Research Article

The Effect Of School-Age Population Decline On Elementary School Closures And Consolidations In South Korea: Evidence From A Municipality-Year Panel Using Double Machine Learning And Causal Forests

Young-Chool Choi

Chair Professor, aSSIST (a Seoul School of Integrated Sciences and Technology, KOREA, email: ycchoi@assist.ac.kr

Abstract

South Korea's exceptionally low fertility and uneven population redistribution are shrinking elementary school-age cohorts, raising concerns about the sustainability of local school networks. This study examines whether municipality-level declines in the population aged 6–11 are associated with elementary school closures and consolidations. We compile a municipality-year panel covering all 229 Si/Gun/Gu municipalities for 2019–2024 (1,374 observations) and link demographic measures to public elementary school structure and annual closure/consolidation actions. We estimate two-way fixed-effects models and apply double machine learning with cross-fitting to flexibly adjust for observed covariates, and we explore heterogeneity using interaction models and forest-based approaches. Consolidation actions are rare but spatially concentrated: 7.35% of municipality-years record at least one action. Average within-municipality effects of short-run cohort change are small and statistically imprecise. However, effects are heterogeneous: municipalities with structurally small baseline school systems exhibit greater sensitivity to cohort decline, and a forest-based severe-decline contrast reveals wide dispersion in conditional effects. The findings suggest that demographic decline acts as a structural pressure that triggers consolidation primarily in vulnerable local systems. Policy responses should prioritize place-based early-warning screening and mitigation measures in high-sensitivity municipalities rather than uniform consolidation rules.

Keywords: school closures; school consolidation; demographic decline; double machine learning; heterogeneous treatment effects.

1. Introduction

1.1 Demographic contraction and pressure on local school networks

South Korea's demographic transition has intensified pressure on local public services, with elementary education among the most directly exposed sectors. Official vital statistics report 238.3 thousand live births in 2024, and the total fertility rate (TFR) varies sharply across regions, ranging from 1.03 in Jeonnam and Sejong to 0.58 in Seoul (Statistics Korea, 2025). International comparative evidence likewise places Korea's fertility rate among the lowest in the OECD; for example, the OECD's Family Database reports a TFR of 0.72 in 2023 (Organisation for Economic Co-operation and Development [OECD], 2024b), and OECD analysis highlights Korea's uniquely severe fertility decline and its structural drivers (Yang et al., 2024). Such sustained low fertility translates into smaller school-age cohorts, and—through both fertility and migration—into uneven local trajectories across municipalities.

Demographic shrinkage affects education systems through both demand and supply channels. Falling cohorts reduce enrollment, while school facilities and staffing structures adjust only partially in the short run. OECD work on school system responsiveness emphasizes that education networks must adapt to shifting demand while balancing efficiency and equity, especially in rural and non-metropolitan contexts where student numbers can fall below operationally sustainable thresholds (OECD, 2018). At the same time, school closures and consolidations can impose substantial access and distributional costs by increasing travel distances, reducing local service availability, and potentially widening rural-urban disparities (OECD, 2024a). These trade-offs make it critical to understand not only whether demographic decline is associated with school consolidation, but also where such links are strongest.

1.2 Research problem and motivation

Despite the prominence of demographic decline in policy debates, evidence on how local cohort change maps into elementary school closures and consolidations remains limited at the municipality level. In practice, consolidation is a discrete administrative decision that is often 'lumpy'—implemented through mergers, closures, or network restructuring—rather than a smooth, proportional response to annual cohort change. Moreover, consolidation risk may be heterogeneous: structurally smaller school systems may reach viability constraints sooner, while larger systems may have greater slack to absorb moderate declines.

This study is motivated by municipality-year administrative data covering all 229 Si/Gun/Gu municipalities over 2019–2024 (1,374 observations). Within this panel, annual consolidation actions occur in 101 municipality-years (7.35%) and are concentrated in 75 municipalities (32.75%) over the six-year period. Meanwhile, school-age cohort change is predominantly negative: across 2020–2024, 85.19% of municipality-years record a year-on-year decline in the population ages 6–11, with a median decline of 3.78%. Taken together, these patterns suggest a policy-relevant tension: demographic contraction is widespread, yet consolidation actions are comparatively rare and spatially concentrated. Section 4 documents these descriptive patterns in detail and shows how they motivate a focus on nonlinearity and heterogeneity.

1.3 Contributions

This study makes three contributions. First, it analyzes a nationally comprehensive municipality-year panel that links local demography (total population and ages 6–11) with the structure of public elementary schooling (numbers of schools, students, and classes) and with annual closure/consolidation actions. Second, it estimates the association between lagged cohort decline and consolidation risk using two-way fixed effects to account for time-invariant municipal characteristics and common year shocks. Third, it applies causal machine learning to address nonlinearities and heterogeneity. Specifically, we use double machine learning (DML) to estimate an average effect under flexible adjustment (Chernozhukov et al., 2018) and forest-based methods to characterize heterogeneous effects across municipalities (Athey et al., 2019; Wager & Athey, 2018). Together, these methods are designed to move beyond a single national average and to identify the municipal contexts in which cohort decline is most strongly associated with consolidation activity.

1.4 Research questions and empirical approach

The analysis addresses two research questions. First, does a decline in the elementary school-age population (ages 6–11) increase the probability that a municipality experiences at least one closure or consolidation action in the subsequent year? Second, is this relationship heterogeneous—stronger in municipalities with structurally smaller baseline school systems (e.g., low students-per-school or students-per-class) and weaker in municipalities with greater slack capacity?

To answer these questions, the empirical strategy proceeds in three steps (Sections 3–4). (i) We estimate two-way fixed-effects models as a benchmark specification for within-municipality changes. (ii) We implement DML with cross-fitting to obtain orthogonalized estimates of the average effect of lagged cohort change on consolidation risk (Chernozhukov et al., 2018). (iii) We examine heterogeneity using forest-based approaches, which are designed to learn conditional average treatment effects and uncover systematic variation in sensitivity across municipal contexts (Athey et al., 2019; Wager & Athey, 2018).

1.5 Roadmap

Section 2 reviews related literature and develops a mechanism-based analytical framework and hypotheses. Section 3 describes the dataset, variable construction, identification strategy, and estimation samples. Section 4 presents the descriptive evidence and the main empirical results, with tables and figures concentrated in that chapter. Section 5 concludes with policy implications, limitations, and directions for future research.

2. Literature Review, Analytical Framework, and Contribution

2.1 Demographic Decline and the Reconfiguration of Local School Networks

Municipal education systems are long-lived infrastructures that must be financed and staffed regardless of short-run fluctuations in cohort size. When fertility falls and age structures shift, the demand for schooling becomes both smaller and more spatially uneven, especially in peripheral and rural areas. A central implication in the comparative literature is that school networks become increasingly difficult to sustain under declining enrolment, prompting governments to pursue restructuring through closures, mergers, and consolidation of catchment areas (Ares Abalde, 2014; Howley et al., 2011). These responses are rarely mechanical. Decisions to close or merge schools occur within institutional constraints (e.g., minimum school-size standards, commuting-time regulations, and local consultation requirements) and political constraints (e.g., community resistance and inter-jurisdictional equity concerns), which implies substantial geographic heterogeneity in how demographic decline translates into organizational change.

The descriptive patterns documented in Section 1—namely, the concentration of consolidation activity in particular municipality-years and the heavy right tail in school-age population decline—mirror what prior work suggests about decline-driven restructuring: closures and mergers cluster where cohort shrinkage is both persistent and compounded by low baseline scale (Ares Abalde, 2014; Lyson, 2002). This motivates an empirical framework that treats the change in the school-age population as a municipality-year shock to the demand for schooling, while recognizing that the translation from demographic pressure to closure outcomes is mediated by fiscal, logistical, and political mechanisms that vary by place.

2.2 Why Close or Consolidate? Economies of Scale, Capacity Utilization, and Administrative Rationalization

A dominant justification for consolidation is an efficiency argument: when enrolment declines, fixed costs (facilities, administration, maintenance) are spread over fewer students, raising per-pupil costs. In this view, closing or merging schools can restore scale efficiencies by increasing average school size and reducing duplication in staffing and capital use (Andrews et al., 2002; Duncombe & Yinger, 2007). Empirical studies of district and school size frequently find nonlinearities—cost functions display economies of size at small scale but flatten as size grows—implying that consolidation may be fiscally attractive primarily for very small jurisdictions (Duncombe & Yinger, 2007). Importantly, these studies also highlight adjustment costs (e.g., transportation, transition expenditures), which can offset operating savings and may be politically salient even when net fiscal savings exist (Ares Abalde, 2014; Howley et al., 2011).

Beyond fiscal efficiency, governments sometimes frame consolidation as a strategy to improve educational quality by expanding curricular breadth, staffing specialization, and peer learning opportunities. However, the literature is not uniform on whether larger schools systematically improve outcomes, and many studies emphasize the context dependence of size effects (Ares Abalde, 2014; Howley et al., 2011). This tension is central to the policy problem in low-fertility settings: maintaining geographically proximate schools can support access and community cohesion but may entail higher per-student costs and staffing constraints; consolidation can improve resource utilization but may impose longer commutes and weaken local ties.

2.3 Consequences and Political Economy of School Closures and Consolidations

The empirical literature on closures and consolidations emphasizes distributional and spatial consequences. First, closures can increase travel distance and time for students, potentially affecting attendance, fatigue, and participation in after-school activities; these channels are often highlighted in policy debates even when direct achievement effects are ambiguous (Ares Abalde, 2014; Nitta et al., 2010). Second, closures can have community-level effects because schools are civic hubs, employers, and anchors for local identity—particularly in rural settlements (Lyson, 2002). The loss of a school may therefore alter residential choice, housing markets, and social capital in ways that reinforce depopulation dynamics.

Recent quasi-experimental evidence supports the plausibility of longer-run community effects. In a Danish setting where multiple village schools were closed simultaneously, Sørensen et al. (2021) estimate negative local population effects that emerge several years after closure, consistent with mechanisms operating through housing-market expectations and social-capital 'lock-in'. While this evidence is context-specific, it underscores an important theoretical point for demographic-decline contexts: closures may be both a response to depopulation and a contributor to it, creating feedback loops that complicate interpretation in observational data.

The literature on student outcomes from closures—largely based on urban U.S. cases—reports mixed average effects and substantial heterogeneity depending on the quality of receiving schools and the reasons for closure (Brummet, 2014; Tieken & Auldridge-Reveles, 2019). More broadly, scholars argue that closure decisions are often contested political processes in which quantitative criteria (e.g., utilization rates) interact with histories of inequality and community power, shaping which schools close and how transitions are managed (Caven, 2019; Tieken & Auldridge-Reveles, 2019). These insights reinforce the importance of analyzing not only average relationships between enrolment decline and closures, but also heterogeneous effects across contexts and baseline conditions.

2.4 South Korea: Policy Context and Existing Evidence on Small-School Consolidation

Korea's experience with school consolidation has been shaped by rapid demographic change, spatial concentration of population in metropolitan regions, and persistent depopulation in many counties and small cities. Korean policy discussions emphasize the trade-off between efficient resource allocation and equitable access to education services in rural and island areas. Domestic scholarship documents that consolidation policies can trigger conflict over decision procedures, closure criteria, and the preservation value of community schools, suggesting that governance arrangements matter for implementation (Lee, 2013; Park, 2014). This work highlights the need for transparent decision rules and for incorporating educational-impact assessments and community voice to mitigate disputes (Park, 2014).

Evaluation evidence in Korea is comparatively limited in the international literature, but policy reports and academic studies have begun to document the outcomes and controversies associated with integrating and abolishing small schools in rural areas (Korean Educational Development Institute [KEDI], 2010). More recent empirical work also investigates how 'optimal school size' policy initiatives relate to closure probabilities, emphasizing that policy effects may differ by regional scale and school level (Kim, 2022). Taken together, the Korean literature indicates that enrolment decline is a necessary but not sufficient condition for closure: local administrative capacity, institutional thresholds, and the political economy of community opposition can shape whether and when consolidation occurs (Lee, 2013; Park, 2014).

2.5 Research Gaps, Analytical Framework, and the Contribution of This Study

Three gaps motivate this study. First, much of the international evidence on school closures comes from single-country case studies—often large urban districts—where closures are frequently driven by accountability policies rather than demographic decline (Brummet, 2014; Tieken & Auldridge-Reveles, 2019). Korea, by contrast, provides a setting where low fertility and spatially uneven cohort shrinkage are central drivers, yet national-scale, municipality-level empirical evidence remains sparse (KEDI, 2010; Park, 2014). Second, even when panel data are available, standard approaches face a high-dimensional confounding problem: local fiscal capacity, migration trends, housing markets, and prior school-network characteristics can jointly influence both cohort change and closure decisions. Traditional fixed-effects strategies help address time-invariant heterogeneity, but do not by themselves solve confounding from time-varying covariates or flexible functional form requirements (Angrist & Pischke, 2009; Wooldridge, 2010). Third, closures and consolidations are often rare events at the municipality-year level; models must therefore be robust to imbalanced outcomes and sensitive to heterogeneous patterns (King & Zeng, 2001).

To address these gaps, this study integrates modern causal machine-learning tools with a municipality-year panel design. Double/debiased machine learning (DML) provides a principled way to estimate causal parameters in the presence of high-dimensional nuisance components by combining cross-fitting and orthogonalized score functions, reducing bias from flexible predictive models (Belloni et al., 2014; Chernozhukov et al., 2018). For effect heterogeneity, causal forests and generalized random forests enable nonparametric estimation of conditional average treatment effects (CATEs) and allow the data to reveal where demographic decline translates into higher closure risk (Athey & Imbens, 2016; Athey et al., 2019; Wager & Athey, 2018). These methods align with an analytical framework in which school-age population decline is treated as an exposure that operates through multiple mechanisms—fiscal pressure,

capacity utilization, and community adaptation—whose strength varies by baseline scale and local context.
Based on the literature, the core expectation is that larger declines in the school-age population increase the likelihood of elementary school closure or consolidation (mechanism: declining enrolment raises per-pupil costs and makes minimum-size thresholds binding) (Andrews et al., 2002; Duncombe & Yinger, 2007). A second expectation is heterogeneity: the same percentage decline should translate into larger closure risk where baseline cohorts are small or where school networks already operate near minimum viable scale (Ares Abalde, 2014; Howley et al., 2011). A third expectation is that community and place-based considerations (e.g., rurality, local identity, and expectations about future population) can amplify or dampen consolidation responses (Lyson, 2002; Sørensen et al., 2021). The empirical strategy in Section 3 and results in Section 4 operationalize these expectations using a unified set of models that deliver both average and heterogeneous estimates.

3. Research Design and Methodology

3.1 Empirical design and identification logic

The goal of the empirical strategy is to estimate how changes in the local school-age population translate into the occurrence of elementary school closures and consolidations at the municipality level. The analysis is explicitly framed as an observational study using annual municipal statistics rather than a randomized policy experiment, which makes careful attention to design and assumptions essential (Angrist & Pischke, 2009; Imbens & Rubin, 2015).
We treat year-to-year changes in the population ages 6-11 as the primary demand-side shock to elementary schooling. To reduce simultaneity between cohort change and administrative restructuring, the baseline specification uses a one-year lag: consolidation activity in year t is related to the change in the 6-11 cohort observed between t-2 and t-1. This timing choice is consistent with the idea that consolidation decisions typically require planning and implementation periods, and it helps separate demographic signals from contemporaneous administrative actions (Wooldridge, 2010).
The core identifying variation comes from within-municipality changes over time, net of common year shocks. Specifically, we include municipality fixed effects to absorb time-invariant local characteristics (e.g., geography, long-run settlement patterns) and year fixed effects to absorb nationwide shocks shared across municipalities (e.g., national demographic trends or policy cycles). This two-way fixed-effects structure is a standard approach in panel-data policy evaluation when treatment varies over time within units (Wooldridge, 2010).

3.2 Data sources, unit of analysis, and estimation sample

The unit of analysis is the Si/Gun/Gu municipality-year. The dataset covers 229 municipalities observed annually from 2019 to 2024 (a balanced 229 x 6 panel). Each municipality-year includes total population, population ages 6-11, and basic indicators of the public elementary school system (number of public elementary schools, students, and classes), along with the annual count of closure/consolidation actions.
Because the main treatment variable is defined as a lagged year-on-year cohort change, the main estimation window starts in 2021 (the earliest year for which the lagged change between t-2 and t-1 is defined). The analysis therefore focuses on 2021-2024 as the baseline estimation sample, while 2019-2020 are retained for constructing lags and for descriptive context (see Section 4).
All variables are used as reported in the annual municipality statistics, with a small number of observations excluded only when key values are non-numeric or missing such that lag construction is not feasible. We report the resulting sample sizes for each specification in the results chapter to ensure transparent accounting.

3.3 Variable construction and operationalization

Outcome. The primary outcome is an indicator for whether a municipality experiences at least one closure or consolidation action in a given year. Let Count_it denote the annual number of actions; the event indicator is Y_it = 1(Count_it >= 1). This choice stabilizes estimation in settings where events are sparse and aligns with policy questions about where consolidation occurs. As robustness checks, we also analyze the action count itself.

Treatment. The main treatment is the lagged year-on-year change in the school-age population. Let $Pop6\_11_{it}$ denote the population ages 6-11. The treatment for year t is defined as $D_{it} = (Pop6\_11_{i,t-1} - Pop6\_11_{i,t-2}) / Pop6\_11_{i,t-2}$. Negative values correspond to cohort decline. We also consider log-difference versions, $\Delta \log Pop6\_11$, to reduce the influence of scale differences and to improve distributional stability (Wooldridge, 2010).

Controls. The baseline covariate set includes total population and school-system scale measures observed in the panel: numbers of public elementary schools, students, and classes. These variables capture capacity and utilization constraints that shape consolidation decisions. To support heterogeneity analysis, we additionally compute derived ratios such as students-per-school and students-per-class. These are transformations of observed variables rather than externally merged data and therefore remain consistent with the study's 'no additional data' design.

Missing values and outliers. Administrative panels can contain blanks in event-count fields when no action occurs. In the main analysis, blank consolidation counts are treated as zero actions, consistent with event-recording practice. For growth rates, a small number of extreme year-on-year changes can arise from denominator effects or reporting artifacts; we therefore present descriptive summaries with winsorization (1st/99th percentiles) and assess robustness to alternative trimming rules (Cameron & Miller, 2015).

3.4 Benchmark model: two-way fixed effects

We begin with a benchmark two-way fixed-effects specification estimated as a linear probability model (LPM):

$$Y_{it} = \beta D_{it} + X'_{i,t-2} \theta + \alpha_i + \gamma_t + \epsilon_{it},$$

where α_i are municipality fixed effects and γ_t are year fixed effects. Controls $X_{i,t-2}$ are measured at t-2 to avoid conditioning on post-treatment variables. The LPM provides coefficients that can be interpreted as percentage-point changes in event probability, which is convenient for policy discussion. As robustness checks, we also estimate alternative models appropriate for sparse outcomes, such as fixed-effects logit and count specifications.

Inference uses municipality-clustered standard errors to account for serial correlation and heteroskedasticity within municipalities over time (Arellano, 1987; Bertrand, Duflo, & Mullainathan, 2004; Cameron & Miller, 2015).

3.5 Double Machine Learning for average effects

To relax functional-form restrictions and reduce omitted-nonlinearity bias in the adjustment for observed covariates, we estimate the average effect of cohort change using Double Machine Learning (DML). DML combines flexible machine-learning prediction of nuisance functions with orthogonalized moment conditions, yielding estimators that are robust to regularization bias under suitable conditions (Chernozhukov et al., 2018).

We adopt a partial-linear setup in which Y_{it} depends linearly on D_{it} after controlling for a potentially nonlinear function of covariates and fixed effects. Implementation proceeds via cross-fitting: the sample is split into folds, nuisance models are trained on held-out folds, and out-of-fold predictions are used to residualize both the outcome and the treatment. The final effect estimate is obtained by regressing residualized outcomes on residualized treatments.

Given the modest sample size of the municipality-year panel, we prioritize stable learners (e.g., regularized linear models and random forests) for nuisance estimation. Cross-fitting is performed with folds defined at the municipality level to reduce leakage from repeated observations of the same municipality across years (Chernozhukov et al., 2018). Standard errors are computed using influence-function-based formulas with municipality clustering where applicable.

3.6 Heterogeneous effects and causal forests

Beyond average effects, the conceptual framework predicts that demographic shocks may have stronger implications in municipalities whose school systems operate closer to minimum viable scale. To examine this heterogeneity, we employ forest-based treatment effect estimators. Causal forests build on random forests (Breiman, 2001) to estimate conditional average treatment effects (CATEs) by recursively partitioning the covariate space and averaging within localized neighborhoods (Wager & Athey, 2018).

Operationally, we use school-system scale indicators (e.g., students-per-school, students-per-class, and school-age population share) as moderators and estimate how the effect of cohort decline varies across these dimensions. We summarize heterogeneity using the distribution of predicted CATEs and by comparing predicted effects across policy-relevant subgroups (e.g., municipalities in the bottom quartile of baseline students-per-school). The causal forest implementation follows the generalized random forests framework, which provides principled splitting rules and inference tools for heterogeneous effects (Athey, Tibshirani, & Wager, 2019; Wager & Athey, 2018).

Because forest-based CATE estimation relies on conditional unconfoundedness assumptions in observational settings, we interpret heterogeneity estimates as conditional patterns after adjustment for observed covariates and fixed effects, rather than as definitive causal effects in the presence of unobserved time-varying confounders (Rosenbaum & Rubin, 1983; Imbens & Rubin, 2015).

3.7 Robustness and sensitivity analyses

To assess the stability of findings, we conduct a structured set of robustness checks. First, we vary the treatment definition (percent change versus log change) and timing (lagged versus contemporaneous cohort change). Second, we consider alternative outcomes, including action counts and changes in the number of public elementary schools. Third, we vary the handling of missing event-count entries (blank-as-zero versus treating blanks as missing) and the handling of extreme cohort-change values (winsorization versus trimming). Finally, we compare results across complete-case samples and samples that allow missing covariates via indicator methods. This bundle of checks is designed to address common concerns in short-T panels with rare events and potential serial correlation (Bertrand et al., 2004; Cameron & Miller, 2015).

3.8 Reproducibility and ethical considerations

The analysis is conducted using aggregated municipality-level statistics and does not involve individual-level records. All transformations (lags, growth rates, and ratios) are deterministic functions of the original variables. To support replication and auditability, we document data-processing rules and estimation choices (sample windows, trimming rules, and model specifications) and report them alongside results in Section 4. When disseminating code and outputs, care is taken to preserve the integrity of administrative units and to avoid any inadvertent disclosure risks, even though the data are aggregated (Imbens & Rubin, 2015).

4. Empirical Results

4.1 Descriptive patterns and empirical variation

This chapter reports the empirical results corresponding to the hypotheses summarized in Table 3 (Section 2). We begin by describing the observed prevalence of elementary school closure/consolidation actions and the distribution of school-age cohort change, and then present benchmark two-way fixed-effects estimates, Double Machine Learning (DML) estimates of average effects, and forest-based heterogeneity analyses. All estimates use the municipality-year panel described in Section 3 and adopt municipality-clustered inference following standard guidance for panel settings (Bertrand, Duflo, & Mullainathan, 2004; Cameron & Miller, 2015; Wooldridge, 2010).

Across the full 2019–2024 panel (229 municipalities per year), consolidation actions are rare but fluctuate by year (Table 1). In the main estimation window (2021–2024), the mean event rate is 6.7% after excluding one municipality with missing demographic history, and the dependent variable is defined as an indicator for whether at least one action occurs in municipality i and year t.

Table 1. Annual prevalence of elementary school closure/consolidation actions (2019–2024)

Year	Municipalities (N)	Municipalities with ≥1 action	Event rate (%)	Total actions
2019	229	25	10.92	30
2020	229	15	6.55	18
2021	229	16	6.99	20
2022	229	11	4.80	15
2023	229	11	4.80	12
2024	229	23	10.04	29

Notes. “Municipalities with ≥1 action” counts municipality-years with at least one recorded closure/consolidation action. Event rate is the share of municipalities with any action in a given year.

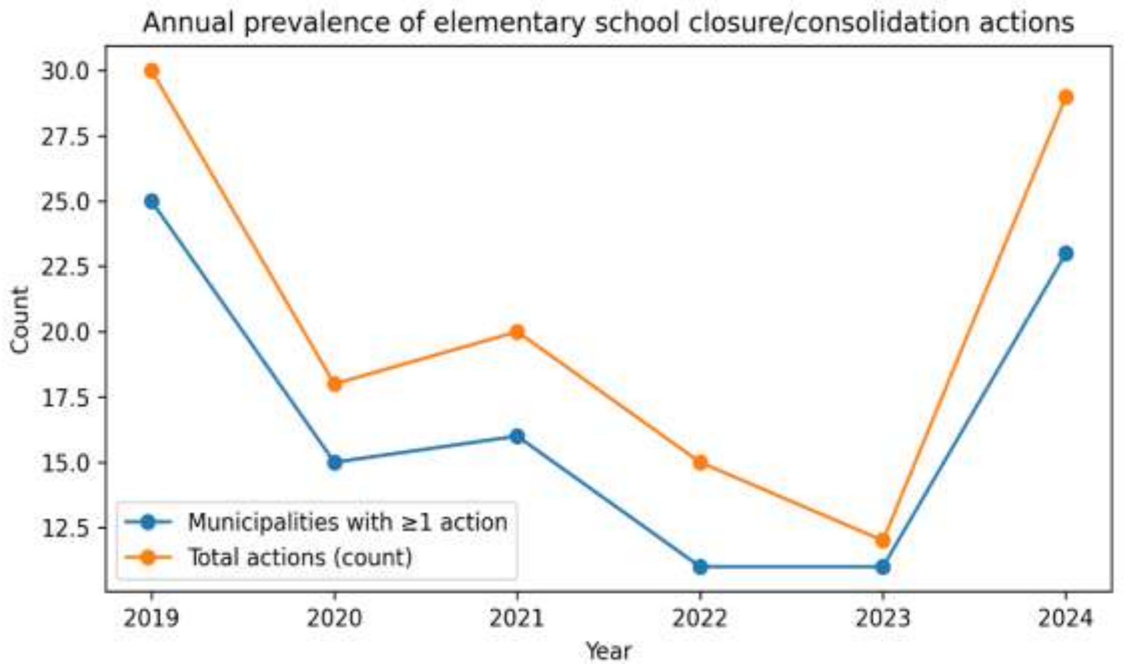

Figure 1. Annual prevalence of closure/consolidation actions: event municipalities and total actions (2019–2024).

The demographic driver used in the empirical strategy is the municipality-level change in the population ages 6–11. Figure 2 shows that year-on-year changes in the 6–11 cohort are predominantly negative in 2020–2024, with a distribution centered around moderate declines. For visualization, changes are winsorized at the 1st and 99th percentiles to reduce the influence of a small number of extreme values that can arise when the baseline cohort size is very small.

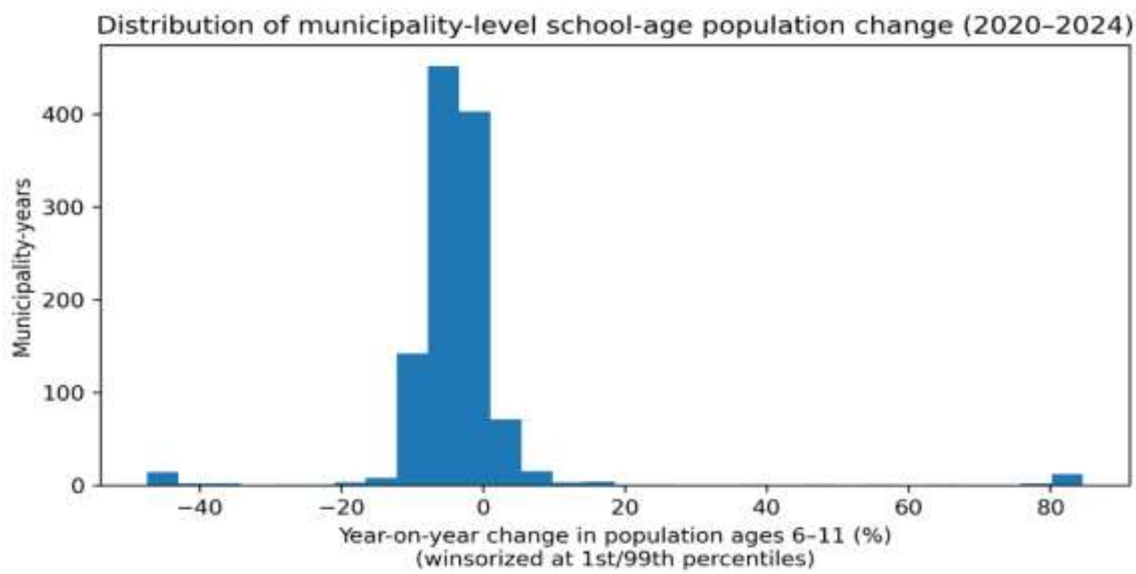

Figure 2. Distribution of municipality-level year-on-year change in population ages 6–11 (2020–2024; winsorized at 1st/99th percentiles).

To connect these demographic changes to consolidation outcomes, Table 2 and Figure 3 summarize consolidation event rates by the severity of the lagged school-age cohort change. Consistent with the motivating descriptive pattern, municipalities experiencing severe lagged declines (less than −5%) exhibit higher event rates than those with stable cohorts (−2% to +2%). Because these are descriptive contrasts, the next sections assess whether these gradients persist once municipality and year fixed effects (and baseline controls) are included.

Table 2. Consolidation event rates by severity of lagged school-age cohort change (2021–2024; trimmed sample)

Lagged change bin (t-1)	N	Mean lagged change (%)	Event rate (%)
<-5%	235	-7.10	7.66
-5% to -2%	338	-3.45	6.80
-2% to +2%	247	-0.50	5.26
>+2%	66	5.47	3.03

Notes. The lagged change is the year-on-year percent change in the population ages 6–11 between t-2 and t-1. The trimmed sample excludes observations with |lagged change| > 20 percentage points.

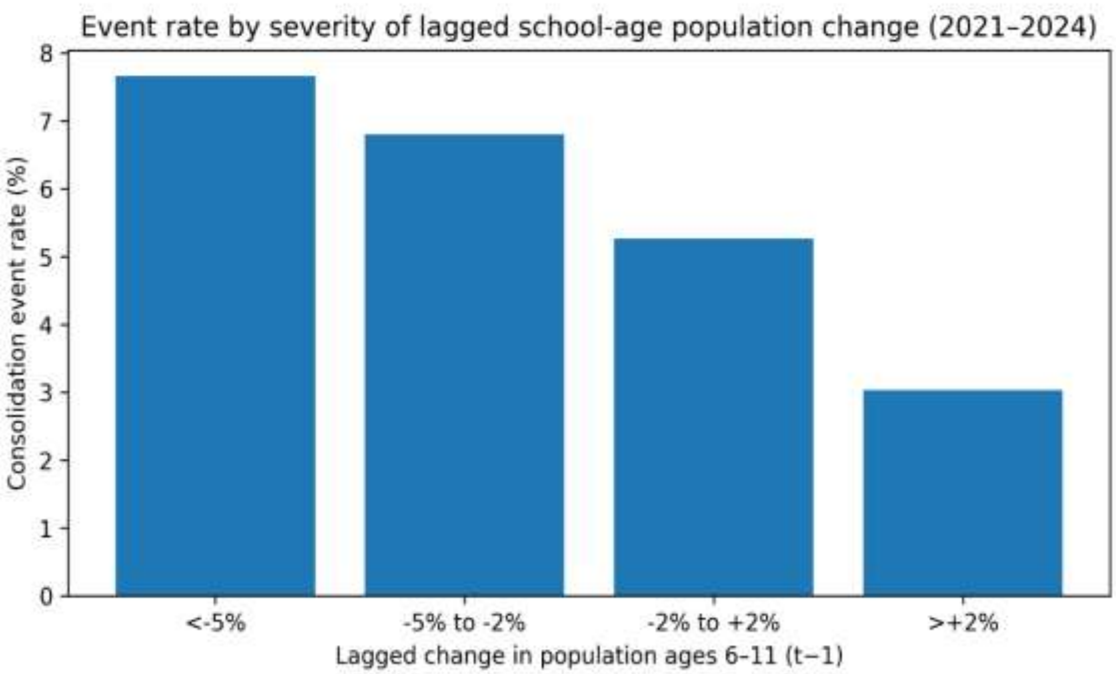

Figure 3. Consolidation event rate by lagged school-age cohort change category (2021–2024; trimmed sample).

4.2 Benchmark two-way fixed-effects estimates

We first estimate a two-way fixed-effects linear probability model (LPM) in which the consolidation event indicator is regressed on the lagged change in the 6–11 population and baseline covariates measured at t-2, with municipality and year fixed effects. The treatment variable is expressed in percentage points (pp), so the coefficient can be interpreted as the pp change in the probability of any consolidation action associated with a 1 pp change in the lagged cohort change. Given the sparse outcome, the LPM is used as

a transparent benchmark; the main conclusions are checked using flexible approaches in later sections (King & Zeng, 2001; Wooldridge, 2010).
To reduce leverage from implausibly large cohort changes, the benchmark results use a trimmed sample that excludes municipality-years with |lagged cohort change| exceeding 20 pp (26 observations; 2.85% of the 2021–2024 panel). Table 3 reports the benchmark estimates. Across specifications, the coefficient on lagged cohort change is small and statistically imprecise. The sign is consistent with H1 (more negative cohort change is associated with higher consolidation probability), but the short-run within-municipality variation over 2021–2024 does not yield precise average effects once fixed effects are included.

Table 3. Benchmark fixed-effects estimates of the effect of lagged school-age cohort change on consolidation events (2021–2024; trimmed sample)

	(1) FE only	(2) + controls	(3) + heterogeneity
Lagged ΔPop(6–11) (pp)	-0.0022	-0.0020	0.0051
	(0.0050)	(0.0054)	(0.0046)
Lagged ΔPop × Small baseline			-0.0187*
			(0.0101)
Baseline controls	No	Yes	Yes
Municipality FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	886	886	886

Notes. Entries report coefficients with cluster-robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$. "Small baseline" is an indicator for municipalities in the bottom quartile of students-per-school in 2019 (≤ 133.28 students per school). All models include municipality and year fixed effects; baseline covariates are measured at t-2.

4.3 Timing and nonlinearity tests

Hypothesis H4 proposes that demographic decline affects consolidation decisions with an administrative lag. To evaluate timing, Table 4 compares the fixed-effects estimate using the lagged cohort change (t-1) to a specification using the contemporaneous change in year t, and to a noted-but-less-preferred specification including both changes. The lagged estimate retains the expected sign (negative), whereas the contemporaneous estimate is not only imprecise but also does not align consistently with the expected direction. Figure 4 visualizes these timing estimates with 95% confidence intervals.

Table 4. Timing comparison: lagged versus contemporaneous cohort change (fixed effects; baseline controls)

	(1) Lagged only	(2) Contemporaneous only	(3) Both included
Lagged change (t−1) (pp)	-0.0020		-0.0002
	(0.0054)		(0.0064)
Contemporaneous change (t) (pp)		0.0026	0.0040
		(0.0048)	(0.0049)
Baseline controls	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	886	873	857

Notes. Entries report coefficients with municipality-clustered standard errors in parentheses. The contemporaneous change is the year-on-year change in population ages 6–11 between t-1 and t. Sample sizes differ slightly due to trimming rules applied to each change measure.

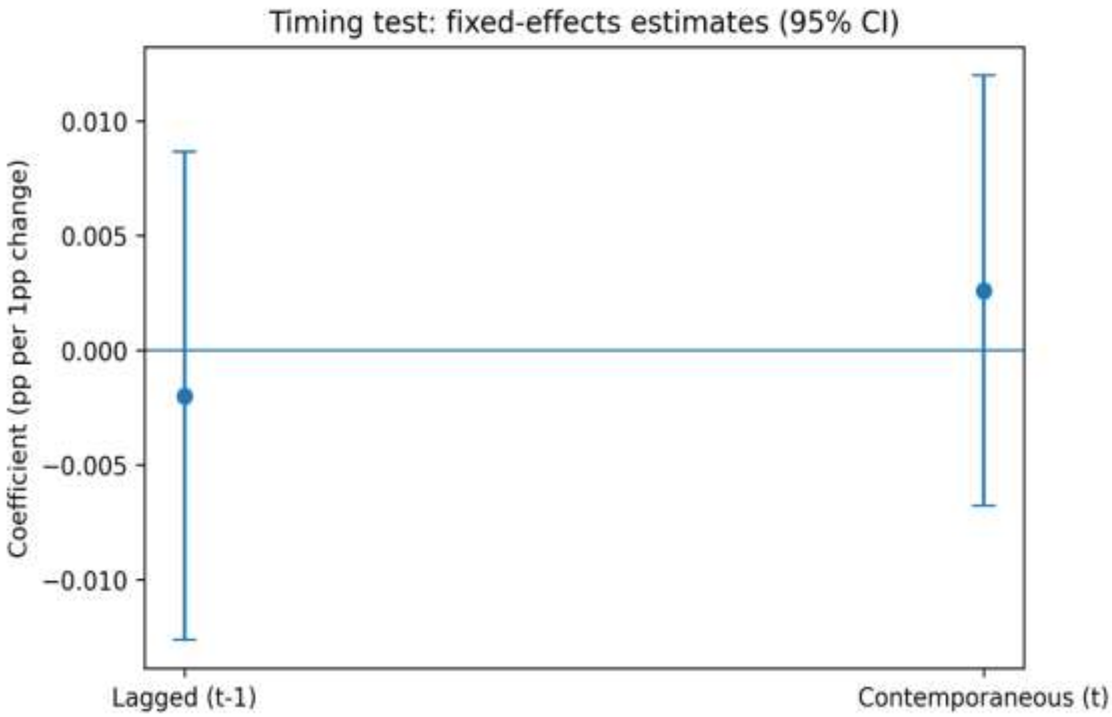

Figure 4. Timing comparison: fixed-effects coefficients for lagged and contemporaneous cohort change (95% CI).

Hypothesis H2 suggests nonlinear or threshold dynamics (e.g., sharp increases in consolidation risk once cohort decline passes a critical level). Table 2 provides descriptive evidence consistent with such thresholds, but a within-municipality FE specification using cohort-change bins does not yield statistically precise differences. Table 6 reports bin coefficients relative to the stable (−2% to +2%) category. Taken together, the results indicate that threshold-like patterns are visible descriptively but are difficult to estimate precisely in a short panel once fixed effects are imposed.

Table 6. Nonlinearity test: fixed-effects estimates using lagged cohort-change bins (reference: −2% to +2%)

Lagged change bin (vs. -2% to +2%)	Coef (pp)	SE (pp)	p-value
<-5%	1.2402	(4.4838)	0.782
-5% to -2%	1.3101	(2.8852)	0.650
>+2%	2.6718	(4.2102)	0.526

Notes. Coefficients are reported in percentage points (pp) relative to the stable category. Municipality and year fixed effects and baseline controls are included. * p < .10, ** p < .05, *** p < .01.

4.4 Double Machine Learning estimates of average effects

To relax functional-form restrictions and reduce regularization bias when adjusting for observed covariates, we estimate the average effect of lagged cohort change using Double Machine Learning (Chernozhukov et al., 2018). In the panel setting, we first remove municipality and year fixed effects using a standard two-way within transformation and then apply cross-fitted machine-learning models to

estimate the nuisance functions for the outcome and treatment. The final-stage coefficient is obtained by regressing the residualized outcome on the residualized treatment, with municipality-clustered inference.
Table 5 reports the DML estimate using random forests as nuisance learners. The estimated average effect is close to zero and statistically imprecise, mirroring the benchmark FE results. This pattern suggests that, in this dataset and time window, short-run within-municipality cohort changes do not translate into a robustly measurable average increase in consolidation-event risk once fixed effects and baseline covariates are accounted for.

Table 5. Double Machine Learning estimate of the average effect of lagged cohort change (2021–2024; trimmed sample)

Estimator	ATE on lagged ΔPop(6–11) (pp)	Cluster-robust SE	p-value	N
DML (partial linear; RF nuisances; 5-fold group cross-fitting)	-0.0006	0.0046	0.893	886

Notes. The treatment is lagged cohort change in percentage points (pp). Cross-fitting is performed using 5-fold group splits at the municipality level. The reported standard error clusters at the municipality level.

4.5 Heterogeneous effects by baseline school-system scale

A central motivation for using causal machine learning is the expectation that consolidation decisions are heterogeneous across municipalities. Hypothesis H3 predicts that the demographic sensitivity of consolidation is stronger where the baseline school system is structurally small (less slack capacity). We first test this hypothesis using an interaction model in which the lagged cohort-change slope is allowed to differ for municipalities in the bottom quartile of baseline students-per-school (measured in 2019).
Column (3) of Table 3 reports the interaction estimate. The interaction term is negative and marginally statistically significant at the 10% level, indicating that the same demographic decline is associated with a larger increase in consolidation-event probability in structurally smaller systems. Figure 5 presents the implied marginal slopes (with 95% confidence intervals) for the bottom-quartile group versus the remaining municipalities.

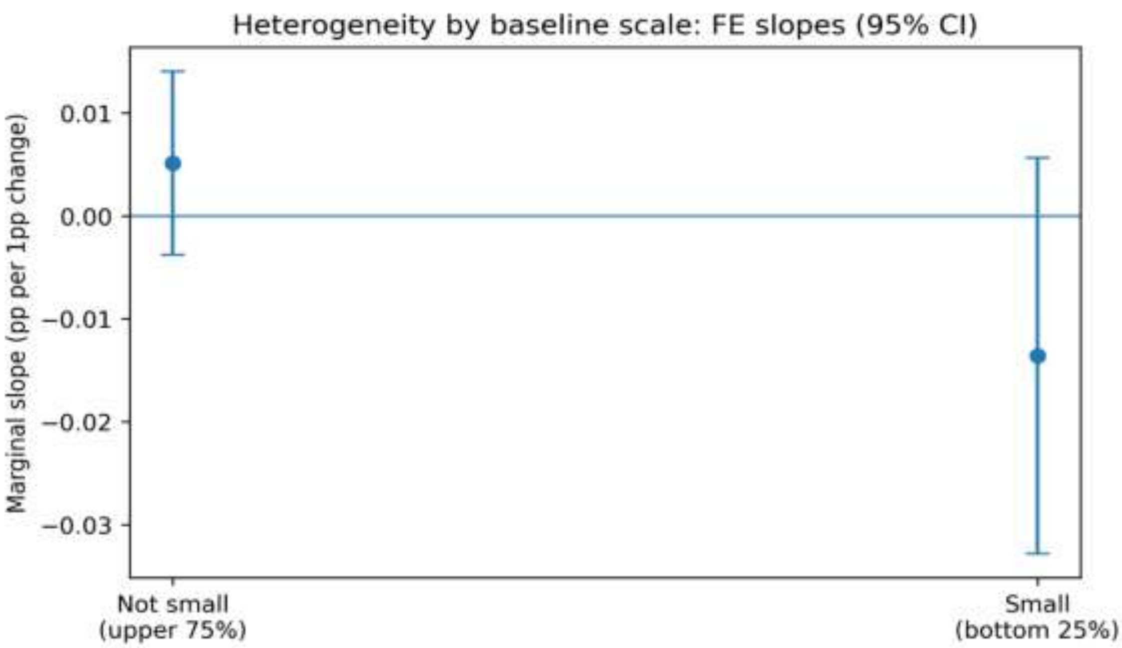

Figure 5. Heterogeneity by baseline system scale: implied FE slopes for small versus non-small baseline municipalities (95% CI).

4.6 Forest-based heterogeneity analysis (severe decline vs stable)

We complement the interaction test with an exploratory forest-based heterogeneity analysis aligned with the causal-forest literature. Following the heterogeneous treatment effects framework developed for tree/forest estimators (Athey & Imbens, 2016; Wager & Athey, 2018) and using random forests (Breiman, 2001), we define a binary treatment that captures whether a municipality-year experiences a severe lagged cohort decline (less than −5%) versus a stable cohort (−2% to +2%). We then construct doubly robust (augmented inverse probability weighted) pseudo-outcomes to reduce sensitivity to model misspecification (Bang & Robins, 2005) and fit a random forest to these pseudo-outcomes to obtain predicted conditional effects.

Table 7 summarizes the distribution of predicted conditional effects (in pp). Although the mean predicted effect is modest, the distribution is wide, suggesting substantial heterogeneity. Figure 6 visualizes the distribution. Figure 7 shows that baseline school-structure measures (students per school and students per class) and demographic scale variables are among the most influential predictors of heterogeneity in the forest model.

Table 7. Distribution of predicted heterogeneous effects (pp): severe decline (<-5%) versus stable (−2% to +2%)

Statistic	Predicted effect (pp)
Mean	1.07
SD	15.03
P10	-18.78
P25	-4.48
P50	1.51
P75	9.48
P90	20.43

Notes. Effects are estimated using cross-fitted doubly robust pseudo-outcomes and a random forest. Reported values summarize predicted conditional effects across the analysis subset (N = 482).

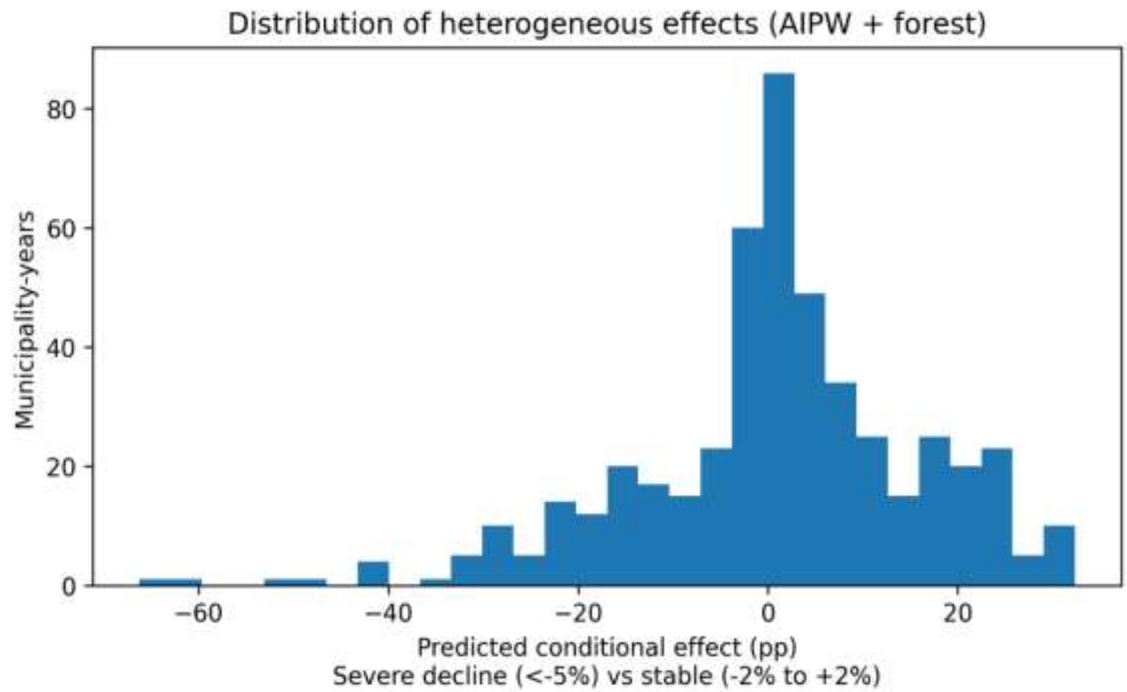

Figure 6. Distribution of predicted heterogeneous effects (pp): severe decline vs stable (AIPW + forest).

Table 8. Feature importance in the forest model (heterogeneity proxy)

Feature	Importance
Students per school (t-2)	0.203
Students per class (t-2)	0.148
School-age share (t-2)	0.122
Log pop ages 6–11 (t-2)	0.122
Log total population (t-2)	0.121
Log number of schools (t-2)	0.109
Log number of students (t-2)	0.093
Log number of classes (t-2)	0.071
Year indicators (2022–2024)	0.011

Notes. Feature importance is computed from the fitted random forest used to model doubly robust pseudo-outcomes. Values sum to 1 across included predictors.

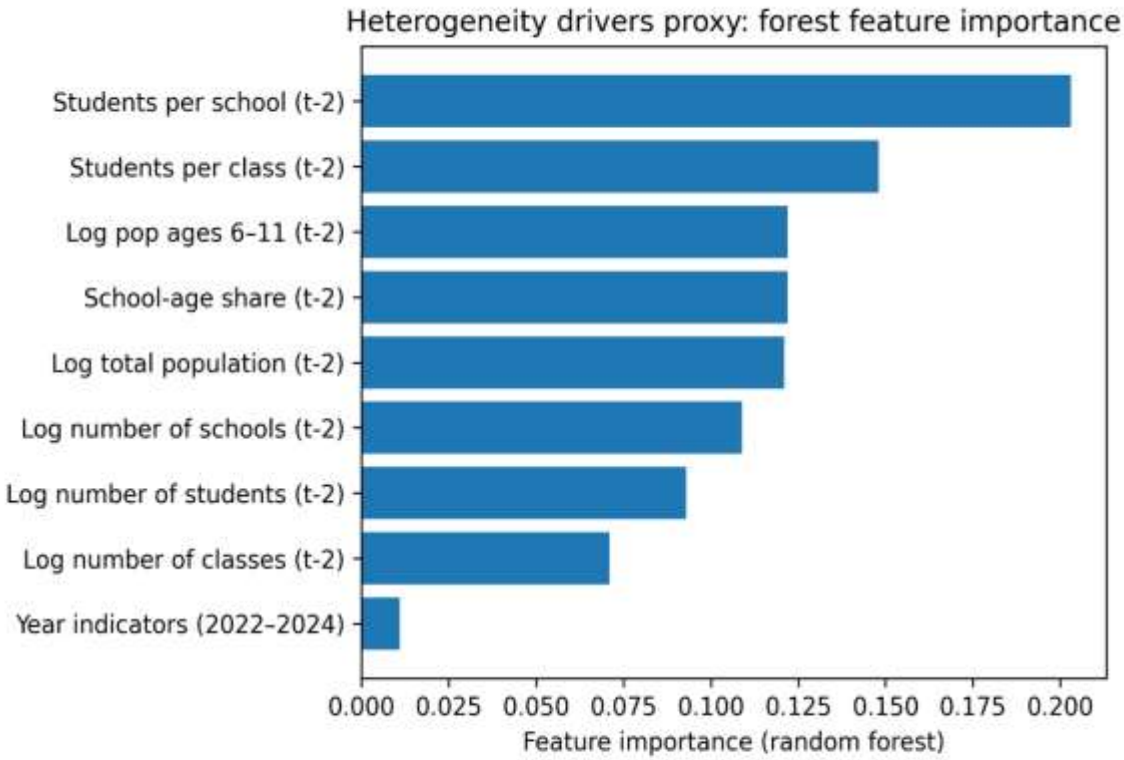

Figure 7. Feature importance in the forest model (heterogeneity proxy).

4.7 Robustness checks

Finally, we assess robustness to alternative treatment and outcome definitions and to alternative outlier handling. Table 9 summarizes four checks: (i) the baseline FE estimate with trimming, (ii) an alternative treatment defined as the lagged log change in population ages 6–11, (iii) an alternative outcome using the count of actions rather than the event indicator, and (iv) an outlier rule based on winsorization rather than trimming. Across these variants, the estimated average effects remain statistically imprecise, reinforcing the conclusion that the primary empirical signal lies in heterogeneity rather than in a strong average effect.

Table 9. Robustness checks summary (key coefficient on cohort change)

Specification	Key coefficient	SE	N
Main (FE + controls; trimmed |Δ|≤20pp)	-0.0020	(0.0054)	886
Alt treatment: lagged Δlog Pop(6–11) (pp)	-0.0025	(0.0056)	886
Alt outcome: action count (OLS FE)	-0.0018	(0.0062)	886
Alt outliers: winsorized ΔPop (5th/95th)	0.0033	(0.0065)	912

Notes. Coefficients are reported in probability units (pp per 1 pp cohort change) for the event outcome and in count units for the count outcome. * p < .10, ** p < .05, *** p < .01.

4.8 Section summary

In summary, benchmark two-way fixed-effects estimates and DML estimates both indicate that the average effect of short-run school-age cohort change on consolidation events is small and imprecisely estimated over 2021–2024. However, interaction tests and forest-based heterogeneity exercises suggest that the relationship is not uniform across municipalities: structurally small school systems exhibit greater sensitivity to cohort decline, and predicted conditional effects vary widely across municipality-years. These results motivate policy responses that are targeted toward structurally vulnerable municipalities, rather than uniform national rules.

5. Conclusion and Policy Implications

5.1 Summary of findings

This study examined whether and how local demographic contraction translates into elementary school closure and consolidation activity in South Korea. Using a municipality-year panel covering all 229 Si/Gun/Gu municipalities from 2019 to 2024, the analysis linked year-on-year changes in the elementary school-age population (ages 6–11) to annual consolidation outcomes and estimated average and heterogeneous associations using two-way fixed effects and causal machine-learning methods.

Three empirical conclusions emerge from the results in Section 4. First, consolidation activity is episodic and uneven across time and space. Even within a six-year window, consolidation events are relatively rare in most years, and a subset of municipalities accounts for a disproportionate share of event-years. This pattern is consistent with a broader international literature showing that school network adjustment is often implemented in discrete policy episodes rather than as a smooth response to annual enrollment fluctuations (Ares Abalde, 2014; Organisation for Economic Co-operation and Development [OECD], 2018).

Second, after accounting for time-invariant municipal characteristics and common year shocks, the average within-municipality association between short-run cohort change and consolidation activity is small and statistically imprecise. In the benchmark fixed-effects model, the point estimate for lagged cohort change is close to zero (Section 4, Table 3). The Double Machine Learning (DML) estimate likewise indicates an average effect close to zero (Section 4, Table 5), suggesting that year-to-year cohort variation over 2021–2024 is not, on average, a strong predictor of whether a municipality experiences a consolidation action in a given year once fixed effects and baseline controls are incorporated (Chernozhukov et al., 2018).

Third, the evidence is more consistent with a heterogeneous-risk interpretation than with a homogeneous national effect. Interaction results indicate that municipalities with structurally small baseline school systems (proxied by low students-per-school in 2019) are more sensitive to cohort decline (Section 4, Table 3; Figure 5). The forest-based heterogeneity analysis also yields a wide distribution of conditional effects for a severe-decline contrast (Section 4, Table 7; Figure 6), highlighting that demographic contraction interacts with baseline system constraints and local context (Wager & Athey, 2018).

Taken together, these results imply that demographic decline is better understood as a background structural pressure that raises consolidation risk primarily where local school systems operate close to viability thresholds, rather than as a uniform, near-mechanical trigger for consolidation across all municipalities and years (Duncombe & Yinger, 2007; OECD, 2018).

5.2 Contributions to scholarship and practice

This study contributes to the scholarship on school network reconfiguration in three ways. First, it shifts the unit of analysis from individual schools to the municipality-year level, which aligns with how consolidation decisions are frequently planned and recorded in administrative systems. Second, by combining fixed-effects panel models with DML and causal forests, it integrates modern causal machine-learning tools into the study of demographic change and public service reconfiguration, while preserving transparent benchmarking against conventional estimators (Chernozhukov et al., 2018; Wager & Athey, 2018).

Third, the analysis emphasizes heterogeneity and vulnerability. Existing research highlights that consolidation can be motivated by economies of scale and capacity utilization, but can also impose community and access costs, especially in rural or sparsely populated areas (Ares Abalde, 2014; Lyson, 2002; OECD, 2018). By showing that sensitivity to cohort decline is stronger in structurally small baseline systems, the paper provides an empirical basis for prioritizing municipalities where these trade-offs are likely to be most acute.

5.3 Policy implications

The results suggest that policy responses to demographic shrinkage should move beyond uniform rules and toward risk-sensitive, place-based planning. Because average within-municipality effects are imprecise in short panels, consolidation planning should not rely exclusively on one-year cohort changes as automatic triggers. Instead, early-warning screening can combine baseline indicators of structural vulnerability (e.g., students-per-school or other capacity proxies) with sustained multi-year cohort decline to identify municipalities where proactive network planning is most needed (Ares Abalde, 2014; OECD, 2018).

Where consolidation becomes likely, the policy objective should be to maintain service continuity and minimize equity losses. International evidence emphasizes that the fiscal savings from consolidation depend on local scale and adjustment costs, and that transportation and transition costs can offset headline efficiency gains (Duncombe & Yinger, 2007; OECD, 2018). Accordingly, consolidation packages in high-risk municipalities should be paired with mitigation tools such as transportation support, multi-campus or shared-services models, and phased implementation tied to facility investment cycles.

Finally, consolidation decisions should explicitly consider community externalities. Schools often function as local anchors, and closures may affect community vitality, social capital, and perceptions of local viability (Lyson, 2002). Although the present study does not estimate downstream impacts, the broader evidence on student and receiving-school effects indicates that closure and reassignment policies can create disruption and spillovers even when average achievement impacts are not uniformly negative (Brummet, 2014). These considerations reinforce the need for transparent decision processes and structured stakeholder engagement, particularly in municipalities facing repeated consolidation pressure.

5.4 Limitations

Several limitations qualify interpretation. First, the consolidation outcome is sparse and measured as annual recorded actions, which may mask heterogeneity in the magnitude and type of restructuring (e.g., mergers versus closures). Sparse outcomes also raise practical challenges for statistical modeling and prediction in short panels, where rare-event bias and small-sample instability can be concerns (King & Zeng, 2001).

Second, the panel horizon is short for a policy process that may respond to multi-year demographic trajectories and administrative planning cycles. If consolidation decisions are driven by cumulative enrollment trends or multi-year fiscal planning, year-to-year cohort changes will only partially capture the relevant decision signal.

Third, although municipality and year fixed effects absorb many confounders, the dataset contains a limited set of time-varying covariates. Time-varying factors such as migration, housing development, or local fiscal capacity could affect both cohort trajectories and consolidation decisions. The DML and forest

estimates therefore should be interpreted as causal only under an unconfoundedness assumption conditional on observed controls and fixed effects (Chernozhukov et al., 2018; Wager & Athey, 2018).
Fourth, inference in panel settings requires careful treatment of within-municipality dependence. While the analysis used clustered standard errors, finite-sample properties can be sensitive to the number of clusters and the correlation structure (Cameron & Miller, 2015).

5.5 Directions for future research

Future research can extend this work in at least four directions. First, expanding the panel window would enable event-history or hazard models that better reflect the discrete nature of consolidation decisions and allow cumulative demographic pressure to be modeled explicitly. Second, linking municipality statistics with school-level administrative records would support a clearer distinction between closures, mergers, and reorganizations, and would permit analyses of how consolidation affects school size distributions and travel distances.

Third, incorporating spatial dependence and inter-municipal interactions is important, as consolidation decisions can shift demand across boundaries and may be coordinated regionally. Fourth, outcome-focused work should evaluate the consequences of consolidation for student access and community well-being, building on evidence that closure policies can generate disruption and spillovers (Brummet, 2014; Lyson, 2002).

Finally, the present study's emphasis on heterogeneity motivates applied visualization and decision-support tools. Mapping municipality-level risk and vulnerability indicators can improve communication with stakeholders and facilitate transparent prioritization, consistent with OECD guidance that school network adaptation be embedded in broader regional service planning (OECD, 2018).

5.6 Concluding remarks

South Korea's demographic transition is reshaping the demand for public education in ways that differ sharply across municipalities. This study finds that short-run cohort changes do not translate into a precisely estimated average increase in consolidation activity once fixed effects are considered, but that consolidation risk is concentrated and appears more sensitive to demographic decline in structurally small baseline school systems. A policy strategy that is risk-sensitive, place-based, and paired with mitigation measures is therefore more likely to balance efficiency, equity, and community sustainability than a uniform consolidation rule.

References

1. Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting economies of size in American education: Are we any closer to a consensus? Economics of Education Review, 21(3), 245–262. https://doi.org/10.1016/S0272-7757(01)00006-1
2. Angrist, J. D., & Pischke, J.-S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton University Press.
3. Arellano, M. (1987). Computing robust standard errors for within-groups estimators. Oxford Bulletin of Economics and Statistics, 49(4), 431–434.
4. Ares Abalde, M. (2014). School size policies: A literature review (OECD Education Working Papers No. 106). OECD Publishing. https://doi.org/10.1787/5jxt472ddkjl-en
5. Athey, S., & Imbens, G. W. (2016). Recursive partitioning for heterogeneous causal effects. Proceedings of the National Academy of Sciences, 113(27), 7353–7360. https://doi.org/10.1073/pnas.1510489113
6. Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. Annals of Statistics, 47(2), 1148–1178. https://doi.org/10.1214/18-AOS1709
7. Bang, H., & Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. Biometrics, 61(4), 962–973.
8. Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. The Review of Economic Studies, 81(2), 608–650. https://doi.org/10.1093/restud/rdt044
9. Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? Quarterly Journal of Economics, 119(1), 249–275.

10. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324
11. Brummet, Q. (2014). The effect of school closings on student achievement. Journal of Public Economics, 119, 108–124. https://doi.org/10.1016/j.jpubeco.2014.06.010
12. Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317–372.
13. Caven, M. (2019). Quantification, inequality, and the contestation of school closures in Philadelphia. Sociology of Education, 92(1), 21–40. https://doi.org/10.1177/0038040718815167
14. Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. The Econometrics Journal, 21(1), C1–C68. https://doi.org/10.1093/ectj/utx023
15. Duncombe, W. D., & Yinger, J. (2007). Does school district consolidation cut costs? Education Finance and Policy, 2(4), 341–375. https://doi.org/10.1162/edfp.2007.2.4.341
16. Howley, C., Johnson, J., & Petrie, J. (2011). Consolidation of schools and districts: What the research says and what it means. National Education Policy Center.
17. Imbens, G. W., & Rubin, D. B. (2015). Causal inference for statistics, social, and biomedical sciences: An introduction. Cambridge University Press.
18. Kim, J. (2022). The effect of the Optimal School Size Development policy on small-school closure rates: Differences by regional scale and school level [In Korean]. Korean Journal of Educational Research, 60(6), 493–522. https://doi.org/10.30916/KERA.60.6.493
19. King, G., & Zeng, L. (2001). Logistic regression in rare events data. Political Analysis, 9(2), 137–163. https://doi.org/10.1093/oxfordjournals.pan.a004868
20. Korean Educational Development Institute. (2010). Analysis of the effects of integrating and abolishing small schools in rural and fishing villages (RR2010-07) [Research report, in Korean]. KEDI.
21. Lee, I. (2013). An Analysis on the Policy of Integrating and Abolishing Small Schools from the Perspective of Educational Governance [In Korean]. Tamla Munhwa (Tamla Culture), 42, 347–375. https://doi.org/10.35221/tamla.2013..42.012
22. Lyson, T. A. (2002). What does a school mean to a community? Assessing the social and economic benefits of schools to rural villages in New York. Journal of Research in Rural Education, 17(3), 131–137. https://jrre.psu.edu/sites/default/files/2019-08/17-3_1.pdf
23. Nitta, K. A., Holley, M. J., & Wrobel, S. L. (2010). A phenomenological study of rural school consolidation. Journal of Research in Rural Education, 25(2), 1–19.
24. Organisation for Economic Co-operation and Development. (2018). Responsive school systems: Connecting facilities, sectors and programmes for student success. OECD Publishing. https://doi.org/10.1787/9789264306707-en
25. Organisation for Economic Co-operation and Development. (2024a). How are demographic changes affecting education systems? Education Indicators in Focus, No. 87. OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/12/how-are-demographic-changes-affecting-education-systems_2de89d67/158d4c5c-en.pdf
26. Organisation for Economic Co-operation and Development. (2024b). SF2.1: Fertility rates (OECD Family Database). https://webfs.oecd.org/els-com/Family_Database/SF_2_1_Fertility_rates.pdf
27. Park, S. (2014). Small school consolidation and policy issues [In Korean]. The Journal of Educational Administration, 32(4), 1–21.
28. Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41–55.
29. Statistics Korea. (2025, August 27). Birth statistics in 2024 [Press release]. Ministry of Data and Statistics. https://www.kostat.go.kr/board.es?act=view&bid=11773&list_no=439008&mid=a20101000000
30. Sørensen, J. F. L., Svendsen, G. L. H., Jensen, P. S., & Schmidt, T. D. (2021). Do rural school closures lead to local population decline? Journal of Rural Studies, 87, 226–235. https://doi.org/10.1016/j.jrurstud.2021.09.016

31. Tieken, M. C., & Auldridge-Reveles, T. R. (2019). Rethinking the school closure research: School closure as spatial injustice. Review of Educational Research, 89(6), 917–953. https://doi.org/10.3102/0034654319877151
32. Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association, 113(523), 1228–1242. https://doi.org/10.1080/01621459.2017.1319839
33. Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data (2nd ed.). MIT Press.
34. Yang, Y., Hwang, H., & Pareliussen, J. (2024). Korea's unborn future: Lessons from OECD experience (OECD Economics Department Working Papers No. 1824). OECD Publishing. https://doi.org/10.1787/75aa749c-en