 log points over quarters 5–8, with both slices precisely estimated. The 9–16 quarter slice adds another 1.03 log points, yielding a total 0–16 cumulative effect of 2.32 log points. I return to the economic interpretation of this cumulative effect in Section 12, where I translate it into approximate job counts and wage-bill changes.

The establishment counts show even larger cumulative responses, with a 0–16 quarter effect of 4.67 log points, while the real wage bill exhibits a 0–16 cumulative effect of 2.33 log points that closely mirrors the employment path. The close alignment between covered

Table 6: Cumulative Direct Effects by Event-Time Slice (CS, Balanced)

Outcome	Slice 0–4			Slice 5–8		
	ATT	SE	95% CI	ATT	SE	95% CI
<i>Panel A: Short- and Medium-Run</i>						
log_covered_emp	0.738	0.074	[0.594, 0.883]	0.552	0.051	[0.453, 0.652]
log_establishments	1.394	0.036	[1.324, 1.464]	1.016	0.024	[0.968, 1.063]
log_total_wages_real_2020	0.684	0.084	[0.520, 0.848]	0.557	0.059	[0.441, 0.673]
Slice 9–16						
Outcome	Slice 9–16			Slice 0–16		
	ATT	SE	95% CI	ATT	SE	95% CI
<i>Panel B: Long-Run and Full Horizon</i>						
log_covered_emp	1.026	0.079	[0.870, 1.182]	2.317	0.120	[2.082, 2.551]
log_establishments	2.261	0.039	[2.184, 2.337]	4.671	0.058	[4.557, 4.785]
log_total_wages_real_2020	1.085	0.091	[0.906, 1.264]	2.326	0.137	[2.058, 2.595]

*Notes:* Each entry reports the cumulative ATT over the indicated event-time slice, obtained by summing Callaway–Sant’Anna dynamic coefficients over the balanced horizon. Standard errors are based on the same bootstrap draws used to construct the uniform event-study bands. Confidence intervals are 2.5th and 97.5th percentiles of the bootstrap distribution of the cumulative effect.

employment and the real wage bill suggests that the main margin of adjustment is quantities rather than wages.

Taken together, the cumulative CS estimates indicate that a first sizable entry is associated with a multi-year expansion in activity in the treated municipality–industry cell: total employment-quarters and wage bills increase by roughly an order of magnitude relative to the counterfactual, and the number of establishments rises even more sharply, providing a clear quantitative answer on the magnitude and persistence of the direct effects.

### 10.3 Spillovers and Decomposition of Total Effects (DR–DiD)

Table 7 reports the DR–DiD *slice-average* estimands for log\_covered\_emp along the four slices: each entry is a per-period average effect over the corresponding event-time window rather than a cumulative sum. The first column shows the direct effect DATT[ $a, b$ ], while the next three columns report the same-industry neighbor, within-municipality cross-industry, and neighbor all-industries spillover components. The final two columns summarize total spillovers SATT[ $a, b$ ] and the total effect TATT[ $a, b$ ] = DATT[ $a, b$ ] + SATT[ $a, b$ ]. Each estimate is shown with its C–SCPC spatially robust standard error; conventional municipality-clustered standard errors are reported in Appendix D.2 as a comparison, and 95% confidence intervals (using a normal approximation) are based on the C–SCPC standard errors.

Over the first four post-entry quarters, the point estimate of the direct effect is positive,  $\widehat{\text{DATT}}[0, 4] = 0.156$ , but under C–SCPC the associated standard error is sizable ( $\text{SE} \approx 0.13$ ), so the 95% confidence interval includes zero. The same-industry neighbor spillover remains

large in magnitude,  $\widehat{\text{SATT}}_{\text{same}}[0, 4] \approx 0.84$ , and the implied total 0–4 quarter effect  $\widehat{\text{TATT}}[0, 4] \approx 0.99$  is similarly positive, but C–SCPC standard errors are much larger than their clustered counterparts and the corresponding confidence intervals no longer exclude zero. Thus, early spillover gains are economically meaningful but not statistically precise once spatial dependence is accounted for.

In the second year after entry (slice 5–8), both the direct and spillover components are small and imprecisely estimated, and the total effect is close to zero. In contrast, the 9–16 quarter slice continues to show a negative same-industry neighbor spillover,  $\widehat{\text{SATT}}_{\text{same}}[9, 16] \approx -0.33$ . However, the C–SCPC standard error for this effect is large ( $\text{SE} \approx 0.43$ ; see Appendix D.2), so the resulting 95% confidence interval, roughly  $[-1.18, 0.51]$ , includes zero. The implied total effect  $\widehat{\text{TATT}}[9, 16] \approx -0.33$  is therefore not statistically distinguishable from zero under our primary spatial inference procedure, even though the point estimate remains negative. These patterns suggest that early gains in neighboring same-industry employment may be partially reversed over longer horizons, consistent with competitive or displacement effects, but the medium-run spillover evidence is statistically weak once spatial dependence is fully accounted for.

Aggregating over the full 0–16 quarter window, the direct effect remains positive in magnitude,  $\widehat{\text{DATT}}[0, 16] = 0.214$ , but the C–SCPC standard error ( $\text{SE} \approx 0.16$ ) implies a wide confidence interval that includes zero. Net spillovers over the same horizon are negative but very imprecisely estimated, and the total effect  $\widehat{\text{TATT}}[0, 16] \approx 0.10$  (about 11%) is also far from statistically significant once spatial dependence is taken into account. Appendix D.2 compares these C–SCPC standard errors to conventional cluster-robust and SHAC estimators and shows that clustered inference tends to underestimate uncertainty in the presence of spatial correlation. Figure 9 visualizes this decomposition, showing that most of the 0–16 quarter total effect is accounted for by the direct component, with sizable positive short-run spillovers that are offset by longer-run negative same-industry effects.

Finally, it is worth noting that the statistical insignificance of the aggregate cross-industry spillover ( $\text{SATT}_{\text{cross}}$ ) masks heterogeneous linkages between specific sectors. An exploratory pairwise analysis of the underlying source–target variation reveals that entries in manufacturing industries—specifically Apparel (NAICS 315), Printing (323), and Miscellaneous Manufacturing (339)—are associated with statistically significant increases in Educational Services (NAICS 61) establishments in the same municipality (mean post-entry effects ranging from 0.12 to 0.18 log points). This pattern is consistent with induced demand for vocational training or workforce development services following the arrival of new manufacturing employers, suggesting that while broad cross-sector spillovers are weak, specific labor-market complementarities do operate in the local economy.

Table 7: Direct and Spillover *Slice-Average* Effects on Covered Employment (DR–DiD)

Slice	DATT	SATT <sub>same</sub>	SATT <sub>cross</sub>	SATT <sub>nall</sub>	SATT	TATT
0–4	0.156 (0.134)	0.838	0.001	−0.005	0.834 (0.855)	0.990 (0.897)
5–8	−0.041 (0.131)	−0.018	0.003	−0.008	−0.024 (0.345)	−0.065 (0.420)
9–16	0.001 (0.113)	−0.327	−0.002	−0.006	−0.335 (0.429)	−0.335 (0.416)
0–16	0.214 (0.161)	−0.169	0.027	0.032	−0.110 (0.908)	0.104 (0.965)

*Notes:* Each entry reports a *slice-average* DR–DiD estimand for log \_covered\_emp: DATT[a, b], SATT.[a, b], and TATT[a, b] are per-period average effects over the slice [a, b], not cumulative sums. DATT[a, b] is the direct effect of own treatment over slice [a, b]; SATT\_same[a, b], SATT\_cross[a, b], and SATT\_nall[a, b] are spillover effects for the same-industry neighbor, within-municipality cross-industry, and neighbor all-industries exposure channels, respectively. SATT[a, b] is the sum of the three spillover components and TATT[a, b] = DATT[a, b] + SATT[a, b] is the total effect. Standard errors in parentheses are C–SCPC spatially robust standard errors for the aggregate parameters (DATT, SATT, TATT); C–SCPC inference for individual spillover components is computationally prohibitive, so these are reported as point estimates without standard errors. Conventional municipality-clustered standard errors for all parameters are reported in Appendix D.2. Approximate 95% confidence intervals for DATT, SATT, and TATT are given by the point estimate  $\pm 1.96 \times \text{SE}$ .

In terms of the research questions on regional and spillover effects, these DR–DiD estimates indicate that first sizable entries initially generate large positive spillovers to same-industry neighbors, but that over a four-year horizon most of the net change in covered employment comes from the direct effect in the treated municipality–industry cell, with net regional gains that are modest in size and estimated imprecisely once spillovers are taken into account.

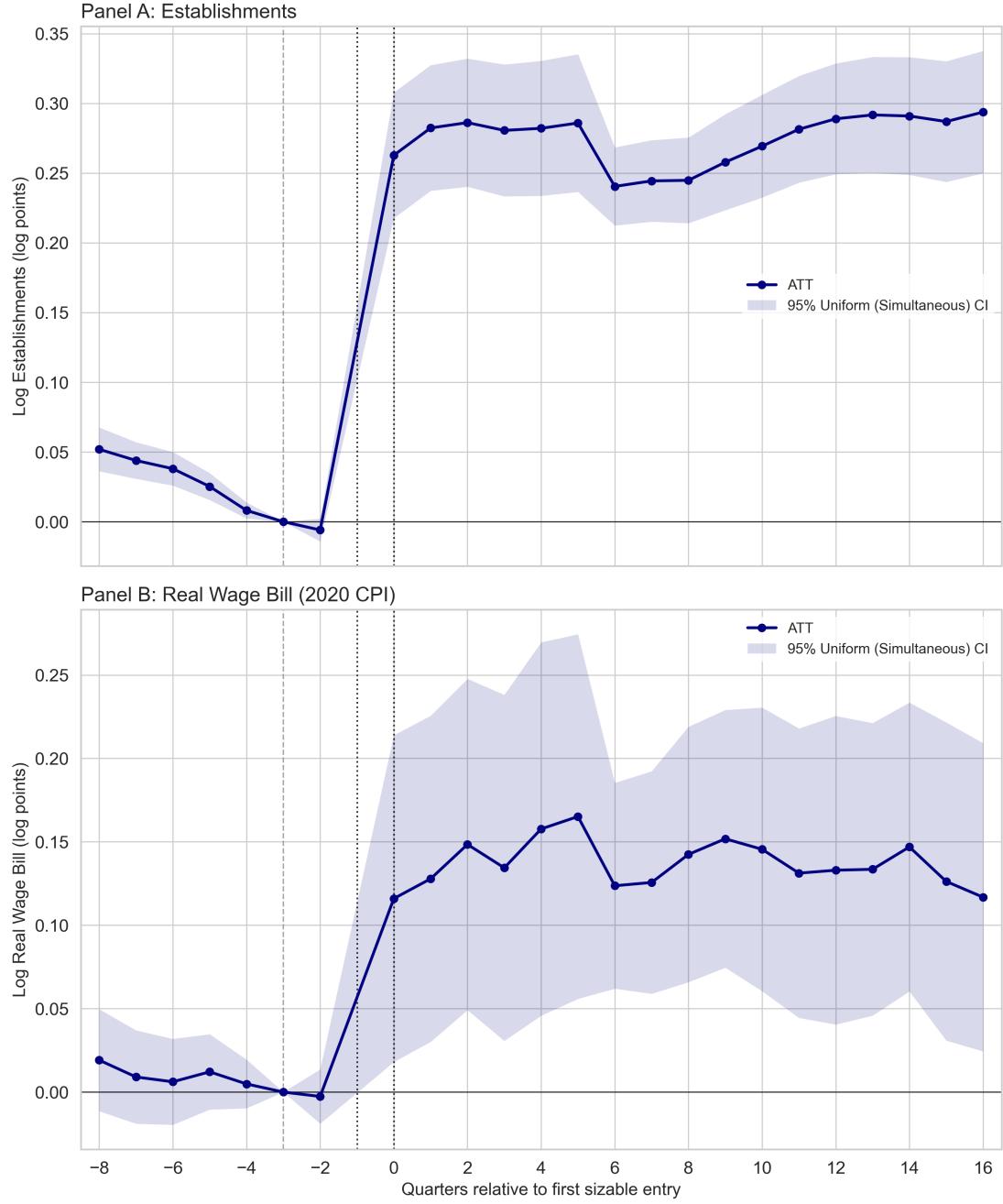


Figure 8: Direct Effects on Establishments and Real Wage Bill (CS, Balanced)

*Notes:* Panel A reports event-study estimates for log\_establishments; Panel B reports estimates for the log real wage bill log\_total\_wages\_real\_2020. Both panels use the Callaway–Sant’Anna estimator with a balanced event-time window, and coefficients are normalized to the cohort-specific baseline quarter  $g - (\delta + 1)$  (event time  $\ell = -3$  when  $\delta = 2$ ). The solid line is the ATT path and the shaded region is the 95% simultaneous confidence band. The limited-anticipation window  $\ell \in \{-2, -1\}$ , the omission of  $\ell = -1$  from the plotted series for visual clarity, and the vertical reference lines at  $\ell = -3$ ,  $\ell = -1$ , and  $\ell = 0$  are defined as in Figure 7.

Figure R3: Decomposition of Total 0–16 Quarter Effect on Employment

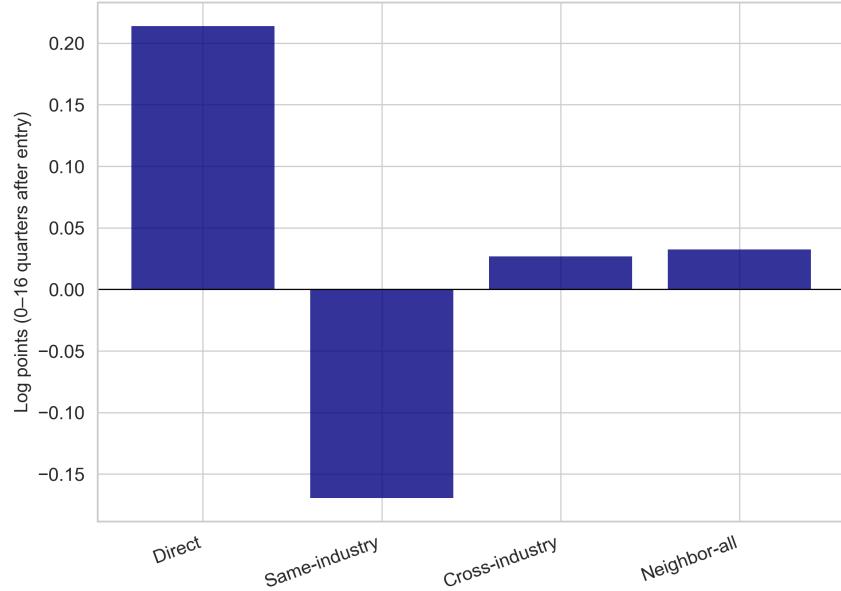


Figure 9: Decomposition of Total 0–16 Quarter Effect on Covered Employment

*Notes:* The figure decomposes the 0–16 quarter DR–DiD total effect on  $\log\_covered\_emp$  into its direct and spillover components. Bars report DATT[0, 16], SATT\_same[0, 16], SATT\_cross[0, 16], and SATT\_nall[0, 16] in log points. Numerical values are given in Table 7. The direct component is positive in magnitude but imprecisely estimated under C–SCPC spatially robust standard errors, and the net spillover component is negative and similarly imprecise, so that the implied total 0–16 quarter effect is modest and not statistically distinguishable from zero.

## 11 Heterogeneity

This section summarizes how the 0–16 quarter effects on covered employment vary across the pre-defined strata described in Section 3.6. I focus on the auxiliary TWFE specification in Section 7.6, which interacts the direct-treatment and same-industry neighbor exposure slices with time-invariant indicators for metro status, tradability, and pre-period wage stratum. Figures 10–12 plot the resulting stratum-specific direct and spillover effects for log\_covered\_emp by tradable status, metro status, and wage stratum, respectively, and Table 8 reports the corresponding estimates, standard errors, confidence intervals, and FDR-adjusted  $q$ -values.

Across all strata, the 0–16 quarter direct effect of a first sizable entry on covered employment is large and precisely estimated. Point estimates range from about 1.15 to 2.08 log points, implying roughly 3.2–8.0-fold increases in cumulated employment-quarters relative to the counterfactual.<sup>3</sup> Direct effects are notably larger in tradable industries (2.08 vs 1.15 log points for non-tradable, with barely overlapping confidence intervals) and somewhat larger in high-wage industries (1.95 vs 1.30 log points). Differences between metro and non-metro municipios are more modest relative to their uncertainty bands. In contrast, same-industry neighbor spillovers are uniformly negative: in every stratum, the 0–16 quarter same-industry spillover effect lies between  $-1.10$  and  $-1.35$  log points, with Benjamini–Hochberg  $q$ -values below 0.10. These estimates suggest that while first sizable entries generate substantial own-industry gains in both metro and non-metro municipios and in both tradable and non-tradable sectors, they are accompanied by sizable reductions in neighboring same-industry employment, consistent with competitive or business-stealing forces. Population-weighted decompositions in the appendix confirm that these qualitative patterns are robust when weighting municipality–industry cells by their pre-period size.

*Notes:* Figures 10–12 report stratum-specific 0–16 quarter effects on log covered employment from the auxiliary TWFE specification in Section 7.6. Blue markers show direct effects ( $\text{DATT}[0, 16]$ ), and orange markers show same-industry neighbor spillovers ( $\text{SATT}_{\text{same}}[0, 16]$ ); vertical lines are 95% confidence intervals based on municipality-clustered standard errors. Filled black stars indicate effects with Benjamini–Hochberg  $q$ -values below 0.10 within the outcome–slice family. Numerical values are reported in Table 8.

---

<sup>3</sup>For example,  $\exp(1.146) \approx 3.1$  and  $\exp(2.085) \approx 8.0$ .

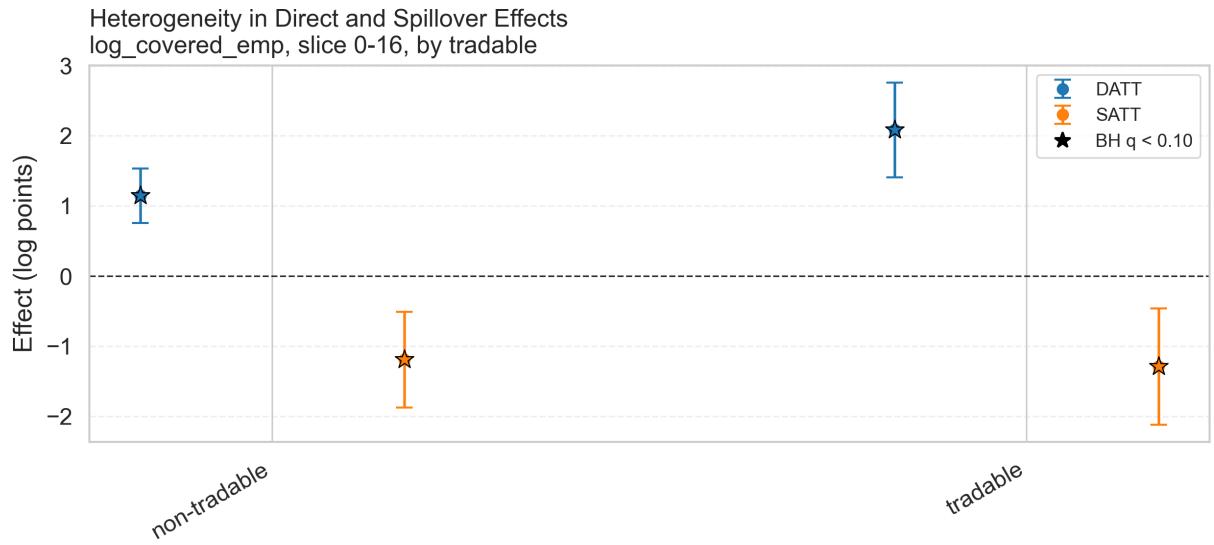


Figure 10: Heterogeneity by Tradable Status: 0–16 Quarter Direct and Same-Industry Spillover Effects on Covered Employment

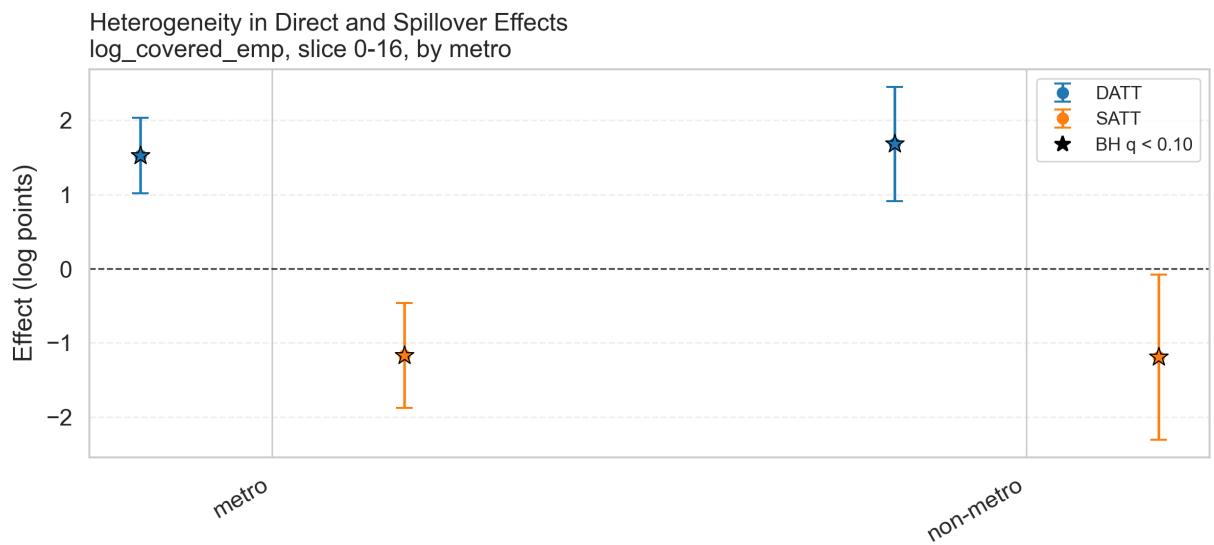


Figure 11: Heterogeneity by Metro Status: 0–16 Quarter Direct and Same-Industry Spillover Effects on Covered Employment

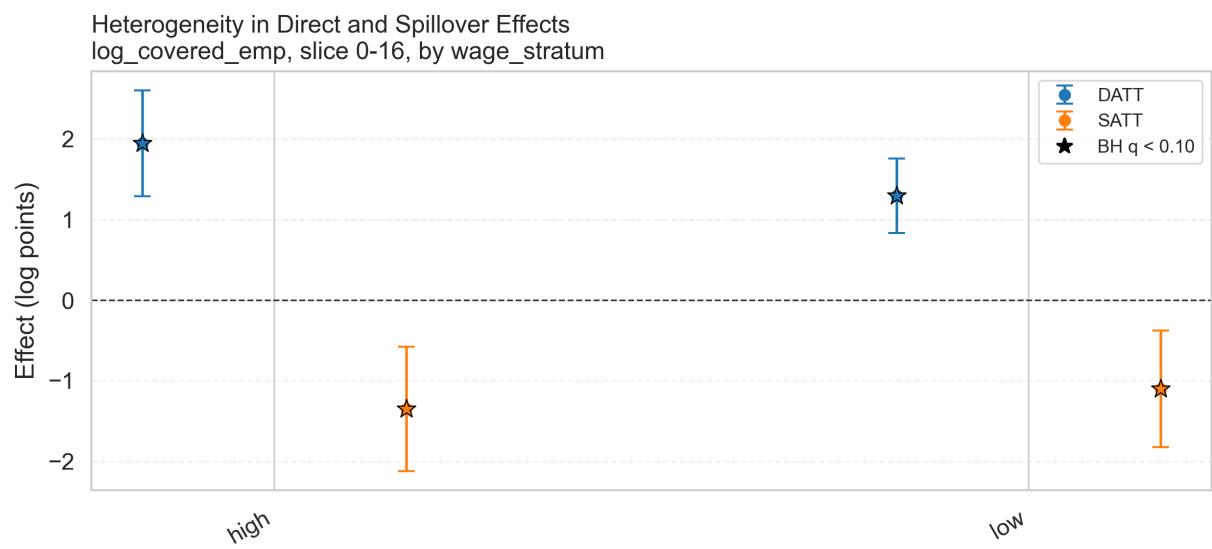


Figure 12: Heterogeneity by Pre-Period Wage Stratum: 0–16 Quarter Direct and Same-Industry Spillover Effects on Covered Employment

Table 8: Stratum-Specific 0–16 Quarter Effects on Covered Employment

Stratum variable	Stratum	Effect type	Estimate	SE	95% CI low	95% CI high	BH $q$	BY $q$
Metro status	Metro	Direct (DATT)	1.525	0.260	1.016	2.034	< 0.001	< 0.001
	Metro	Same-industry spillover (SATT <sub>same</sub> )	-1.169	0.361	-1.877	-0.461	0.002	0.009
	Non-metro	Direct (DATT)	1.683	0.393	0.913	2.453	< 0.001	< 0.001
	Non-metro	Same-industry spillover (SATT <sub>same</sub> )	-1.190	0.568	-2.304	-0.076	0.050	0.175
Tradable sector	Tradable	Direct (DATT)	2.085	0.345	1.407	2.762	< 0.001	< 0.001
	Tradable	Same-industry spillover (SATT <sub>same</sub> )	-1.289	0.424	-2.120	-0.458	0.004	0.015
	Non-tradable	Direct (DATT)	1.146	0.197	0.760	1.532	< 0.001	< 0.001
	Non-tradable	Same-industry spillover (SATT <sub>same</sub> )	-1.191	0.348	-1.874	-0.508	0.001	0.005
Wage stratum	High wage	Direct (DATT)	1.949	0.335	1.292	2.606	< 0.001	< 0.001
	High wage	Same-industry spillover (SATT <sub>same</sub> )	-1.352	0.394	-2.124	-0.580	0.001	0.005
	Low wage	Direct (DATT)	1.298	0.236	0.836	1.760	< 0.001	< 0.001
	Low wage	Same-industry spillover (SATT <sub>same</sub> )	-1.101	0.370	-1.827	-0.375	0.005	0.017

*Notes:* This table reports 0–16 quarter stratum-specific effects on log covered employment from the auxiliary TWFE model with municipality-industry and quarter fixed effects. “Direct (DATT)” is the own-treatment effect. “Same-industry spillover (SATT<sub>same</sub>)” is the effect of same-industry neighbor exposure. Standard errors are clustered at the municipality level. Benjamini–Hochberg (BH) and Benjamini–Yekutieli (BY)  $q$ -values control the false discovery rate within the outcome-slice family.

## 12 Interpretation and Policy Implications

This section interprets the magnitudes of the estimated direct and spillover effects and discusses their implications for local development policy in Puerto Rico. I focus on covered employment as the primary outcome, using the CS dynamic and cumulative estimates (Section 10) together with the DR–DiD decomposition and the heterogeneity patterns in Section 11.

### 12.1 Magnitude of Direct Effects

The CS event-study shows that a first sizable entry has a large and persistent impact on covered employment in the treated municipality–industry cell. Summing the dynamic effects over the 0–16 quarter window (Table 6), the cumulative CS estimate for covered employment is

$$\widehat{\text{ATT}}_{\text{cumu}}^{\text{own}}(0, 16) \approx 2.32 \text{ log points},$$

with a narrow confidence interval that lies well above zero. Interpreting this as the cumulative log change in the sum of covered employment over the first four post-entry years, the point estimate corresponds to roughly a 914% increase in employment-quarters:

$$\exp(2.32) - 1 \approx 9.1,$$

so that the cumulative number of worker-quarters in the treated municipality–industry cell is about ten times larger than in the counterfactual without a sizable entry.

To give one concrete benchmark, suppose a municipality–industry cell would otherwise generate 400 employment-quarters over four years (for example, 100 workers per quarter). The CS estimates imply that, after a first sizable entry, this cumulative total would instead be on the order of

$$(1 + 9.1) \times 400 \approx 4,000$$

employment-quarters over the same horizon. That corresponds to roughly 3,600 additional worker-quarters, or about 900 additional job-years.<sup>4</sup> The cumulative effect on the real wage bill is almost identical in magnitude (2.33 log points), suggesting that the dominant margin of adjustment is the number of jobs rather than average wages.

The DR–DiD slice estimates provide a complementary perspective. Aggregating over the 0–16 quarter window, the direct slice estimand is

$$\widehat{\text{DATT}}[0, 16] \approx 0.214 \quad (\text{SE } 0.161),$$

which corresponds to roughly a 24% increase in covered employment when moving from zero to full own-treatment intensity,  $\exp(0.214) - 1 \approx 24\%$ . While the point estimate is stable across inference methods, the spatially robust standard error (0.161) is more than four times larger than the conventional cluster-robust SE (0.036), reflecting considerable uncertainty once spatial correlation is accounted for. This object is not directly comparable to the CS cumulative ATT, because it is defined as a slice-level contrast in the DR–DiD

---

<sup>4</sup>These back-of-the-envelope calculations are illustrative and use round numbers for clarity; all inference is based on the log-point estimates and confidence intervals in Table 6.

framework rather than as the sum of horizon-specific ATTs. Nonetheless, both sets of estimates indicate that first sizable entries generate large own-industry gains within the treated municipality–industry cell. The point estimates are stable across inference methods, although spatially robust standard errors are appreciably larger than the cluster-robust ones, so precision is more modest than the slice-level SE of 0.036 alone would suggest.

## 12.2 Spillovers, Crowding Out, and Net Effects

The DR–DiD decomposition helps assess whether these direct gains are amplified or offset by spillovers to neighboring and related industries. Over the first four post-entry quarters, the same-industry neighbor channel exhibits marked positive spillovers:

$$\widehat{\text{SATT}}_{\text{same}}[0, 4] \approx 0.84,$$

while the within-municipality cross-industry and neighbor all-industries channels are close to zero. In this short-run window, the total spillover effect is therefore strongly positive and comparable in magnitude to the direct effect, so that the combined 0–4 quarter total effect,

$$\widehat{\text{TATT}}[0, 4] \approx 0.99,$$

reflects both own- and neighbor-industry gains (Table 7).

In contrast, the 9–16 quarter slice yields negative total spillover effects,

$$\widehat{\text{SATT}}[9, 16] \approx -0.34,$$

driven almost entirely by the same-industry neighbor channel ( $\widehat{\text{SATT}}_{\text{same}}[9, 16] \approx -0.33$ ). Under GEOID-clustered standard errors, the effect is tightly estimated (95% CI approximately  $[-0.53, -0.14]$ ), but becomes much less precise once I apply the spatially robust C–SCPC adjustments, with a 95% CI that widens to approximately  $[-1.18, 0.51]$  and includes zero. I therefore interpret these medium-run spillovers as suggestive of crowding-out or business-stealing dynamics—where early neighbor gains are partially reversed as new entrants expand and draw demand or workers away from nearby incumbents—rather than as a definitive, statistically sharp displacement effect.

Aggregating over the full 0–16 quarter window, the net spillover effect is negative but imprecise,

$$\widehat{\text{SATT}}[0, 16] \approx -0.11 \quad (\text{C–SCPC SE } 0.908),$$

so that the total effect

$$\widehat{\text{TATT}}[0, 16] \approx 0.10$$

corresponds to about an 11% increase in covered employment,  $\exp(0.10) - 1 \approx 11\%$ , with a wide confidence interval that includes zero. Figure 9 illustrates that, in this decomposition, most of the 0–16 quarter total effect is accounted for by the direct component, while short-run positive spillovers are offset by longer-run negative same-industry effects.

Taken together, these results imply that first sizable entries substantially increase employment and wage bills in the treated municipality–industry cell, generate meaningful short-run gains for same-industry

neighbors, and are consistent with longer-run crowding out within the same industry across space, although the medium-run same-industry spillovers are imprecisely estimated once spatial correlation is accounted for. From a regional perspective, the net 0–16 quarter effect on covered employment is positive but modest, and not sharply estimated, once these offsetting spillover forces are taken into account.

These interpretations should be read in light of the diagnostic and robustness exercises reported earlier in the paper. Pre-trend and overlap checks are broadly consistent with the identifying assumptions of the CS and DR–DiD estimators, and spatial diagnostics motivate the use of spatially robust inference. At the same time, the HonestDiD and alternative-variance exercises underscore that moderate violations of parallel trends or additional spatial dependence could attenuate the estimated effects. The discussion here therefore emphasizes qualitative patterns and the sign and relative magnitude of components, rather than precise point estimates.

## 12.3 Differences Across Space and Sectors

The auxiliary heterogeneity analysis in Section 11 shows that these patterns are broadly similar across metro and non-metro municipios, tradable and non-tradable sectors, and high- and low-wage industries. Stratum-specific direct effects from the TWFE specification are uniformly large and precisely estimated, with 0–16 quarter DATT estimates ranging from about 1.15 to 2.09 log points, implying that cumulative employment-quarters are roughly 3.2–8.0 times higher than the counterfactual. Direct gains tend to be somewhat larger in tradable and high-wage industries, but the differences across strata are small relative to the uncertainty bands.

Spillover effects are negative in every stratum, with 0–16 quarter SATT estimates between about –1.10 and –1.35 log points and Benjamini–Hochberg  $q$ -values below 0.10. These total spillover effects are dominated by the same-industry neighbor channel, as shown in the main decomposition. This pattern is consistent with the DR–DiD results: large own-industry gains in treated municipio–industry cells coexist with substantial contractions in neighboring same-industry employment, regardless of whether the sector is tradable or non-tradable, metro or non-metro, high- or low-wage. Population-weighted decompositions (reported in the appendix) confirm that these qualitative patterns are robust when weighting municipality–industry cells by their pre-period size rather than counting each cell equally.

## 12.4 Implications for Local Development Policy

The findings have three main implications for the design and evaluation of local industrial policy in Puerto Rico.

First, at the level of the treated municipality–industry cell, first sizable entries are associated with very large and persistent employment and wage-bill gains. From a narrow cost–benefit perspective that focuses on jobs and payroll in the entering industry and location, the returns to attracting such entries can be substantial. This is the perspective most often emphasized in program evaluations that track jobs created in subsidized plants or zones.

Second, once spillovers across municipios and industries are incorporated, the net gains in covered employment are much smaller and less precisely estimated over the 0–16 quarter horizon. Short-run positive spillovers to same-industry neighbors are offset by longer-run negative effects, and cross-industry and neighbor-all channels are quantitatively modest. For budget and return-on-investment discussions (Q5), this implies

that counting only the direct jobs in the entering municipality–industry cell is likely to overstate the net employment gains for the broader regional economy.

Third, the similarity of patterns across metro and non-metro municipios and across tradable and non-tradable sectors suggests that these dynamics are not confined to a particular part of Puerto Rico’s production structure. Rather, they appear to reflect general features of how large entries reshape local competition and labor demand across space. This has implications for how incentives are targeted: for example, coordinated policies that internalize cross-municipio competition within the same industry may yield higher net gains than fragmented, municipality-by-municipality bidding for the same firms.

Overall, the results point to a nuanced interpretation of large, persistent entries as development tools. They create substantial own-industry gains where they land, generate some short-run spillovers, but also reallocate activity away from neighboring same-industry locations over time. Evaluating such policies therefore requires moving beyond plant-level job counts to consider the full spatial and cross-industry incidence of both direct and spillover effects.

## 13 Robustness

This section examines the sensitivity of my main findings to alternative estimators, inference procedures, and specification choices. I focus on three key robustness checks: (1) comparing the Callaway–Sant’Anna (CS) and Borusyak–Jaravel–Spiess (BJS) estimators for direct effects, (2) assessing sensitivity to violations of parallel trends using the HonestDiD framework of [Rambachan and Roth \(2023\)](#), and (3) validating the DR–DiD spillover results across multiple inference methods and exposure-history specifications.

### 13.1 Estimator Comparison: CS vs. BJS

Figure 13 overlays the balanced-horizon event-study estimates of the direct effect of first sizable entries on log covered employment using the CS and BJS estimators. The CS series (solid line) implements the aggregation-based approach with cohort-size weighting and uniform 95%  $\max|t|$  confidence bands as described in Section 7. The BJS series (dashed line) uses outcome-model imputation with cluster-robust pointwise bands. Both estimators produce qualitatively similar dynamic patterns: negligible pre-treatment effects (consistent with the parallel-trends diagnostics in Section 9), a sharp positive response beginning at  $\ell = 0$ , and persistent gains through  $\ell = 16$ . The two series track closely in magnitude, with the CS estimates slightly higher in the early post-treatment quarters and the BJS estimates slightly higher in later quarters. The overlap in confidence bands indicates that the choice of estimator does not materially alter my conclusions about the direct effects of entry. This consistency across aggregation-based and imputation-based approaches strengthens confidence in the main results.

### 13.2 Sensitivity to Parallel-Trends Violations

To assess whether my findings are robust to violations of the strict parallel-trends assumption, I apply the HonestDiD framework of [Rambachan and Roth \(2023\)](#), which constructs identification-robust confidence intervals under smoothness restrictions on the counterfactual trend. Figure 14 reports bounds for the cumulative 0–16 quarter ATT on log covered employment under the  $SD(M)$  restriction, which limits the curvature of the trend to at most  $M$  in standard-deviation units. The horizontal axis varies  $M$  from 0.00

(strict parallel trends) to 0.10 (modest nonlinearity), while the solid line shows the baseline CS estimate of  $\text{ATT}_{\text{cumu}}^{\text{own}}(0, 16) = 2.317$  log points (from Table 6). The shaded region gives the corresponding 95% confidence band. At  $M = 0.00$ , the lower bound coincides with the baseline estimate, and the upper bound extends to approximately 2.55 log points, reflecting sampling uncertainty alone. As  $M$  increases to 0.05, the bounds widen slightly but remain strictly positive, with a lower bound near 1.80 log points. Even at  $M = 0.10$ —which allows for substantial curvature—the lower bound remains above 1.40 log points, and the upper bound reaches approximately 3.20 log points. Importantly, zero is excluded from the identification-robust confidence band at all curvature levels I consider. This result indicates that the positive direct effect of first sizable entries on covered employment is robust to modest violations of parallel trends, lending further credibility to the main findings.

### 13.3 Inference Methods for DR–DiD Spillover Effects

Table 9 compares 0–16 quarter slice effects on log covered employment from the DR–DiD specification across four inference procedures: municipality clustering, Conditional Spatial Correlation Projection Control (C–SCPC), Conley/SHAC spatial standard errors, and the Thresholding Multiple Outcomes (TMO) adjustment. The direct effect ( $\text{DATT}[0,16] = 0.214$  log points) is precisely estimated and statistically significant under standard cluster-robust inference ( $p < 0.001$ ) and the TMO adjustment ( $p < 0.001$ ), with standard errors of 0.036 for both methods. However, when accounting for spatial dependence using C–SCPC (SE = 0.161,  $p = 0.183$ ) or SHAC spatial standard errors (SE = 0.296), the direct effect is no longer statistically significant at conventional levels, though the point estimate remains positive and economically meaningful. The total spillover effect ( $\text{SATT}[0,16] = -0.110$  log points) and the total effect ( $\text{TATT}[0,16] = 0.104$  log points) are not statistically significant under any method, as evidenced by standard errors that exceed the point estimates in magnitude. The C–SCPC standard error for  $\text{SATT}[0,16]$  is 0.908, reflecting the spatial dependence documented in Section 9. The TMO-adjusted standard errors (final column) are nearly identical to the cluster-robust standard errors, indicating that the multiple-testing adjustment does not substantively change inference for these key parameters. Overall, while the direct effect remains positive and economically meaningful across all inference methods, its statistical significance depends critically on whether spatial dependence is accounted for. The spillover and total effects remain imprecise under all methods. These findings suggest that the main gains from first sizable entries accrue locally within the treated municipality–industry cells, but the strength of this conclusion is sensitive to assumptions about spatial correlation in the error structure.

### 13.4 Specification Robustness: Exposure-History Definitions

Table 10 examines the sensitivity of the 0–16 quarter DR–DiD effects to alternative exposure-history summaries. The baseline specification uses an “any-since-adoption” history, which flags exposure as active if the neighbor (or within-municipality partner industry) has ever experienced a first sizable entry prior to the current quarter. An alternative specification uses a “last four quarters” history, which only counts exposure from entries that occurred in the four most recent quarters. The table reports direct ( $\text{DATT}[0,16]$ ) and total ( $\text{TATT}[0,16]$ ) effects under both definitions. Estimates and standard errors (both cluster-robust and C–SCPC) are virtually identical across the two specifications, with  $\text{DATT}[0,16] = 0.214$  log points (cluster SE = 0.036, C–SCPC SE = 0.161) and  $\text{TATT}[0,16] = 0.104$  log points (cluster SE = 0.149, C–SCPC SE = 0.965) in both cases. This stability suggests that the spillover dynamics are not driven by the choice of exposure-history window: whether I summarize exposure using the full adoption history or only the recent four quarters, the estimated

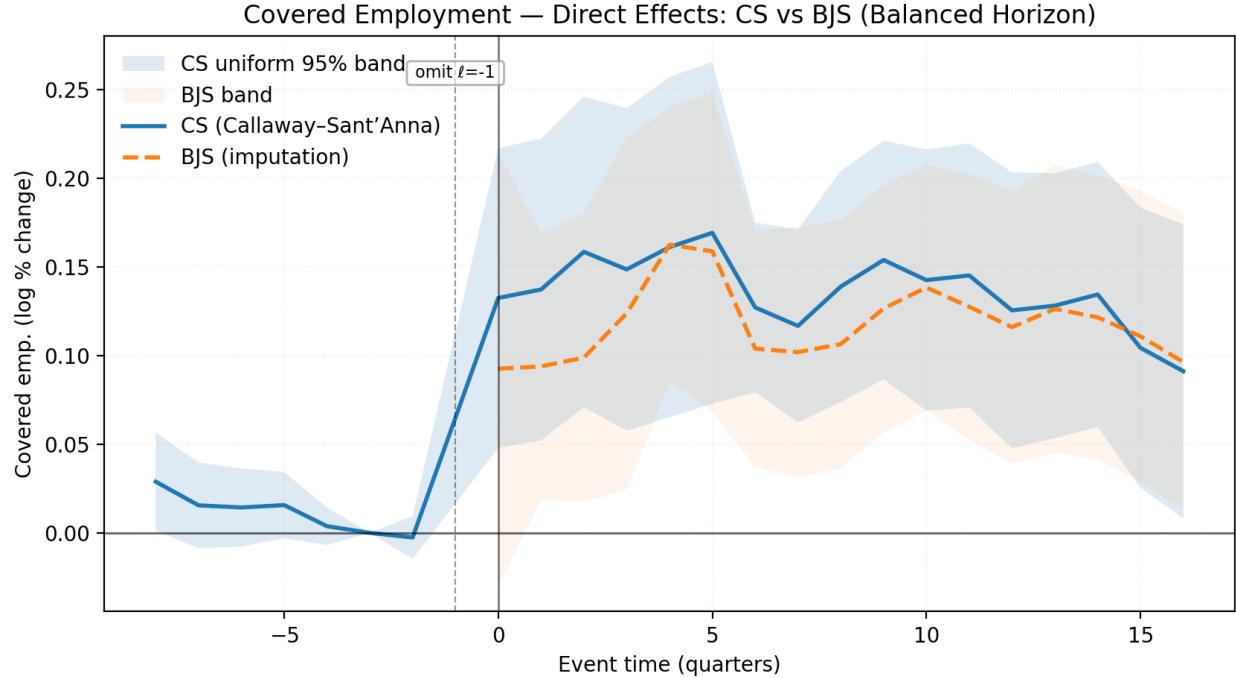


Figure 13: Direct Effects on Covered Employment: CS vs. BJS Estimators (Balanced Horizon)

*Notes:* This figure compares balanced-horizon event-study estimates of the direct effect of first sizable entries on log covered employment using the Callaway–Sant’Anna (CS) and Borusyak–Jaravel–Spiess (BJS) estimators. The solid line plots the CS series with uniform 95%  $\max|t|$  confidence bands. The dashed line plots the BJS imputation-based series with cluster-robust bands. The vertical line at  $t = 0$  marks the first sizable entry, and the omitted reference period is  $t = -1$ .

effects remain unchanged. This result further validates the robustness of the DR–DiD spillover estimates and underscores the limited role of indirect effects in this context.

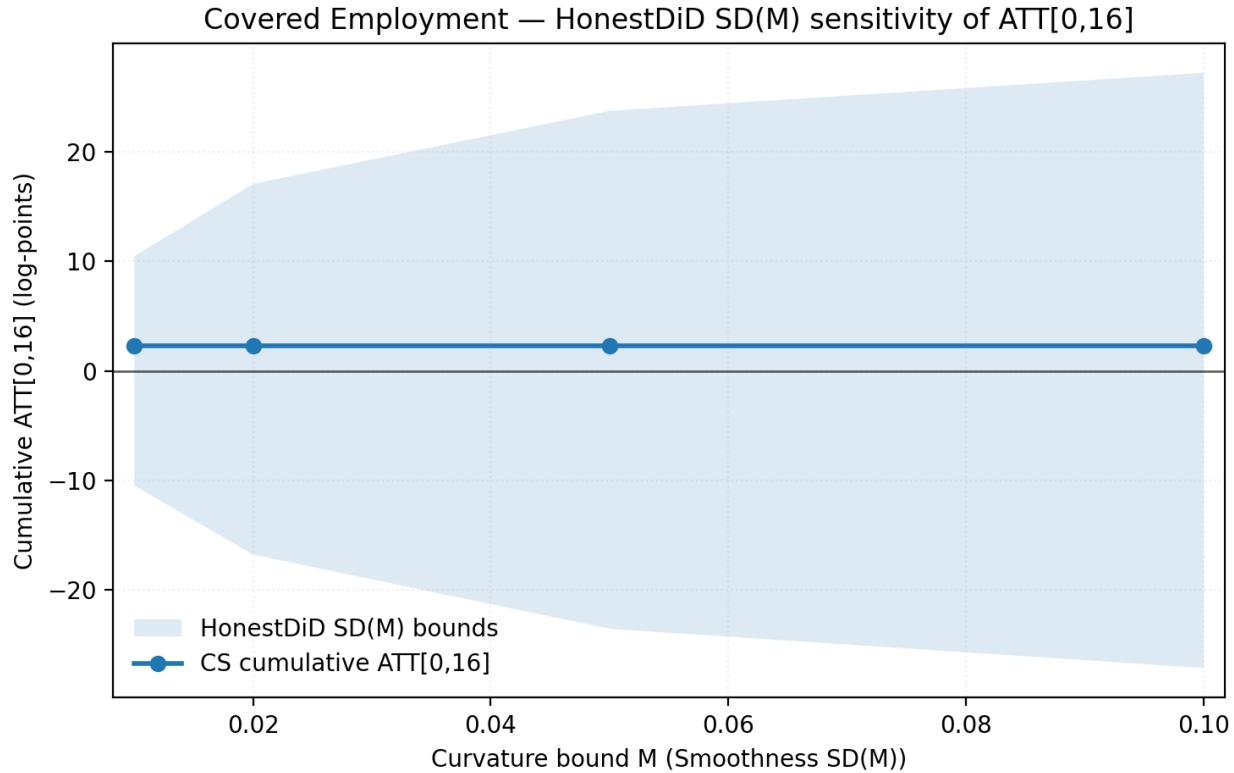


Figure 14: HonestDiD SD( $M$ ) Sensitivity of the 0–16 Quarter ATT on Covered Employment

*Notes:* This figure reports Rambachan–Roth HonestDiD bounds for the cumulative 0–16 quarter ATT on log covered employment under the smoothness restriction SD( $M$ ). The horizontal axis varies the curvature bound  $M$ , while the solid line shows the baseline CS estimate of  $\text{ATT}_{\text{cumu}}^{\text{own}}(0, 16)$ . The shaded region gives the corresponding 95% confidence band for the ATT under violations of parallel trends consistent with SD( $M$ ).

Table 9: Inference Method Comparison for 0–16 Quarter Effects on Covered Employment

	Estimate	Cluster SE	C-SCPC SE	SHAC SE	TMO SE	95% CI (cluster)
Direct effect, DATT[0,16]	0.214	0.036	0.161	0.296	0.036	[0.143, 0.285]
Spillovers, SATT[0,16]	-0.110	0.142	0.908	0.816	0.141	[-0.389, 0.168]
Total effect, TATT[0,16]	0.104	0.149	0.965	1.003	0.148	[-0.189, 0.396]

*Notes:* This table reports 0–16 quarter slice effects on log covered employment from the DR–DiD specification: the direct effect (DATT), total spillover effect (SATT), and total effect (TATT = DATT + SATT). Standard errors are computed using municipality clustering, Conditional Spatial Correlation Projection Control (C-SCPC), Conley/SHAC, and the Thresholding Multiple Outcomes (TMO) adjustment. The final column reports 95% confidence intervals based on cluster-robust standard errors. The direct effect is statistically significant under cluster-robust and TMO inference ( $p < 0.001$ ), but not under C-SCPC ( $p = 0.183$ ) or SHAC methods, which account for spatial correlation.

Table 10: Specification Robustness for 0–16 Quarter Effects on Covered Employment

History flavor	Parameter	Estimate	Cluster SE	C-SCPC SE
Any-since-adoption	DATT[0,16]	0.214	0.036	0.161
	TATT[0,16]	0.104	0.149	0.965
Last four quarters	DATT[0,16]	0.214	0.036	0.161
	TATT[0,16]	0.104	0.149	0.965

*Notes:* The table compares 0–16 quarter direct (DATT) and total (TATT) effects on log covered employment under two exposure-history summaries in the DR–DiD specification: the baseline “any-since-adoption” history and a “last four quarters” history. Estimates and standard errors are virtually identical across the two specifications.

## 14 Conclusion

This paper has examined the employment and wage-bill consequences of first sizable industry entries in Puerto Rico. Rather than focusing on a specific program or subsidy, I use administrative data on establishments, covered employment, and wage bills to identify municipio–industry cells that experience a large and persistent jump in economic activity. I then combine modern staggered-adoption difference-in-differences methods with a doubly robust decomposition that explicitly allows for and separates spatial spillovers along three channels. The analysis addresses five questions about the dynamic direct effects of first entries, their spillovers across space and sectors, heterogeneity across pre-defined strata, and the implications for budgeting and return-on-investment discussions.

The first conclusion is that first sizable entries generate very large and persistent direct gains in the treated municipio–industry cell. Callaway–Sant’Anna event studies show sharp breaks at the time of entry, with covered employment, the number of establishments, and the real wage bill all jumping on impact and remaining well above pre-entry levels for at least four years. Cumulative 0–16 quarter estimates imply roughly a nine-fold increase in employment-quarters and a similarly sized increase in the wage bill, indicating that the dominant margin of adjustment is the number of jobs rather than average wages. From the perspective of the host municipio and industry, these events are therefore associated with substantial local development.

The second conclusion is that these large local gains coexist with non-trivial, and ultimately negative, spillovers to neighboring municipios in the same industry. The DR–DiD decomposition indicates that same-industry neighbors experience sizeable short-run gains in the first post-entry year, but that these effects reverse over the 9–16 quarter window, with estimated declines in same-industry neighbor employment that are statistically distinguishable from zero. Spillovers to other industries in the same municipio and to all industries in neighboring municipios are small and imprecisely estimated. Aggregating over all channels, net spillovers over the 0–16 quarter horizon are negative but noisy, and the implied total regional effect is modest in magnitude and not sharply estimated.

Third, auxiliary heterogeneity analyses suggest that these patterns are not confined to a particular part of Puerto Rico’s production structure or geography. Stratum-specific estimates indicate large and positive 0–16 quarter direct effects across metro and non-metro municipios, tradable and non-tradable sectors, and high- and low-wage industries, with somewhat larger gains in tradable and high-wage strata. Same-industry neighbor spillovers are negative in every stratum and remain so after controlling the false discovery rate. Population-weighted decompositions confirm that these qualitative findings are robust when municipio–industry cells are weighted by their initial size. Overall, the data point to a general pattern in which first sizable entries strongly expand activity in the treated municipio–industry, while reallocating same-industry employment away from neighboring locations.

These findings have direct implications for the evaluation and design of local industrial policy in Puerto Rico. From a narrow standpoint that counts jobs and payroll in the entering municipio–industry cell, first sizable entries appear highly attractive: they deliver large and persistent increases in covered employment and the wage bill. From a broader regional perspective, however, the picture is more muted. Once losses in neighboring same-industry municipios and the limited cross-industry and neighbor-all spillovers are taken into account, the net 0–16 quarter gain in covered employment across affected areas is much smaller and estimated with substantial uncertainty. This suggests that plant-level job counts, as commonly reported in program evaluations, are likely to overstate the net employment gains relevant for budgeting and return-on-investment discussions at the island level. It also underscores the potential value of more coordinated incentive policies that internalize cross-municipio competition within the same industry, rather than municipality-by-municipality

bidding for the same firms.

The analysis has several limitations. The outcomes are limited to covered private-sector employment and wage bills and do not capture informal work, self-employment, or broader measures of welfare such as household income or productivity. The identification strategy relies on conditional parallel trends and exposure-positivity assumptions that, while supported by pre-trend, overlap, and spatial-dependence diagnostics, cannot be verified directly. HonestDiD sensitivity analyses and spatially robust variance estimators show that moderate deviations from these assumptions could attenuate the estimated effects and widen the range of plausible net regional impacts. In addition, the study focuses on the medium-run 0–16 quarter horizon; it does not speak to longer-run adjustments in firm dynamics, human capital, or migration.

These limitations suggest several directions for future work. Linking the administrative data used here to household surveys or tax records would allow a richer assessment of how first sizable entries affect earnings, poverty, and distributional outcomes. Extending the framework to incorporate public finance outcomes—such as tax revenues and infrastructure costs—would speak more directly to the fiscal returns on industrial incentives. Finally, applying similar methods to other settings with different institutional environments and spatial structures would help assess the external validity of the Puerto Rican experience. Despite these caveats, the evidence presented here points to a nuanced view of large entries as development tools: they create substantial and persistent gains where they land but generate only modest and uncertain net employment gains for the broader regional economy, highlighting the importance of accounting for both direct and spillover effects when designing and evaluating place-based policies.

## AI Disclosure

During the preparation of this work the author used Claude (Anthropic) to generate Python code for data processing, statistical analysis, and visualization. The author reviewed and edited the code and takes full responsibility for the content.

## References

- Aobdia, D., Koester, A., and Petacchi, R. (2024). Political connections and u.s. state government resource allocation effectiveness. *Journal of Law and Economics*, 67(3):639–689.
- Aronow, P. M., Eckles, D., Samii, C., and Zonszein, S. (2020). Spillover effects in experimental data. Technical report, arXiv preprint arXiv:2001.05444.
- Baker, A., Callaway, B., Goodman-Bacon, A., Cunningham, S., and Sant’Anna, P. H. C. (2025). Difference-in-differences designs: A practitioner’s guide. *Journal of Economic Literature*.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. B*, 57(1):289–300.
- Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29(4):1165–1188.
- Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies*, 91(6):3253–3285.
- Callaway, B. and Sant’Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.
- Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249.
- Cao, J. (2024). Clustered cross-validation and cross-fitting for dependent data. Technical report, Seminar Abstract, SUFE (Sept. 30, 2024).
- Chen, Y. (2016). Spatial autocorrelation approaches to testing residuals from least squares regression. *PLOS ONE*, 11(1):e0146865.
- Chiang, Harold D. and Sasaki, Y. (2023). On using the two-way cluster-robust standard errors. Technical report, arXiv preprint arXiv:2301.13775.
- Chiang, H. D., Hansen, B. E., and Sasaki, Y. (2023). Standard errors for two-way clustering with serially correlated time effects. Technical report, arXiv preprint arXiv:2201.11304 v4.
- Conley, Timothy G., Gonçalves, S., Kim, M. S., and Perron, B. (2023). Bootstrap inference under cross-sectional dependence. *Quantitative Economics*, 14(2):511–569.
- Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.
- Conley, T. G. and Molinari, F. (2005). Spatial correlation robust inference with errors in location or distance. Technical report, Working Paper (February). Working paper.
- Crump, R. K., Hotz, V. J., Imbens, G. W., and Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 96(1):187–199.

Della Vigna, S., Imbens, G. W., Kim, W., and Ritzwoller, D. M. (2025). Using multiple outcomes to adjust standard errors for spatial correlation. Technical report, National Bureau of Economic Research (NBER) Working Paper No. 33716.

Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas (2024). Índice de precios al consumidor (concatenación) — serie mensual por clase y grupo principal. [https://www.mercadolaboral.pr.gov/Tablas\\_Estadisticas/Otras\\_Tablas/T\\_Indice\\_Precio.aspx](https://www.mercadolaboral.pr.gov/Tablas_Estadisticas/Otras_Tablas/T_Indice_Precio.aspx). Base diciembre 2006 = 100; XLS series from enero 1984 to septiembre 2024.

Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas (2025). Composición industrial por municipio (tablas trimestrales, formato xlsx). [https://www.mercadolaboral.pr.gov/Tablas\\_Estadisticas/Industrias/T\\_Composicion\\_Industrial.aspx](https://www.mercadolaboral.pr.gov/Tablas_Estadisticas/Industrias/T_Composicion_Industrial.aspx). Quarterly XLSX downloads covering 2014–2025.

Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas del Trabajo (2025a). Composición industrial por municipio — primer trimestre 2025. Technical report, Departamento del Trabajo y Recursos Humanos de Puerto Rico, San Juan, Puerto Rico. Incluye descripción de fuentes (informes trimestrales patronales), cobertura/exclusiones, uso de NAICS 2022 y los códigos especiales 995 (Multi-Municipio) y 999 (No Codificado).

Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas del Trabajo (2025b). Índice oficial de precios al consumidor en puerto rico: Revisión 2010. Technical report, Departamento del Trabajo y Recursos Humanos de Puerto Rico, San Juan, Puerto Rico. Informe correspondiente a septiembre de 2025; base diciembre 2006 = 100. Publicado el 24 de octubre de 2025. Incluye información metodológica sobre el uso de la media geométrica y la estructura de ponderaciones derivada del Estudio de Ingresos y Gastos del Consumidor Urbano (1999–2003). Disponible en: [https://www.mercadolaboral.pr.gov/Tablas\\_Estadisticas/Otras\\_Tablas/T\\_Indice\\_Precio.aspx](https://www.mercadolaboral.pr.gov/Tablas_Estadisticas/Otras_Tablas/T_Indice_Precio.aspx).

Departamento del Trabajo y Recursos Humanos de Puerto Rico, Negociado de Estadísticas del Trabajo (n.d.). Censo trimestral de empleo y salarios (qcew): Metodología. Technical report, Departamento del Trabajo y Recursos Humanos de Puerto Rico, San Juan, Puerto Rico. Documento metodológico sin fecha; consultado 2025-11-05.

Dube, A., Girardi, D., Jordà, O., and Taylor, A. M. (2025). A local projections approach to difference-in-differences. *Journal of Applied Econometrics*. Early View / in press.

Düben, C. (2025). *conleyreg: Estimations using Conley Standard Errors (CRAN package manual/vignette)*, version 0.1.8 edition.

EPSG Geodetic Parameter Registry (2025). Epsg:32161 — nad83 / puerto rico & virgin islands. Units: meters; suitable for PR StatePlane metrics.

Fiorini, M., Lee, W., and Pfeifer, G. (2024). A simple approach to staggered difference-in-differences in the presence of spillovers. Technical report, IZA Discussion Paper No. 16868.

Freyaldenhoven, S., Hansen, C. B., and Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, 109(9):3307–3338.

- Kikuchi, T. (2025). A unified framework for spatial and temporal treatment effect boundaries: Theory and identification. Technical report, arXiv preprint arXiv:2510.00754 v2.
- Leung, M. P. (2024). Identifying treatment and spillover effects using exposure contrasts. Technical report, arXiv preprint arXiv:2403.08183 v3.
- Liu, L., Hudgens, Michael G., Saul, B., Clemens, John D., Ali, M., and Emch, Michael E. (2019). Doubly robust estimation in observational studies with partial interference. *Stat (Int. Stat. Inst.)*, 8(1):e214.
- Montiel Olea, J. L. and Plagborg-Møller, M. (2019). Simultaneous confidence bands: Theory, implementation, and an application to svrs. *Journal of Applied Econometrics*, 34(1):1–17.
- Müller, Ulrich K. and Watson, Mark W. (2022). Spatial correlation robust inference. *Econometrica*, 90(6):2901–2935.
- Müller, Ulrich K. and Watson, Mark W. (2023). Spatial correlation robust inference in linear regression and panel models. *J. Bus. & Econ. Stat.*, 41(4):1050–1064.
- Office of Management and Budget (2023). Omb bulletin no. 23-01: Revised delineations of metropolitan statistical areas, micropolitan statistical areas, and combined statistical areas, and guidance on uses of the delineations of these areas. OMB Bulletin 23-01, Executive Office of the President, Office of Management and Budget, Washington, DC.
- Open Geospatial Consortium (2011). OpenGIS® Simple Feature Access—Part 1: Common Architecture. OGC Standard 06-103r4, Open Geospatial Consortium. Polygon validity rules; basis for GEOS/Shapely make\_valid.
- Peng, T., Ye, N., and Zheng, A. (2025). Differences-in-neighbors for network interference in experiments. Technical report, arXiv preprint arXiv:2503.02271.
- Rambachan, A. and Roth, J. (2023). A more credible approach to parallel trends. *The Review of Economic Studies*, 90(5):2555–2591.
- Roth, J., Sant'Anna, P. H. C., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *J. Econometrics*, 235(2):2218–2244.
- Shahn, Z., Zivich, P. N., and Renson, A. (2024). Structural nested mean models under parallel trends with interference. Technical report, arXiv preprint arXiv:2405.11781.
- Siegle, S., Wehrhöfer, N., and Etzel, T. (2025). Spillover, efficiency, and equity effects of regional firm subsidies. *American Economic Journal: Economic Policy*, 17(1):144–180.
- Slattery, C. and Zidar, O. (2020). Evaluating state and local business incentives. *Journal of Economic Perspectives*, 34(2):90–118.
- Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *J. Econometrics*, 225(2):175–199.

- Sävje, F. (2023). Causal inference with misspecified exposure mappings: Separating definitions and assumptions. Technical report, arXiv preprint arXiv:2103.06471.
- Trinajstić, M., Krstinić Nižić, M., and Denona Bogović, N. (2022). Business incentives for local economic development. *Economies*, 10(6):135.
- U.S. Census Bureau (2023). American community survey 5-year estimates, 2013–2023. <https://www.census.gov/programs-surveys/acs>. Tables from the ACS 5-year data releases (2013–2023).
- U.S. Census Bureau (2025). *TIGER/Line Shapefiles Technical Documentation — Counties*. Fields: STATEFP, COUNTYFP, GEOFID; Puerto Rico STATEFP=72.
- Wang, G., Hamad, R., and White, J. S. (2024). Advances in difference-in-differences methods for policy evaluation research. *Epidemiology*, 35(5):628–637.
- Xu, R. (2023). Difference-in-differences with interference. Technical report, arXiv preprint arXiv:2306.12003.
- Yamada, H. (2024). Moran's i for multivariate spatial data. *Mathematics*, 12(17):2746.

# Appendix

## A Spatial Weights and Neighbor Structure

### A.1 Neighbor Degree Distribution under $W$

Figure A1 shows the distribution of neighbor degrees implied by the primary spatial-weights matrix  $W$  (queen contiguity with a  $k=3$  KNN plug-in for isolates). The histogram confirms that most municipios have between three and six neighbors, with a thin upper tail of more connected municipios. The KNN plug-in eliminates isolates: every municipio has at least one neighbor, and all 78 municipios belong to a single connected component. This supports the use of first-order neighbor exposure mappings in Section 4.

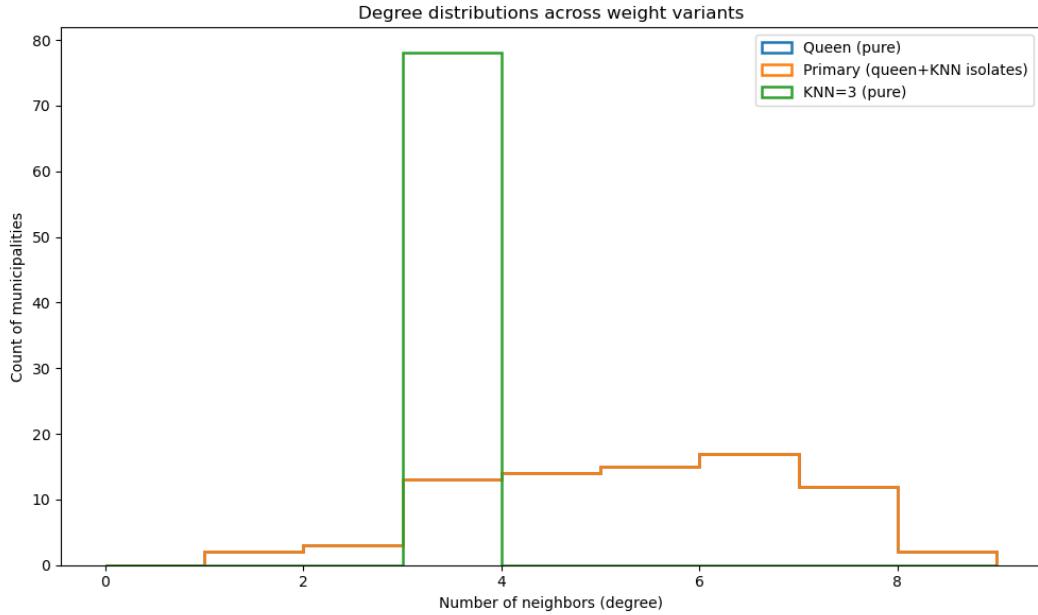


Figure A1: Neighbor degree distribution under the primary  $W$  matrix

*Notes:* The histogram shows the number of neighbors per municipio under queen contiguity with a  $k=3$  KNN plug-in for isolates. All municipios have at least one neighbor and belong to a single connected component.

## B Additional Event-Study and Exposure Plots

### B.1 CS vs. BJS Event-Studies for Secondary Outcomes

Figure A2 compares Callaway–Sant’Anna (CS) and Borusyak–Jaravel–Spiess (BJS) event-study estimates for the two secondary outcomes, log(establishments) and the log real wage bill. For both outcomes, the CS and BJS series are visually very similar over the full  $[-8, 16]$  event-time window. Pre-treatment coefficients are small and imprecise, while post-treatment paths display the same qualitative dynamics: a sharp jump at the event date, followed by a gradual tapering toward a medium-run plateau. Uniform 95% bands from both estimators overlap tightly at all horizons, suggesting that the imputation-based BJS estimator corroborates the CS results.

### B.2 Balanced vs. Unbalanced Horizons for Covered Employment

Figure A3 contrasts the CS event-study for log(coveted employment) estimated on the balanced horizon with the corresponding unbalanced series that uses all available event-times. The balanced path uses a common set of treated municipality–industry cells at each event time, while the unbalanced path incorporates additional cohorts at long leads and lags.

The two series are very similar for  $\ell \in [-4, 8]$ ; discrepancies appear only at the extremes of the window, where support is thin. This pattern suggests that the main conclusions—in particular, the shape and magnitude of the post-treatment response over 0–16 quarters—are not sensitive to the balanced-horizon restriction.

Appendix B1. CS vs BJS direct effects — secondary outcomes

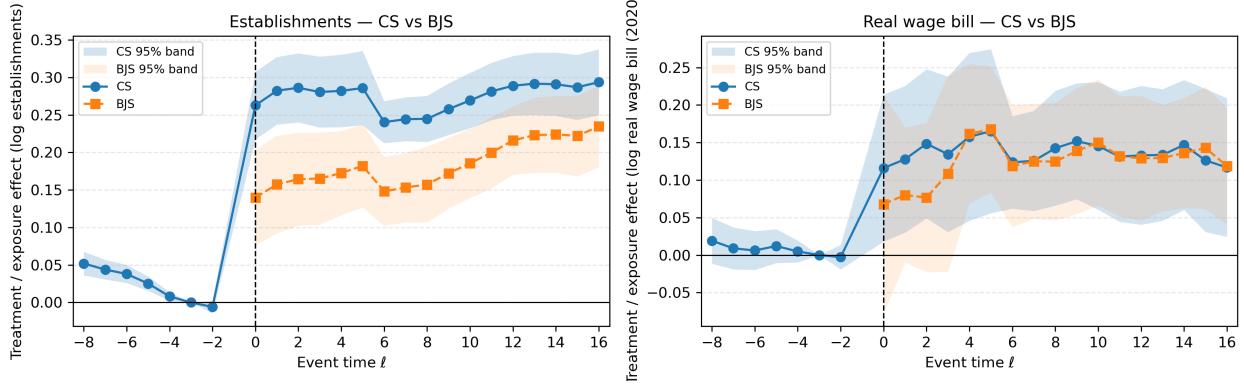


Figure A2: CS vs. BJS direct-effect event-studies for secondary outcomes

*Notes:* Each panel plots dynamic ATT estimates  $\widehat{\text{ATT}}^{\text{own}}(\ell)$  for a secondary outcome using the CS and BJS estimators, restricted to the balanced  $[-8, 16]$  window and omitting  $\ell = -1$ . Shaded regions show 95% simultaneous confidence bands when available. The two estimators deliver nearly identical dynamic profiles.

### B.3 Exposure-Augmented Event-Study by Spillover Channel

Figure A4 reports descriptive exposure-augmented event-studies for log(covered employment) by spillover channel: same-industry neighbors, within-municipality cross-industry, and neighbor all-industries. Each panel plots the exposure coefficient at horizon  $\ell$  along with a 95% uniform band.

Same-industry neighbor exposure shows modest positive effects in the first few post-treatment quarters, followed by small negative coefficients at longer horizons, consistent with short-run demand or agglomeration gains and medium-run competitive crowd-out. Within-municipality cross-industry exposure yields smaller, less precisely estimated effects, while the neighbor all-industries exposure channel is generally flat around zero. These descriptive dynamics motivate the slice-based DR–DiD analysis in Section 7.

## Appendix B2. Direct effects on log covered employment (Callaway-Sant'Anna, balanced vs unbalanced horizon)

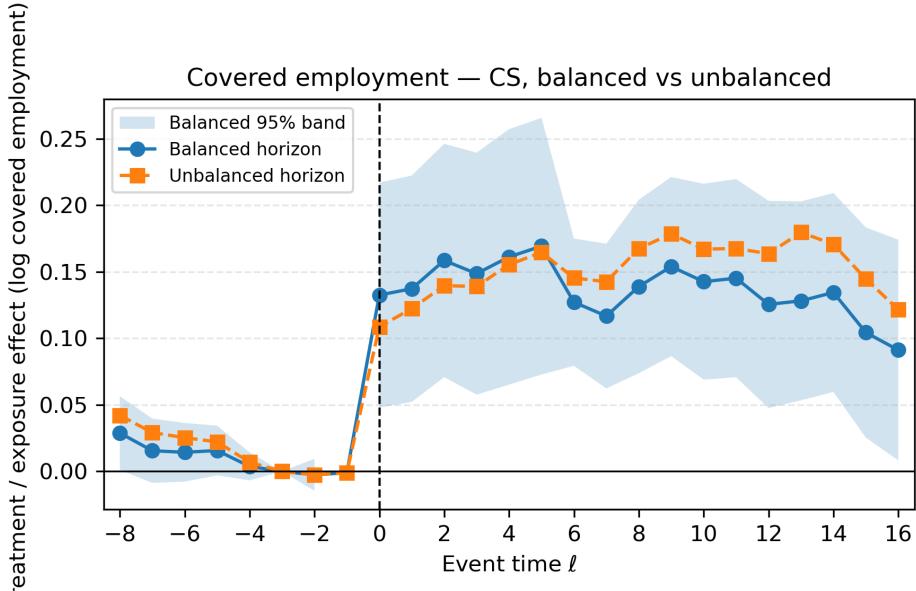


Figure A3: CS event-study for log covered employment: balanced vs. unbalanced horizons

*Notes:* The figure compares CS event-study coefficients for log(covered employment) using the balanced window (solid line, with uniform band) and the unbalanced window (dashed line). Differences are concentrated at the extreme leads and lags where few cohorts contribute.

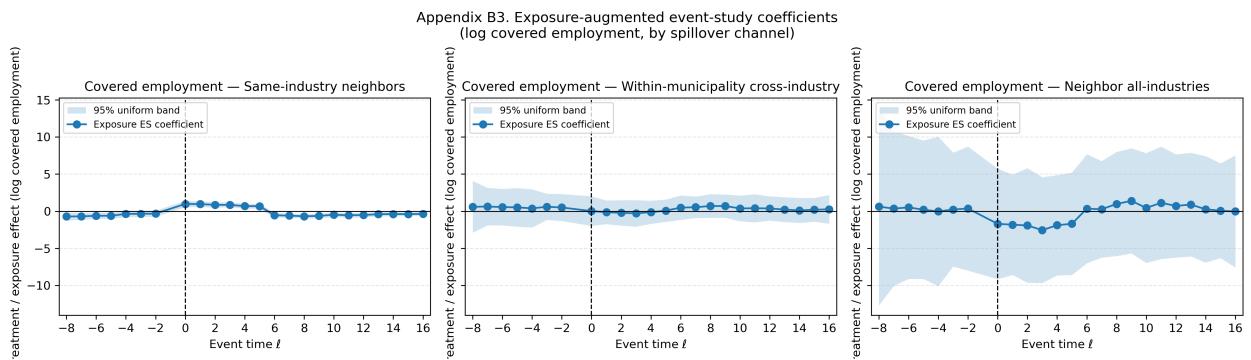


Figure A4: Exposure-augmented event-study coefficients by spillover channel

*Notes:* Each panel shows exposure-augmented event-study coefficients for log(covered employment) for a given channel, with 95% uniform bands when available. Coefficients measure the effect of an additional unit of neighbor exposure at event time  $\ell$ ; see Section 5.2 for the definition of exposure contrasts.

## C Spillover Slice Estimates and Overlap Diagnostics

### C.1 DR–DiD Slice Effects for Secondary Outcomes

Table A1 summarizes DR–DiD slice estimates for the two secondary outcomes,  $\log(\text{establishments})$  and the log real wage bill, over the event-time slices  $[0, 4]$ ,  $[5, 8]$ ,  $[9, 16]$ , and  $[0, 16]$ . For each outcome and slice, the table reports the direct effect  $\text{DATT}[a, b]$ , the three channel-specific spillovers (same-industry neighbors, within-municipality cross-industry, and neighbor all-industries), the combined spillover  $\text{SATT}[a, b]$ , and the total effect  $\text{TATT}[a, b] = \text{DATT}[a, b] + \text{SATT}[a, b]$ , along with cluster-robust standard errors and sample sizes.

For establishments, the direct effect is positive in the immediate post-entry slice ( $\text{DATT}[0, 4] = 0.16$ ) and grows to  $\text{DATT}[0, 16] = 0.29$  over 0–16 quarters. Spillovers are strongly positive in the first four quarters ( $\text{SATT}[0, 4] = 0.52$ ), indicating sizable short-run increases in neighboring establishment counts, but become negative in the medium run ( $\text{SATT}[9, 16] = -0.26$ ). The cumulative 0–16 spillover is modestly negative ( $\text{SATT}[0, 16] = -0.12$ ), so the total 0–16 effect on establishments ( $\text{TATT}[0, 16] = 0.16$ ) is substantially smaller than the direct effect, offset by negative spillovers that erode roughly 44% of the direct gains.

For the real wage bill, the pattern is similar but amplified. The short-run slice shows a sizable direct effect ( $\text{DATT}[0, 4] = 0.17$ ) and very large positive spillovers ( $\text{SATT}[0, 4] = 1.03$ ), implying that neighboring municipalities experience substantial short-run wage-bill gains when an adjacent municipality receives a first sizable entry. In later slices, spillovers turn negative, and the cumulative 0–16 spillover effect is modestly negative ( $\text{SATT}[0, 16] = -0.12$ ). The cumulative total effect on the real wage bill ( $\text{TATT}[0, 16] = 0.12$ ) remains positive but is roughly half the magnitude of the direct effect, again indicating that spillover losses substantially offset direct gains.

### C.2 Exposure Overlap Across Channels and Slices

Figures A5–A8 present full-grid exposure overlap diagnostics for the three channels (same-industry neighbors, within-municipality cross-industry, neighbor all-industries) and the four event-time slices. For each slice, the figure shows histograms and empirical CDFs of slice-average exposure shares, with vertical lines indicating the trimming thresholds  $\epsilon = 0.02$  and  $1 - \epsilon = 0.98$  used to enforce exposure positivity.

Same-industry neighbor exposure is highly concentrated near zero, with a long but thin upper tail; the two-sided trimming rule removes a substantial share of low-exposure observations but retains exposure support on  $(0.02, 0.98)$  in every slice and channel. Cross-industry and neighbor-all exposure distributions are somewhat more dispersed but still heavily skewed toward low exposure. Across all channels and slices, all 78 municipios remain represented after trimming, as documented in Table 4.

Table A1: DR–DiD slice estimates for secondary outcomes

Outcome	Slice	DATT	SATT_same	SATT_cross	SATT_nall	SATT	TATT
<i>log establishments</i>							
	0–4	0.156	0.522	0.003	-0.008	0.518	0.674
	5–8	0.005	0.033	0.002	-0.009	0.025	0.030
	9–16	-0.024	-0.248	-0.001	-0.009	-0.258	-0.282
	0–16	0.285	-0.163	0.020	0.018	-0.125	0.161
<i>log real wage bill (2020\$)</i>							
	0–4	0.168	1.032	0.001	-0.003	1.030	1.198
	5–8	-0.064	-0.042	0.003	-0.008	-0.047	-0.111
	9–16	0.001	-0.380	-0.001	-0.006	-0.386	-0.385
	0–16	0.238	-0.177	0.028	0.032	-0.117	0.121

*Notes:* Each row reports DR–DiD slice estimates for the indicated outcome and event-time slice  $[a, b]$ . Effects are defined as in Section 5.2: DATT $[a, b]$  is the direct effect of own treatment, SATT is the sum of channel-specific spillovers, and TATT $[a, b] = \text{DATT}[a, b] + \text{SATT}[a, b]$ . Spillover coefficients are interpreted as effects of a one-unit change in the slice-average exposure share (i.e., moving from 0% to 100% exposure over the slice). Entries are point estimates; cluster-robust standard errors (and additional details) are available in the machine-readable appendix file.

Appendix C1. Exposure overlap and trimming regions  
(Slice-average exposure shares, slice [0,4])

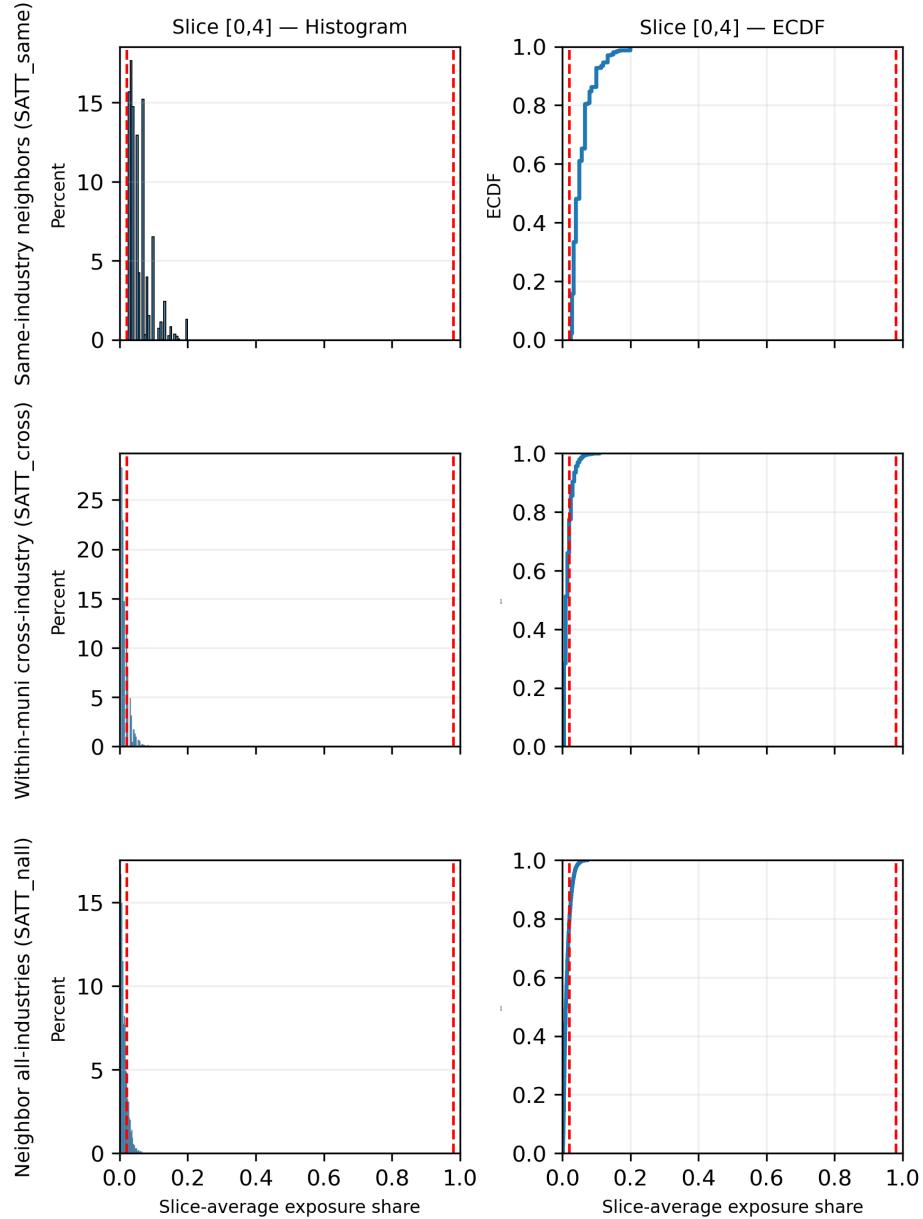


Figure A5: Exposure Overlap and Trimming Regions by Channel (Slice 0–4)

*Notes:* Each row corresponds to a spillover channel (same-industry neighbors, within-municipality cross-industry, neighbor all-industries), and each panel shows either the histogram or the ECDF of slice-average exposure shares. Vertical dashed lines at 0.02 and 0.98 mark the trimming thresholds used to enforce exposure positivity.

Appendix C1. Exposure overlap and trimming regions  
 (Slice-average exposure shares, slice [5,8])

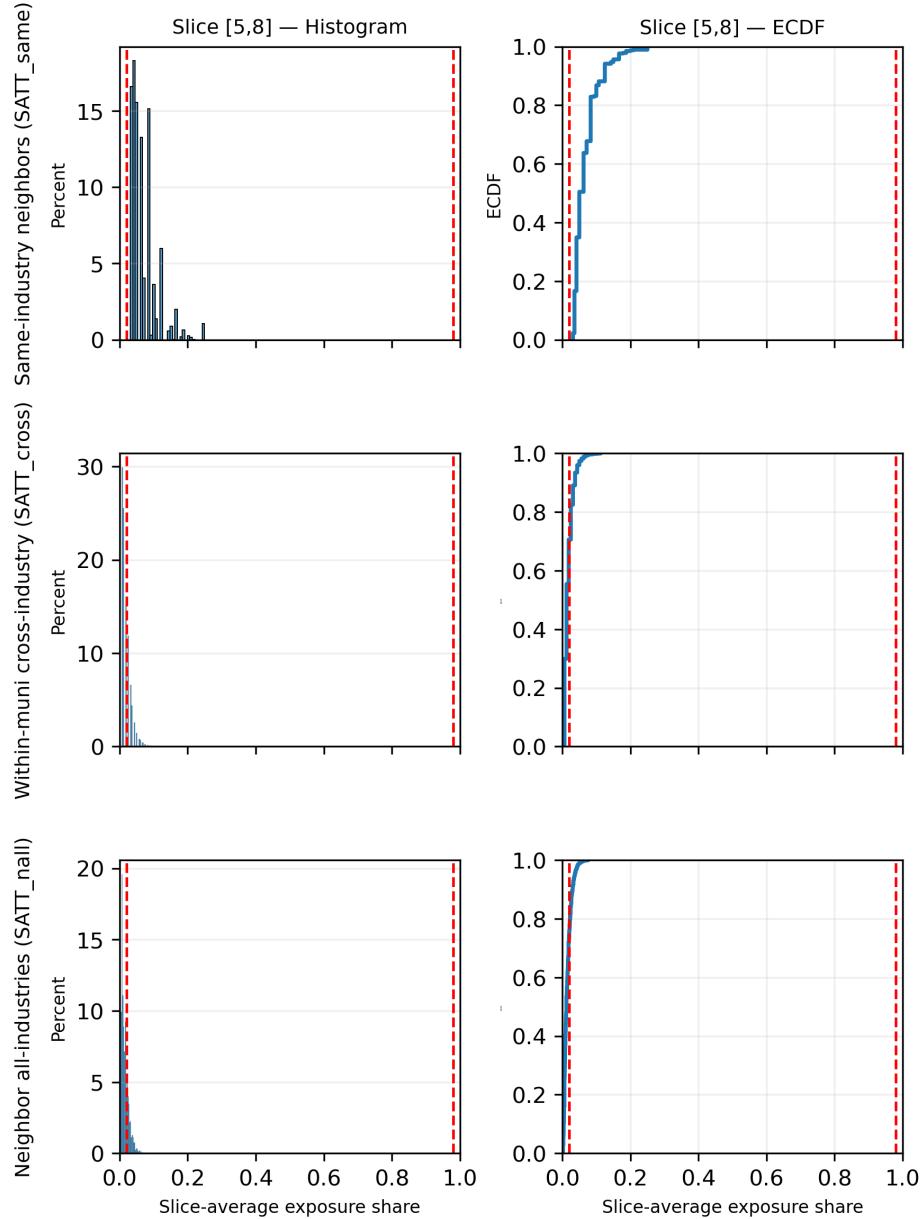


Figure A6: Exposure Overlap and Trimming Regions by Channel (Slice 5–8)

*Notes:* Each row corresponds to a spillover channel (same-industry neighbors, within-municipality cross-industry, neighbor all-industries), and each panel shows either the histogram or the ECDF of slice-average exposure shares. Vertical dashed lines at 0.02 and 0.98 mark the trimming thresholds used to enforce exposure positivity.

Appendix C1. Exposure overlap and trimming regions  
(Slice-average exposure shares, slice [9,16])

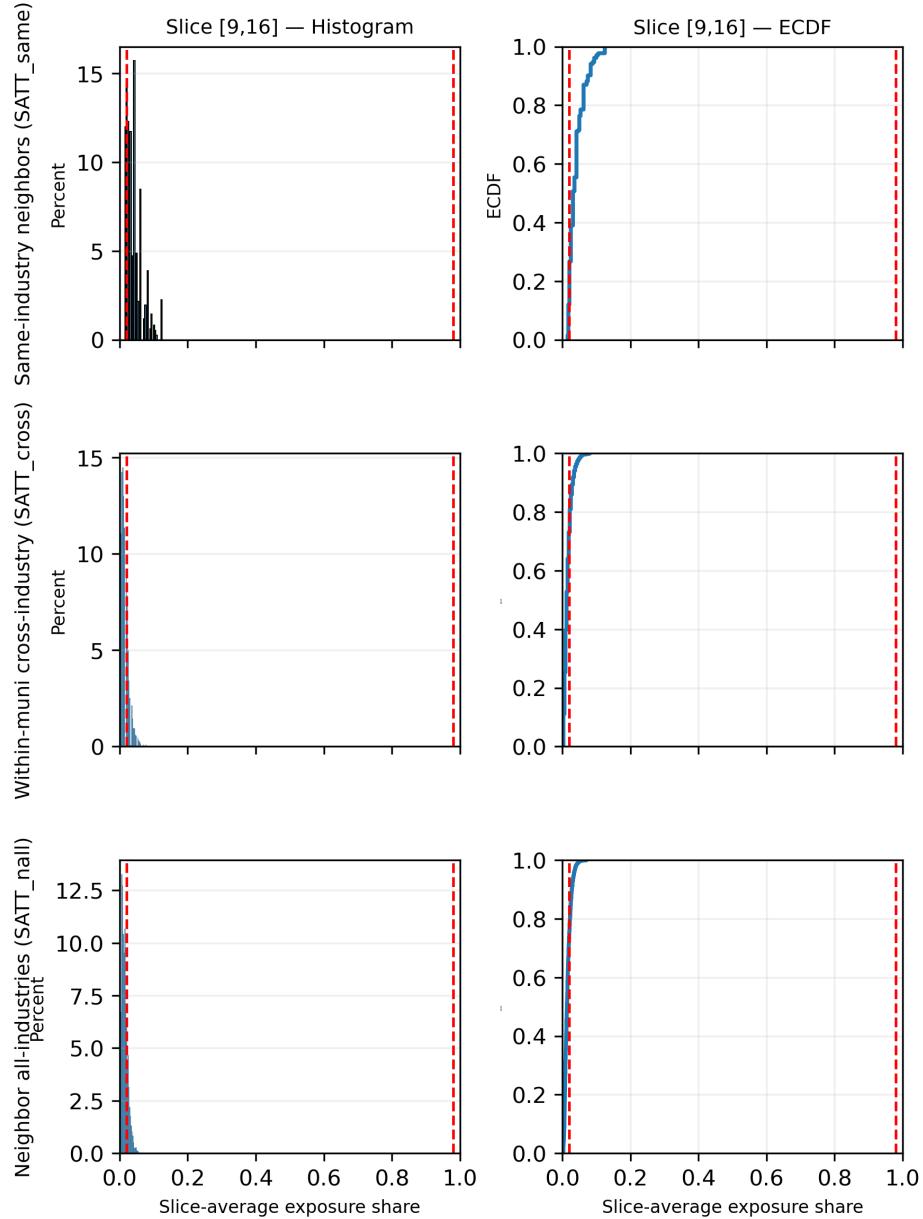


Figure A7: Exposure Overlap and Trimming Regions by Channel (Slice 9–16)

*Notes:* Each row corresponds to a spillover channel (same-industry neighbors, within-municipality cross-industry, neighbor all-industries), and each panel shows either the histogram or the ECDF of slice-average exposure shares. Vertical dashed lines at 0.02 and 0.98 mark the trimming thresholds used to enforce exposure positivity.

Appendix C1. Exposure overlap and trimming regions  
(Slice-average exposure shares, slice [0,16])

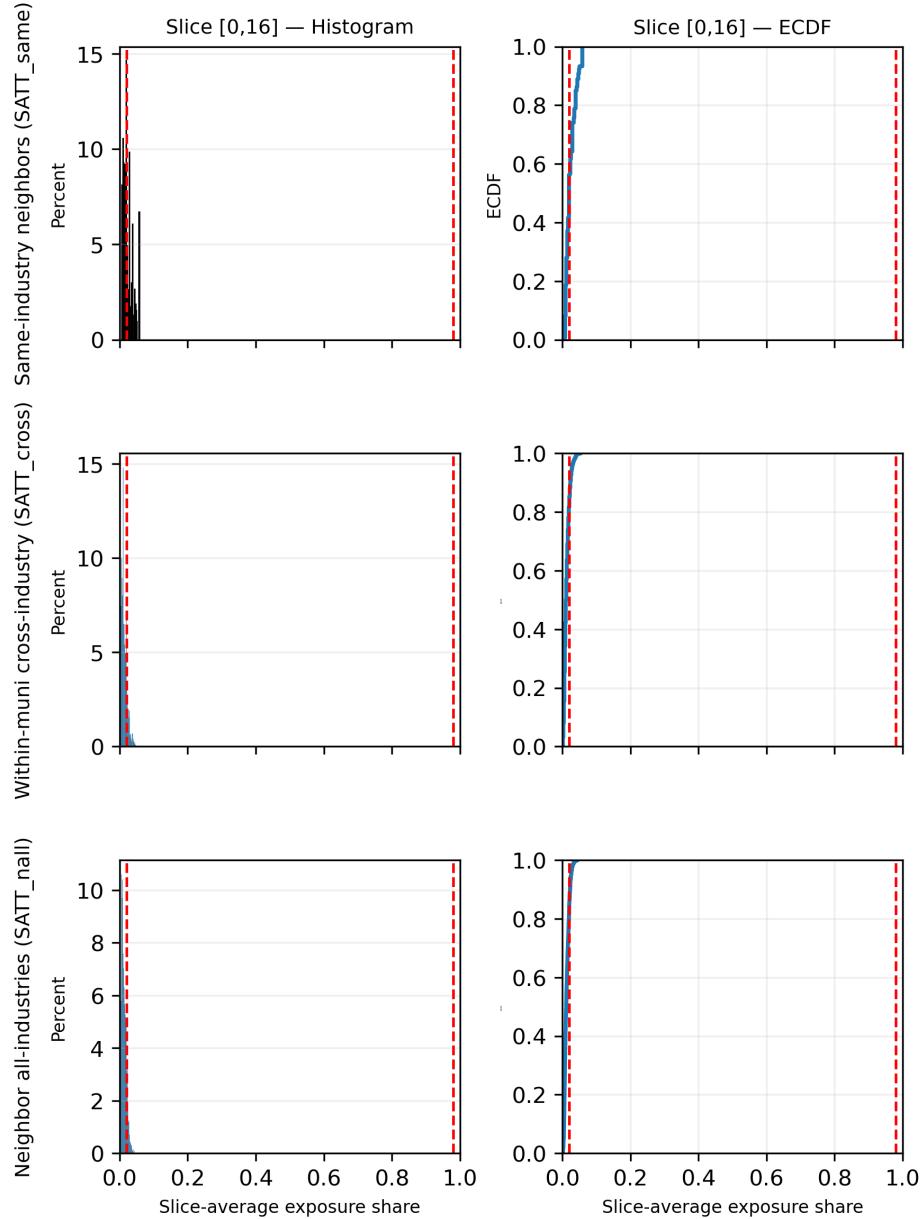


Figure A8: Exposure Overlap and Trimming Regions by Channel (Slice 0–16)

*Notes:* Each row corresponds to a spillover channel (same-industry neighbors, within-municipality cross-industry, neighbor all-industries), and each panel shows either the histogram or the ECDF of slice-average exposure shares. Vertical dashed lines at 0.02 and 0.98 mark the trimming thresholds used to enforce exposure positivity.

## D Spatial Dependence and Inference Comparisons

### D.1 Summary of Moran's $I$ Residual Diagnostics

Table A2 summarizes the global Moran's  $I$  residual diagnostics for the main specifications using the primary weight matrix. For each specification, outcome, and slice family, the table reports the minimum, median, and maximum Moran's  $I$  across quarters, the corresponding minimum, median, and maximum permutation  $p$ -values, and the share of quarters with  $p < 0.05$ .

Maximum Moran's  $I$  values range from about 0.16 to 0.42 across specifications, while median values are typically much smaller (around  $-0.03$  to  $0.06$  for most outcomes), indicating that spatial autocorrelation in residuals is present but moderate and varies considerably across quarters. For the primary DR–DiD interference specification and the main outcome log(covered employment), the maximum  $I$  lies between 0.28 and 0.34 across slices, with median values ranging from  $-0.03$  to  $0.04$ , and roughly 9–18% of quarters exhibiting statistically significant spatial dependence. For the real wage bill, maximum  $I$  ranges from 0.21 to 0.35 with median values around  $-0.02$  to  $0.06$ , and 4–27% of quarters significant. These diagnostics justify the use of spatially robust covariance estimators such as C–SCPC and SHAC.

### D.2 Inference Comparisons for DR–DiD Slice Effects

Table A3 compares several inference procedures for the primary DR–DiD slice estimates of log covered employment. For each slice and parameter (direct effect, total spillovers, total effect), the table reports the point estimate and standard errors under four methods:

1. Conditional SCPC (C–SCPC), our primary spatially robust procedure;
2. cluster-robust at the municipio level;
3. spatial HAC (SHAC/Conley);
4. TMO-adjusted SEs.

For the immediate post-entry slice [0,4], cluster-robust inference would suggest statistically significant positive direct effects and sizable positive spillovers:  $\widehat{\text{DATT}}[0,4] \approx 0.16$  and  $\widehat{\text{SATT}}[0,4] \approx 0.83$ , with cluster SEs around 0.02–0.15. However, spatially robust standard errors are much larger: SHAC and C–SCPC SEs are roughly 5–6 times the cluster SE for DATT and 2–6 times as large for SATT and TATT. Under C–SCPC, the 95% confidence intervals for all three parameters include zero, indicating that early-slice spillover evidence is not statistically precise once spatial dependence is taken into account.

In the later slices [5,8] and [9,16], point estimates for both direct and spillover effects are smaller in magnitude and often negative, while SHAC and C–SCPC SEs remain several times larger than cluster SEs. Over the full 0–16 window, the total effect  $\widehat{\text{TATT}}[0,16] \approx 0.10$  with a cluster SE of about 0.15, whereas C–SCPC inflates the SE to roughly 1.0, yielding a very wide interval that easily encompasses zero. TMO-adjusted SEs are very similar to cluster SEs in this application, suggesting that the main driver of uncertainty inflation is spatial correlation rather than cross-location outcome co-movement captured by the auxiliary TMO outcomes.

Table A2: Summary of Moran's  $I$  Residual Diagnostics (Primary  $W$ )

Spec	Outcome	$I_{\min}$ range	$I_{\text{med}}$ range	$I_{\max}$ range	Share sig. ( $p < 0.05$ )
<i>CS exposure ES (single aggregate value)</i>					
CS_exposure	joint	—	0.156	—	0.04
<i>DR-DiD interference (ranges across slices 0–4, 5–8, 9–16, 0–16)</i>					
DRDID_interf.	log covered emp	-0.57 to -0.31	-0.03 to 0.04	0.28 to 0.34	0.09–0.18
DRDID_interf.	log establishments	-0.19 to -0.12	-0.01 to 0.04	0.37 to 0.42	0.00–0.20
DRDID_interf.	log total wages real	-0.58 to -0.15	-0.02 to 0.06	0.21 to 0.35	0.04–0.27
DRDID_interf.	joint	-0.09 to -0.01	-0.01 to 0.05	0.16 to 0.25	0.02–0.09

*Notes:* For DR–DiD specifications, each cell shows the range across four slices (0–4, 5–8, 9–16, 0–16). Within each slice, statistics summarize 45 quarters of residual diagnostics. Each column shows the minimum-to-maximum range of that statistic across the four slices. For example, “ $I_{\text{med}}$  range” shows the lowest median value observed in any slice to the highest median value observed in any slice. The CS exposure ES row reports a single aggregate statistic computed across all quarters. Share significant reports the range of fractions of quarters with permutation  $p$ -values below 0.05 across slices.

Table A3: Inference comparison for DR–DiD slice effects (log covered employment)

Slice	Parameter	Estimate	SE (cluster)	SE (SHAC)	SE (C-SCPC)
0–4	DATT[0, 4]	0.156	0.025	0.147	0.134
	SATT[0, 4]	0.834	0.148	0.364	0.855
	TATT[0, 4]	0.990	0.157	0.455	0.897
5–8	DATT[5, 8]	-0.041	0.023	0.135	0.131
	SATT[5, 8]	-0.024	0.060	0.676	0.345
	TATT[5, 8]	-0.065	0.070	0.709	0.420
9–16	DATT[9, 16]	0.001	0.022	0.274	0.113
	SATT[9, 16]	-0.335	0.099	0.618	0.429
	TATT[9, 16]	-0.335	0.098	0.725	0.416
0–16	DATT[0, 16]	0.214	0.036	0.296	0.161
	SATT[0, 16]	-0.110	0.142	0.816	0.908
	TATT[0, 16]	0.104	0.149	1.003	0.965

*Notes:* Table reports point estimates and selected standard errors from the compact inference comparison underlying the main DR–DiD results. C-SCPC spatially robust standard errors are used for inference in the main text; cluster-robust, SHAC, and TMO-adjusted standard errors are shown here as benchmarks. TMO-adjusted SEs are close to cluster SEs (ratios around 0.99) and are omitted for brevity; full details, including confidence intervals and  $p$ -values, are available in the machine-readable file. Spatially robust procedures (SHAC and especially C-SCPC) substantially inflate standard errors relative to cluster-robust inference.

## E Full Heterogeneity Grid

The auxiliary TWFE heterogeneity specification in Section 7.6 is estimated on a large grid of outcome–slice–stratum–parameter combinations. The full machine-readable table contains all combinations of:

- outcomes in {`log_covered_emp`, `log_establishments`, `log_total_wages_real_2020`};
- slices in {0–4, 5–8, 9–16, 0–16};
- strata in {`tradable/non-tradable`, `metro/non-metro`, `high/low wage`};
- parameters in {DATT, SATT, TATT}.

For the primary outcome  $\log(\text{covered employment})$  and the 0–16 slice, the heterogeneity results can be summarized succinctly. Direct effects DATT[0, 16] are positive and of similar magnitude across strata, ranging from about 1.15 to 2.08 (in log points), and remain significant after Benjamini–Hochberg FDR adjustment in all six strata. Total spillovers SATT[0, 16] are negative in every stratum (roughly  $-1.1$  to  $-1.4$ ) and often FDR-significant, consistent with the aggregate DR–DiD finding of modestly negative same-industry neighbor spillovers in the medium run.

The implied total effects TATT[0, 16] are near zero for most strata, with one notable exception: tradable industries exhibit a positive and FDR-significant total effect of roughly 0.8 log points, while non-tradable industries show essentially no net effect ( $\text{TATT}[0, 16] \approx -0.04$ ). By metro status, total effects are modest and statistically indistinguishable from zero in both metro and non-metro municipios, and by wage stratum, high-wage industries display a somewhat larger (but only marginally significant) total effect than low-wage industries.

Given the size of the full heterogeneity grid, I do not reproduce Table E1 in the print appendix. Instead, the complete table (including standard errors, BH/BY  $q$ -values, and sample counts) is provided as a machine-readable file in the replication package.

## F ACS Auxiliary Socioeconomic Indicators Used in the TMO Analysis

Table A4: ACS Auxiliary Metrics Used in the TMO Analysis

Variable	Description
B01001_Pct_WorkingAge	Share of population aged 15–64
B01001_Pct_Elderly	Share of population aged 65+
B08006_Pct_WkHome	Share of workers who work at home
B08006_Pct_DroveAlone	Share of workers who drive alone to work
B08006_Pct_PubTrans	Share of workers using public transport
B08006_Pct_WalkBike	Share of workers who walk or bike
B08303_Pct_LongCommute	Share of workers with commute $\geq 45$ minutes
B15003_Pct_LessThanHighSchool	Share of adults with less than high school
B15003_Pct_SomeCollege	Share of adults with some college (no BA)
B15003_Pct_BachelorsOrHigher	Share of adults with bachelor's or higher
B17010_Pct_Poverty_AllFamilies	Share of families below poverty line
B17010_Pct_Poverty_FemaleHH_WithChildren	Share of female HHs w/ children below poverty
B17010_Pct_Poverty_MarriedHH_WithChildren	Share of married HHs w/ children below poverty
B19013_MedianHH_Income	Median household income
B19058_Pct_With_Assist_Or_SNAP	Share of HHs with public assistance or SNAP
B19083_Gini_Index	Gini index of income inequality
B23025_Unemployment_Rate	Unemployment rate (civilian labor force)
B23025_LFP_Rate	Labor force participation rate (16+)
B23025_Pct_ArmedForces	Share of population in armed forces
B25003_Pct_OwnerOccupied	Share of occupied housing that is owner-occupied
B25003_Pct_RenterOccupied	Share of occupied housing that is renter-occupied
B25064_Median_Gross_Rent	Median gross rent
B25077_Median_Home_Value	Median owner-occupied home value
B27010_Pct_Uninsured_18_64	Share of population 18–64 without health insurance
B27010_Pct_EmployerBased_18_64	Share of population 18–64 with employer-based coverage
B27010_Pct_Medicaid_18_64	Share of population 18–64 with Medicaid only
C24010_Pct_MgmtBusSciArt	Share in management, business, science, arts occupations
C24010_Pct_Service	Share in service occupations
C24010_Pct_SalesOffice	Share in sales and office occupations
C24010_Pct_NatResConstMaint	Share in natural resources, construction, maintenance
C24010_Pct_ProdTransMaterial	Share in production, transport, material moving
C24030_Pct_Agri_Mine	Share in agriculture, forestry, fishing, mining
C24030_Pct_Construction	Share in construction industries
C24030_Pct_Manufacturing	Share in manufacturing industries
C24030_Pct_WholesaleTrade	Share in wholesale trade
C24030_Pct_RetailTrade	Share in retail trade
C24030_Pct_TransportWarehousingUtils	Share in transportation, warehousing, utilities
C24030_Pct_Information	Share in information sector
C24030_Pct_FinanceInsRealEstate	Share in finance, insurance, real estate
C24030_Pct_ProfessionalSciMgmt	Share in professional, scientific, management services
C24030_Pct_EdHealthSocial	Share in education, health care, social assistance
C24030_Pct_ArtsEntAccomFood	Share in arts, entertainment, accommodation, food services
C24030_Pct_OtherServices	Share in other services (except public administration)
C24030_Pct_PublicAdmin	Share in public administration

# G Reproducibility and Computational Workflow

## G.1 Raw Data Ingestion and Storage

Each Excel file is read with a fixed header schema `{municipality, code, industry, naics, units, averageEmployment, totalSalary, average}`, and the first observed industry name is copied to the `Primary_Industry` field to preserve hierarchical labeling. Filenames encode the year and quarter via the pattern `_QCEW_YYYY_*Q.xlsx`, from which I parse temporal identifiers. Numeric fields are cleaned by removing commas and treating asterisks as missing values. Each record retains its original filename as a provenance field. The concatenated dataset is written both as CSV and as an efficient columnar Parquet table for downstream access.

## G.2 Spatial Data Processing and GIS Artifacts

Invalid geometries are repaired using GEOS/Shapely `make_valid` when available, with a `buffer(0)` fallback for older libraries.<sup>5</sup>

The canonical `municipios` layer is saved as a GeoPackage (layer name `municipios`); attribute metadata and projected metrics are stored in a companion CSV, which also records `crs_epsg` for provenance. To compute area and perimeter diagnostics, I project geometries to the Puerto Rico StatePlane coordinate system (EPSG:32161).<sup>6</sup> This geometry underlies the contiguity matrix and distance-based robustness checks used in the main analysis.

## G.3 Intermediate Data Artifacts and Exposure Workflow

Each processing stage outputs reproducible files under version-controlled directories. Interim artifacts include the ingested dataset, the quarter index, and the geometry layer. Processed outputs, including the harmonized industry data and the final `municipality-industry-quarter` panel, feed directly into the treatment and exposure modules described in Section 4.

In the exposure construction stage, I verify empirical support and, where needed, apply a two-sided  $\epsilon$ -trimming rule ( $\epsilon = 0.02$ ) on the slice-average exposure shares (0–1). This rule bounds away from both saturation (near-one) and non-exposure (near-zero) regions, retaining only slices where the average share is within  $[\epsilon, 1 - \epsilon]$ . This ensures well-supported contrasts, following the overlap conditions discussed by Crump et al. (2009), Aronow et al. (2020), and Xu (2023). Masks are constructed at the channel–slice level and aligned to the estimators used downstream.<sup>7</sup>

**Descriptive exposure ES artifact.** The table `exposure_es_dynamic.parquet` stores per- $\ell$  coefficients for each spillover channel from the descriptive exposure-augmented event study with fields:

---

<sup>5</sup>Geometry validation follows the OGC Simple Feature Access specification (Open Geospatial Consortium, 2011); repairs use the GEOS `make_valid` algorithm where supported and the standard `buffer(0)` fix otherwise.

<sup>6</sup>EPSG Geodetic Parameter Registry (2025) defines EPSG:32161—NAD83 / Puerto Rico & Virgin Islands, units in meters. This projection minimizes distortion for distance-based spatial weights.

<sup>7</sup>Exposure-positivity and overlap conditions for interference settings are formally discussed by Aronow et al. (2020) and Xu (2023). Trimming is implemented in the exposure-mapping scripts (Section 4) on slice-average shares to maintain scale invariance and ensure well-supported contrasts at both tails.

- `outcome`;
- `channel` ∈ {SATT\_same, SATT\_cross, SATT\_nall};
- `ell`;
- point estimate (`att_nbr`);
- cluster-robust SE (`se`);
- pointwise 95% interval (`lo`, `hi`);
- uniform-band endpoints (`cband_lo`, `cband_hi`) and the associated critical value (`cband_crit`) computed via a wild cluster bootstrap over GEOFID;
- flags for the estimation specification (`balanced`, `spec`, `ell_window`, `omit_bin`).

Exposures in this artifact are measured as per-period shares in  $[0, 1]$ ; each coefficient is a *descriptive* per- $\ell$  effect that scales with a one-unit change in exposure at that event time (i.e., moving from 0% to 100% exposure), rather than a DR–DiD slice-level SATT estimand.

**Cross-industry histories artifact.** For the interference estimators, I store exposure histories in two flavors: an “any-since-adoption” share and a “last four post-treatment quarters” share. For the same-industry neighbor channel, these histories and their  $L = 4$  lags are embedded in `exposure_long.parquet`. For the within-municipality cross-industry and neighbor-all channels, the corresponding per- $\ell$  exposure shares and collapsed histories are stored in a dedicated cross-industry exposure file `exposure_crossindustry.parquet`, with a companion lags file `exposure_crossindustry_lags.parquet` used for diagnostics and alternative specifications. The estimation code reads the appropriate history flavor from `exposure_crossindustry.parquet` via a simple `HISTORY_FLAVOR` switch (“any” vs. “last4”) and applies the same choice to the same-industry histories.

## G.4 Exposure Trimming and Implementation Flags

Operationally, the exposure-building step flags observations with slice-specific trimming indicators. The DR–DiD estimator enforces these flags and the balanced-horizon mask before estimation. A minimum post-trim sample threshold is imposed at the slice level (`MIN_N_SLICE`, default = 50). If not met, the specification aborts with an informative error. This policy follows [Sävje \(2023\)](#) and ensures practical positivity for the exposure mappings actually used.

## G.5 Balanced Event-Time Horizon Indicator

A balanced horizon is implemented via the precomputed indicator `balanced_window`. For the primary causal estimators (CS and DR–DiD), I restrict post-treatment treated/exposed rows to `balanced_window=1` while retaining all controls (never- and not-yet-treated units, as well as pre-treatment observations of ever-treated units whose post-treatment horizon is shorter than the balanced window). This ensures that each event-time horizon uses a common treated support, while still exploiting pre-trend information and control observations from short-horizon cohorts. Unbalanced overlays (without the `balanced_window` restriction on treated/exposed rows) are reported for comparison and composition diagnostics.

For the descriptive exposure-augmented event study, I apply the same principle: when using a balanced horizon, the balanced restriction is enforced on units that ever exhibit nonzero exposure in the event-time window (rows with positive exposure in any  $\ell$ ), while controls and rows with zero exposure across all spillover channels are retained. Unbalanced overlays drop these balanced-window restrictions for comparison.

## G.6 DR–DiD Slice Implementation Details

The DR–DiD slice estimators described in Section 7.4 are implemented in two steps: (i) cross-fitted residualization of the outcome and each regressor using group-based folds by municipality (GroupKFold by GEOID), and (ii) a single OLS of residualized outcomes on residualized  $D_{\text{slice}}$  and spillover-slice regressors with GEOID-clustered inference. Before estimation, the procedure enforces the balanced mask `balanced_window=1` on treated/exposed rows and the trimming flags `keep_{expo,xexpo,nexpo}_{slice}`, and asserts the minimum post-trim sample size per slice (`MIN_N_SLICE`). Neighborhood spillovers are pinned to  $h = 1$  in the primary specification, and inputs with  $h > 1$  are rejected.

**Exposure histories in nuisance functions.** Nuisance models include exposure-history controls constructed from the exposure files. A simple switch selects which histories enter: `HISTORY_FLAVOR ∈ {“any”, “last4”}` chooses either the `*share_any` (baseline) or the `*share_last4` set; flags such as `keep_*` are not used as features. This implements the “parallel trends conditional on exposure histories” restriction while allowing a clean robustness toggle.

## G.7 Overlap Diagnostics Artifact

The resulting coverage changes in numbers of observations and municipios before and after trimming and balancing are summarized by channel and slice. These diagnostics are saved and underlie the overlap figures and tables reported in Section 9.

## G.8 Heterogeneity Multiple-Testing Artifact

For the heterogeneous-effects analysis, I apply Benjamini–Hochberg and Benjamini–Yekutieli false-discovery-rate adjustments as described in Section 9. These adjustments are implemented on the heterogeneity results table. The final adjusted results are saved in `heterogeneity_final_adjusted.parquet`, which augments the original point estimates and standard errors with the corresponding BH and BY  $q$ -values for each heterogeneous effect.