

Composition Is Not Adjustment

Regional Aging and Growth in Germany

Michael Schymura

18 August 2026 · Referee-redesign draft 0.8.1

ABSTRACT

An older regional population, a region undergoing an aging transition, and the economic adjustment to that transition are different empirical objects. This paper separates them for all 16 German Länder. During periods with a 10% higher average old-to-young ratio, annual real workplace GDP growth per resident aged 20–79 is 0.317 percentage points lower, close to Hayashi’s 0.397-point cross-country analogue. Under a separate common- X design, a 10% larger adjacent-average aging transition is associated with 0.289 points higher annual output growth per resident and 0.388 points higher output growth per hour. Only the output-per-hour component survives the single-step max- t family gate. Capital per hour rises by 0.490 points, but primitive capital-stock index growth is 0.027 points while actual hours growth is -0.464 points. The capital ratio turns negative when the eastern group is removed or the sample is restricted to western Länder. The full-sample association is therefore consistent with East-specific convergence or restructuring, not a nationwide aging-induced investment or automation response. Resident employment and bilateral commuting remain unavailable. All results are conditional regional and accounting associations, not an identified aging effect or causal mediation.

JEL: J11, O40, R11, R23, O33

Keywords: population aging; regional growth; productivity; migration; Germany

Data statement: official population, output, employment, hours, capital-stock, and sectoral value-added releases underlie the Länder analysis. Raw files remain local. The package registers sources, hashes, licenses, transformations, supported cells, and explicit evidence gaps.

Reproducibility statement: the 16-Land profiles, 56-run estimand map, six legacy accounting runs, 396-region Kreis descriptive layer, and subsequent evidence dispositions are frozen for offline reconstruction. Paper 0.8.1 binds corrected annualization, direct-null WCR- t inference, primitive capital/hours outcomes, descriptive sensitivities, and standardized influence to paper-081-correctness-e58951758cd9de21. That append-only release retains the frozen Spec 035 and 036 products as reviewed parents. Resident/workplace evidence remains bound to hayashi-spec-037-resident-workplace-bridge-08fa7b6df1473e1f; Kreis descriptions remain bound to kreis-descriptive-evidence-de30cfa44327e03d.

Contents

1	The Question Behind the Coefficient	3
2	Why Composition and Adjustment Differ	4
2.1	Composition, Transition, and Pressure	4
2.2	The Accounting Backbone	4
2.3	Regional Equilibrium and Testable Propositions	5
2.4	From Hayashi to the German Design	5
2.5	How the Framework Changes the Literature	6
2.5.1	Age composition is not one treatment	6
2.5.2	Migration makes regional demography endogenous	6
2.5.3	Adjustment can run through labor and technology	6
3	Data and the German Regional Setting	7
3.1	Sources and Measurement	7
3.2	Recovered Historical Inputs and Their Boundary	7
3.3	Residence and Workplace Denominators	8
3.4	What the Regional Profiles Show	11
3.5	First Kreis Descriptive Evidence	11
3.6	What the Extension Audits Rule Out	13
3.7	From Annual Levels to Growth Observations	13
3.8	Boundary of the Current Evidence	15
4	Empirical Design and Identification Ladder	15
4.1	Growth and Timing	15
4.2	Weights, Scaling, and Inference	15
4.3	Accounting Associations	16
4.4	Common-X Accounting and Joint Inference	16
4.5	Identification Ladder	16
5	Composition, Transition, and Denominators	17
5.1	The German Benchmark and the Hayashi Scale	17
5.2	Same Sample, Different Demographic Object	18
5.3	Within-Land Evidence and the Denominator	18
5.4	Horizon and Pandemic Boundaries	18
6	How Regions Adjust: What the Accounts Show	22
6.1	Outcome-Specific Accounting Regressions	22
6.2	Output per Hour	22
6.3	Capital Deepening Is a Ratio	22
6.4	The Residual Does Not Name the Mechanism	22
6.5	The Common-X Decomposition	22
6.6	The Capital Numerator and Hours Denominator	24
6.7	Direct Tests, Power, and Sensitivity	24
7	Threats, Power, and Decisive Tests	26
7.1	Sorting and Post-Unification Convergence	26
7.2	Sixteen Clusters and Many Correlated Questions	26
7.3	Census, Price, and Location Concepts	26
7.4	What Would Change the Interpretation	26
8	Conclusion: Composition Is Not Adjustment	27

SECTION 1

The Question Behind the Coefficient

A regional pattern can survive even when its proposed mechanism does not. Across the German Länder, periods with older average age structures coincide with slower growth in real workplace GDP per resident aged 20–79 during the same outcome windows. The primary estimate is close to Hayashi’s mature-economy coefficient in both sign and scale. Yet a larger realized aging transition is positively associated with output growth under a separate common- X design. Composition is not adjustment. More precisely, demographic composition and demographic transition are not interchangeable treatments, and neither is by itself demographic adjustment.

The distinction matters because regional age structure is partly an equilibrium outcome. Younger adults move toward universities, jobs, and expanding labor markets; commuters produce in places where they do not reside; and post-unification convergence links demographic and economic change across decades. Germany holds law, currency, and many national institutions in common, but it does not hold migration, commuting, sector structure, or regional history constant. A negative composition coefficient may therefore describe demographic pressure, selective sorting, persistent regional type, or several channels at once.

Hayashi (2026) relates real GDP growth per resident aged 20–79 to the age composition of that same broad adult population. The denominator is demographic, not an employment count. Paper 0.8.1 therefore avoids “working-age adult” as a default label: ages 20–79 include many residents outside the labor force, while workplace GDP can also embody incoming commuters. GDP per resident aged 20–79 and GDP per workplace-employed person answer different accounting questions. Their disagreement is an economic result to explain, not a terminology problem to conceal.

The closest German regional evidence finds a negative but imprecise aggregate relationship between workforce aging and county productivity growth (Bode et al., 2023). This paper makes a different contribution. First, it reconstructs Hayashi’s exact resident old-to-young ratio and separates period-average composition from the change between adjacent equal-length period averages. Second, it reports the complete pre-specified grid for two GDP denominators, allowing the denominator to reveal an adjustment margin rather than act as a hidden modeling choice. Third, it estimates exact common- X accounting identities for output, hours, workplace employment, and capital on one sample. Fourth, it tests same-system coefficients and substantive contrasts with direct-null studentized $WCR-t$, separately reporting pointwise, single-step, and stepdown procedures. Finally, it preserves the line the available data cannot cross: resident employment and bilateral commuting remain unobserved on the required common geography and timing.

SEQUENTIAL FREEZE

Definitions, sources, windows, and inference were frozen before estimation. The complete 56-run grid and six-run accounting extension remain unchanged. Paper 0.8.1 consumes an append-only correction release over the immutable common- X and joint-inference parents.

The results establish one benchmark and a sharper tension. In the primary all-Länder comparison, a 10% higher period-average old-to-young ratio is contemporaneously associated with 0.317 percentage points lower annual growth in real GDP per resident aged 20–79. Hayashi’s comparable cross-country point estimate is 0.397 points. Under one common- X transition design, however, a 10% larger adjacent-average aging change is associated with 0.289 points higher annual output growth per resident and 0.388 points higher output growth per hour. The two estimates are similar in absolute scale and opposite in sign, but they are different endogenous estimands and are not linked by a valid equality test.

Those findings do not establish that faster aging caused higher output growth or induced investment. Output per hour survives the corrected single-step component-family gate; capital per hour does not. Primitive capital-stock index growth is near zero, actual-hours growth is negative, and the positive capital ratio reverses outside eastern Germany. The resident/workplace bridge still emits zero numerical rows. Common- X accounting narrows the mechanism question; it does not answer it. The mechanism still has to earn its name.

SECTION 2

Why Composition and Adjustment Differ

2.1 Composition, Transition, and Pressure

Let P_{rat} denote the population of region r , age a , in year t . The annual log old-to-young ratio is

$$a_{rt} = \log \left(\frac{\sum_{a=50}^{79} P_{rat}}{\sum_{a=20}^{49} P_{rat}} \right). \quad (2.1)$$

For a horizon of H years, the period-average level uses exactly H annual observations:

$$\bar{a}_{r,t,H} = \frac{1}{H} \sum_{s=0}^{H-1} a_{r,t+s}. \quad (2.2)$$

This is the logarithm of the geometric mean of annual ratios. It is not the log ratio of arithmetic-average population stocks.

The frozen realized-aging field compares adjacent equal-length period averages:

$$\Delta \bar{a}_{r,t,H} = \bar{a}_{r,t,H} - \bar{a}_{r,t-H,H}. \quad (2.3)$$

It is not the endpoint change $a_{r,t+H} - a_{rt}$. The two objects load different years, smooth measurement error differently, and can rank regions differently. Paper 0.8.1 retains Equation (2.3) because it matches the frozen field `observed_aging_change`; endpoint change remains a separate, unestimated sensitivity.

A third object is predetermined cohort-transition pressure. For horizon h , a leave-focal-region-out stock-transition factor $\widehat{\tau}_{-r,t,h,a}$ applies historical donor transitions to the focal region's baseline cohorts. The notation is deliberate. The factor can contain mortality, donor migration, census revisions, and unresolved boundary changes; it is neither a pure survival probability nor automatically an excluded instrument.

2.2 The Accounting Backbone

Let Q^W denote real workplace GDP, $\mathcal{H}^W$ workplace actual hours, E^W workplace employment, E^R resident employment, and N^R residents aged 20–79. The current data support the exact identity

$$\frac{Q^W}{N^R} = \frac{Q^W}{\mathcal{H}^W} \times \frac{\mathcal{H}^W}{E^W} \times \frac{E^W}{N^R}. \quad (2.4)$$

Its components are hourly labor productivity, hours per workplace-employed person, and workplace employment relative to the resident broad-adult population. The final term is a mixed utilization-and-commuting margin. Once resident employment is observed on a common concept and geography, it can be split further:

$$\frac{E^W}{N^R} = \frac{E^W}{E^R} \times \frac{E^R}{N^R}. \quad (2.5)$$

The first factor then measures workplace concentration or net commuting; the second measures resident employment utilization.

Taking log changes makes Equation (2.4) additive. If every component is estimated on the same observations with the same demographic regressor, predetermined controls, fixed effects, weights, and resampling sequence, linearity also implies

$$\widehat{\beta}_{Q/N} = \widehat{\beta}_{Q/\mathcal{H}} + \widehat{\beta}_{\mathcal{H}/E^W} + \widehat{\beta}_{E^W/N}. \quad (2.6)$$

The restriction does not apply to the original headline coefficients because their lagged outcome controls differ by denominator. A subtraction of those coefficients is therefore a diagnostic contrast, not a common- X

mechanism result. Spec 035 estimates a separate system that holds observations, regressors, effects, weights, parameter order, and bootstrap draws fixed across all seven accounting outcomes.

The growth-accounting layer is

$$g^{Q/H} = \alpha g^{K/H} + g^R, \quad (2.7)$$

where K is the frozen gross-fixed-asset stock chain index, α is the endpoint-average capital share from the adjusted labor-share rule, and g^R is the accounting residual. Moreover,

$$g^{K/H} = g^{K/N^R} - g^{H/N^R}. \quad (2.8)$$

Thus, capital deepening can rise because capital grows, hours fall, or both. The residual can contain utilization, labor quality, allocation, markups, sector composition, and measurement error. The accounting backbone does not identify causal mediation, technology, or automation.

2.3 Regional Equilibrium and Testable Propositions

A parsimonious effective-labor block is sufficient to organize the regional setting:

$$L_{rt}^e = e_{rt}^y N_{rt}^y + e_{rt}^o N_{rt}^o + M_{rt}, \quad (2.9)$$

$$Q_{rt}^W = A_{rt} K_{rt}^\alpha (L_{rt}^e)^{1-\alpha}, \quad (2.10)$$

where e^y and e^o collect age-specific participation, employment, hours, education, experience, occupation, and task mix. The term M_{rt} represents labor supplied across residence boundaries. In a spatial equilibrium, productivity and job opportunities can attract younger residents and commuters, lowering the observed old-to-young ratio while raising workplace output. The same negative composition coefficient is therefore compatible with a direct demographic response, selective sorting, commuting, or persistent regional type (Mitze and Reinkowski, 2011; Monte et al., 2018).

The framework yields four testable propositions. First, if aging pressure operates mainly through utilization, its association should be more negative for employment and hours than for output per hour. Second, a capital-deepening association should be decomposed into capital and hours before it is interpreted as investment. Third, commuting may offset resident employment decline, so E^W/N^R should be split before it is called an employment mechanism. Finally, if selective sorting biases composition coefficients downward, baseline or cohort-transition estimates should be less negative than contemporaneous period-average estimates. Specs 035–036 now test the first two propositions as conditional accounting associations. Spec 037 leaves the third blocked, and Spec 034 does not estimate the fourth.

2.4 From Hayashi to the German Design

The August 2026 paper supplies both a canonical average-level equation and a transformed aging-speed projection. Table 1 prevents the two from acquiring one ambiguous label.

Table 1 Version and estimand crosswalk

Source object	Outcome	Demographic object	German role
Hayashi (2026), eq. (2)	Growth in GDP per population aged 20–79	Period-average log ratio	Canonical benchmark
Hayashi (2026), eq. (8)	Momentum-adjusted growth	Realized aging speed	Horizon adaptation
Hayashi (2026), eq. (17)	Projected growth differential	Forecast-vintage aging speed	Information-set benchmark
Acemoglu et al. (2026), App. C	Growth in GDP per capita	Average level versus end-point change	Estimand audit

The complete machine-readable mapping is stored in `metadata/hayashi_version_crosswalk.csv`.

Source: Hayashi (2026), supplied August 2026 paper; Acemoglu et al. (2026).

Sample: Source equations retained in the Spec 002 foundation crosswalk.

The projection design introduces the convergence-weighted momentum coefficient $m(H, \kappa) = \exp(-H\kappa)$. The pre-specified grid is $\kappa \in \{0.02, 0.04, 0.06\}$. Spec 007 audited an authentic 12th coordinated projection package published in 2009 with base 31 December 2008, variants W1/W2, and annual totals for all 16 Länder in 2011 and 2021. Its exact-age sheets, however, cover coordinated aggregate groups rather than individual Länder; individual-Land ages 20–49 and 50–79 cannot be reconstructed from the surviving broad bands. The live GENESIS table now contains the 16th projection and cannot be backdated. Forecast-vintage momentum therefore remains blocked, while cohort-predicted aging is neither a substitute for that information set nor estimated in the current release.

2.5 How the Framework Changes the Literature

2.5.1 Age composition is not one treatment

The empirical literature does not imply one mechanical mapping from age structure to economic performance. Cross-country work relates population composition to growth and productivity while emphasizing age profiles, initial conditions, and the horizon over which adjustment occurs (Lindh and Malmberg, 1999; Feyrer, 2007). Regional evidence adds another complication: demographic composition, human capital, sector structure, and spatial sorting vary together, and the resulting associations are heterogeneous across places (Daniele et al., 2019).

German studies make that regional problem concrete. Brunow and Hirte (2006) study age structure and regional growth; Brunow and Hirte (2009) connect workforce age and human capital to productivity across labor-market regions; Bönnte et al. (2009) examine entrepreneurship; and Gregory and Patuelli (2015) document persistent spatial polarization in demographic and economic conditions. The closest outcome study, Bode et al. (2023), relates workforce aging to county productivity growth and finds an aggregate relationship that is negative but imprecise. These papers do not estimate one common object. Workforce age differs from the resident old-to-young ratio; administrative counties differ from functional labor markets; and GDP per resident aged 20–79 differs from GDP per employed person. The present paper constructs Hayashi’s exact resident ratio for the Länder and reports both economic denominators under one registered estimand map.

2.5.2 Migration makes regional demography endogenous

Regional age structure is partly an equilibrium outcome. Age-specific migration responds to labor-market conditions in Germany (Mitze and Reinkowski, 2011); a productive region can attract younger adults, expand the denominator of Equation (2.1), and appear demographically younger for reasons that also predict subsequent output. This makes a negative level association compatible with selective sorting rather than a demographic mechanism.

Cohort-shift designs offer one way to remove contemporaneous local migration mechanically. Maestas et al. (2023) use baseline cohort structure and national transitions to study aging across U.S. states. Yet baseline exposure can proxy for historical fertility, industry, education, reunification, or prior migration. The shift-share literature therefore requires explicit exposure diagnostics, leave-out construction, and an account of which shocks carry identifying content (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022; Adão et al., 2019). Those diagnostics are not estimated in the current release. The Länder evidence below remains a set of conditional associations, and cohort-predicted aging remains a separately governed extension.

2.5.3 Adjustment can run through labor and technology

The mechanism literature also cautions against reading a demographic coefficient as a single productivity channel. Labor scarcity can reduce employment and entrepreneurship, but it can also induce capital deepening, automation, and organizational adjustment (Acemoglu and Restrepo, 2017, 2022; Hoshi et al., 2025). German robot exposure shows that technology shocks reallocate tasks and employment across workers and regions rather than operating as a frictionless productivity increment (Dauth et al., 2021). The Japanese discussion has likewise moved below the national aggregate and toward regional migration and adjustment (Hoshi, 2026).

This literature motivates the paper’s denominator discipline. GDP per resident aged 20–79 combines workplace output, realized workplace employment, commuting, and the size of the resident broad-adult population; GDP per employed person removes the final mixed margin. Divergence between them is therefore

informative, not an invitation to choose the more convenient outcome. The six accounting regressions observe output per hour, capital deepening, and a residual, but they do not observe automation, ICT adoption, or causal mediation. The stronger technology literature likewise requires direct adoption or numerator evidence rather than a convenient ratio (Acemoglu and Restrepo, 2022; Dauth et al., 2021).

Few-cluster inference imposes a related discipline. WCR procedures, influence diagnostics, and complete specification maps are useful precisely because 16 Länder do not support casual asymptotics (MacKinnon et al., 2023; Canay et al., 2021). Yet separate intervals still do not test a coefficient difference, and many correlated targets require a pre-specified familywise procedure (Romano and Wolf, 2005). Paper 0.8.1 therefore reports directly imposed studentized WCR- t tests from the Spec 040R child release and retains cross-sample differences as descriptive diagnostics.

Taken together, the literature makes the contribution narrower than a new regional-aging correlation and more demanding than a sign comparison. The paper asks whether the same demographic pattern survives a change in estimand, denominator, sample, weighting rule, and horizon. Germany holds many national institutions constant, but it does not resolve sorting or identify a technology channel by itself. The complete specification map is therefore evidence about where the association changes, not a menu from which one preferred coefficient is selected.

SECTION 3

Data and the German Regional Setting

The empirical core is a balanced annual panel of all 16 German federal states (Bundesländer). The unit of observation is a Land-year, not Germany as a whole and not a county. The panel runs from 1991 through 2024 and contains $16 \times 34 = 544$ observations. Berlin, Bremen, and Hamburg remain in the main sample. Because their city-state boundaries make workplace output especially sensitive to commuting, every result is also reported for the 13 non-city Länder.

3.1 Sources and Measurement

Three official source families supply the core outcome panel. Destatis GENESIS table 12411-0012 provides year-end resident population by Land and exact age. VGRdL provides nominal and chain-linked real GDP by place of work. ETR provides employed persons and actual hours, also by place of work. The accounting extension additionally uses the VGRdL gross-fixed-asset stock chain index and nominal manufacturing and total gross value added. Every object retains its literal statistical concept; in particular, a capital-stock index is not capital services.

3.2 Recovered Historical Inputs and Their Boundary

The evidence-recovery release adds 32 manufacturing-share controls: one same-vintage VGRdL ratio for each Land in 2000 and 2010. Each ratio retains the literal manufacturing numerator and total-GVA denominator source cells. The control is therefore fixed before its corresponding ten-year outcome window; no contemporaneous sector composition is inserted after the outcome is known.

Historical exact-age recovery is deliberately partial. Across 496 Land-years from 1981 through 2011, 451 contain complete single-age observations from 0 to 89, yielding 40,590 supported cells. The remaining 45 eastern Land-years in 1981–1989 contain 4,050 unsupported cells. They are recorded as unavailable, not converted from zeros and not backfilled from totals, births, or later age distributions. Consequently, the complete 1991–2024 core panel remains available, while an all-Länder exact-age design beginning in 1981 remains blocked. More historical rows do not remove a geographical hole.

The annual population series also crosses census and revision regimes. The 2011 census affects the population basis inside the main windows, while the 2022 census anchors a later retrospective Kreis bridge and the 2022–2024 Länder observations. The frozen annual and window products retain revision flags rather than smoothing

Table 2 Observed Länder data used in the empirical analysis

Object	Official source and concept	Construction	Coverage
Population	Destatis, resident population at year-end	Exact ages summed to 0–19, 20–49, 50–79, and 20–79	16 Länder, 1991–2024
Real GDP	VGRdL, workplace GDP	Official chain-volume index anchored to nominal GDP in 2020	16 Länder, 1991–2024
Employment	ETR, workplace employed persons	Official employment count; not resident workers and not hours	16 Länder, 1991–2024
Actual hours	ETR, workplace hours worked	Annual hours used to construct output per actual hour	Registered 2001–2021 windows
Capital stock	VGRdL, gross fixed assets	Chain index used as a stock proxy; not capital services	Registered 2001–2021 windows
Sector control	VGRdL, nominal gross value added	Manufacturing GVA divided by total GVA in 2000 and 2010	32 Land-baseline values

Source: Destatis GENESIS 12411-0012, VGRdL Länder results, and ETR Länder results; frozen source identities are recorded in the evidence-recovery release.

Sample: Balanced core panel of 544 Land-year observations plus the registered accounting windows and predetermined controls.

the breaks away. Paper 0.8.1 treats those flags as measurement boundaries. Spec 036 records the 2011 stable-vintage sensitivity as blocked because no alternative immutable vintage is admitted; the 2022 census change lies outside the registered 2001–2021 windows.

The demographic variable is the number of residents aged 50–79 divided by the number aged 20–49. A ratio of one means that the two groups are equally large. The empirical models use its natural logarithm, either as an annual level, as the average log ratio within a period, or as the realized change between two periods. This is an age-composition measure. It is not the usual old-age dependency ratio, which would use a retirement-age numerator and a different denominator.

3.3 Residence and Workplace Denominators

Both headline outcomes use the same chain-linked real-GDP numerator. They differ only in the denominator:

$$y_{rt}^N = \frac{\text{real workplace GDP}_{rt}}{\text{resident population aged 20–79}_{rt}}, \quad (3.1)$$

$$y_{rt}^E = \frac{\text{real workplace GDP}_{rt}}{\text{workplace employed persons}_{rt}}. \quad (3.2)$$

GDP per resident aged 20–79, y_{rt}^N , captures real workplace output relative to a broad resident population. It can change because output, employment, commuting, or the resident denominator changes. It is neither GDP per employed person nor a residence-based welfare measure.

Hayashi describes ages 20–79 as working age, but this is not an official German labor-force category. The band includes people above statutory retirement age and other residents who are not employed. Paper 0.8.1 therefore uses “resident aged 20–79” or “broad adult population” except when the inherited terminology itself is under discussion.¹

GDP per employed person, y_{rt}^E , divides the same output by realized workplace employment. It is a labor-productivity proxy, although it does not adjust for hours worked. Comparing the two outcomes reveals whether a regional pattern is tied to output per potential worker or remains after conditioning on the number of people actually employed.

¹This terminology change does not alter the frozen variable or any estimate. It prevents a demographic denominator from being mistaken for an observed labor-force count.

Demographic change by Land, 1991–2024

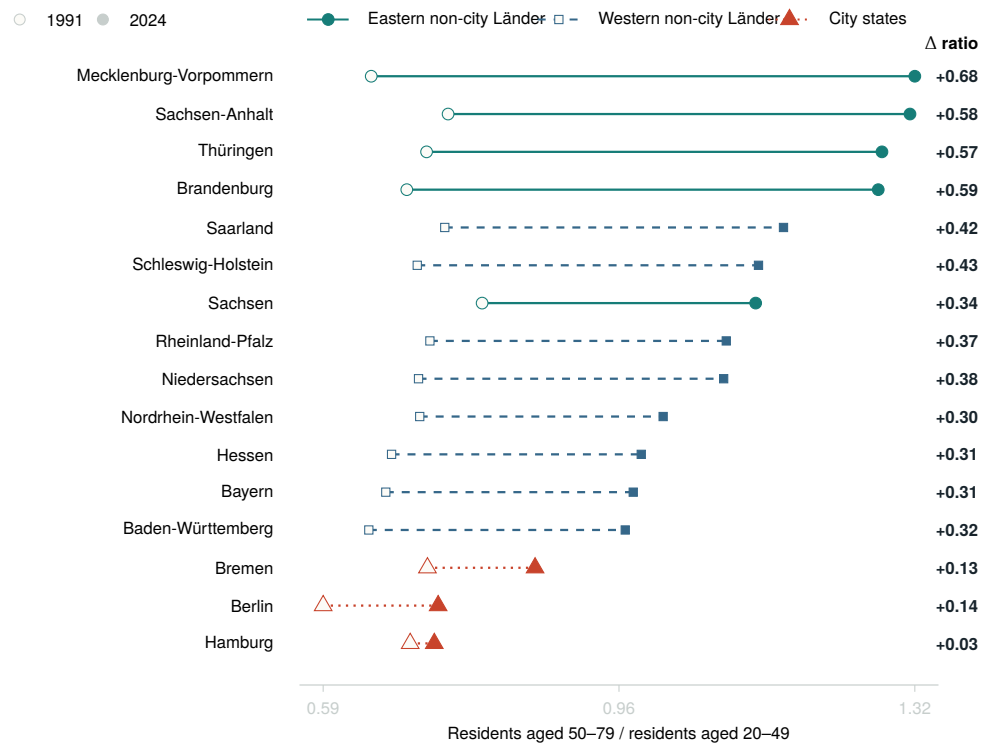


Figure 1 Regional demographic change. Each row connects a Land's observed 1991 and 2024 old-to-young ratios; rows are ordered by the 2024 endpoint and the right column reports the signed change.

Source: Frozen Länder pre-analysis annual panel.

Sample: All 16 Länder. Open and filled endpoints identify 1991 and 2024; shape and line style identify regional groups without relying on color.

Demographic change and long-run economic growth

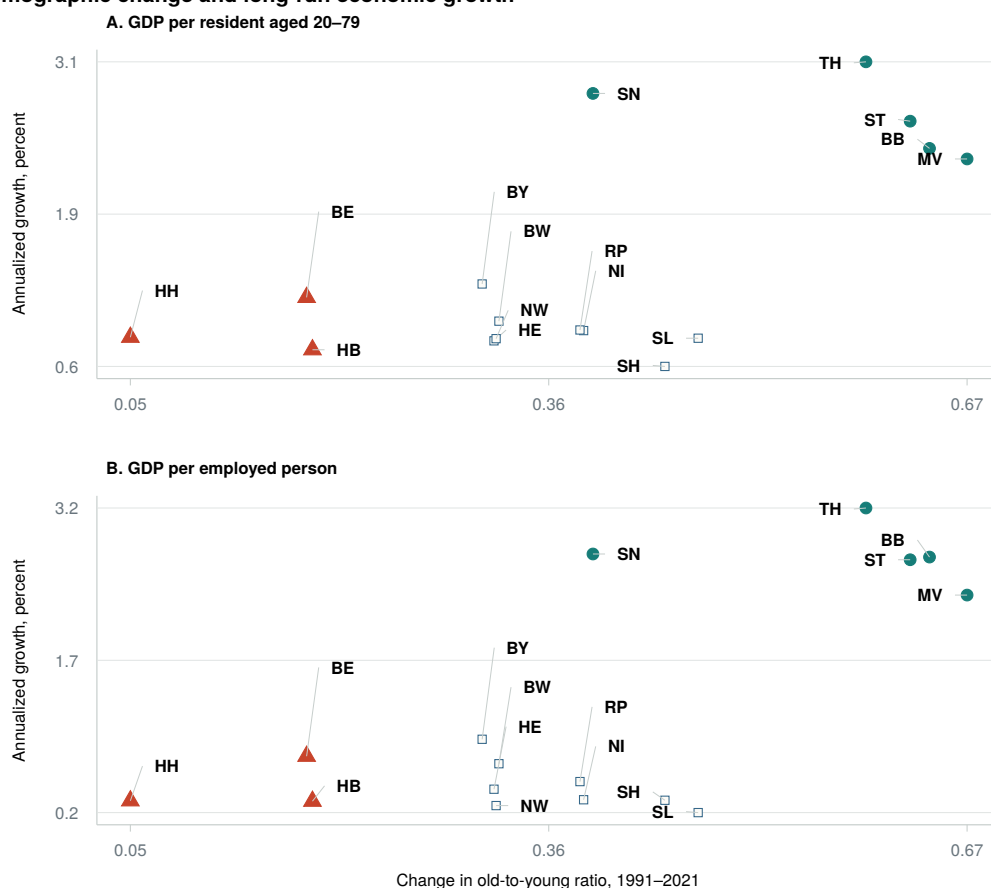


Figure 2 Observed matched-window changes in demography and economic growth. Each panel contains one 1991–2021 observation per Land on a shared demographic-change scale. Points and labels are observed values, not fitted relationships.

Source: Frozen Länder pre-analysis annual and long-window panels.

Sample: All 16 Länder. GDP is workplace based; the ages-20–79 denominator is residence based. Marker shape and direct labels preserve identity without relying on color.

Table 3 Basic descriptive statistics for the annual Länder panel

Variable	Unit	Mean	Std. dev.	Minimum	Maximum
Residents aged 50–79 / residents aged 20–49	ratio	0.879	0.192	0.587	1.373
Real GDP per resident aged 20–79	2020 euros	47,449	14,564	16,450	94,781
Real GDP per employed person	2020 euros	70,540	13,080	24,823	106,640

Source: Frozen Länder pre-analysis panel; unweighted calculations from registered annual values.

Sample: $N = 544$ Land-year observations, 16 Länder, 1991–2024. Minima and maxima pool regional and time variation.

3.4 What the Regional Profiles Show

The profiles reveal a distribution, not one representative German region. By 2024, the observed old-to-young ratio ranges from 0.73 in Hamburg to 1.32 in Mecklenburg-Vorpommern. The endpoints in [Figure 1](#) also show that this dispersion emerged along different paths rather than through one common demographic shift.

Among the five eastern non-city Länder, the mean increase in the old-to-young ratio is 0.59, compared with 0.37 among the eight western non-city Länder. Their observed 1991–2021 growth averages are 2.66 versus 0.93 percentage points for GDP per resident aged 20–79 and 2.71 versus 0.47 for GDP per employed person. These contrasts span post-unification convergence and large differences in initial economic levels. They are descriptive, not causal: faster growth in the east cannot be assigned to its simultaneous demographic change.

Among the city states, Hamburg records the highest 2024 GDP per resident aged 20–79, at 94.0 thousand euros; its GDP per employed person is 96.6 thousand euros. The comparison is especially sensitive to commuting because workplace output is paired with a resident denominator in the first measure. A high level therefore describes the location of production relative to residents, not a city-state demographic premium.

The two main scatter panels contain one aligned 1991–2021 observation for each of the 16 Länder in [Figure 2](#), while the separate appendix edition retains all 544 annual observations for each displayed series. The main panels expose the raw long-run association without allowing common time trends to masquerade as additional regional evidence. The appendix then shows how each Land reached its endpoints. Neither display isolates an effect of aging or fits a new model.

3.5 First Kreis Descriptive Evidence

The Kreis layer now crosses a descriptive, but not an estimative, boundary. Spec 021 consumes the frozen 3,960-row clean core for 396 stable Kreise and kreisfreie Städte from 2012 through 2021. It does not reopen the source workbooks, alter the December 2022 stable-region key, or construct a spatial matrix. The resulting profiles expose regional levels, changes, ranks, and distributions under the parent concepts.

The demographic shift is broad. The median old-to-young ratio rises from 0.966 in 2012 to 1.145 in 2021, corresponding to a median annualized log increase of 1.85%. The 2021 distribution is also wide: its 10th and 90th percentiles are 0.845 and 1.447. Heidelberg has the lowest endpoint ratio at 0.571, while Spree-Neiße has the highest at 1.804. These are observed age compositions on one stable analytical geography. They are not estimated effects of aging.

Economic scale grows more quickly than population in the median region. Median annualized population growth is 0.19%, compared with 2.79% for nominal workplace GDP and 0.60% for workplace employment. Consequently, the median current-price GDP ratio rises by 2.94% per year when divided by residents aged 20–79, by 2.23% when divided by employed persons, and by 2.73% when divided by standard hours. Inflation, commuting, and regional price differences remain inside these objects. Calling them real regional productivity would not make those components disappear.

The dispersion is substantively useful even before estimation. In 2021, current-price GDP per resident aged 20–79 ranges from about EUR 24,301 in Südwestpfalz to EUR 197,535 in Wolfsburg; GDP per employed person ranges from about EUR 57,110 in Erzgebirgskreis to EUR 167,883 in Mainz. Such extrema mix industrial concentration, commuting, urban functions, and prices. Spec 022 confirms that governed geometry and spatial structure are not yet available. The paper therefore draws no choropleth and does not interpret neighboring regions.

Table 4 Regional profiles: observed demography, output levels, and growth

Land	Group	O/Y 1991	O/Y 2024	Δ O/Y	GDP/A	GDP/E	Growth A	Growth E
Brandenburg	East	0.69	1.28	0.59	42.3	69.3	2.41	2.69
Mecklenburg-Vorpommern	East	0.65	1.32	0.68	41.8	65.0	2.32	2.32
Sachsen	East	0.78	1.13	0.34	45.8	65.1	2.85	2.72
Sachsen-Anhalt	East	0.74	1.32	0.58	39.7	63.5	2.63	2.66
Thüringen	East	0.72	1.28	0.57	41.3	63.4	3.11	3.16
Baden-Württemberg	West	0.64	0.96	0.32	66.2	85.7	1.01	0.69
Bayern	West	0.66	0.97	0.31	67.8	85.0	1.31	0.93
Hessen	West	0.67	0.98	0.31	68.0	87.8	0.85	0.44
Niedersachsen	West	0.71	1.09	0.38	53.3	74.4	0.93	0.34
Nordrhein-Westfalen	West	0.71	1.01	0.30	54.5	74.2	0.87	0.29
Rheinland-Pfalz	West	0.72	1.09	0.37	49.0	73.1	0.94	0.52
Saarland	West	0.74	1.16	0.42	45.8	67.0	0.87	0.22
Schleswig-Holstein	West	0.70	1.13	0.43	45.7	67.6	0.64	0.34
Berlin	City	0.59	0.73	0.14	64.8	81.5	1.20	0.76
Bremen	City	0.72	0.85	0.13	67.5	78.9	0.78	0.33
Hamburg	City	0.70	0.73	0.03	94.0	96.6	0.88	0.33

Source: Frozen Länder pre-analysis annual and long-window observations.

Sample: All 16 Länder. O/Y is residents aged 50–79 divided by residents aged 20–49; GDP levels are thousands of 2020 euros in 2024; growth is annualized 1991–2021 log growth in percentage points.

Table 5 Kreis descriptive evidence: levels and changes

Metric	2021 median	2021 p10–p90	Median annualized change
Residents aged 50-79 per resident aged 20-49	1.145	[0.845, 1.447]	1.85
Resident population	154,816	[72,378, 344,518]	0.19
Workplace GDP	5,805	[2,646, 16,396]	2.79
Workplace employment	76.1	[37.7, 200.5]	0.60
GDP per resident aged 20-79	50,293	[36,978, 84,483]	2.94
GDP per employed person	73,189	[64,489, 89,339]	2.23
GDP per standard hour	55.3	[47.1, 66.8]	2.73
Workplace employment per resident	0.494	[0.403, 0.764]	0.46

Source: Frozen Kreis descriptive release kreis-descriptive-evidence-de30cfa44327e03d; parent kreis-pilot-c55bb5b3b86db3f0.

Sample: 396 stable regions, 2012–2021. Levels retain current-price, workplace, residence, and standard-hours concepts. No regression is estimated.

3.6 What the Extension Audits Rule Out

The subsequent releases do not add a coefficient. They make the absence of one auditable. Each branch retains its literal object, failed gate, and reopening condition.

Table 6 Frozen evidence boundary after Specs 018–024

Branch	Literal object	Frozen decision
Patents (Spec 018)	Applicant-location, filing-year applications with fractional counting	No portable 2001–2021 panel under one route, status, method, and geography contract; no estimate.
Broadband (Spec 019)	Fixed-network household availability at a named threshold or technology	No 2001 product and a documented 2021–2022 collection break; no estimate.
Automation (Spec 020)	IFR installations by country/industry with predetermined Länder weights	Licensed inputs, detailed pre-window weights, classifications, aggregation, and leave-out rules are not jointly admitted; no exposure.
Kreis geography (Spec 022)	Functional memberships, reconciled polygons, contiguity/distance weights	Stable region identities pass; mapping, geometry, matrices, maps, and spatial inference remain blocked.
Kreis economics (Spec 023)	Residence income, workplace compensation/wages, sectors, commuting flows	Six intended outcome/control products remain blocked by source and geography evidence; no values.
Kreis reduced forms (Spec 024)	Change/level and extended economic designs with spatial sensitivities	Two parent gates pass and two fail; seven designs remain blocked and emit no coefficient.

Source: Frozen release manifests for Specs 018–020 and 022–024 in the replication package.

Sample: Unavailable does not mean zero patents, broadband, automation, income, spatial dependence, or effect.

A post-review source audit adds one correction without adding data. Official classification identifies ‘12611’ as marriage statistics, not births. The append-only source overlay instead registers ‘12612-01-01-4’ for Kreis live births and ‘12411-05-01-4’ for average annual population. It also registers ‘12711-04-02-4’ and ‘12711-02-02-4’ as distinct migration-flow families. Their published age bins cannot reproduce exact ages 20–49, and pooled execution must address the documented 2016 comparability break. These corrections sharpen the acquisition contract; they do not admit historical bytes or produce a cohort or migration estimate.

3.7 From Annual Levels to Growth Observations

The growth outcomes are annualized log changes between a window’s endpoints, as defined formally in [Equation \(4.1\)](#). Thus, 2.0 means two percentage points of annualized log growth (approximately 2% per year); it does not mean a two-percent level difference. The main adjacent-window data contain three decade comparisons for each Land: 1991–2001, 2001–2011, and 2011–2021. The matched estimand audit uses the latter two decades because they also have a preceding period from which realized aging can be calculated. Two additional comparisons use one observation per Land: the 1991–2021 long window and the 2009–2019 pre-pandemic window.

The 2011–2021 comparison spans the pandemic years and is retained in the main design; the 2009–2019 window shows the corresponding pre-pandemic boundary. The 2022–2024 observations enter the annual fixed-effects models but not the registered endpoint-growth windows, which stop in 2021.

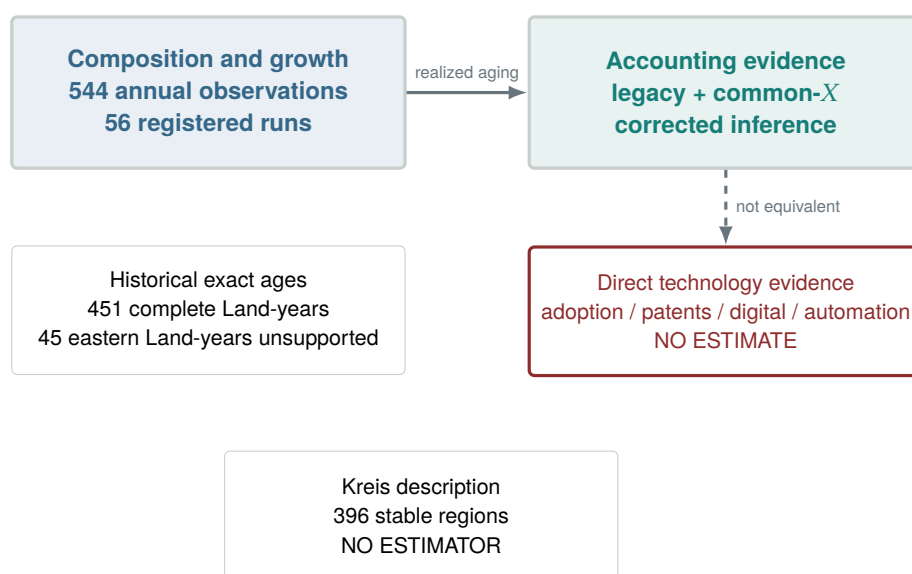
These observed growth rates must be separated from the estimates reported in [Section 5](#). A result such as –0.317 percentage points is not a Land’s growth rate. It is the estimated difference in annual growth associated with a 10% higher old-to-young ratio, after the model’s controls and rescaling. The annual fixed-effects rows use euros rather than percentage points because they relate demographic composition to annual GDP levels, not endpoint growth.

Table 7 Observed annualized real-GDP growth by Land-window

Registered window	Endpoints	N	GDP / ages 20–79	GDP / employed
First adjacent decade	1991–2001	16	2.10 [0.16, 5.63]	2.48 [−0.08, 7.16]
Second adjacent decade	2001–2011	16	1.20 [0.43, 1.90]	0.66 [−0.12, 1.28]
Third adjacent decade	2011–2021	16	1.12 [0.04, 1.88]	0.38 [−0.35, 1.10]
Long comparison	1991–2021	16	1.47 [0.64, 3.11]	1.17 [0.22, 3.16]
Pre-pandemic	2009–2019	16	1.96 [1.09, 2.94]	0.91 [0.03, 1.69]

Source: Frozen Länder pre-analysis windows; unweighted means with minima and maxima in brackets.

Sample: Values are observed annualized growth rates in percent, not regression coefficients.

**Figure 3** Evidence architecture after the recovery release

Source: Frozen Länder release chain through `laender-evidence-recovery-a1e5a165266ffd3b`.

Sample: Composition, accounting, technology, historical-age, and Kreis claims remain separate evidence classes. Arrows denote design progression, not causality.

Table 8 Measurement contracts by regional layer

Dimension	Länder	Kreise / kreisfreie Städte
Demography	Exact single ages	Exact ages with census and territorial controls
Output	Nominal and chain-linked real	Current-price workplace output
Labor input	Employed persons and actual hours	Employed persons and standard hours
Capital	Official investment and capital-stock series	No official capital stock
Primary claim	Real growth and bounded accounting associations	Description; reduced-form association blocked
Prohibited inference	Direct technology adoption, causality, or large-sample certainty with 16 clusters	Official real output or TFP

Source: GENESIS, Regionalstatistik, VGRdL, and the frozen source-release registry.

Sample: Länder contracts are populated through 2024 for the real-output core and through the registered accounting windows; the Kreis descriptive contract is populated and the reduced-form contract is blocked.

3.8 Boundary of the Current Evidence

The current evidence layer is therefore narrower than the full research architecture. The Kreis product supports frozen descriptive profiles and distributions, but Specs 022–024 establish that no regression or spatial claim is currently eligible. A national deflator can remove common inflation from current-price output, but it cannot recover relative regional prices.

SECTION 4

Empirical Design and Identification Ladder

4.1 Growth and Timing

For any positive outcome k_{rt} , annualized endpoint growth is

$$g_{r,t,H}^k = \frac{100}{H} [\log k_{r,t+H} - \log k_{rt}] . \quad (4.1)$$

The period-average composition $\bar{a}_{r,t,H}$ is defined in Equation (2.2). Realized aging compares adjacent equal-length period averages as defined in Equation (2.3). It is not the endpoint change $a_{r,t+H} - a_{rt}$, which remains a separate registered sensitivity without a frozen estimate.

The canonical adjacent-window equation for outcome k is

$$g_{r,w}^k = \alpha^k + \beta^k \bar{a}_{r,w} + \gamma^k \log k_{r,w,0} + \delta^k \bar{c}_{r,w} + \lambda_w^k + u_{r,w}^k, \quad (4.2)$$

where $\log k_{r,w,0}$ is the start-of-window outcome level and $\bar{c}_{r,w}$ is the period-average log child-dependency ratio. The primary sample contains 48 observations: 16 Länder across 1991–2001, 2001–2011, and 2011–2021. Window effects absorb common decade shocks.

The matched transition audit uses the latter two windows and replaces $\bar{a}_{r,w}$ with $\Delta \bar{a}_{r,w}$:

$$g_{r,w}^k = \bar{\alpha}^k + \bar{\beta}^k \Delta \bar{a}_{r,w} + \bar{\gamma}^k \log k_{r,w,0} + \bar{\delta}^k \bar{c}_{r,w} + \bar{\lambda}_w^k + \bar{u}_{r,w}^k. \quad (4.3)$$

The 32 observations are identical to the corresponding matched composition runs. Sample alignment therefore removes one source of disagreement, but the two treatments remain different empirical objects.

The annual fixed-effects audit estimates levels rather than growth:

$$y_{rt}^k = \mu_r^k + \tau_t^k + \beta_{FE}^k a_{rt} + \delta_{FE}^k c_{rt} + \varepsilon_{rt}^k, \quad (4.4)$$

where y_{rt}^k is measured in 2020 euros, c_{rt} is the annual log child-dependency ratio, and Land and year effects use all 544 observations. This is a within-Land level association, not the endpoint-growth equation.

4.2 Weights, Scaling, and Inference

Unweighted estimation is primary. The secondary window weight is resident population aged 20–79 at the start of the outcome window; the secondary annual weight is the contemporaneous population aged 20–79. Weights are normalized to mean one within each run. Every model is also repeated after jointly excluding Berlin, Bremen, and Hamburg.

Treatment coefficients are estimated per one log point. Main growth tables and figures multiply coefficients and intervals by $\log(1.1)$, so the displayed unit is percentage points of annual growth associated with a 10% higher old-to-young ratio. Annual fixed-effects displays use the corresponding euro level difference. All Länder inference clusters by Land and uses restricted, null-imposed WCR11 bootstrap- t inference with Rademacher weights, 9,999 draws, seed 20260813, and equal-tailed percentile- t intervals. Leave-one-Land and observation-level influence diagnostics are reported for every run (MacKinnon et al., 2023; Canay et al., 2021).

4.3 Accounting Associations

For accounting outcome m and Land-window w , the frozen design is

$$m_{rw} = \alpha_m + \beta_m \Delta \bar{a}_{rw} + \rho_m m_{r,w-1} + \theta_m s_{r,w-1} + \lambda_{m,w} + u_{m,rw}, \quad (4.5)$$

where $m_{r,w-1}$ is the outcome-specific pre-window level and $s_{r,w-1}$ is the strictly predetermined manufacturing share. The outcomes are annualized growth in real output per actual hour, the gross-fixed-asset stock chain index per actual hour, and the growth-accounting residual defined by Equation (2.7). The pooled/full variant contains both 2001–2011 and 2011–2021 and includes a period effect. The early 2001–2011 window contains 16 observations and no period effect. Both are unweighted and use the same WCR11 policy.

Because $m_{r,w-1}$ differs across outcomes, these six regressions do not form the common- X system in Equation (2.6). They locate conditional covariation in accounting outcomes; they do not identify adaptation, causal mediation, compensating innovation, direct adoption, or an identified aging effect.

4.4 Common- X Accounting and Joint Inference

For each common- X accounting outcome M , paper 0.8.1 uses the annualized ten-year change

$$g_{r,w}^M = \frac{100}{H} [\log M_{r,w,H} - \log M_{r,w,0}], \quad H = 10. \quad (4.6)$$

Displayed effects multiply the resulting coefficient by $\log(1.1)$. Spec 035 estimates a separate seven-outcome system on the two matched 2001–2011 and 2011–2021 windows. Every equation uses the same 32 Land-window observations, adjacent-average aging treatment, initial manufacturing share, period effect, parameter order, Land clusters, and restricted WCR11 draws. The primary system is unweighted; population weighting and exclusion of the three city states are pre-specified secondary systems. This design removes outcome-specific lagged controls so that the coefficient restrictions in Equations (2.6) and (2.8) hold exactly. All 64 observation-level identities and all six coefficient identities reconcile within 10^{-10} .

Spec 040R corrects the inference layer without changing those parent products. Each component or same-system linear contrast is represented as a transformed outcome and tested under its own null with studentized restricted WCR- t draws. Symmetric pointwise intervals invert the pointwise test. Single-step max- t p-values invert the displayed simultaneous intervals. Romano–Wolf stepdown p-values are reported separately because they do not invert the single-step intervals. Minimum detectable effects are simulation-calibrated from the same max- t draws at 80% power. Sample and window variants report literal coefficients as descriptive concentration diagnostics; no corrected cross-sample difference inference is claimed. Stable-vintage census and annual dynamics remain blocked where required inputs do not exist.

4.5 Identification Ladder

Migration is both confounder and mechanism and is not inserted automatically as a contemporaneous control. Forecast-vintage momentum, timing-specific cohort execution, direct technology outcomes, the resident/workplace split, and Kreis reduced forms remain separate gates. A sign difference or one interval excluding zero is not a covariance-aware difference test; Specs 035–036 provide those direct tests only for the admitted common- X system.

FROZEN All 56 core variants and six outcome-specific accounting runs remain unchanged. Spec 040R freezes 14 direct-null WCR- t tests, 12 family rows, 12 simulation-calibrated MDES rows, 54 descriptive sensitivities, and 144 standardized influence rows.

Table 9 Demographic objects, evidence states, and claim ceilings

Object	Treatment	Current state	Interpretation
Period-average composition	$\bar{a}_{r,t,H}$	Frozen estimates	Conditional older-composition association
Baseline composition	a_{rt} at window start	Registered; not estimated	Predetermined regional state
Adjacent-average transition	$\Delta \bar{a}_{r,t,H}$	Frozen estimates	Realized transition correlation
Endpoint transition	$a_{r,t+H} - a_{rt}$	Registered; not estimated	Separate sensitivity
Cohort transition	$Z_{r,t,h}(\widehat{\tau}_r)$	Timing gates pending	Reduced form only after gates pass
Common-X components	Exact identities in Equation (2.4)	Frozen in Specs 035–036	Accounting location of an association

Source: Spec 033 estimand and accounting overlays; frozen Specs 004, 005, 015, 035, and 036.

Sample: Länder and Kreise remain separate regional layers. Registered but unestimated objects carry no numerical result.

SECTION 5

Composition, Transition, and Denominators

The German evidence begins with a result that resembles Hayashi’s international benchmark. It does not end there. The primary comparison relates annualized growth in GDP per resident aged 20–79 to the period-average log old-to-young ratio across three adjacent decades and all 16 Länder. The model contains 48 Land-window observations, initial log output, period-average child dependency, and window effects. Länder are the clusters; the primary estimate is unweighted and uses null-imposed WCR11 bootstrap- t inference with 9,999 draws. Population weighting and removal of the three city states are registered secondary views rather than selection rules. Each model estimates a common conditional slope, not 16 Land-specific effects.

The answer to the title question depends on what “aging” means. Average age structure produces a stable negative association. Faster realized aging does not reproduce it with comparable precision, and annual within-Land estimates change sign when the GDP denominator changes. The complete grid is therefore an ordered audit: first the closest German version of Hayashi’s level equation, then changes in weighting, demographic object, denominator, horizon, and accounting outcome. None of those changes is a cosmetic robustness check. A sign comparison is not a covariance-aware difference test.

5.1 The German Benchmark and the Hayashi Scale

For GDP per resident aged 20–79, all 4 adjacent-window point estimates are negative. In the primary all-Länder unweighted run, a 10% higher old-to-young ratio corresponds to 0.317 percentage points lower annual growth, with a 95% WCR11 interval of $[-0.565, -0.068]$. Across all registered variants, the estimates range from -0.573 to -0.317 percentage points; 3 intervals lie below zero and 1 spans it.

Hayashi’s column (C) estimate for OECD post-transitioners, the Asian Tigers, and China is -4.17 on the log old-to-young ratio. Multiplication by $\log(1.1)$ places it on the paper’s reporting scale: -0.397 percentage points for a 10% higher ratio. The German population-weighted estimate is -0.376, while the primary unweighted estimate is -0.317. The proximity is informative because the demographic ratio and GDP denominator are aligned. It is not a replication identity: Hayashi compares countries over 1990–2020, whereas the German estimate pools three Land-level decades, adds period effects and child dependency, and relies on 16 clusters.

For GDP per employed person, all 4 adjacent-window point estimates are also negative. The primary all-Länder unweighted estimate is -0.367 percentage points with interval $[-0.616, -0.116]$; the registered range is -0.585 to -0.367. Again, 3 intervals lie below zero and 1 spans it.

Table 10 Hayashi and the primary German level benchmark on a common scale

Estimate	Unit of comparison	Effect of 10% higher ratio	Statistical contract
Hayashi (2026), column (C)	37 mature economies, 1990–2020	−0.397 pp	OLS slope −4.17 (SE 0.68)
Germany, all Länder, un-weighted	48 Land-decades, 1991–2021	−0.317 pp	WCR11 95% CI [−0.565, −0.068]
Germany, all Länder, population-weighted	48 Land-decades, 1991–2021	−0.376 pp	WCR11 95% CI [−0.672, −0.084]

Source: Hayashi (2026), Table 1, column (C); frozen Länder target estimates in `laender-results-e03dd8e13726e7b4`.

Sample: Point estimates are multiplied by $\log(1.1)$. Different samples, controls, and inference prevent a pooled equality test.

Thus, the closest German counterpart to Hayashi’s level equation yields a clear negative association under both denominators. That is the benchmark result. Its limitation is equally clear: it compares periods with different average age structures, conditional on initial output and child dependency. It does not isolate the adjustment produced by a demographic change.

5.2 Same Sample, Different Demographic Object

On the matched resident-aged-20–79 sample, all 8 canonical and adjacent-average transition point estimates are negative, between −0.202 and −0.015 percentage points. However, all 8 WCR11 intervals span zero.

The matched employed-person audit does not preserve even the common sign: 4 point estimates are negative and 4 are positive, across a range from −0.178 to 0.121 percentage points. All 8 intervals span zero.

Restricting the composition and adjacent-average transition estimands to the same two decades removes sample composition as an explanation. For the working-age-adult outcome, all point estimates remain negative but every interval spans zero. For the employed-person outcome, the realized-change estimates are positive while the matched level estimates are negative. An older average age structure and a faster change in age structure are therefore not two labels for the same empirical object. Spec 036 retains the direct period-average versus adjacent-average contrast as blocked because the two targets do not share an admitted joint estimator. The separate intervals do not establish that the coefficients differ.

5.3 Within-Land Evidence and the Denominator

Within Länder over time, the four resident-aged-20–79 level estimates are negative. On the 10% demographic scale they range from −1220.583 to −431.240 euros, and every interval spans zero.

For GDP per employed person, the four corresponding fixed-effects estimates are positive, between 458.576 and 1091.384 euros, while every interval again spans zero.

Land and year effects remove persistent Land levels and common annual shocks, but they do not resolve reverse migration or other time-varying confounding. The result is imprecise under both outcomes and changes sign with the denominator. GDP per resident aged 20–79 combines workplace output, employment, commuting, and a resident demographic denominator; GDP per employed person conditions on realized workplace employment. Their opposite point-estimate signs show that the within-Land result is not a denominator-invariant productivity relationship. They do not by themselves prove a statistically resolved denominator difference. In the later common- X system, $Q/N - Q/H = H/N$ is an exact accounting identity rather than a separate denominator-mechanism test. The corrected contrasts in Section 6 therefore test substantive transformed outcomes instead of relabelling that identity.

5.4 Horizon and Pandemic Boundaries

All four long-period resident-aged-20–79 estimates are negative, from −0.785 to −0.263 percentage points. Two intervals lie below zero and two span it.

All four long-period employed-person estimates are negative, between -0.532 and -0.250 percentage points, and all four intervals lie below zero.

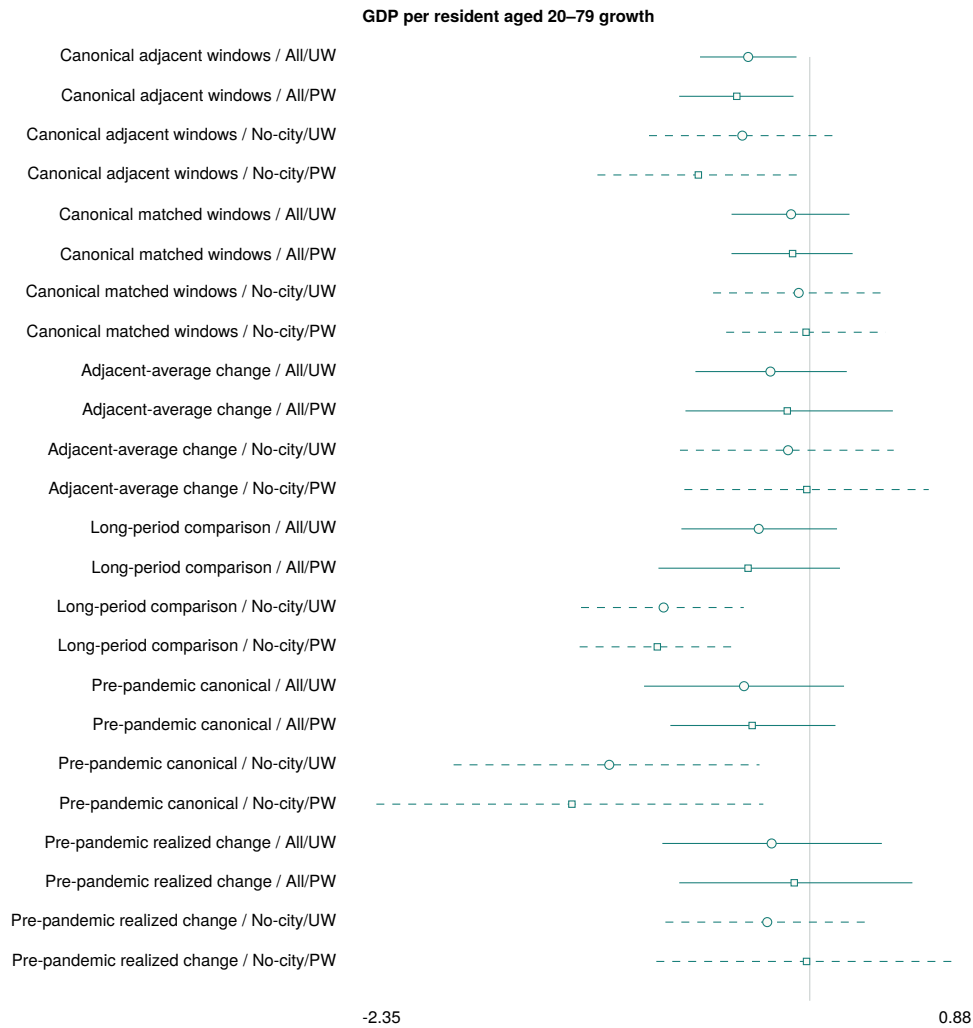
The long comparison resembles Hayashi's cross-sectional horizon most closely. It also compresses each Land into one observation. The uniformly negative point estimates are economically legible, especially for GDP per employed person, but persistent regional differences remain inseparable from average demographic composition.

Before the pandemic endpoint, all eight resident-aged-20–79 composition and adjacent-average transition estimates remain negative, ranging from -1.224 to -0.017 percentage points. Two intervals lie below zero and six span it.

The corresponding employed-person estimates split evenly: 4 are negative and 4 are positive, with a range from -0.684 to 0.155 percentage points. One interval lies below zero and seven span it.

Removing the 2020–2021 endpoint does not erase the negative resident-denominator pattern, but it does not reconcile the employed-person change estimates either. The pandemic is a relevant boundary condition. It is not a general explanation for the difference between composition and change.

Taken together, the composition benchmark survives changes in weights and the city-state sample and resembles Hayashi's coefficient in scale. The transition and annual evidence are less stable. That is evidence for distinguishing the objects, while the unavailable treatment contrast prevents a formal equality claim. The next section asks where the realized-transition association appears under both outcome-specific and common- X accounting designs.

**Figure 4** Länder specification map: GDP per resident aged 20–79 growth

Source: Spec 005 frozen target estimates and WCR11 intervals.

Sample: All 24 registered panel runs. Circles are unweighted; squares are population-weighted; dashed intervals exclude city states.

Table 11 Länder estimates: resident ages 20–79 denominator

Design	Unit	All/UW	All/PW	No-city/UW	No-city/PW
Canonical adjacent windows	pp	-0.317 [-0.565, -0.068]	-0.376 [-0.672, -0.084]	-0.347 [-0.827, 0.141]	-0.573 [-1.093, -0.062]
Canonical matched windows	pp	-0.096 [-0.402, 0.204]	-0.088 [-0.403, 0.221]	-0.057 [-0.498, 0.380]	-0.019 [-0.430, 0.391]
Adjacent-average change	pp	-0.202 [-0.588, 0.190]	-0.116 [-0.640, 0.428]	-0.112 [-0.668, 0.432]	-0.015 [-0.645, 0.612]
Annual Land and year fixed effects	EUR	-974 [-2364, 407]	-1221 [-2767, 329]	-431 [-1740, 876]	-792 [-3195, 1569]
Long-period comparison	pp	-0.263 [-0.661, 0.140]	-0.318 [-0.779, 0.155]	-0.752 [-1.177, -0.340]	-0.785 [-1.184, -0.396]
Pre-pandemic canonical	pp	-0.338 [-0.853, 0.176]	-0.296 [-0.717, 0.132]	-1.032 [-1.833, -0.259]	-1.224 [-2.230, -0.239]
Pre-pandemic realized change	pp	-0.196 [-0.759, 0.371]	-0.080 [-0.672, 0.528]	-0.219 [-0.743, 0.315]	-0.017 [-0.790, 0.760]

Source: Spec 005 frozen target estimates; WCR11 intervals.

Sample: All registered sample and weight variants. Values scale a 10% higher old-to-young ratio.

Table 12 Länder estimates: employed-person denominator

Design	Unit	All/UW	All/PW	No-city/UW	No-city/PW
Canonical adjacent windows	pp	-0.367 [-0.616, -0.116]	-0.425 [-0.717, -0.135]	-0.402 [-0.820, 0.012]	-0.585 [-1.020, -0.148]
Canonical matched windows	pp	-0.042 [-0.129, 0.044]	-0.060 [-0.147, 0.029]	-0.145 [-0.343, 0.047]	-0.178 [-0.417, 0.059]
Adjacent-average change	pp	0.041 [-0.063, 0.146]	0.121 [-0.065, 0.309]	0.006 [-0.156, 0.161]	0.080 [-0.197, 0.354]
Annual Land and year fixed effects	EUR	534 [-539, 1629]	459 [-1016, 1932]	1091 [-816, 3006]	977 [-2943, 4857]
Long-period comparison	pp	-0.250 [-0.499, -0.001]	-0.292 [-0.559, -0.021]	-0.504 [-0.902, -0.104]	-0.532 [-0.796, -0.274]
Pre-pandemic canonical	pp	-0.051 [-0.213, 0.111]	-0.147 [-0.432, 0.139]	-0.526 [-1.074, 0.000]	-0.684 [-1.319, -0.054]
Pre-pandemic realized change	pp	0.116 [-0.022, 0.257]	0.155 [-0.005, 0.318]	0.000 [-0.081, 0.082]	0.094 [-0.144, 0.332]

Source: Spec 005 frozen target estimates; WCR11 intervals.

Sample: All registered sample and weight variants. Values scale a 10% higher old-to-young ratio.



SECTION 6

How Regions Adjust: What the Accounts Show

The accounting evidence asks where the association with adjacent-average realized aging appears in the available workplace accounts. It does not observe the technology, migration, participation, or organizational response that may produce those accounts. The first layer retains six outcome-specific regressions from paper 0.7.0. The second applies one common- X design and shared covariance to seven exact-identity components.

6.1 Outcome-Specific Accounting Regressions

Figure 6 reports the original six targets on the same $\log(1.1)$ scale as the growth results: percentage points of annual growth associated with a 10% higher old-to-young ratio. Each regression conditions on its own pre-window outcome and the initial manufacturing share.

6.2 Output per Hour

Output per actual hour remains close to zero. The pooled/full estimate is -0.019 percentage points with a 95% WCR11 interval of $[-0.188, 0.148]$. The early 2001–2011 window estimate is -0.010 with interval $[-0.127, 0.105]$. Both intervals span zero. Thus, the negative period-composition benchmark does not reappear as evidence that faster adjacent-average aging reduced output per hour in these accounting equations.

This comparison is informative but not decisive. The composition and accounting treatments differ, and these equations use outcome-specific baseline controls. They therefore remain separate from the common- X system below.

6.3 Capital Deepening Is a Ratio

Capital deepening is positive in both point estimates, but only the pooled/full-window interval excludes zero. Across 2001–2011 and 2011–2021, the estimate is 0.498 percentage points with interval $[0.141, 0.870]$ and $p = 0.014$. In the early 2001–2011 window, the estimate is 0.407 with $[-0.160, 0.977]$ and $p = 0.139$. The intervals are wide, overlap materially, and do not constitute a covariance-aware difference test.

The early-window estimate is also highly sensitive to Hamburg. On the unscaled log-point coefficient, deleting Hamburg moves the estimate from 4.268 to 14.698, and Hamburg's Cook distance is 4.30. More fundamentally, the capital object is a gross-fixed-asset stock chain index, not capital services. Equation (2.8) adds a second boundary: $K/\mathcal{H}$ can rise because K rises, $\mathcal{H}$ falls, or both. The common- X system estimates those terms separately, but the result remains a conditional capital-deepening association rather than evidence that aging produced investment or induced automation.

6.4 The Residual Does Not Name the Mechanism

The growth-accounting residual does not preserve its sign. The pooled/full estimate is -0.105 with interval $[-0.321, 0.114]$, whereas the early 2001–2011 estimate is 0.169 with $[-0.282, 0.624]$. Both intervals span zero. The residual may contain utilization, labor quality, allocation, markups, sector composition, and measurement error. Calling it total factor productivity would make the label more ambitious, not the evidence more informative.

6.5 The Common- X Decomposition

Spec 035 now applies the same observations, regressors, effects, weights, and bootstrap draws to all seven components. The system changes the question. It does not ask whether separately controlled equations happen to have different signs; it asks where one shared conditional association is located while enforcing the accounting identities.

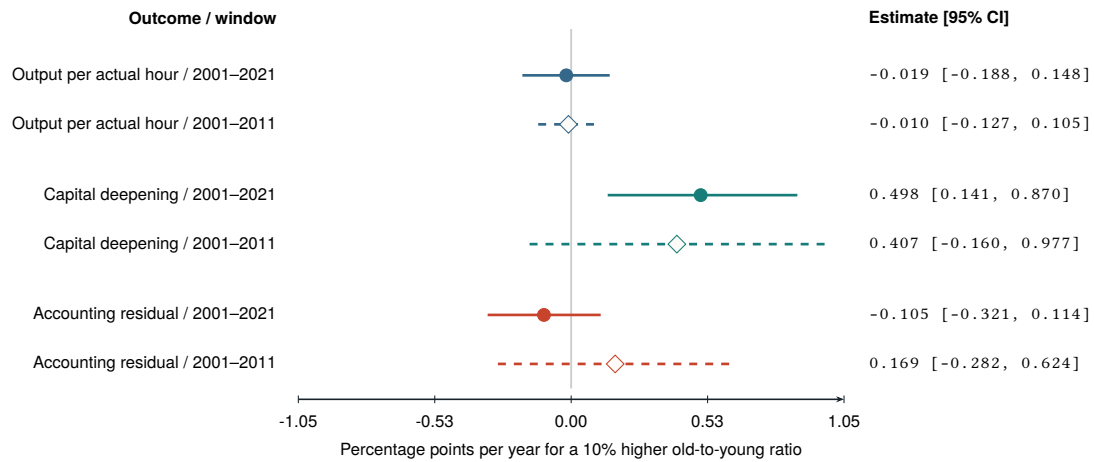


Figure 6 Conditional accounting associations with regional aging. Points are OLS estimates scaled by $\ln(1.10)$; horizontal lines are equal-tailed 95% WCR11 intervals. Filled circles with solid intervals use both ten-year windows. Open diamonds with dashed intervals use 2001–2011 only.

Source: Frozen evidence-recovery release laender-evidence-recovery-a1e5a165266ffd3b; content digest

a1e5a165266ffd3b9bab8c955ebc220e93731a552ddb37b7a835ed341e1f4cc3.

Sample: Six accounting-only targets for 16 Länder. The pooled/full variant has 32 Land-window observations; the early 2001–2011 variant has 16. Estimates condition on baseline outcome, initial manufacturing share, and period effects where applicable. These are accounting associations, not direct technology-adoption, innovation, automation, efficiency, mediation, or causal estimates.

Table 13 Corrected Common-X accounting associations with regional aging

Outcome	Estimate	Pointwise CI	Pointwise p	Simultaneous CI	Single-step p	RW stepdown p
GDP / resident ages 20–79	0.289	[0.007, 0.572]	0.045	[-0.125, 0.704]	0.172	0.161
GDP / actual hour	0.388	[0.136, 0.640]	0.016	[0.010, 0.766]	0.045	0.045
Actual hours / workplace-employed person	-0.064	[-0.158, 0.031]	0.133	[-0.181, 0.054]	0.337	0.290
Workplace employment / resident ages 20–79	-0.035	[-0.142, 0.072]	0.519	[-0.263, 0.193]	0.991	0.519
Capital-stock index / actual hour	0.490	[0.152, 0.829]	0.011	[-0.153, 1.134]	0.133	0.128
Capital-stock index / resident ages 20–79	0.392	[0.036, 0.747]	0.028	[-0.299, 1.082]	0.302	0.290
Actual hours / resident ages 20–79	-0.099	[-0.238, 0.041]	0.157	[-0.276, 0.079]	0.316	0.290

Source: Frozen correctness release paper-081-correctness-e58951758cd9de21.

Sample: Primary unweighted common-X system: 32 Land-window observations and 16 Land clusters. Estimates and intervals are annual log percentage points for a 10% larger adjacent-average old-to-young transition. Pointwise WCR-t and Single-step max-t p -values invert their displayed intervals; RW stepdown p -values are a separate multiplicity result.

For a 10% larger adjacent-average old-to-young transition, the primary unweighted association is 0.289 annual log percentage points for output per resident aged 20–79 and 0.388 for output per hour. Hours per workplace-employed person contribute -0.064 points and workplace employment per resident contributes -0.035 . Under the single-step component-family test, only output per hour excludes zero: its simultaneous interval is $[0.010, 0.766]$ and $p = 0.045$. Capital per hour is 0.490 points, but its single-step interval $[-0.153, 1.134]$ spans zero and its single-step p -value is 0.133.

The arithmetic is exact rather than rhetorical. Output per resident equals output per hour plus hours per employed person plus workplace employment per resident at both the observation and coefficient levels. Capital per hour equals capital per resident minus hours per resident. Reconciliation errors are below 10^{-10} . The decomposition locates a conditional association; it does not identify why the capital stock or hours changed.

6.6 The Capital Numerator and Hours Denominator

Table 14 Primitive capital-stock and actual-hours associations

Target	Estimate	Pointwise 95% CI	Pointwise p
Capital-stock index growth	0.027	$[-0.333, 0.386]$	0.860
Actual-hours growth	-0.464	$[-0.759, -0.169]$	0.014
Capital growth plus hours growth	-0.437	$[-0.982, 0.108]$	0.110

Source: Frozen correctness release paper-081-correctness-e58951758cd9de21.

Sample: Annual log percentage points for a 10% larger aging transition. The capital object is a gross-fixed-asset stock chain index. These rows are accounting associations, not investment or automation estimates.

The primitive estimates narrow the ratio's interpretation. Capital-stock index growth is 0.027 points with pointwise interval $[-0.333, 0.386]$ and $p = 0.860$. Actual-hours growth is -0.464 points with interval $[-0.759, -0.169]$ and $p = 0.014$. Their signed contribution contrast is -0.437 points and is not resolved under either family procedure. Thus, the positive capital-per-hour association is accounted for mainly by an actual-hours decline, not measured capital-stock growth. It is not evidence of investment, capital services, or automation.

6.7 Direct Tests, Power, and Sensitivity

The corrected contrast family excludes four tests that merely renamed an accounting component. Under the single-step gate, output per hour minus the workplace-employment-per-resident margin is resolved at 0.423 points, with simultaneous interval $[0.099, 0.748]$ and $p = 0.016$. Output per hour minus hours per worker has stepdown $p = 0.042$ but single-step $p = 0.058$ and a simultaneous interval that crosses zero. The procedures answer different multiplicity questions and are not interchangeable. Simulation-calibrated MDES rows retain the same 16-cluster family and make imprecision visible without asserting equality.

The sensitivity coefficients reveal geographic concentration. Capital per hour is 0.490 points in the full sample but -0.181 after deleting the five eastern non-city Länder and -0.191 in the western-Länder sample. Output per resident falls from 0.289 to 0.027 and 0.016, respectively. These literal coefficients are descriptive concentration diagnostics, not corrected cross-sample difference tests. The positive full-sample capital and output transition associations are concentrated in eastern Germany and are consistent with convergence, restructuring, or other East-specific forces. Stable-vintage census and annual variants remain blocked rather than approximated.

Spec 037 then tests whether the final E^W/N^R term can be split. Only the workplace lineage is admitted. Resident employment, bilateral commuting, dated legal geography, and functional-region memberships remain blocked, and the release emits zero numerical rows. The term therefore remains a mixed utilization-and-commuting margin.

Direct technology adoption, patent innovation, digital infrastructure, and automation exposure also remain unestimated because their required historical panels do not pass the source and comparability gates. Accounting can show where a change appears. It does not identify causal mediation, and it does not turn a residual into a technology measure.

Table 15 Direct-null tests of substantive accounting contrasts

Contrast	Estimate	Pointwise CI	Pointwise p	Simultaneous CI	Single-step p	RW stepdown p
Output/hour minus hours/worker	0.452	[0.125, 0.779]	0.019	[-0.015, 0.919]	0.058	0.042
Output/hour minus workplace employment/resident	0.423	[0.191, 0.655]	0.006	[0.099, 0.748]	0.016	0.016
Hours/worker minus workplace employment/resident	-0.029	[-0.167, 0.110]	0.695	[-0.331, 0.274]	1.000	0.781
Capital growth plus hours growth	-0.437	[-0.982, 0.108]	0.110	[-1.240, 0.365]	0.440	0.320
Capital/hour minus output/hour	0.102	[-0.535, 0.740]	0.588	[-0.700, 0.904]	0.998	0.781

Source: Frozen correctness release paper-081-correctness-e58951758cd9de21.

Sample: Each contrast is estimated as a transformed outcome under the same 32-observation system and bootstrapped under its own null. Units are annual log percentage points for a 10% larger aging transition. Accounting relabelings and cross-sample difference tests are excluded.

Table 16 Common- X coefficient concentration by sample and window

Variant	Q/N	Q/H	K/H	K/N	K	H
All Lander	0.289	0.388	0.490	0.392	0.027	-0.464
No city states	0.460	0.486	0.672	0.646	0.084	-0.588
Eastern group deleted	0.027	0.162	-0.181	-0.316	-0.460	-0.279
Western Lander only	0.016	0.109	-0.191	-0.284	-0.428	-0.237
2001–2011	0.269	0.515	0.386	0.140	-0.106	-0.492
2011–2021	0.306	0.326	0.521	0.502	0.080	-0.441

Source: Frozen correctness release paper-081-correctness-e58951758cd9de21.

Sample: Annual log percentage points for a 10% larger aging transition. These are descriptive concentration diagnostics. Different samples are not linked by corrected difference-test inference. K is the capital-stock index and H is aggregate actual hours.

SECTION 7

Threats, Power, and Decisive Tests

7.1 Sorting and Post-Unification Convergence

The main composition regressor is contemporaneous with the outcome window. Young adults can move toward productive regions, older residents can remain in declining regions, and workplace output can grow through incoming commuters. The 1991–2001 window adds a distinct historical transition: reunification, capital reallocation, and East–West convergence are not ordinary decade effects. Initial output, child dependency, period effects, and manufacturing shares absorb some observable structure, but they do not make regional age composition exogenous.

For that reason, the identification ladder moves toward baseline composition and timing-specific cohort-transition exposure. The exposure must be evaluated timing by timing, use leave-focal-region-out composite transition factors, and pass concentration, balance, placebo, and weak-identification gates. Even a strong first stage would not establish an exclusion restriction by itself. Reduced forms come first.

7.2 Sixteen Clusters and Many Correlated Questions

Sixteen Länder provide institutional comparability but limited independent regional variation. The frozen analysis therefore uses WCR11 inference, leave-one-Land refits, and observation-level influence diagnostics. The appendix reports compact named extrema for every run; the complete products contain 812 deletion estimates and 5,104 influence records. They show where a small sample is sensitive. They do not authorize deleting an inconvenient Land or promoting another model.

The same limitation applies to cross-model interpretation. A negative estimate in one equation and a positive estimate in another do not establish heterogeneity. Nor does one interval excluding zero while another includes it. Spec 040R therefore tests same-system coefficients and transformed outcomes with direct-null studentized WCR- t . Pointwise, single-step, and stepdown procedures remain distinct; cross-sample differences are descriptive. Output per hour survives the single-step component-family gate. Capital per hour does not. The direct tests narrow the conclusion rather than decorate the sign pattern.

7.3 Census, Price, and Location Concepts

The population data cross the 2011 census and later Census 2022 revisions. The Länder outcomes use chain-linked real workplace GDP, whereas the Kreis layer uses current-price workplace output. Workplace employment and hours do not become resident labor utilization merely because the denominator is a resident population. These boundaries are retained in the variable and source contracts. A stable-vintage 2011 sensitivity, resident employment, bilateral commuting, and functional labor markets remain blocked extensions rather than unreported assumptions. The 2022 census change lies outside the registered 2001–2021 windows.

7.4 What Would Change the Interpretation

Three results would materially change the interpretation. First, eligible cohort-transition estimates would show whether predetermined demographic pressure predicts outcomes differently from contemporaneous composition. Second, resident employment and bilateral commuting would split the mixed E^W/N^R margin into utilization and workplace concentration. Finally, direct technology, investment-flow, or organizational evidence would test the mechanism suggested by the capital-stock association. Corrected common-X accounting and joint inference now answer their narrower questions; they do not supply these missing objects.

SECTION 8

Conclusion: Composition Is Not Adjustment

The German benchmark is clear within its limits. Regions and periods with an older average age composition exhibit slower contemporaneous growth in real workplace GDP per resident aged 20–79 during the same windows, and the primary coefficient resembles Hayashi’s cross-country estimate in scale. The pattern also appears when GDP is divided by workplace employment. If the question concerns conditional average composition, the answer is a qualified negative association.

The stronger claim does not follow. Under a separate realized-transition design, output per resident and output per hour are positive, and output per hour survives the corrected single-step family gate. Capital per hour is positive in the full sample but not under that family gate. Primitive capital-stock index growth is near zero, while the actual-hours decline is large; the capital ratio also reverses outside eastern Germany. This is an East-concentrated accounting pattern. The apparent capital deepening is driven mainly by the hours denominator, not a nationwide aging-induced capital response. Capital remains a stock index, and an accounting location does not name a causal mechanism.

The contribution of paper 0.8.1 is therefore a distinction with sharper empirical consequences. Composition describes a regional equilibrium state. Transition describes how that state changes. Adjustment describes the margins through which output, employment, hours, capital, migration, and commuting respond. The current evidence identifies conditional composition and accounting associations, an East-concentrated transition pattern, and a bounded Kreis description. Cohort reduced forms and the residence–workplace bridge remain the decisive next evidence. A regional pattern is not a demographic law. Even after the accounts reconcile, its mechanism still has to earn its name.

References

- Acemoglu, D., Autor, D., Beirne, K., and Scott, A. (2026). Baby Busts and Growth Booms: Demographic Change and the Macroeconomy. Technical Report 35401, National Bureau of Economic Research.
- Acemoglu, D. and Restrepo, P. (2017). Secular Stagnation? The Effect of Aging on Economic Growth in the Age of Automation. *American Economic Review*, 107(5):174–179.
- Acemoglu, D. and Restrepo, P. (2022). Demographics and Automation. *Review of Economic Studies*, 89(1):1–44.
- Adão, R., Kolesár, M., and Morales, E. (2019). Shift-Share Designs: Theory and Inference. *Quarterly Journal of Economics*, 134(4):1949–2010.
- Bode, E., Dohse, D., and Stolzenburg, U. (2023). Aging and Regional Productivity Growth in Germany. *Review of Regional Research*, 43:409–432.
- Bönte, W., Falck, O., and Heblich, S. (2009). The Impact of Regional Age Structure on Entrepreneurship. *Economic Geography*, 85(3):269–287.
- Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-Experimental Shift-Share Research Designs. *Review of Economic Studies*, 89(1):181–213.
- Brunow, S. and Hirte, G. (2006). Age Structure and Regional Economic Growth. *Jahrbuch für Regionalwissenschaft*, 26(1):3–23.
- Brunow, S. and Hirte, G. (2009). The Age Pattern of Human Capital and Regional Productivity: A Spatial Econometric Study on German Regions. *Papers in Regional Science*, 88(4):799–823.
- Canay, I. A., Santos, A., and Shaikh, A. M. (2021). The Wild Bootstrap with a “Small” Number of “Large” Clusters. *Review of Economics and Statistics*, 103(2):346–363.
- Daniele, F., Honiden, T., and Lembcke, A. C. (2019). Aging and Productivity Growth in OECD Regions: Combatting the Economic Impact of Aging through Productivity Growth? Technical Report 2019/08, OECD.
- Dauth, W., Findeisen, S., Suedekum, J., and Woessner, N. (2021). The Adjustment of Labor Markets to Robots. *Journal of the European Economic Association*, 19(6):3104–3153.
- Feyrer, J. (2007). Demographics and Productivity. *Review of Economics and Statistics*, 89(1):100–109.
- Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8):2586–2624.
- Gregory, T. and Patuelli, R. (2015). Demographic Ageing and the Polarization of Regions—An Exploratory Space–Time Analysis. *Environment and Planning A*, 47(5):1192–1210.
- Hayashi, F. (2026). The Root Cause of Japan’s Reversal of Fortune: Implications for China, the U.S., and Other Mature Economies. Technical Report 4553699, Social Science Research Network. First draft September 2023; August 2026 revision.
- Hoshi, T. (2026). Demographic Challenges and Economic Stagnation in Japan. *The Japanese Economic Review*, 77:411–433.
- Hoshi, T., Kodama, N., and Li, H. (2025). The Effect of Aging on Entrepreneurship and Aggregate Productivity. *The Journal of the Economics of Ageing*, 32:100605.
- Lindh, T. and Malmberg, B. (1999). Age Structure Effects and Growth in the OECD, 1950–1990. *Journal of Population Economics*, 12(3):431–449.

- MacKinnon, J. G., Nielsen, M. Ø., and Webb, M. D. (2023). Cluster-Robust Inference: A Guide to Empirical Practice. *Journal of Econometrics*, 232(2):272–299.
- Maestas, N., Mullen, K. J., and Powell, D. (2023). The Effect of Population Aging on Economic Growth, the Labor Force, and Productivity. *American Economic Journal: Macroeconomics*, 15(2):306–332.
- Mitze, T. and Reinkowski, J. (2011). Testing the Neoclassical Migration Model: Overall and Age-Group Specific Results for German Regions. *Journal for Labour Market Research*, 43(4):277–297.
- Monte, F., Redding, S. J., and Rossi-Hansberg, E. (2018). Commuting, Migration, and Local Employment Elasticities. *American Economic Review*, 108(12):3855–3890.
- Romano, J. P. and Wolf, M. (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica*, 73(4):1237–1282.