TI 2026-056/III
Tinbergen Institute Discussion Paper

When Is the Mover-Design Event Study Coefficient a Place-Effect Share?

Vahid Moghani[1]

1 Erasmus University Rotterdam and Tinbergen Institute

When Is the Mover-Design Event Study Coefficient a Place-Effect Share?

Vahid Moghani*

August 11, 2026

Abstract

Mover-design event studies are a leading approach to separating place effects from sorting. The coefficient is often interpreted as the share of cross-sectional variation in location means due to places. I show the coefficient depends on how movers connect locations: two economies can be identical in place effects, sorting, and location means, yet deliver different mover coefficients. Only under *directional isotropy*, a testable condition, does the coefficient permit a cross-sectional reading: it averages the variance share of place effects and the share that equalizing them removes. An illustration in Dutch employer–employee wage data rejects the condition.

*Erasmus School of Economics, Tinbergen Institute, and Erasmus Center for Health Economics Rotterdam. I am especially grateful to Patrick Kline for insightful comments and observations. I also thank Hans van Kippersluis, Owen O'Donnell, Sonia Bhalotra, Olivier Marie, Jonathan Roth, Peter Hull, Jonathan Rossi, Lisa Voois, and Joaquim Vidiella-Martin for their helpful comments. I gratefully acknowledge financial support from the Erasmus Initiative Smarter Choices for Better Health. I declare that there are no conflicts of interest.

I Introduction

Outcomes differ greatly across firms, regions, hospitals, and schools, but separating genuine place effects from sorting of people into places is difficult. The mover-design event study is a leading approach to this problem (Finkelstein et al., 2016). By tracking individuals who change locations, this design relates the change in a mover's outcome to the difference in average outcomes between their destination and origin. The resulting coefficient captures how much of the gap between locations passes to the individual upon arrival. This estimate is often interpreted as the contribution of place effects to the variation in mean outcomes across locations. The design is widely used in health, public economics, political economy, and labor economics.[1]

This paper provides the first formal link between the mover-design coefficient and cross-sectional variance components. I study the design in the standard Abowd–Kramarz–Margolis (AKM) setting (Abowd et al., 1999)—the structured and tractable data-generating process commonly assumed in mover-design event studies (Finkelstein et al., 2016), in which outcomes are additively separable in individual and place effects—with large locations and a stationary cross-section, as in most applications,[2] while remaining agnostic about the underlying sorting and mobility process. The main result is that the mover-design coefficient is a transition-weighted projection: its moments are determined by where movers come from and where they go, not by the cross-sectional distribution of places. Two economies can share the same AKM decomposition—the same place-effect dispersion, individual-effect dispersion, sorting, and location-mean dispersion—and still produce different mover coefficients. The cross-section describes the map. The mover coefficient depends on the particular routes taken through it.

In a mover-design event study, the regressor is the change in location-level mean outcomes. A location mean combines the place effect with *composition*: the average individual effect of the people who sort into that location. Composition has two components. One is correlated with the place effect, as when higher-paying firms employ better workers. The other, a residual, is orthogonal to the place effect, as when two firms with the same pay premium hire different types of workers because they differ in amenities, industry, hours, or career structure. How much a typical move changes each part depends on who moves where, and need not mirror how the two parts vary across locations. In employer–employee settings, a worker who climbs a job ladder and moves to a higher-paying firm with similar amenities mainly moves along the place-effect dimension. A worker who switches careers and moves to a similar-paying firm with very different coworkers

[1] Applications include studies of geographic variation in health care (Finkelstein et al., 2016; Godøy and Huitfeldt, 2020; Salm and Wübker, 2020; Zeltzer et al., 2021; Luan et al., 2025), physician and organizational behavior (Molitor, 2018; Huang and Ullrich, 2024; Doyle and Staiger, 2025; Agha et al., 2023; Encinosa and Dor, 2025; Badinski et al., 2023), health behaviors and outcomes (Hinnosaar and Liu, 2022; Hut, 2020; Finkelstein et al., 2025; Ding, 2026; Chyn and Shenhav, 2025), insurance and program choice (Cabral et al., 2025), employer and research institution effects (Ahammer et al., 2024; Chandra and Xu, 2025), political behavior (Cantoni and Pons, 2022), and carbon emissions (Lyubich, 2025).

[2] Throughout, the *cross-section* is the location-level snapshot the researcher observes: each location's place effect, mean individual effect, and hence mean outcome. *Stationarity* requires location sizes, mover shares, and mean mover composition to be stable across periods. Substantial individual churning is compatible with this condition when count and average individual inflows and outflows balance.

mainly moves along the residual-composition dimension. These moves contribute differently to the mover coefficient, even if the cross-sectional distribution of firms is unchanged. In simulations, I construct economies with identical place effects, sorting, location means, and essentially identical pooled AKM decompositions, and vary only the realized moves. The mover-design event study coefficient varies substantially.

I show that the mover-design event-study coefficient has a cross-sectional interpretation only if observed moves are representative of cross-sectional variation in place effects and composition. In particular, the changes in place effects and composition generated by moves—and their covariance—must reflect the corresponding variation in levels in the same proportions.[3] If mobility disproportionately reflects one source of variation or particular combinations of the two—if moves disproportionately climb the job ladder but rarely cross between career types, for example—the mover coefficient will not recover the cross-sectional variation share. Call this requirement *directional isotropy*: the mover-design coefficient recovers the cross-sectional variation share only when mobility is directionally isotropic. This requirement is distinct from the familiar concern that movers are selected individuals. It can fail even when movers are randomly selected because it restricts the routes through which movers travel, not who travels through them. Directional isotropy is therefore a restriction on mobility patterns, and its necessity is not an artifact of the AKM model. Allowing for richer heterogeneity or nonseparability does not, by itself, remove the need for such a restriction.

Under directional isotropy, the coefficient equals the cross-sectional regression of place effects on location means, and this object has a sharp interpretation. It is the average of two shares of the cross-sectional variance of location means—the share attributable to place effects, and the share that would disappear if place effects were equalized. Under isotropy, two limiting configurations recover familiar readings. At one extreme, if composition is uncorrelated with place effects, the coefficient is the variance share of place effects. This case corresponds to sorting on factors unrelated to the outcome under study: in a setting like that of Finkelstein et al. (2016), individuals who settled where they do for family, retirement, or housing—rather than for health care—would generate composition orthogonal to the place effect. At the other extreme, if composition is *affine*—linear and residual-free in the place effect—the coefficient magnitude is the standard-deviation share. This case corresponds to matching on the outcome dimension itself, as when workers value jobs only through pay and a location's average worker type rises linearly, with no residual component, in its premium. Under such affine sorting, a coefficient of 0.20—often described as "places explain twenty percent of the variation"—corresponds to a variance contribution of four percent.[4] The affine reading is special in a second way: it is robust to *any* mobility pattern.

Whether mobility satisfies isotropy is an empirical question. The condition restricts how cross-sectional and mobility moments relate in the data. I provide a test for isotropy—a cross-fitted

[3] Formally, the covariance matrix of changes in place effects and composition across moves must be a scalar multiple of the covariance matrix of their levels in the mover-weighted cross-section.

[4] E.g., "[l]ocation explains 37 percent of the cross-state variation in turnout" (Cantoni and Pons, 2022, abstract); similarly, Molitor (2018) reads a mover coefficient of 0.6–0.8 as environment factors explaining "between 60–80 percent of regional disparities" in physician behavior.

diagnostic with an individual-level bootstrap—and illustrate it on the mobility of one-time movers across large firms in Dutch matched employer–employee data. Sorting on the firm wage premium is statistically undetectable in the cross-section and negative along moves; movers, however, traverse the residual-composition dimension about sixty percent more intensively than its cross-sectional dispersion implies, and isotropy is rejected. The mover-design coefficient therefore cannot be mapped to a share of the variation of location means.[5]

The mover-design event study originates with Bronnenberg et al. (2012), where the share of the origin–destination gap that closes upon moving is a structural constant; the jump then measures a cross-market share by construction, and the interpretive question is settled by the model. Finkelstein et al. (2016) bring the design to the two-way fixed-effects setting, and it is there that the cross-sectional reading becomes a focal point. They read the jump as a weighted average of pair-level place-effect shares and bridge to the cross-sectional variation share through the observed linearity of movers' outcome changes in the origin–destination gap. I show that this bridge is not generic—it holds only under directional isotropy—and in its place I derive an exact mapping from the coefficient to the cross-sectional variance components.[6] Hull (2018) asks what the mover regression recovers *causally*, showing it aggregates difference-in-differences comparisons with convex weights under a parallel-trends assumption. Card et al. (2025) confront a similar contamination as the one in this paper—a move across regions also changes the firm, so firm sorting biases the regional (unscaled) movers design—which they resolve by estimating a model with worker and firm effects and aggregating to the commuting-zone level.[7] This paper instead asks what the mover-design event study coefficient estimates, and the characterization is complete under the AKM structure and stationarity: the estimand depends only on the cross-sectional sorting geometry and the undirected mobility graph, and every admissible combination of the two is generated by some assignment process in the AKM class. Whether the cross-sectional reading is valid is therefore a property of the reduced form alone—no individual-location assignment model is needed to test for it.[8]

[5]Throughout, "cross-sectional" moments are weighted by mover incidence: the cross-section taken over the locations movers actually visit. Section IV makes the weighting explicit.

[6]For example, if the place-effect and composition changes are jointly elliptically distributed across moves, the average outcome change is exactly linear in the origin–destination gap whatever the sorting geometry: the geometry fixes the slope of the mover plot. A straight plot is thus consistent with every geometry, and cannot on its own separate a valid cross-sectional reading from an invalid one.

[7]In labor economics, mover designs often serve a different role. In their unscaled, co-worker-quartile form, they assess exogenous mobility and additive separability (Card et al., 2013, 2016), while variance shares come from the AKM decomposition (Abowd et al., 1999).

[8]The object of interest in this paper is also distinct from limited-mobility bias (Andrews et al., 2008; Kline et al., 2020; Bonhomme et al., 2023)—noise in estimated place effects when movers are sparse—which vanishes in the large-location regime of this paper.

II Model and Setup

Consider a balanced panel of individuals $i = 1, \dots, N$ observed in two periods, $t \in \{1, 2\}$[9], each residing in a location $j \in \mathcal{J}_0$, where $J_0 := |\mathcal{J}_0|$; $j(i,t)$ denotes i's location at t. I study an asymptotic regime with fixed J_0 and large location sizes $n_j \to \infty$—the benchmark regime for mover-design event studies—which separates the mobility mechanisms studied here from limited-mobility bias.

Outcomes follow the standard AKM,

$$Y_{it} = \alpha_i + \psi_{j(i,t)} + \varepsilon_{it}, \qquad t \in \{1, 2\}. \tag{1}$$

Equivalently, stacking observations gives

$$\mathbf{Y} = \mathsf{U}_N \boldsymbol{\alpha} + \mathsf{Z}_N \boldsymbol{\psi} + \boldsymbol{\varepsilon},$$

where U_N and Z_N are the individual- and location-incidence matrices. Conditional on $\mathcal{F}_N := \sigma(\boldsymbol{\alpha}, \mathsf{Z}_N, \boldsymbol{\psi})$, this is the usual fixed-design AKM representation.

The individual and location effects are time invariant, and idiosyncratic shocks satisfy the exogenous-mobility restriction $\mathbb{E}[\varepsilon_{it} \mid \mathcal{F}_N] = 0$. This rules out mobility driven by realized or anticipated idiosyncratic shocks but places no restriction on the joint distribution of $(\alpha_i, \psi_{j(i,t)})$. Individuals may sort across locations without restriction. Error vectors are independent across individuals with bounded second moments; within an individual, ε_{i1} and ε_{i2} may be arbitrarily correlated. For population statements and inference, $\boldsymbol{\psi}$ and $\mathcal{J}_0$ are held fixed while complete worker histories—including α_i, assignments, and shocks—are i.i.d. draws; thus α_i remains a fixed effect in the structural sense although worker types and paths vary across repeated samples. The AKM model serves two purposes. First, it is a structured and tractable data-generating process assumed to be underlying mover-design event studies (Finkelstein et al., 2016). Second, I show that even under such an additively separable data-generating process, mover-design event studies do not generally admit a cross-sectional interpretation and require additional assumptions to establish that link. The need for additional assumptions to justify a cross-sectional interpretation should therefore not be viewed as an artifact of this specification.[10]

Write $o(i) := j(i,1)$ and $d(i) := j(i,2)$ for origin and destination, $M := \{i : o(i) \neq d(i)\}$ for movers, S for stayers, and $\Delta\psi_i := \psi_{d(i)} - \psi_{o(i)}$. Movers and stayers share the outcome process (1) but may differ in their effect distributions and sorting.

Let $\pi_{j,t} := \Pr\{j(i,t) = j\}$ and $w_{j,t} := \Pr\{i \in M \mid j(i,t) = j\}$ denote the population location mass and mover share.

[9] In Supplemental Appendix S1, I show that, under plausible assumptions, the canonical mover-design event study of Finkelstein et al. (2016) is asymptotically equivalent to a two-period first-difference regression in which the dependent variable is the change in the mover's outcome and the independent variable is the change in the leave-one-out average outcome.

[10] When the AKM specification includes time-varying controls, every object is defined on the control-adjusted outcome $Y_{it} - X_{it}'\gamma$, and all recovered shares are shares of *control-adjusted* variation; see Supplemental Appendix S1.

Assumption 1 (Mean stationarity)**.** Location masses and mover shares are constant across periods, $\pi_{j,1} = \pi_{j,2} \equiv \pi_j$ and $w_{j,1} = w_{j,2} \equiv w_j$, and, for each location j with positive mover mass,

$$\mathbb{E}[\alpha_i \mid o(i) = j,\ i \in M] = \mathbb{E}[\alpha_i \mid d(i) = j,\ i \in M].$$

Assumption 1 is an assumption purely on the mobility patterns—not on the AKM components. It assumes locations are in a demographic steady: each keeps its size, its mover share, and its mean composition stable across periods—arriving and departing movers have the same average individual effect. It places no restriction on the object the design exploits, who moves where: origins and destinations may pair arbitrarily, and movers with different α_i may take different routes, provided each location's population inflows and outflows balance in mass and mean composition. This is a population restriction; realized counts and mover mean compositions may differ by sampling noise. Supplemental Appendix S2 characterizes these restrictions on mover flows.

Assumption 1 is a premise, not a necessity. Everything that follows—the estimand, the recovery characterization, and the isotropy test—survives without it, read against the average of the two periods' cross-sections, with the location means in the regressor averaged over the window; the one exception is the reduction of mobility to an undirected flow (Supplemental Appendix S8). Mean-stationarity is also the *least* favorable case for the thesis here: relaxing it only enlarges the admissible flows—drifting configurations now qualify—so the spread of coefficients consistent with a given cross-section can only grow. The assumption is nonetheless the natural benchmark for the settings where mover designs are applied. Realized mobility can be thought of as a circulation that preserves the cross-section plus a residual that carries its drift, and the circulation dominates. A location's composition moves only through the net of offsetting inflows and outflows—second-order where gross flows dwarf net flows. For example, employee–employer wage-data cross-sections drift over decadal horizons, workers churn annually (cf. Card et al., 2013). Supplemental Appendix S8 records, separately, the centered directed-flow formulas for a common full-window benchmark when this assumption is relaxed.[11]

III The Mover-Design Estimand

Among movers, the researcher regresses the individual outcome change on the leave-one-out change in location mean outcomes:

$$\Delta Y_i = c + \theta\, \delta_i^{(-i)} + u_i, \qquad i \in M, \qquad \delta_i^{(-i)} := \overline{Y}^{(-i)}_{d(i),2} - \overline{Y}^{(-i)}_{o(i),1}, \tag{2}$$

where $\overline{Y}^{(-i)}_{j,t}$ is the leave-one-out mean outcome in cell (j, t). Let $\bar{\alpha}_j$ denote location j's composition: the mean individual effect of its occupants, time-invariant under Assumption 1. Under (1), $\Delta Y_i = \Delta\psi_i + \Delta\varepsilon_i$, and exogenous mobility makes $\Delta\varepsilon_i$ orthogonal to the regressor. The regressor,

[11] In a dynamic system of locations, an intertemporally defined object is the only coherent way to measure the cross-sectional variation shares.

however, is not the change in location effects; in the large-location limit,

$$\delta_i^{(-i)} = \Delta\psi_i + \Delta\bar{\alpha}_i + o_p(1), \qquad \Delta\bar{\alpha}_i := \bar{\alpha}_{d(i)} - \bar{\alpha}_{o(i)}.$$

The dependent variable changes only because the mover changes location effects; the regressor changes because the mover changes both location effects and location composition. Write $\Delta\bar{Y}_i := \Delta\psi_i + \Delta\bar{\alpha}_i$ and assume $\mathrm{Var}_M(\Delta\bar{Y}_i) > 0$. The large-location estimand is therefore

$$\theta_\infty = \frac{\mathrm{Var}_M(\Delta\psi_i) + \mathrm{Cov}_M\big(\Delta\psi_i, \Delta\bar{\alpha}_i\big)}{\mathrm{Var}_M\big(\Delta\psi_i + \Delta\bar{\alpha}_i\big)}, \tag{3}$$

a *transition* object: $\mathrm{Var}_M(\Delta\psi_i)$ depends on how far movers travel in the location-effect distribution, and can be small even when location effects are highly dispersed in the cross-section, if movers mostly move between similar locations; the covariance and the composition variance depend on whether movements in location effects and in composition reinforce or offset each other along realized routes.

Projecting the composition change on the location-effect change,

$$\Delta\bar{\alpha}_i = \beta_\Delta \Delta\psi_i + R_i, \qquad \beta_\Delta := \frac{\mathrm{Cov}_M(\Delta\bar{\alpha}_i, \Delta\psi_i)}{\mathrm{Var}_M(\Delta\psi_i)}, \qquad \lambda_\Delta := \frac{\mathrm{Var}_M(R_i)}{\mathrm{Var}_M(\Delta\psi_i)}, \tag{4}$$

and substituting into (3) gives

$$\boxed{\theta_\infty = f(\beta_\Delta, \lambda_\Delta), \qquad f(\beta, \lambda) := \frac{1+\beta}{(1+\beta)^2 + \lambda}.} \tag{5}$$

The slope β_Δ measures the component of composition change that moves with location-effect changes; λ_Δ measures the composition change orthogonal to $\Delta\psi_i$, relative to the dispersion of location-effect changes. The pair $(\beta_\Delta, \lambda_\Delta)$ is the realized *mobility anatomy*: pooled AKM variance components do not pin it down.

A transition-network representation. The transition moments have a compact representation on the mobility network. Let $F_{jk} := \Pr\{o(i) = j, d(i) = k \mid i \in M\}$, with $\sum_{j \neq k} F_{jk} = 1$, and $\bar{F}_{jk} := \frac{1}{2}(F_{jk} + F_{kj})$ the symmetrized flow. For any location-level f, g, the mass balance implied by Assumption 1 gives $\mathbb{E}_M[\Delta f_i] = 0$, so

$$\mathrm{Cov}_M(\Delta f_i, \Delta g_i) = \sum_{j<k} 2\bar{F}_{jk}\,(f_j - f_k)(g_j - g_k) =: \mathcal{L}(f, g), \tag{6}$$

a weighted-graph Laplacian quadratic form whose edges are the moves movers make and whose weights count how often they make them; θ_∞ is a ratio of such forms in $(\psi, \bar{\alpha})$.

Proposition 1 (Undirected flow). *θ_∞ depends on the mobility process only through the symmetrized flow*

$\bar{F}$. *Two processes with the same* $\bar{F}$ *but different directed flows—hence different origin-dependent destination rules and different history dependence—induce the identical estimand.*

Proof. In (6) the summand is symmetric in (j, k), so only $\bar{F}$ enters; the antisymmetric part of F cancels. □

Remark 1 (The reduced form is without loss of generality). Proposition 1 reduces mobility to the symmetrized flow $\bar{F}$; the reduction is also exhaustive. Lemma S1 in Supplemental Appendix S4 shows that for *every* admissible cross-section $(\psi_j, \bar{\alpha}_j)$ and every symmetric flow consistent with location sizes and mover shares, there is an assignment process in the AKM class that induces it while satisfying exogenous mobility and Assumption 1. The pair (cross-section, undirected flow) is therefore a complete and unrestricted description of the environments the theory covers.

IV When the Mover Design Recovers the Cross-Section

Cross-sectional moments are taken over the locations movers visit: weight each location j by its mover incidence μ_j, the share of mover endpoints located at j. This mover-incidence cross-section is the one the design can speak to.[12] For location-level f, g, write $\mathrm{Cov}_M(f, g)$ for the μ-weighted covariance, and define the cross-sectional projection of composition on the location effect,

$$\bar{\alpha}_j = a_M + \beta_M \psi_j + q_j, \qquad \beta_M = \frac{\mathrm{Cov}_M(\bar{\alpha}_j, \psi_j)}{\mathrm{Var}_M(\psi_j)}, \qquad \lambda_M := \frac{\mathrm{Var}_M(q_j)}{\mathrm{Var}_M(\psi_j)}, \tag{7}$$

the cross-sectional analogue of the transition pair (4): β_M is the sorting slope and λ_M is the normalized dispersion of residual composition. With $\bar{Y}_j := \psi_j + \bar{\alpha}_j$ and $\mathrm{Var}_M(\bar{Y}_j) > 0$, define three cross-sectional benchmarks,

$$\theta_{\text{var}} := \frac{\mathrm{Var}_M(\psi_j)}{\mathrm{Var}_M(\bar{Y}_j)}, \qquad \theta_{\text{sd}} := \sqrt{\theta_{\text{var}}}, \qquad \theta_{\text{iso}} := \frac{\mathrm{Cov}_M(\psi_j, \bar{Y}_j)}{\mathrm{Var}_M(\bar{Y}_j)} = f(\beta_M, \lambda_M). \tag{8}$$

The first two are the variance and standard-deviation shares; θ_{iso} is the coefficient from regressing the location effect on the location mean across locations.

Theorem 1 (Recovery of the cross-section)**.** *Maintain the AKM model, exogenous mobility, the large-location limit, and Assumption 1, with* $\mathrm{Var}_M(\Delta\psi_i) > 0$*,* $\mathrm{Var}_M(\psi_j) > 0$*,* $\mathrm{Var}_M(\Delta\bar{Y}_i) > 0$*, and* $\mathrm{Var}_M(\bar{Y}_j) > 0$*. The following are equivalent.*

(i) ***(Recovery.)*** *The mover design reproduces the cross-sectional sorting projection:* $\beta_\Delta = \beta_M$ *and* $\lambda_\Delta = \lambda_M$.

[12] If the movers sample is not representative of the economy, this does not need to coincide with the person- or match-weighted cross-section an applied researcher might be interested in. The wedge between μ-weighted and individual-weighted benchmarks is separate from the isotropy question studied here.

(ii) ***(Directional isotropy.)*** *There is a scalar $\kappa > 0$ with*

$$\mathrm{Var}_M(\Delta\psi_i) = \kappa\,\mathrm{Var}_M(\psi_j), \quad \mathrm{Cov}_M(\Delta\psi_i, \Delta\bar{\alpha}_i) = \kappa\,\mathrm{Cov}_M(\psi_j, \bar{\alpha}_j), \quad \mathrm{Var}_M(\Delta\bar{\alpha}_i) = \kappa\,\mathrm{Var}_M(\bar{\alpha}_j).$$

Either condition implies

$$\theta_\infty = \theta_{\mathrm{iso}} = (1 + \beta_M)\,\theta_{\mathrm{var}}. \tag{9}$$

The proof (Appendix A) is ratio algebra. Isotropy says movers reproduce the cross-section's sorting geometry: they traverse the location-effect dimension at the same slope and cross the orthogonal composition dimension at the same intensity as the cross-section—in the network form, the Laplacian $\mathcal{L}$ is proportional to the cross-sectional covariance operator on $\mathrm{span}\{\psi, \bar{\alpha}\}$. The link is two-sided at the level of the pair (β, λ) but cannot be two-sided at the level of the scalar: f maps two coordinates to one, so distinct mobility anatomies can deliver the same θ_∞, and scalar equality $\theta_\infty = \theta_{\mathrm{iso}}$ can hold accidentally without recovery.

IV.1 The Mover Coefficient as a Midpoint

Write $V_\psi := \mathrm{Var}_M(\psi_j)$, $V_\alpha := \mathrm{Var}_M(\bar{\alpha}_j)$, $C := \mathrm{Cov}_M(\psi_j, \bar{\alpha}_j)$, $V_Y := V_\psi + 2C + V_\alpha$. Beyond the variance share $\theta_{\mathrm{var}} = V_\psi / V_Y$, a distinct object asks how much cross-sectional dispersion in location means would disappear if location effects were equalized while composition were held fixed,

$$S_\psi^{\mathrm{eq}} := \frac{\mathrm{Var}_M(\bar{Y}_j) - \mathrm{Var}_M(\bar{Y}_j - \psi_j)}{\mathrm{Var}_M(\bar{Y}_j)} = \frac{V_\psi + 2C}{V_Y}. \tag{10}$$

Corollary 1 (Midpoint interpretation and ordering)**.** *Under directional isotropy and $V_Y > 0$,*

$$\boxed{\theta_\infty = \theta_{\mathrm{iso}} = \tfrac{1}{2}\big(\theta_{\mathrm{var}} + S_\psi^{\mathrm{eq}}\big), \qquad S_\psi^{\mathrm{eq}} - \theta_\infty = \theta_\infty - \theta_{\mathrm{var}} = \frac{C}{V_Y}.} \tag{11}$$

If $C \geq 0$ then $\theta_{\mathrm{var}} \leq \theta_\infty \leq S_\psi^{\mathrm{eq}}$, and if $C < 0$ then $S_\psi^{\mathrm{eq}} \leq \theta_\infty \leq \theta_{\mathrm{var}}$; in either case $S_\psi^{\mathrm{eq}} \leq 1$. Moreover $\theta_\infty \leq \theta_{\mathrm{sd}}$ for either sign of C, and if $C \geq 0$ then $\theta_\infty^2 \leq \theta_{\mathrm{var}} \leq \theta_\infty$.

The mover coefficient recovers neither the variance share of location effects nor the full accounting effect of equalizing them; under isotropic mobility it lies exactly halfway between.

IV.2 Two Benchmark Readings

Corollary 2 (The two familiar readings)**.** *Maintain $\mathrm{Var}_M(\psi_j) > 0$ and $\mathrm{Var}_M(\bar{Y}_j) > 0$.*

(i) ***Orthogonal composition.*** *Under directional isotropy, if $\beta_M = 0$ then $\theta_\infty = \theta_{\mathrm{iso}} = \theta_{\mathrm{var}} = S_\psi^{\mathrm{eq}}$.*

(ii) ***Affine composition.*** *If $\lambda_M = 0$ and $\beta_M \geq 0$—so that $\bar{\alpha}_j = a_M + \beta_M\psi_j$ across the locations movers* *visit—then* every *flow with* $\mathrm{Var}_M(\Delta\psi_i) > 0$ *satisfies*

$$\theta_\infty = \theta_{\mathrm{sd}} = \frac{1}{1+\beta_M}, \qquad \theta_{\mathrm{var}} = \theta_\infty^2, \qquad S_\psi^{\mathrm{eq}} = 2\theta_\infty - \theta_\infty^2.$$

Each corner has a concrete economic case. Composition is orthogonal ($\beta_M = 0$) when individuals choose locations for reasons unrelated to the outcome under study and an isotropic mover design then recovers the variance share. The affine benchmark ($\lambda_M = 0,\ \beta_M \geq 0$) arises when matching runs on the outcome dimension itself and the coefficient is the SD share.

The two readings differ in their robustness. The orthogonal reading relies on isotropy. The affine benchmark with $\beta_M \geq 0$ is *mobility-robust*: once composition is perfectly collinear with the location effect there is no second dimension for mover routes to sample differently, so $\theta_\infty = \theta_{\mathrm{sd}}$ for every flow with $\mathrm{Var}_M(\Delta\psi_i) > 0$. This is a sufficient mobility-robust case. When $\lambda_M > 0$, mobility geometry generally matters and the coefficient is the transition projection $f(\beta_\Delta, \lambda_\Delta)$.

IV.3 Anisotropic Mobility

Outside isotropy the coefficient remains a well-defined transition object but loses the cross-sectional interpretation. Two examples illustrate the possibilities and organize the simulations.

Example 1 (Composition-preserving mobility)**.** If movers cross the location-effect dimension without crossing composition classes—a job ladder in which workers move to firms that hire similar coworkers, or relocation within a fixed tier of neighborhoods—then $\Delta\bar{\alpha}_i \approx 0$ on every move, $\beta_\Delta, \lambda_\Delta \to 0$, and $\theta_\infty \to 1$ regardless of the cross-sectional slope $\beta_M > 0$. A mover design can report full passthrough even when composition is highly dispersed across locations.

Example 2 (Balanced mobility with anisotropic exposure)**.** Consider four equally weighted locations indexed by $(p, z) \in \{-1, 1\}^2$, with $\psi_{p,z} = p$ and $\bar{\alpha}_{p,z} = cz$, $c > 0$, so composition is orthogonal to the location effect and $\theta_{\mathrm{iso}} = \theta_{\mathrm{var}} = 1/(1 + c^2)$. Let a balanced flow place probability ρ on moves that switch p holding z fixed, and $1 - \rho$ on moves that switch z holding p fixed. Then $\theta_\infty = \rho/[\rho + c^2(1-\rho)]$: at $\rho = \frac{1}{2}$ movers expose the two dimensions in the cross-sectional proportion and $\theta_\infty = \theta_{\mathrm{var}}$; for $\rho < \frac{1}{2}$ they over-sample residual composition and $\theta_\infty < \theta_{\mathrm{var}}$; as $\rho \to 1$ they avoid it and $\theta_\infty \to 1$. Balanced inflows and outflows alone do not deliver a cross-sectional reading; what matters is the geometry of the realized routes.

V Testing the Cross-Sectional Interpretation

Theorem 1 turns the cross-sectional interpretation into two testable restrictions. With $Q_M := \mathrm{Var}_M(q_j)$, $Q_\Delta := \mathrm{Var}_M(R_i)$, $V_{\psi,M} := \mathrm{Var}_M(\psi_j)$, and $V_{\psi,\Delta} := \mathrm{Var}_M(\Delta\psi_i)$, directional isotropy is equivalent to

$$g_\beta := \beta_\Delta - \beta_M = 0, \qquad g_Q := Q_\Delta V_{\psi,M} - Q_M V_{\psi,\Delta} = 0. \tag{12}$$

The first asks whether movers traverse the cross-sectional sorting slope; the second whether they cross the residual-composition dimension with the same normalized intensity. The null is that the transition and cross-sectional projection pairs coincide.

Location effects and means are estimated with finite-cell noise, and squaring noisy estimates inflates second moments—the same plug-in inflation that drives limited-mobility bias in sparse networks (Andrews et al., 2008; Kline et al., 2020); in the dense regime here it is a finite-sample nuisance, which cross-fitting removes to first order. I therefore assign each individual to one of the two data-independent halves and estimate the first stage separately within each half $s \in \{1,2\}$—$\widehat{\psi}_j^s$ from the half's movers, and $\widehat{\bar{Y}}_j^s$ as the half's equally weighted two-period location mean—and form every quadratic moment as a symmetrized *cross-split* product, one factor from each half, so that cross-split first-stage errors create no leading-order plug-in bias. Collecting the six cross-fitted primitive moments into $\widehat{\vartheta}$, the restrictions $\widehat{g} = (\widehat{g}_\beta, \widehat{g}_Q)' = g(\widehat{\vartheta})$ follow through smooth maps. Because every first-stage estimator is linear in outcomes, each individual's influence on $\widehat{g}$—including the dependence among movers sharing origins and destinations—is available in closed form, and an individual-level multiplier bootstrap requires no re-estimation (Supplemental Appendix S3).

Two further objects appear in the following proposition. Split s estimates location effects on a connected estimation universe $\mathcal{J}_s^E \supseteq \mathcal{J}$ of size J_s, where $\mathcal{J}$ is the common diagnostic core, and L_s denotes the Laplacian of the split's mover network—the graph whose vertices are locations and whose edges are the split's movers. The split is treated as an independent fair coin per individual.

Proposition 2 (Inference)**.** *Fix $\mathcal{J}$ and $\mathcal{J}_s^E$, or suppose their data-selected versions equal the fixed sets with probability approaching one, and let $N \to \infty$ with J and J_s fixed. Suppose the two-period individual observations are independent and identically distributed across individuals, the baseline split is an independent Bernoulli$(1/2)$ assignment independent of the individual data, and $\max_{t\in\{1,2\}} \mathbb{E}|Y_{it}|^{4+\eta} < \infty$ for some $\eta > 0$. Suppose an individual is a retained diagnostic mover with positive probability, every retained location–period cell has positive probability mass in each split, and*

$$N^{-1}L_s \xrightarrow{p} L_s^0, \qquad \operatorname{rank}(L_s^0) = J_s - 1, \qquad \lambda_2(L_s^0) > 0, \quad s \in \{1,2\}.$$

Let $\vartheta = (a,b,c,d,e,f)'$ (the mover-weighted cross-sectional variances and covariance of place effects and composition, and their transition counterparts), be the population primitive-moment vector, with $a = \operatorname{Var}_M(\psi_j) > 0$ and $d = \operatorname{Var}_M(\Delta\psi_i) > 0$. Let Σ_ϑ be the asymptotic covariance of $\sqrt{N}(\widehat{\vartheta} - \vartheta)$ characterized in Supplemental Appendix S3, set $G_g := \partial g(\vartheta)/\partial\vartheta'$, and suppose $\Omega := G_g \Sigma_\vartheta G_g' > 0$. Then

$$\sqrt{N}\{\widehat{g} - g(\vartheta)\} \xrightarrow{d} \mathcal{N}(0,\Omega).$$

As $N, B_v \to \infty$, where B_v counts the bootstrap draws $\widehat{g}^{*(b)} := g(\widehat{\vartheta}^{*(b)})$ whose a- and d-coordinate analogues are positive and whose restrictions are finite, the individual-level multiplier covariance satisfies $N\widehat{V}_g \xrightarrow{p} \Omega$; under $H_0\colon g(\vartheta) = 0$, $\mathcal{W} := \widehat{g}'\widehat{V}_g^{-1}\widehat{g} \xrightarrow{d} \chi_2^2$.

The proof is a standard fixed-J smooth-map argument: the graph and cell-mean estimators are smooth functions of finitely many individual averages, so asymptotic linearity, the delta method, and the conditional Gaussian multiplier central limit theorem deliver the result; Supplemental Appendix S3 provides the influence functions and details. $\widehat{g}_\beta$ shows whether mover routes load on a different sorting slope than the cross-section, $\widehat{g}_Q$ whether they over- or under-sample residual composition. Under the null, the midpoint reading (11) applies; under rejection, the coefficient remains a valid transition estimand but cannot be mapped to a cross-sectional share, and $(\widehat{g}_\beta, \widehat{g}_Q)$ reveal which failure drives the wedge.[13]

VI Simulations

I evaluate the theory on a simulated two-period AKM economy, $Y_{it} = \alpha_i + \psi_{j(i,t)} + \varepsilon_{it}$. The cross-section is fixed by design: $J_0 = 256$ locations with unit-variance place effects and composition $\bar{\alpha}_j = \beta_M \psi_j + q_j$, sorting slope $\beta_M = 0.6$, residual intensity $\lambda_M = 0.7$; the orthogonal and affine designs set $\beta_M = 0$ and $\lambda_M = 0$, respectively. Each location has a stayer core and one-time movers ($\approx$ 332,000 workers, mover share 7.5 percent, within-location dispersion of α_i equal to one, $\sigma_\varepsilon = 0.2$). Stayers carry $\bar{\alpha}_j$; mover types are centred on the symmetric edge potentials of the representation lemma (Supplemental Appendix S4), so leavers, arrivers, and stayers at every location share the same mean composition in both periods and Assumption 1 holds under every flow.

Mobility enters only through the symmetrized flow. Random reshuffling draws destinations independently of origins and reproduces the cross-sectional geometry exactly, $\beta_\Delta = \beta_M$ and $\lambda_\Delta = \lambda_M$: directional isotropy. The anisotropic designs mix the two move types of Example 2—with probability ρ a job-ladder step, crossing the place-effect dimension at fixed coworker type; with probability $1 - \rho$ a career switch, crossing coworker type at fixed place effect—which holds $\beta_\Delta = \beta_M$ and traces $\lambda_\Delta = \frac{1-\rho}{\rho}\lambda_M$: isotropic at $\rho = \frac{1}{2}$, oversampling the residual dimension for $\rho < \frac{1}{2}$, undersampling it for $\rho > \frac{1}{2}$. Each of 500 replications runs the Section V procedure end to end.

Table 1 reports three findings. First, under isotropic mobility the mover coefficient equals the cross-sectional midpoint of Corollary 1 (0.49), distinct from both the variance share (0.31) and the SD share (0.55); the orthogonal and affine rows reproduce the two corner readings of Corollary 2. Second—the headline—the four Panel B economies share an *identical* cross-section and an essentially identical pooled AKM decomposition, so an AKM variance study would call them one economy; yet the mover coefficient ranges from 0.25 to 0.61 purely because movers traverse the same map differently. As the flow tilts from career switches ($\rho = 0.15$, $\widehat{\lambda}_\Delta \approx 4$) toward near-pure job-ladder steps ($\rho = 0.92$, $\widehat{\lambda}_\Delta \to 0$), $\widehat{\theta}$ sweeps from far below θ_{iso} to well above it, while every cross-sectional benchmark—θ_{iso}, θ_{var}, λ_M—stays put. Third, the isotropy test holds nominal size in

[13] The formal joint-inference result is interior and excludes exact affine composition, where $\lambda_M = \lambda_\Delta = 0$ and the first-order covariance of the residual restriction is singular. No boundary law is derived; the usual two-degree-of-freedom Wald test under-rejects in the affine simulation designs. When $\beta_M \geq 0$ and $\text{Var}_M(\Delta\psi_i) > 0$, however, exact affinity makes the SD-share reading mobility-robust by Corollary 2(ii). Remark S3 gives deterministic near-affinity bounds and a practical diagnostic sequence.

the interior isotropic economies, under-rejects at the affine boundary, where the SD-share reading is mobility-robust in any case, and rejects every anisotropic flow with probability one; Panel C confirms the robustness—an affine cross-section under a residual-heavy flow still delivers $\widehat{\theta} = \theta_{\text{sd}}$ and no rejection.

VII Empirical Illustration: Dutch Administrative Data

I use the Statistics Netherlands (CBS) linked employer–employee panel covering 2009–2016 (). The outcome is the log gross hourly wage[14] of full-time male workers, residualized on age and calendar-year indicators in a preliminary step; the Supplemental Appendix repeats the analysis on the unadjusted wage with the same conclusions. Following the two-period convention of Section II, I retain stayers and one-time movers, and I keep firms above the 50th percentile of total worker-period observations; Supplemental Appendix S7 reports a denser large-firm sensitivity. Moves are inferred from the firm identifier: a one-time mover changes employer exactly once, the move year being the first year observed at the destination.

For the empirical illustration, the outcome change is spell-based: ΔY_i is the worker's average outcome over destination-spell years minus the average over origin-spell years. All firm attributes use the common 2009–2016 window. Workers are first averaged within firm–year, and the eight resulting firm–year means are then averaged equally. Following the aggregation convention of Finkelstein et al. (2016), movers' transition-year observations are excluded when constructing firm means, while stayers remain; the transition-year outcome itself remains in the event study. Only the canonical mover regression and event study use the pooled full-sample leave-one-out destination–origin difference in these full-window means. The six diagnostic moments use the corresponding split-specific common means directly. Using the same fixed window for both geometries is essential: otherwise differences in the timing of firm attributes could masquerade as mobility anisotropy. The estimation sample comprises 112,889 one-time movers and 697,802 stayers at 1,598 firms, with a mean firm-year size of 443 workers; full counts are in the notes to Table 2.

Workers are split into two random halves; within each, AKM recovers $\widehat{\psi}_j^s$ on the largest connected component of mover transitions, and firm-year means use all eligible workers in the half. Diagnostics are evaluated on the diagnostic core—firms inside both split connected components with nonempty cells in every window year in both halves—which retains 1,205 of 1,598 firms and 31,453 of 112,889 movers (27.9 percent).[15] The resulting target is movers between stable, continuously observed large firms, not all Dutch movers. Inference uses the worker-level multiplier bootstrap of Supplemental Appendix S3 with 1,000 draws.

Table 2 reports the results. The mover-design coefficient is $\widehat{\theta} = 0.150$, and the transition

[14] The hourly wage is annual gross contract earnings divided by contracted annual hours, computed as the job's full-time factor (FTE) times a 52-week, 40-hour full-time year; the sample is restricted to FTE ≥ 0.9.

[15] The split connected components retain all but one mover, so essentially all of the reduction comes from the complete-window endpoint requirement; every retained mover also has a complete canonical leave-one-out gap.

Table 1: Simulated Mover Designs: the Same AKM Decomposition Can Hide Different Estimands

Design	Pooled AKM decomposition				Cross-sectional readings					Residual share		
	Var α	Var ψ	Corr	$\widehat{\theta}$	θ_{iso}	θ_{var}	θ_{sd}	S_{ψ}^{eq}	midpt. $\frac{1}{2}(\theta_{\text{var}}+S^{\text{eq}})$	$\widehat{\lambda}_M$	$\widehat{\lambda}_\Delta$	Rej. H_0
Panel A: Random reshuffling (directionally isotropic); vary the cross-section												
General ($\beta > 0,\ \lambda_M > 0$)	2.14	1.00	0.41	0.491	0.491	0.307	0.554	0.675	0.491	0.701	0.700	0.054
Orthogonal ($\beta = 0$)	1.75	1.00	0.00	0.588	0.588	0.588	0.767	0.588	0.588	0.700	0.700	0.040
Affine ($q_j \equiv 0$)	1.39	1.00	0.51	0.625	0.625	0.391	0.625	0.859	0.625	0.000	0.000	0.020[a]
Panel B: Same cross-section and same AKM decomposition; vary only mobility ρ												
Balanced ($\rho = 0.50$)	2.09	1.00	0.42	0.491	0.491	0.307	0.554	0.675	0.491	0.700	0.699	0.048
Job ladder, ψ-heavy ($\rho = 0.85$)	2.09	1.00	0.42	**0.596**	0.491	0.307	0.554	0.675	0.491	0.699	0.123	1.000
Career switching, residual-heavy ($\rho = 0.15$)	2.11	1.00	0.41	**0.245**	0.491	0.307	0.554	0.675	0.491	0.701	3.969	1.000
Job ladder, near-pure ($\rho = 0.92$)	2.09	1.00	0.42	**0.610**	0.491	0.307	0.554	0.675	0.491	0.700	0.061	1.000
Panel C: Affine composition is mobility-robust												
Affine + residual-heavy flow	1.36	1.00	0.51	0.624	0.625	0.391	0.625	0.859	0.625	0.000	0.000	0.002[a]

Notes. Monte Carlo means over 500 replications of the individual-level economy: $J_0 = 256$ locations, $\approx 1{,}200$ stayers per location and 25,000 movers ($\approx 332{,}000$ workers, mover share 7.5%), $Y_{it} = \alpha_i + \psi_{j(i,t)} + \varepsilon_{it}$ with $\sigma_\alpha = 1$ (the within-location dispersion of α_i; the total Var α adds the between-location composition), $\sigma_\varepsilon = 0.2$, $\beta_M = 0.6$, $\lambda_M = 0.7$. Mover effects are pinned on symmetric edge potentials so that each location's mean composition is identical in both periods (Section VI). Each replication estimates $\widehat{\psi}_j$ by AKM on the largest connected set within each of two worker splits, forms leave-one-out coworker means, cross-fits the six μ-weighted primitive moments, and applies the 1,000-draw worker-level multiplier bootstrap of Supplemental Appendix S3, propagating first-stage $\widehat{\psi}$ and cell-mean noise in closed form. The first three columns are the pooled population AKM moments; $\widehat{\theta}$ is the leave-one-out mover-regression coefficient (2); the midpoint column verifies $\theta_{\text{iso}} = \frac{1}{2}(\theta_{\text{var}} + S_{\psi}^{\text{eq}})$. "Rej. H_0" is the 5% rejection frequency of the joint isotropy test: the test's *size* in the isotropic rows and its *power* in the anisotropic rows. In Panel B the cross-section—and hence every cross-sectional reading and λ_M—is held *exactly* fixed across rows; only ρ changes, and the pooled Var α and Corr coincide up to about a percent.

[a] At the affine boundary $\lambda_M = \lambda_\Delta = 0$ the parameter sits on the edge of the parameter space and formal joint inference is not derived there (Remark S3); in these designs the test under-rejects rather than spuriously rejecting the one-dimensional configuration, where the SD-share reading is mobility-robust in any case (Corollary 2). The interior rows show correct size.

projection reproduces it:

$$f(\widehat{\beta}_\Delta, \widehat{\lambda}_\Delta) = 0.155.$$

[16] The μ-weighted benchmarks are tightly estimated: $\theta_{\text{var}} = 0.256$, $\theta_{\text{sd}} = 0.506$, $S_\psi^{\text{eq}} = 0.253$, and $\theta_{\text{iso}} = 0.254$—the midpoint of Corollary 1 exactly. With the cross-sectional sorting slope essentially zero, θ_{var}, θ_{iso}, and S_ψ^{eq} nearly coincide near 0.25; the mover coefficient is roughly forty percent below these benchmarks and roughly seventy percent below θ_{sd}.

Panel B reports the sorting geometry, which fails isotropy on both coordinates. The cross-sectional sorting slope is statistically indistinguishable from zero ($\widehat{\beta}_M = -0.006$), whereas the transition slope is negative ($\widehat{\beta}_\Delta = -0.165$). Within the estimated AKM representation, the best linear prediction of the coworker-composition change for a premium-gaining move is negative. Residual composition is substantial in the cross-section ($\widehat{\lambda}_M = 2.92$), ruling out the affine benchmark of Corollary 2(ii) and hence the mobility-robust SD-share reading. Transition routes load on the residual-composition dimension about sixty percent more heavily than the cross-section does ($\widehat{\lambda}_\Delta = 4.68$). Both restrictions reject individually, and the joint Wald statistic is 156.8 (bootstrap $p < 0.001$). Under the maintained assumptions and the fixed-window bridge of Supplemental Appendix S1, this spell-average contrast targets θ_∞.

With isotropy rejected, $\widehat{\theta}$ remains a transition projection but inherits no cross-sectional share interpretation. The estimated geometry resembles the residual-heavy, career-switching anatomy of Example 2 and the corresponding row of Table 1: relative to the cross-section, transition routes load more heavily on residual composition, pushing the mover coefficient below each of the cross-sectional benchmarks. Supplemental Appendix S7 repeats the analysis on the denser universe of firms above the 75th percentile of worker-period observations: the full universe contains 220 firms with a mean firm-year size of roughly 1,604, of which 176 enter the diagnostic core. The sorting geometry there differs—the cross-sectional slope turns positive and significant—yet the residual restriction still rejects isotropy decisively; the mover coefficient lies below θ_{iso}, S_ψ^{eq}, and θ_{sd}, while slightly exceeding θ_{var}.

Figure 1 reports the multi-period FGW scaled event study. Its regressor uses the Finkelstein et al. (2016) window-average convention: eligible workers are averaged within firm–year, those firm–year means are averaged equally over the fixed window, and the origin mean is subtracted from the destination mean. Pre-event coefficients from $r = -6$ onward are economically negligible—at most 0.017 in absolute value—although some are statistically distinguishable from zero at this sample size; the outermost coefficient (0.062 at $r = -7$) rests on the thin balanced cell at the window edge. Post-event coefficients jump to 0.135 in the move year, mechanically attenuated because that year straddles the origin and destination firms, and rise to about 0.18 by $r = 3$, flattening thereafter; the implied multi-period scaled coefficient is 0.16, in line with the two-period $\widehat{\theta}$. This slight post-move drift is a departure from the static benchmark and is consistent with gradual wage adjustment,

[16]Finite-cell noise in the leave-one-out regressor attenuates $\widehat{\theta}$, so under the static model the mover regression tends to lie slightly below the cross-fitted projection—as it does here.

including the accumulation of returns at the destination.[17]

Table 2: Mover-Design Diagnostics, Dutch Administrative Data 2009–2016

	Estimate	SE	95% CI
Panel A. Mover-design coefficient and cross-sectional benchmarks			
Mover regression $\widehat{\theta}$	0.150	0.005	[0.139, 0.160]
Transition projection $f(\widehat{\beta}_\Delta, \widehat{\lambda}_\Delta)$	0.155	0.006	[0.144, 0.167]
θ_{iso} (midpoint)	0.254	0.009	[0.236, 0.272]
θ_{var}	0.256	0.020	[0.218, 0.296]
θ_{sd}	0.506	0.020	[0.467, 0.544]
S_ψ^{eq}	0.253	0.016	[0.220, 0.282]
Panel B. Sorting geometry: cross-section versus transitions			
Cross-sectional slope $\widehat{\beta}_M$	−0.006	0.060	[−0.110, 0.122]
Transition slope $\widehat{\beta}_\Delta$	−0.165	0.061	[−0.266, −0.039]
Cross-sectional residual intensity $\widehat{\lambda}_M$	2.922	0.205	[2.562, 3.356]
Transition residual intensity $\widehat{\lambda}_\Delta$	4.683	0.363	[4.077, 5.491]
Panel C. Isotropy test			
Slope restriction $\widehat{g}_\beta$	−0.160	0.038	[−0.232, −0.086]
Residual restriction $\widehat{g}_Q$ ($\times 10^4$)	3.648	0.432	[2.809, 4.521]
Joint Wald $\mathcal{W}$ (χ^2_2 $p = 8.9 \times 10^{-35}$)	156.800	—	—

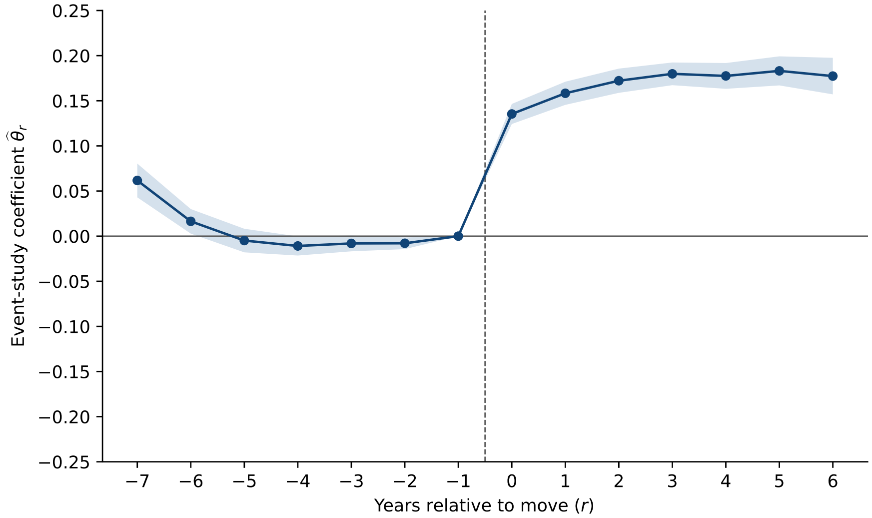
Notes. CBS linked employer–employee data, 2009–2016, full-time male workers; outcome is log gross hourly wage residualized on age and calendar-year indicators. Stayers and one-time movers; firms at or below the 50th percentile of worker-period observations dropped. N = 5,309,689 worker-years, 810,691 workers, 697,802 stayers and 112,889 one-time movers (a worker-level mover share of 13.9 percent), 1,598 firms; diagnostics on the core of firms inside both split connected components with nonempty cells in every window year in both halves, which retains 1,205 firms (75.4 percent) and 31,453 movers, all with complete canonical leave-one-out gaps. The mover regression, the transition moments, and the cross-sectional benchmarks are all computed on this common core sample (n = 31,453). SEs and percentile CIs from the worker-level multiplier bootstrap of Supplemental Appendix S3 with 1,000 draws, conditioning on the preliminary age- and calendar-year residualization coefficients; the SE for $\widehat{\theta}$ is the conventional OLS SE, and its CI the corresponding analytic normal interval. These conventional quantities treat the generated leave-one-out location averages as fixed and do not incorporate their sampling variation under the paper's asymptotics. θ_{iso} equals $\frac{1}{2}(\theta_{\text{var}} + S_\psi^{\text{eq}})$ exactly, and the two formulas for the transition projection agree by construction; both identities hold in the estimates as reported. The reported χ^2_2 and bootstrap joint-test p-values are interior results; the bootstrap value is below 0.001.

VIII Conclusion

Mover-design event study coefficients identify a transition-weighted projection on changes in location means, not a generic share of the cross-sectional variance of location means due to places. Under stationarity, the projection depends on mobility only through the undirected flow between locations, and it admits a cross-sectional interpretation exactly when mobility reproduces the cross-section's sorting geometry—a testable condition. When the condition holds, the recovered

[17] The drift enters the mover coefficient directly because the post-move outcome averages incorporate these later gains. It does not enter the isotropy diagnostic through movers' wage changes; the diagnostic is constructed from the estimated firm effects. Dynamic wage adjustment can therefore affect the test only insofar as it is absorbed into those estimates, particularly in ways that vary systematically across types of moves. The empirical rejection should accordingly be interpreted as evidence against isotropy within the static AKM representation.

Figure 1: Event Study of Outcomes around the Mover Transition



Notes. Coefficients $\widehat{\theta}_r$ from the multi-period FGW event-study regression of Supplemental Appendix S1, with base period $r = -1$ normalized to zero. The regressor is the pooled full-sample leave-one-out destination–origin difference in the common 2009–2016 firm means, held fixed across horizons; the transition-year outcome remains in the event-study panel. Estimates use the mover sample (241,098 mover-year observations). The shaded band is the pointwise 95 percent analytical confidence interval.

object is the midpoint between the variance share of place effects and the equalization share; positively affine sorting is a sufficient case in which the standard-deviation-share reading is robust to arbitrary mobility. When the condition fails—as it does in the Dutch wage-data illustration, where movers traverse far more residual composition than the cross-section contains—the coefficient cannot be mapped to a cross-sectional share.

A Proofs

Proof of Theorem 1. (ii)⇒(i): dividing the three displayed equalities gives

$$\begin{aligned}
\beta_\Delta &= \frac{\mathrm{Cov}_M(\Delta\psi_i, \Delta\bar{\alpha}_i)}{\mathrm{Var}_M(\Delta\psi_i)} \\
&= \frac{\mathrm{Cov}_M(\psi_j, \bar{\alpha}_j)}{\mathrm{Var}_M(\psi_j)} = \beta_M, \\
\lambda_\Delta &= \frac{\mathrm{Var}_M(\Delta\bar{\alpha}_i)}{\mathrm{Var}_M(\Delta\psi_i)} - \beta_\Delta^2 \\
&= \frac{\mathrm{Var}_M(\bar{\alpha}_j)}{\mathrm{Var}_M(\psi_j)} - \beta_M^2 = \lambda_M.
\end{aligned}$$

(i)⇒(ii): set $\kappa := \mathrm{Var}_M(\Delta\psi_i)/\mathrm{Var}_M(\psi_j) > 0$; then $\beta_\Delta = \beta_M$ delivers the cross-covariance equality and $\lambda_\Delta = \lambda_M$ the $\bar{\alpha}$ equality. For (9), divide the numerator and denominator of (3) by $\mathrm{Var}_M(\Delta\psi_i)$ to obtain $\theta_\infty = f(\beta_\Delta, \lambda_\Delta)$; under (i) this is $f(\beta_M, \lambda_M) = \theta_{\text{iso}} = (1 + \beta_M)\theta_{\text{var}}$, using

$$\begin{aligned}\text{Cov}_M(\psi_j, \bar{Y}_j) &= \text{Var}_M(\psi_j)(1 + \beta_M),\\ \text{Var}_M(\bar{Y}_j) &= \text{Var}_M(\psi_j)[(1 + \beta_M)^2 + \lambda_M].\end{aligned}$$

□

Proof of Corollary 1. Under isotropy, Theorem 1 gives $\theta_\infty = \theta_{\text{iso}} = (V_\psi + C)/V_Y$. With $\theta_{\text{var}} = V_\psi/V_Y$ and $S_\psi^{\text{eq}} = (V_\psi + 2C)/V_Y$, the average of the two is $(V_\psi + C)/V_Y = \theta_\infty$; subtracting gives the distance identity. If $C \geq 0$ the ordering follows from $1 - S_\psi^{\text{eq}} = V_\alpha/V_Y \geq 0$; the case $C < 0$ is analogous. For the secondary bounds: when $1 + \beta_M > 0$, $\theta_\infty/\theta_{\text{sd}} = (1 + \beta_M)\sqrt{\theta_{\text{var}}} \leq 1$ is equivalent to $\lambda_M \geq 0$, and the bound is immediate when $1 + \beta_M \leq 0$; and if $C \geq 0$, $\theta_\infty^2 \leq \theta_{\text{var}} \leq \theta_\infty$ follows from the midpoint identity by substitution. □

Proof of Corollary 2. (i) $\beta_M = 0$ gives $C = 0$; isotropy then yields $\theta_\infty = \theta_{\text{iso}} = V_\psi/V_Y = \theta_{\text{var}} = S_\psi^{\text{eq}}$. (ii) $\lambda_M = 0$ forces $\text{Var}_M(q_j) = 0$, so $\bar{\alpha}_j = a_M + \beta_M\psi_j$ at every location movers visit and every move obeys $\Delta\bar{\alpha}_i = \beta_M\Delta\psi_i$, $\Delta\bar{Y}_i = (1 + \beta_M)\Delta\psi_i$; hence for any flow with $\text{Var}_M(\Delta\psi_i) > 0$, $\theta_\infty = 1/(1 + \beta_M) = \theta_{\text{sd}}$, and the remaining identities follow by substitution. □

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Agha, L., K. M. Ericson, and X. Zhao (2023). The impact of organizational boundaries on health care coordination and utilization. *American Economic Journal: Economic Policy 15*(3), 184–214.

Ahammer, A., A. Packham, and J. Smith (2024). Firms and worker health. National Bureau of Economic Research, Working Paper No. w32011.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008). High wage workers and low wage firms: Negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society Series A: Statistics in Society 171*(3), 673–697.

Badinski, I., A. Finkelstein, M. Gentzkow, and P. Hull (2023). Geographic variation in healthcare utilization: The role of physicians. Technical report, National Bureau of Economic Research.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2023). How much should we trust estimates of firm effects and worker sorting? *Journal of Labor Economics 41*(2), 291–322.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies 91*(6), 3253–3285.

Bronnenberg, B. J., J.-P. H. Dubé, and M. Gentzkow (2012). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review 102*(6), 2472–2508.

Cabral, M., C. Carey, and J. Son (2025). Partial outsourcing of public programs: Evidence on determinants of choice in medicare. *Review of Economics and Statistics*. Forthcoming.

Cantoni, E. and V. Pons (2022). Does context outweigh individual characteristics in driving voting behavior? evidence from relocations within the united states. *American Economic Review 112*(4), 1226–1272.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Card, D., A. R. Cardoso, and P. Kline (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics 131*(2), 633–686.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly Journal of Economics 128*(3), 967–1015.

Card, D., J. Rothstein, and M. Yi (2025). Location, location, location. *American Economic Journal: Applied Economics 17*(1), 297–336.

Chandra, A. and C. Xu (2025). Where discovery happens: Research institutions and fundamental knowledge in the life-sciences. Working Paper 33996, National Bureau of Economic Research.

Chyn, E. and N. Shenhav (2025). Place effects and geographic inequality in health at birth. *American Economic Journal: Economic Policy*. Forthcoming.

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American economic review 110*(9), 2964–2996.

Ding, H. (2026). Geographic variation in mental health treatment utilization: Evidence from migration. *American Economic Journal: Economic Policy*. Forthcoming.

Doyle, J. J. and B. Staiger (2025). Physician group influences on treatment intensity and health: Evidence from physician switchers. *American Economic Journal: Economic Policy 17*(2), 471–505.

Encinosa, W. and A. Dor (2025). Does physician-hospital vertical integration signal care-coordination? evidence from mover-stayer analysis of commercially insured enrollees. *Journal of Health Economics 101*.

Finkelstein, A., M. Gentzkow, and D. Li (2025). What drives risky prescription opioid use? evidence from migration. *The Quarterly Journal of Economics 140*(4).

Finkelstein, A., M. Gentzkow, and H. Williams (2016). Sources of geographic variation in health care: Evidence from patient migration. *The Quarterly Journal of Economics 131*(4), 1681–1726.

Godøy, A. and I. Huitfeldt (2020). Regional variation in health care utilization and mortality. *Journal of Health Economics 71*, 102254.

Hinnosaar, M. and E. M. Liu (2022). Malleability of alcohol consumption: Evidence from migrants. *Journal of Health Economics 85*.

Huang, S. and H. Ullrich (2024). Provider effects in antibiotic prescribing: Evidence from physician exits. *Journal of Human Resources*. Forthcoming.

Hull, P. (2018). Estimating treatment effects in mover designs. *arXiv preprint arXiv:1804.06721*.

Hut, S. (2020). Determinants of dietary choice in the US: Evidence from consumer migration. *Journal of Health Economics*.

Kline, P., R. Saggio, and M. Sølvsten (2020). Leave-out estimation of variance components. *Econometrica 88*(5), 1859–1898.

Kline, P. M. (2025). Branching fixed effects: A proposal for communicating uncertainty. Technical report, National Bureau of Economic Research.

Luan, W. P., R. Leal, J. S. Zhou, and J. S. Skinner (2025). Place-based variation in health care: Evidence from mandatory movers in the us military health system. Technical report, National Bureau of Economic Research.

Lyubich, E. (2025). The role of people versus places in individual carbon emissions. *American Economic Review 115*(5), 1439–1484.

Molitor, D. (2018). The evolution of physician practice styles: Evidence from cardiologist migration. *American Economic Journal: Economic Policy 10*(1), 326–356.

Salm, M. and A. Wübker (2020). Sources of regional variation in healthcare utilization in Germany. *Journal of Health Economics 69*.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics 225*(2), 175–199.

Zeltzer, D., L. Einav, A. Chasid, and R. D. Balicer (2021). Supply-side variation in the use of emergency departments. *Journal of Health Economics*.

Supplemental Appendix

When Is the Mover-Design Event Study Coefficient a Place-Effect Share?

Vahid Moghani

This Supplemental Appendix relates the multi-period Finkelstein et al. (2016) event study to the two-period first difference used in the paper (Section S1), characterizes the flow restrictions implied by mean stationarity (Section S2), develops the asymptotic theory and multiplier bootstrap behind the isotropy test (Section S3), proves a representation lemma showing that the paper's reduced-form objects exhaust the AKM class (Section S4), and reports descriptive statistics for the estimation sample (Section S5), the robustness of the empirical results to the unadjusted log hourly wage (Section S6), and a denser large-firm sensitivity (Section S7). The paper applies its recovery and inference results under mean stationarity. The final section records, as a separate deviation, recovery against a common full-window cross-section when mean stationarity is dropped (Section S8). The supplement is referenced from, but—except for the inference theory of Section S3—not required for, the results in the paper.

S1 The FGW Event Study and the Two-Period Target

The applied mover-design event study is the multi-period Finkelstein et al. (2016) (FGW) specification with base period $k = -1$. This section makes its relation to the paper's two-period objects precise in two steps. First, on a balanced event-time window the FGW coefficients coincide, by projection algebra alone, with a sequence of per-event-time long-difference regressions (Proposition S1). Second, under the AKM outcome model with residual exogeneity, those per-horizon coefficients converge—pre-event to zero and post-event to a common first-difference limit (Theorem S1). Under mean stationarity for the compressed origin–destination endpoints and the eligible-window bridge below, that limit equals the paper's two-period estimand.

Consider movers $i \in M$ observed in periods $t \in \{1, \dots, T\}$, $T \geq 4$, each changing location exactly once at t_i^*, with event time $\tau_{it} := t - t_i^*$ and $o(i) := j(i, t_i^* - 1)$, $d(i) := j(i, t_i^*)$.

Assumption S1 (AKM residual exogeneity)**.** $Y_{it} = \alpha_i + \psi_{j(i,t)} + \varepsilon_{it}$ with $\psi_{j(i,t)} = \psi_{o(i)}$ for $t < t_i^*$ and $\psi_{d(i)}$ thereafter. Let $\mathcal{F}^T$ be the sigma-field generated by all workers' individual effects and realized location paths and by the location effects. Then $\mathbb{E}[\varepsilon_{it} \mid \mathcal{F}^T] = 0$; the residual vectors $(\varepsilon_{i1}, \dots, \varepsilon_{iT})'$ are conditionally uncorrelated across workers, with unrestricted dependence across t within a worker (Card et al., 2013, 2018).

Assumption S2 (Balanced event-time window)**.** Fix $\mathcal{K} := \{-K_-, \dots, -2\} \cup \{0, \dots, K_+\}$ and restrict move dates to $K_- + 1 \leq t_i^* \leq T - K_+$, so that every mover's event window $\mathcal{K} \cup \{-1\}$ lies inside the calendar panel. Estimation uses only the event-window observations, $\tau_{it} \in \mathcal{K} \cup \{-1\}$.

Let $\mathcal{T} := \{1, \dots, T\}$ denote the fixed calendar window and let $A_{\ell t}$ indicate that worker ℓ's observation is eligible for the location mean at t. In particular, $A_{\ell t} = 0$ for a mover's transition-year observation, while stayers remain eligible. Define

$$\overline{Y}_j^{(-i)} := \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \frac{\sum_{\ell \neq i} A_{\ell t} \mathbf{1}\{j(\ell, t) = j\} Y_{\ell t}}{\sum_{\ell \neq i} A_{\ell t} \mathbf{1}\{j(\ell, t) = j\}}, \qquad \delta_i^{(-i)} := \overline{Y}_{d(i)}^{(-i)} - \overline{Y}_{o(i)}^{(-i)}. \tag{S1}$$

Thus workers are averaged within location–year first and the resulting location–year means are then averaged equally over the fixed years, as in Finkelstein et al. (2016). The superscript $(-i)$ is the leave-one-out refinement used here; FGW use the same aggregation order. If any required post-LOO endpoint-year denominator is zero, the complete gap is left undefined rather than averaging a shorter window. Write $\bar{Y}_j$ for the large-cell probability limit of this full-window location attribute and $\Delta \bar{Y}_i := \bar{Y}_{d(i)} - \bar{Y}_{o(i)}$.

The common-window construction is a measurement convention within the maintained structural regime. Maintain the eligible-window condition

$$\bar{Y}_j := \frac{1}{T} \sum_{t \in \mathcal{T}} \mathbb{E}[Y_{it} \mid A_{it} = 1,\, j(i,t) = j] = \psi_j + \bar{\alpha}_j, \qquad j \in \mathcal{J}_0. \tag{S2}$$

Although the FGW event study itself requires a longer panel, the aggregation formula specializes to the paper's two-period location mean when $T = 2$ and $A_{it} \equiv 1$, the eligibility convention for its structural origin and destination periods. For $T > 2$, the empirical eligibility convention targets the same location object under (S2).

Following equation (6) of Finkelstein et al. (2016)[18], the event-study specification is

$$Y_{it} = \tilde{a}_i + \theta_{\tau_{it}} \delta_i^{(-i)} + \rho_{\tau_{it}} + e_{it}, \tag{S3}$$

estimated by OLS on the mover sample over the window of Assumption S2: individual effects $\tilde{a}_i$, event-time main effects ρ_k, and the common full-window leave-one-out destination–origin gap $\delta_i^{(-i)}$ interacted with event time, normalized at the base period ($\theta_{-1} = \rho_{-1} = 0$).[19] The coefficients of interest are the θ_k.

The reference point is the per-k long difference: for $k \in \mathcal{K}$ define $Y_{ik}^* := Y_{i,t_i^*+k} - Y_{i,t_i^*-1}$ and estimate $Y_{ik}^* = c_k + \widetilde{\theta}_k \delta_i^{(-i)} + e_{ik}$ by OLS. The intercept is retained, so $\widetilde{\theta}_k$ is the centered sample slope $\widehat{\mathrm{Cov}}_M(\delta_i^{(-i)}, Y_{ik}^*)/\widehat{\mathrm{Var}}_M(\delta_i^{(-i)})$. The equivalence below does not require the mean gap to vanish. These per-k long differences are the algebraic reference for the event-study coefficients; the next

[18] FGW's specification additionally includes calendar-time effects and age controls; both are covered by the time-varying-controls remark below. Estimating calendar effects *jointly* within equation (S3) requires *cohort exchangeability*—that the joint distribution of $(\Delta\psi_i, \delta_i^{(-i)})$ not vary with the move date t_i^*: after partialling the individual effects, calendar dummies are functions of (k, t_i^*), so without exchangeability they load on cohort-varying moments of $(\Delta\psi_i, \delta_i^{(-i)})$ and the θ_k mix cohorts in the manner of the staggered-adoption critiques (Sun and Abraham, 2021; De Chaisemartin and d'Haultfoeuille, 2020; Borusyak et al., 2024). The per-k population projections remain defined cohort by cohort either way, and Theorem S1 applies within each cohort under its conditions.

[19] Following FGW, a mover's transition-year observation is excluded when constructing the location–year means because it may combine exposure to the origin and destination. This exclusion concerns only the construction of location means: the transition-year outcome remains in the event-study regression, so the coefficient at $k = 0$ is reported.

result shows the FGW coefficients coincide with them.

Proposition S1 (The FGW coefficients are per-k long differences)**.** *If* $\widehat{\mathrm{Var}}_M(\delta_i^{(-i)}) > 0$, *then the sample OLS coefficients in equation* (S3) *on the balanced window of Assumption S2 satisfy* $\theta_k = \widetilde{\theta}_k$ *and* $\rho_k = c_k$ *exactly for every* $k \in \mathcal{K}$.

Proof. Partial out the individual effects. On the balanced window, within-mover demeaning acts identically across movers, and the specification is saturated in event time—main effects and interactions at every k—so the projection separates into the per-k contrasts of event time k against the base period. Each contrast is the regression of Y_{ik}^* on $(1, \delta_i^{(-i)})$, delivering $\rho_k = c_k$ and $\theta_k = \widetilde{\theta}_k$. □

Theorem S1. *Under Assumptions S1 and S2, suppose the T-period worker panels are independently and identically sampled across workers,* $|M|/N \xrightarrow{p} \pi_M > 0$, $\sup_t \mathbb{E}[Y_{it}^2] < \infty$, *every eligible location–year cell used in* (S1) *has a positive limiting share, and* $\mathrm{Var}_M(\Delta\bar{Y}_i) > 0$, *in the fixed-$J_0$, large-location regime of the paper, as $N \to \infty$: (i) for every $k \in \mathcal{K}$ with $k \le -2$,* $\widetilde{\theta}_k \xrightarrow{p} 0$; *and (ii) for every $k \in \mathcal{K}$ with $k \ge 0$,*

$$\widetilde{\theta}_k \xrightarrow{p} \theta_{\mathrm{FD}} := \frac{\mathrm{Cov}_M(\Delta\psi_i, \Delta\bar{Y}_i)}{\mathrm{Var}_M(\Delta\bar{Y}_i)},$$

independently of k. If Assumption 1, applied to the compressed origin–destination endpoints, and the fixed-window bridge (S2) *are also maintained, then* $\bar{Y}_j = \psi_j + \bar{\alpha}_j$ *and* $\theta_{\mathrm{FD}} = \theta_\infty$ *in equation* (3).

Proof. Under Assumption S1, $Y_{it} = a_i + \Delta\psi_i \mathbf{1}\{\tau_{it} \ge 0\} + \varepsilon_{it}$, with $a_i := \alpha_i + \psi_{o(i)}$, so $Y_{ik}^* = \Delta\psi_i \mathbf{1}\{k \ge 0\} + \Delta\varepsilon_{ik}$. Because J_0 and T are fixed and every eligible location–year cell used in (S1) grows proportionally with N, the cell-mean law of large numbers and the exact leave-one-out identity give

$$\frac{1}{|M|} \sum_{i \in M} \left(\delta_i^{(-i)} - \Delta\bar{Y}_i\right)^2 \xrightarrow{p} 0.$$

Conditional mean zero and cross-worker residual orthogonality in Assumption S1, together with the omission of the focal worker from every cell mean, make the sample covariance between $\Delta\varepsilon_{ik}$ and $\delta_i^{(-i)}$ $o_p(1)$. Substitution in the centered OLS slope yields zero before the move and θ_{FD} after it. The fixed-window bridge gives $\bar{Y}_j = \psi_j + \bar{\alpha}_j$; together with Assumption 1, this yields $\theta_{\mathrm{FD}} = \theta_\infty$. □

Time-varying controls. Applied AKM specifications include time-varying controls, $Y_{it} = \alpha_i + \psi_{j(i,t)} + X_{it}'\gamma + \varepsilon_{it}$, with calendar-time effects the leading case. Every object in the paper is then defined on the control-adjusted outcome $\widetilde{Y}_{it} := Y_{it} - X_{it}'\gamma$: location means, the mover regression, the transition and cross-sectional moments, and the benchmark shares. The interpretation adjusts accordingly—the recovered shares are shares of the variance of *control-adjusted* location means (age- and year-adjusted wages in the application of Section VII of the paper)—and Theorem S1 holds unchanged with $\widetilde{Y}$ in place of Y.

Unbalanced event windows. Assumption S2 delivers the exact finite-sample projection identity above. In a conventional unbalanced event window, equation (S3) instead forms cohort–horizon-weighted projections. If the cohort-exchangeability condition stated above holds and, conditional on move cohort, availability at each horizon is mean-independent of $(\Delta\psi_i, \Delta\bar{Y}_i)$ and the corresponding outcome innovation, those weights do not change the limits in Theorem S1: pre-move coefficients converge to zero and post-move coefficients to the common θ_{FD}. Without these conditions, the coefficients remain well-defined weighted projections.

S2 Stationarity and Restrictions on Mover Flows

Assumption 1 of the paper requires the average mover composition at each location to be the same whether viewed as an origin or a destination. Let f_{jk} be the population mover flow from j to k, μ_{jk} the average individual effect along that flow, and $\Phi_{jk} := f_{jk}\mu_{jk}$ the effect-weighted flow. Constant location masses and mover shares give mass balance $\sum_k f_{jk} = \sum_l f_{lj}$ at every j; combining with the composition condition yields

$$\sum_k \Phi_{jk} = \sum_l \Phi_{lj}, \qquad \forall j : \tag{S4}$$

the total individual effect flowing out of a location must equal the total flowing in. Conversely, conditional on mass balance, (S4) is exactly equality of mean mover composition between departures and arrivals. It imposes at most $J_0 - 1$ independent linear restrictions on the effect-weighted flows, leaving origin–destination pairings, within-pair sorting, and all higher moments otherwise unrestricted. In the special case where destination assignment is independent of α_i conditional on origin, (S4) forces the average mover type to be constant within each connected component of the mover network; with dependence between mobility and α_i, a large class of mobility patterns satisfies mean stationarity while generating very different mover-design estimands. Flows satisfying mass balance and (S4) are the *circulations*: the mobility that leaves location sizes and mean composition unchanged. A general flow splits into such a circulation and a residual that carries mean-composition drift; Assumption 1 keeps the circulation, and the final Supplemental Appendix S8 records separately how the residual changes the centered covariance algebra.

S3 Inference for the Isotropy Diagnostic

This section states first-order inference for a fixed calendar window $\mathcal{T}$. The main paper's two-period estimator is the case $T = 2$; for $T > 2$, the result is its fixed-window implementation extension under (S2) matching the empirical illustration. The length T is fixed as $N \to \infty$, the static AKM relation (1) is understood period by period over the window, and the fixed-window residual-exogeneity condition of Assumption S1 is maintained. A one-time mover enters the first stage through $\Delta Y_i = \Delta\psi_i + \bar{\varepsilon}_{i,d} - \bar{\varepsilon}_{i,o}$: the structural signal is the paper's two-period contrast, while its residual is averaged over the two spells. Exogenous mobility and Assumption 1, for the limiting retained analysis population, are maintained throughout. The sampling unit is the worker: worker histories are independent and identically distributed draws from a common population—location effects are fixed parameters, and a draw comprises the worker's type, assignment path, eligibility indicators, and outcomes—and observations may be arbitrarily dependent within a worker. The split is data-independent: the analysis takes it to be an independent fair coin per worker, and Remark S1 covers the exactly balanced partition used in the implementation as well as user-supplied fixed splits. All transition moments are centered over retained movers, matching the intercept convention of the mover regression(Finkelstein et al., 2016).

Split estimation graphs and diagnostic core. Let R_{is} indicate that the complete history of worker i is assigned to split $s \in \{1,2\}$; the split is assigned independently of the worker data. The retained diagnostic core is $\mathcal{J} = \{1,\ldots,J\}$. Let M_i indicate a retained diagnostic mover, write $o(i)$ and $d(i)$ for its core endpoints, and define

$$D_i := \mathbf{e}_{d(i)} - \mathbf{e}_{o(i)}, \qquad \iota_i := \tfrac{1}{2}\{\mathbf{e}_{o(i)} + \mathbf{e}_{d(i)}\}, \qquad m := \sum_{i=1}^{N} M_i.$$

The estimated mover-incidence weights are

$$\widehat{\mu} := \frac{1}{m}\sum_{i=1}^{N} M_i \iota_i, \qquad \mathbf{1}'\widehat{\mu} = 1. \tag{S5}$$

The split-s location-effect regression may use a larger connected estimation universe $\mathcal{J}_s^E \supseteq \mathcal{J}$, of size J_s. Let $D_{is}^E \in \mathbb{R}^{J_s}$ be mover i's incidence vector there, write $o_s^E(i), d_s^E(i)$ for its endpoint coordinates, let E_{is} indicate inclusion in that split regression, and let $S_s \in \{0,1\}^{J \times J_s}$ select the diagnostic-core coordinates. With ΔY_i equal to the destination-spell outcome average minus the origin-spell average, set

$$L_s := \sum_{i=1}^{N} E_{is} D_{is}^E D_{is}^{E\prime}, \qquad b_s := \sum_{i=1}^{N} E_{is} D_{is}^E \Delta Y_i, \qquad \widehat{\psi}_E^s := L_s^+ b_s, \qquad \widehat{\psi}^s := S_s \widehat{\psi}_E^s. \tag{S6}$$

The Moore–Penrose normalization is used for notation only; every reported object is invariant to an additive constant in $\widehat{\psi}_E^s$.

Common full-window means and the canonical leave-one-out gap. Fix a calendar window $\mathcal{T}$ with $T := |\mathcal{T}|$ fixed. Let A_{it} indicate eligibility of observation (i,t) for a location mean. In the paper's two-period benchmark, $A_{it} \equiv 1$: its two observations are the structural origin and destination periods. In the multi-period implementation, the convention excludes a mover's transition-period observation from the location means while stayers are always retained, independently of whether that observation appears in the event-study outcome sample. For each core location, period, and split, define

$$n^s_{jt} := \sum_i R_{is} A_{it} \mathbf{1}\{j(i,t) = j\}, \qquad \overline{Y}^s_{jt} := \frac{\sum_i R_{is} A_{it} \mathbf{1}\{j(i,t) = j\} Y_{it}}{n^s_{jt}},$$

and the equally weighted window average

$$\widehat{z}^s_j := \frac{1}{T} \sum_{t \in \mathcal{T}} \overline{Y}^s_{jt}, \qquad \widehat{\bar{\alpha}}^s_j := \widehat{z}^s_j - \widehat{\psi}^s_j. \tag{S7}$$

By equation (S2), the population counterpart of $\widehat{z}^s_j$ is the same $\bar{Y}_j = \psi_j + \bar{\alpha}_j$ used by the two-period target. The order of aggregation follows Finkelstein et al. (2016): individuals are averaged within location-period first, and the periods are then averaged with equal weights. A location is retained in the core only if $n^s_{jt} > 0$ for every $t \in \mathcal{T}$ and both splits; otherwise the location and its incident diagnostic movers are removed. The window is never shortened or renormalized. All six diagnostic moments therefore use common, non-leave-one-out, split-specific full-window means.

Only the canonical mover regression and the event study use leave-one-out. With pooled full-sample counts

$$n_{jt} := \sum_\ell A_{\ell t} \mathbf{1}\{j(\ell,t) = j\}, \qquad \delta_{ijt} := A_{it} \mathbf{1}\{j(i,t) = j\},$$

define

$$\overline{Y}_{jt,-i} := \frac{\sum_\ell A_{\ell t} \mathbf{1}\{j(\ell,t) = j\} Y_{\ell t} - \delta_{ijt} Y_{it}}{n_{jt} - \delta_{ijt}}, \qquad \widehat{z}_{j,-i} := \frac{1}{T} \sum_{t \in \mathcal{T}} \overline{Y}_{jt,-i},$$

and $\widehat{\delta}^{(-i)}_i := \widehat{z}_{d(i),-i} - \widehat{z}_{o(i),-i}$. If any of the $2T$ endpoint-period denominators is nonpositive, the entire canonical gap is missing; no averaging over the remaining periods is performed. Writing $\overline{Y}_{jt}$ for the pooled mean,

$$\overline{Y}_{jt,-i} - \overline{Y}_{jt} = \frac{\delta_{ijt}(\overline{Y}_{jt} - Y_{it})}{n_{jt} - \delta_{ijt}}.$$

Let $n_{\min} := \min_{j \in \mathcal{J}, t \in \mathcal{T}} n_{jt}$. Under a uniform $4+\eta$ moment bound, $\max_{i,t} |Y_{it} - \overline{Y}_{j(i,t),t}| = O_p(N^{1/(4+\eta)})$, so $n_{\min}/N^{1/2+1/(4+\eta)} \to \infty$ suffices for the leave-one-out perturbation to be $o_p(N^{-1/2})$ uniformly in i. Under the fixed-J conditions of Proposition S2, retained cell counts grow proportionally with N, so this requirement holds automatically.

Six cross-fitted moments. Let

$$\widehat{P}_\mu := I - \mathbf{1}\widehat{\mu}', \qquad p_s := \widehat{P}_\mu \widehat{\psi}^s, \qquad r_s := \widehat{P}_\mu \widehat{\bar{\alpha}}^s.$$

For retained mover i, define

$$x_{is} := D_i' \widehat{\psi}^s, \qquad v_{is} := D_i' \widehat{\bar{\alpha}}^s,$$

and let x^c_{is} and v^c_{is} denote centering over the m retained movers. The primitive estimators are

$$\begin{aligned}
\widehat{a} &:= \sum_j \widehat{\mu}_j p_{1j} p_{2j}, & \widehat{b} &:= \frac{1}{2}\sum_j \widehat{\mu}_j (p_{1j} r_{2j} + p_{2j} r_{1j}), \\
\widehat{c} &:= \sum_j \widehat{\mu}_j r_{1j} r_{2j}, & \widehat{d} &:= \frac{1}{m}\sum_i M_i x^c_{i1} x^c_{i2}, \\
\widehat{e} &:= \frac{1}{2m}\sum_i M_i (x^c_{i1} v^c_{i2} + x^c_{i2} v^c_{i1}), & \widehat{f} &:= \frac{1}{m}\sum_i M_i v^c_{i1} v^c_{i2}.
\end{aligned} \tag{S8}$$

Every quadratic moment multiplies one split-1 estimate by one split-2 estimate, so first-stage estimation errors, independent across splits, contribute no leading-order plug-in bias(Kline, 2025). Write $\widehat{\vartheta} = (\widehat{a}, \widehat{b}, \widehat{c}, \widehat{d}, \widehat{e}, \widehat{f})'$. For $a > 0$ and $d > 0$, let

$$Q_M := c - \frac{b^2}{a}, \qquad Q_\Delta := f - \frac{e^2}{d},$$

and define

$$g(\vartheta) := \begin{pmatrix} g_\beta \\ g_Q \end{pmatrix} = \begin{pmatrix} e/d - b/a \\ (f - e^2/d)a - (c - b^2/a)d \end{pmatrix}. \tag{S9}$$

Direct contribution and gradients. For a retained mover, set

$$h_i := \begin{pmatrix} \frac{1}{2}\sum_{\ell \in \{o(i), d(i)\}} p_{1\ell} p_{2\ell} \\ \frac{1}{4}\sum_{\ell \in \{o(i), d(i)\}} (p_{1\ell} r_{2\ell} + p_{2\ell} r_{1\ell}) \\ \frac{1}{2}\sum_{\ell \in \{o(i), d(i)\}} r_{1\ell} r_{2\ell} \\ x^c_{i1} x^c_{i2} \\ \frac{1}{2}(x^c_{i1} v^c_{i2} + x^c_{i2} v^c_{i1}) \\ v^c_{i1} v^c_{i2} \end{pmatrix}. \tag{S10}$$

Then

$$\widehat{\vartheta} = \frac{1}{m}\sum_i M_i h_i, \qquad \widehat{\phi}^{\text{edge}}_i := \frac{M_i}{m}(h_i - \widehat{\vartheta}). \tag{S11}$$

This contribution includes estimation of $\widehat{\mu}$ and the mover centering: derivatives of the weighted centers vanish because the opposite split's centered vector has weighted mean zero.

Suppress hats temporarily, let $D_\mu := \operatorname{diag}(\mu)$, and for a mover-centered scalar u_i define

$$\tau(u) := \frac{1}{m}\sum_i M_i D_i u_i.$$

The gradients with respect to the core vectors (ψ^1, z^1) are

$$G_\psi^1 = \begin{pmatrix} (D_\mu p_2)' \\ \frac{1}{2}\{D_\mu(r_2 - p_2)\}' \\ -(D_\mu r_2)' \\ \tau(x_2^c)' \\ \frac{1}{2}\{\tau(v_2^c) - \tau(x_2^c)\}' \\ -\tau(v_2^c)' \end{pmatrix}, \qquad G_z^1 = \begin{pmatrix} 0' \\ \frac{1}{2}(D_\mu p_2)' \\ (D_\mu r_2)' \\ 0' \\ \frac{1}{2}\tau(x_2^c)' \\ \tau(v_2^c)' \end{pmatrix}. \tag{S12}$$

The matrices G_ψ^2, G_z^2 interchange subscripts 1 and 2. Their rows sum to zero, which is the normalization invariance. Rows five and six of G_z^s are the transition response to the same common full-window mean used in the cross-sectional moments. A hat on any gradient below means evaluation at the corresponding estimates.

The restriction Jacobian is

$$G_g(\vartheta) = \begin{pmatrix} b/a^2 & -1/a & 0 & -e/d^2 & 1/d & 0 \\ f - e^2/d - b^2 d/a^2 & 2bd/a & -d & ae^2/d^2 - c + b^2/a & -2ae/d & a \end{pmatrix}, \qquad \widehat{G}_g := G_g(\widehat{\vartheta}). \tag{S13}$$

For reported shares, write $V_M := a + 2b + c$ and $V_\Delta := d + 2e + f$. On $V_M, V_\Delta > 0$, define

$$\theta_{\text{var}} := a/V_M, \qquad \theta_{\text{iso}} := (a+b)/V_M, \qquad S_\psi^{\text{eq}} := (a+2b)/V_M, \qquad \theta_\Delta := (d+e)/V_\Delta,$$

with Jacobian, in this order,

$$G_\theta(\vartheta) = \begin{pmatrix} (2b+c)/V_M^2 & -2a/V_M^2 & -a/V_M^2 & 0 & 0 & 0 \\ (b+c)/V_M^2 & (c-a)/V_M^2 & -(a+b)/V_M^2 & 0 & 0 & 0 \\ c/V_M^2 & 2c/V_M^2 & -(a+2b)/V_M^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & (e+f)/V_\Delta^2 & (f-d)/V_\Delta^2 & -(d+e)/V_\Delta^2 \end{pmatrix}. \tag{S14}$$

When $\theta_{\text{var}} > 0$, the derivative of $\theta_{\text{sd}} = \sqrt{\theta_{\text{var}}}$ is the first row of (S14) divided by $2\sqrt{\theta_{\text{var}}}$.

Worker influence. For a split-s estimation mover, define

$$\widehat{u}_i^s := \Delta Y_i - D_{is}^{E\prime} \widehat{\psi}_E^s, \qquad \dot{\psi}_{E,i}^s := E_{is} L_s^+ D_{is}^E \widehat{u}_i^s.$$

The full-window cell-mean contribution on the core is

$$\dot{z}_i^s := R_{is} \sum_{t \in \mathcal{T}} \frac{1}{T} A_{it} \mathbf{1}\{j(i,t) \in \mathcal{J}\} \mathbf{e}_{j(i,t)} \frac{Y_{it} - \overline{Y}_{j(i,t),t}^s}{n_{j(i,t),t}^s}.$$

The estimated first-order contribution of worker i is

$$\widehat{\phi}_i = \widehat{\phi}_i^{\text{edge}} + \sum_{s=1}^{2} \left(\widehat{G}_\psi^s S_s \dot{\psi}_{E,i}^s + \widehat{G}_z^s \dot{z}_i^s \right). \tag{S15}$$

Equivalently, with

$$W_s := L_s^+ S_s' \widehat{G}_\psi^{s\prime},$$

the location-effect term is

$$\widehat{G}_\psi^s S_s \dot{\psi}_{E,i}^s = E_{is} \{ W_s[d_s^E(i), \cdot] - W_s[o_s^E(i), \cdot] \}' \widehat{u}_i^s,$$

so one six-right-hand-side Laplacian solve is required per split. The normal equations, the within-cell residual identities, and (S11) imply $\sum_i \widehat{\phi}_i = 0$. The finite-sample covariance estimate and its root-N scaling are

$$\widehat{\text{Var}}(\widehat{\vartheta}) := \sum_i \widehat{\phi}_i \widehat{\phi}_i', \qquad \widehat{\Sigma}_\vartheta := N \sum_i \widehat{\phi}_i \widehat{\phi}_i'. \tag{S16}$$

Finite-dimensional reduction. The proposition below exploits that, with J, J_s, and T fixed, the entire construction is one smooth map of a fixed-dimensional vector of sample means. Collect for each worker the graph block, the cell block, and the endpoint block,

$$\begin{aligned} W_i^{\text{g}} &:= \left\{ E_{is} \mathbf{1}\{ o_s^E(i) = j,\ d_s^E(i) = k \} (1,\ \Delta Y_i) \right\}_{j \neq k,\ s}, \\ W_i^{\text{c}} &:= \left\{ R_{is} A_{it} \mathbf{1}\{ j(i,t) = j \} (1,\ Y_{it}) \right\}_{j \in \mathcal{J},\, t \in \mathcal{T},\ s}, \qquad W_i^{\text{e}} := \left\{ M_i \mathbf{1}\{ o(i) = j,\ d(i) = k \} \right\}_{j \neq k}, \end{aligned} \tag{S17}$$

and set $W_i := (W_i^{\text{g}}, W_i^{\text{c}}, W_i^{\text{e}})$, $\overline{W} := N^{-1} \sum_i W_i$, and $\varpi := \mathbb{E}[W_i]$, common to all workers under the sampling frame below. Then $\widehat{\vartheta} = \Phi(\overline{W})$, where Φ is infinitely differentiable on the open set $\mathcal{O}$ on which all retained cell counts are positive, the diagnostic mover mass is positive, and both split estimation graphs are connected: on $\mathcal{O}$ the split Laplacian has constant rank $J_s - 1$, so L_s^+ is a smooth function of the edge shares, and every remaining operation is a ratio, product, or centering. Define the population target and influence

$$\vartheta := \Phi(\varpi), \qquad \phi_i := \frac{1}{N} \nabla \Phi(\varpi) (W_i - \varpi), \tag{S18}$$

so that $\sum_i \phi_i = \nabla \Phi(\varpi)(\overline{W} - \varpi)$ and $\mathbb{E} \phi_i = 0$. Evaluating the three blocks of $\nabla \Phi$—the endpoint shares, the split first stages, and the cell sums and counts—yields exactly the three terms of (S15) with population values in place of estimates: the endpoint block gives the edge kernel, the chain rule through $\widehat{\psi}_E^s = L_s^+ b_s$ gives the $E_{is} L_s^+ D_{is}^E u_i^s$ term, and the ratio expansion of each cell mean gives $\dot{z}_i^s$, with the gradient matrices (S12).

Proposition S2 (Fixed-J asymptotic linearity)**.** *Fix $\mathcal{J}$, $\mathcal{J}_s^E$, S_s, and $\mathcal{T}$; if any of these is selected from the data, it equals the fixed object with probability approaching one. Suppose:*

(i) worker histories—including the split indicator, drawn as an independent Bernoulli($\frac{1}{2}$) assignment per worker, independently of the remaining worker data—are independent and identically distributed across workers, with $\sup_t \mathbb{E}|Y_{it}|^{4+\eta} < \infty$ *for some* $\eta > 0$*;*

(ii) the diagnostic mover mass and every retained cell mass are positive: $\pi_M := \Pr(M_i = 1) > 0$ *and* $\pi^s_{jt} := \Pr\left(R_{is}A_{it}\mathbf{1}\{j(i,t) = j\} = 1\right) > 0$*;*

(iii) $L^0_s := \mathbb{E}\left[E_{is}D^E_{is}D^{E\prime}_{is}\right]$ *is the Laplacian of a connected graph on* $\mathcal{J}^E_s$*:* $\operatorname{rank} L^0_s = J_s - 1$ *and* $\lambda_2(L^0_s) > 0$*.*

Set $\vartheta := \Phi(\varpi)$ *and* $\Sigma_\vartheta := \operatorname{Var}\{\nabla\Phi(\varpi)\, W_i\} = \mathbb{E}[(N\phi_i)(N\phi_i)']$*, which is finite under (i). Then*

$$\widehat{\vartheta} - \vartheta = \sum_{i=1}^{N} \phi_i + O_p(N^{-1}), \qquad \sqrt{N}(\widehat{\vartheta} - \vartheta) \xrightarrow{d} \mathcal{N}(0, \Sigma_\vartheta), \tag{S19}$$

and moreover

$$N \max_i \|\widehat{\phi}_i - \phi_i\| \xrightarrow{p} 0, \qquad N \sum_i \widehat{\phi}_i \widehat{\phi}_i' \xrightarrow{p} \Sigma_\vartheta.$$

Proof. By the law of large numbers, $\overline{W} \xrightarrow{p} \varpi$ and $N^{-1}L_s \xrightarrow{p} L^0_s$; by (ii)–(iii), $\varpi \in \mathcal{O}$. Fix compacts $K \subset K' \subset \mathcal{O}$ with ϖ interior to K, so that $\Pr\{\overline{W} \in K'\} \to 1$ and $\nabla\Phi$ is bounded and Lipschitz with a bounded Hessian on K'. The entries of W_i are indicators or indicators times Y_{it} or ΔY_i, so they inherit finite $4 + \eta$ moments from (i), and $\|\overline{W} - \varpi\| = O_p(N^{-1/2})$. On the event $\{\overline{W} \in K'\}$, a second-order Taylor expansion gives

$$\Phi(\overline{W}) - \Phi(\varpi) = \nabla\Phi(\varpi)(\overline{W} - \varpi) + O_p\left(\|\overline{W} - \varpi\|^2\right) = \sum_i \phi_i + O_p(N^{-1}),$$

which is the linear representation; the identification of $\nabla\Phi(\varpi)(W_i - \varpi)/N$ with the population counterpart of (S15) is the block computation described above. The scaled scores $N\phi_i$ are independent and identically distributed, mean zero, with finite variance, so the classical central limit theorem gives the normal limit, the $O_p(N^{-1})$ remainder being $o_p(N^{-1/2})$.

For the influence accuracy, write $\widehat{\phi}_i = N^{-1}\nabla\Phi(\overline{W})(W_i - \overline{W})$, which is the estimated-gradient, estimated-centering version of ϕ_i and coincides term by term with (S15). Under the sampling frame (i) there is no centering wedge—$\mathbb{E}W_i = \varpi$ for every worker—so

$$N(\widehat{\phi}_i - \phi_i) = \{\nabla\Phi(\overline{W}) - \nabla\Phi(\varpi)\}(W_i - \overline{W}) + \nabla\Phi(\varpi)(\varpi - \overline{W}).$$

The gradient difference is $O_p(N^{-1/2})$ by the Lipschitz property, while $\max_i \|W_i\| = O_p(N^{1/(4+\eta)})$ under the moment bound, so $N \max_i \|\widehat{\phi}_i - \phi_i\| = O_p(N^{1/(4+\eta)-1/2}) + O_p(N^{-1/2}) = o_p(1)$. Consequently $N\sum_i \|\widehat{\phi}_i - \phi_i\|^2 \le \left(N \max_i \|\widehat{\phi}_i - \phi_i\|\right)^2 \to_p 0$, and by Cauchy–Schwarz it suffices that $N\sum_i \phi_i\phi_i' = N^{-1}\sum_i (N\phi_i)(N\phi_i)' \to_p \Sigma_\vartheta$, which is the law of large numbers for independent and identically distributed matrices with a finite mean. □

Two consequences are worth recording. First, no remainder condition is assumed anywhere: at fixed dimension the $O_p(N^{-1})$ remainder is delivered by Taylor's theorem, and the cross-split structure is needed only for the *centering* of the quadratic moments, not for their rates. Second, the leave-one-out refinement of the canonical regressor is covered by the earlier bound, since (ii) implies $n_{\min} \asymp N$.

Remark S1 (Balanced and conditioned splits). The implementation assigns an exactly balanced split without replacement, and a user-supplied split is fixed rather than random. Neither departure alters the influence representation, for a structural reason: Φ is degree-zero homogeneous in each split block of $\overline{W}$ separately—$\widehat{\psi}_E^s = L_s^+ b_s$ is invariant to a common scaling of the split-s edge counts and sums, the cell means are within-split ratios, and the endpoint block enters only through m-normalized shares—so by Euler's identity $\nabla\Phi(\varpi)$ annihilates the mean direction of each block. The difference between an exactly balanced partition and independent fair coins, and likewise the between-stratum mean wedge that arises when conditioning on the realized balanced partition, lie exactly along those block-mean directions and are therefore removed by the gradient: the influence functions, the asymptotic covariance, and the bootstrap are unchanged. The same holds for any fixed partition whose stratum shares converge to one half. For general fixed shares the argument applies verbatim with the block masses of ϖ set to those shares: the target $\Phi(\varpi)$ is unchanged by block homogeneity, while Σ_ϑ then reflects the split proportions.

Interior test and Gaussian multiplier bootstrap. The restriction test requires the population moments to satisfy $a, d > 0$; inference for the reported share ratios additionally requires V_M and V_Δ to be bounded away from zero along the asymptotic sequence. Define

$$\Omega := G_g(\vartheta)\,\Sigma_\vartheta\,G_g(\vartheta)',$$

and assume $\Omega > 0$ for the joint test. Draw independent $\xi_i^{(b)} \sim \mathcal{N}(0,1)$ and form

$$\widehat{\vartheta}^{*(b)} := \widehat{\vartheta} + \sum_i \xi_i^{(b)} \widehat{\phi}_i, \qquad \widehat{g}^{*(b)} := g(\widehat{\vartheta}^{*(b)}). \tag{S20}$$

Conditional on the data, the primitive perturbation is Gaussian with covariance $\sum_i \widehat{\phi}_i \widehat{\phi}_i'$; the implementation draws this six-dimensional Gaussian directly, which is distributionally identical to using Gaussian worker multipliers.

Let $\mathcal{B}_v$ index draws for which the positive a, d checks pass and $\widehat{g}^{*(b)}$ is finite, let $B_v := |\mathcal{B}_v|$, and let $\overline{g}^*$ be their mean. If $B_v < 3$, the joint test is reported missing; otherwise define

$$\widehat{V}_g := \frac{1}{B_v - 1} \sum_{b \in \mathcal{B}_v} (\widehat{g}^{*(b)} - \overline{g}^*)(\widehat{g}^{*(b)} - \overline{g}^*)',$$

so that, as $B_v \to \infty$, $\widehat{V}_g = \widehat{G}_g\big(\sum_i \widehat{\phi}_i\widehat{\phi}_i'\big)\widehat{G}_g' + o_p(N^{-1})$. Set $\widehat{g} := g(\widehat{\vartheta})$ and

$$\mathcal{W} := \widehat{g}'\widehat{V}_g^{-1}\widehat{g}, \qquad \mathcal{W}^{*(b)} := (\widehat{g}^{*(b)} - \widehat{g})'\widehat{V}_g^{-1}(\widehat{g}^{*(b)} - \widehat{g}).$$

Theorem S2 (Interior validity)**.** *Under Proposition S2, the interior conditions, and $\Omega > 0$, as $N, B_v \to \infty$,*

$$N\widehat{V}_g \xrightarrow{p} \Omega.$$

If, in addition, $H_0 : g(\vartheta) = 0$,

$$\mathcal{W} \xrightarrow{d} \chi_2^2,$$

and, conditionally on the data, the multiplier law of $\mathcal{W}^{*(b)}$ consistently estimates the null law. The finite-draw value

$$\widehat{p}_B := \frac{1 + \sum_{b\in\mathcal{B}_v} \mathbf{1}\{\mathcal{W}^{*(b)} \geq \mathcal{W}\}}{B_v + 1}$$

is asymptotically valid as $N, B_v \to \infty$. Scalar restriction tests use the analogous plus-one rule for the absolute centered perturbation.

Proof. Proposition S2 and the delta method give $\sqrt{N}\{\widehat{g} - g(\vartheta)\} \xrightarrow{d} \mathcal{N}(0, \Omega)$. For the conditional part, the two inputs of the Gaussian multiplier central limit theorem are consequences of the proposition rather than assumptions: covariance consistency is $N\sum_i \widehat{\phi}_i\widehat{\phi}_i' \to_p \Sigma_\vartheta$, and maximal negligibility follows from $\sqrt{N}\max_i \|\widehat{\phi}_i\| \leq \sqrt{N}\max_i \|\phi_i\| + \sqrt{N}\max_i \|\widehat{\phi}_i - \phi_i\| = N^{-1/2}O_p(N^{1/(4+\eta)}) + o_p(N^{-1/2}) \to_p 0$. Hence, conditionally on the data, $\sqrt{N}\sum_i \xi_i^{(b)}\widehat{\phi}_i \xrightarrow{d} \mathcal{N}(0, \Sigma_\vartheta)$ in probability, the smooth map g and the interior conditions transfer the limit to $\widehat{g}^{*(b)} - \widehat{g}$, and the continuous mapping theorem gives the two quadratic-form limits. The plus-one correction delivers validity of the finite-draw p-value as $B_v \to \infty$, with Monte Carlo error of order $B_v^{-1/2}$ if the number of draws is held fixed. □

Numerical interior rule. Let ϵ_{mach} denote machine precision and set $s_M := |a| + 2|b| + |c|$, $s_\Delta := |d| + 2|e| + |f|$. The implementation requires

$$a > \sqrt{\epsilon_{\text{mach}}}\, s_M, \qquad d > \sqrt{\epsilon_{\text{mach}}}\, s_\Delta,$$

and applies the analogous conditions to V_M, V_Δ for share ratios. Invalid bootstrap maps are left missing rather than truncated; in particular, negative variance-share draws are not truncated before taking a square root. Under the interior assumptions their share is $o_p(1)$. The joint covariance is accepted only when its diagonal is finite and positive and the smallest eigenvalue of its correlation matrix exceeds 10^{-10}; otherwise the joint statistic and its p-values are reported missing. No generalized inverse is used for the joint test.

Remark S2 (Growing number of locations). Proposition S2 is a fixed-dimension result: with J fixed, every estimator is a smooth function of finitely many sample means and the second-order remainder is $O_p(N^{-1})$ by Taylor's theorem. When J grows with N, that argument no longer applies. Each $\widehat{\psi}_j$ then carries estimation noise of order one over the square root of its degree, and the six quadratic forms accumulate J such terms; the cross-split construction still removes the leading plug-in bias, but controlling the *variance* of the second-order remainder at the $\sqrt{N}$ scale requires conditions on the mover network—uniformly vanishing effective resistance (leverage) of mover edges, an expansion condition on the degree-normalized split Laplacians, and no dominant mover-incidence weight—of the kind developed for variance components in sparse two-way models by Kline et al. (2020); see also Andrews et al. (2008) and Bonhomme et al. (2023).

Remark S3 (Affine boundary). The theorem does not cover exact affine composition, where the first-order covariance of the residual restriction is singular. No boundary law or coverage result is derived here; in the affine simulation designs reported in Table 1, the ordinary joint procedure under-rejects.

For applied work, one can first distinguish the boundary from the interior. For any fixed $\lambda_M > 0$, residual intensity is an interior parameter; provided the other interior conditions hold, a consistent interval excludes zero with probability approaching one and the joint test applies. At exact affinity, if $\beta_M \geq 0$ and $d = \mathrm{Var}_M(\Delta\psi_i) > 0$, Corollary 2(ii) gives $\theta_\infty = \theta_{\mathrm{sd}}$ for every mobility pattern. Conditional on exact affinity, no isotropy test is needed for that SD-share interpretation.

Near-affinity has deterministic content. Under Assumption 1, the symmetrized-flow representation and $(q_j - q_k)^2 \leq 2q_j^2 + 2q_k^2$ give $\mathrm{Var}_M(\Delta q_i) \leq 4Q_M$, hence, by projection optimality and Cauchy–Schwarz,

$$Q_\Delta \leq 4Q_M, \qquad |\beta_\Delta - \beta_M| \leq 2\sqrt{Q_M/d}.$$

Let $U \geq Q_M$, $s := 1 + \beta_M$, and $r := 2\sqrt{U/d}$. Then $Q_\Delta/d \leq r^2$ and $|\beta_\Delta - \beta_M| \leq r$. If $\beta_M \geq 0$ and $r < s$, the denominators of f are bounded away from zero and

$$|\theta_\infty - \theta_{\mathrm{sd}}| \leq \frac{r}{s(s-r)} + \frac{r^2}{(s-r)^3} + \frac{U}{2as^3}, \qquad a = \mathrm{Var}_M(\psi_j).$$

The first two terms bound the transition projection relative to $1/s$; the last uses $\theta_{\mathrm{sd}} = \{s^2 + Q_M/a\}^{-1/2}$. This is a deterministic population sensitivity bound.

A practical diagnostic sequence therefore has three outcomes. If a reported U satisfies $r < s$ and the displayed bound is substantively small, report the SD-share benchmark with that bound and do not interpret the joint p-value. If the lower endpoint of an interval for λ_M is positive and the remaining interior and numerical covariance conditions hold, report the joint isotropy test. If neither condition holds, report the evidence as inconclusive between the near-affine and interior regimes.

S4 A Representation Lemma: The Reduced Form Is Without Loss of Generality

The analysis conditions on three reduced-form objects: the cross-section $(\psi_j, \bar{\alpha}_j)$, location sizes and mover shares (n_j, w_j), and the symmetrized mover flow $\bar{F}$. The trivial direction is that any assignment process in the AKM class induces such a tuple. This section proves the converse: every admissible tuple is induced by some assignment process satisfying the paper's maintained assumptions, so conditioning on the reduced form entails no loss of generality.

Lemma S1 (Representation)**.** *Let $\mathcal{J}_0$ be a finite set of locations with attributes $(\psi_j, \bar{\alpha}_j)_{j\in\mathcal{J}_0}$, sizes $n_j > 0$, and mover shares $w_j \in (0, 1)$, and let $N_M := \sum_j w_j n_j$. Let $\bar{F} = (\bar{F}_{jk})$ be any nonnegative symmetric flow with $\bar{F}_{jj} = 0$ and $\sum_k \bar{F}_{jk} = w_j n_j / N_M$ for all j. Then there exists an assignment process in the AKM class—a period-1 assignment rule, an origin-dependent destination rule, type distributions, and an error process—that satisfies exogenous mobility and Assumption 1 and induces exactly the cross-section $(\psi_j, \bar{\alpha}_j)$, the sizes and shares (n_j, w_j), and the symmetrized flow $\bar{F}$.*

Proof. The proof is constructive. (i) *Population.* At each location j, place $(1 - w_j)n_j$ stayers and $w_j n_j$ movers in period 1. (ii) *Mover types.* Every mover, at every location, draws α_i i.i.d. from a fixed distribution D with mean $\bar{m}$ (say $\bar{m} = 0$), independent of origin and destination. (iii) *Destinations.* A mover at j moves to k with probability $K^2(k \mid j) = N_M \bar{F}_{jk} / (w_j n_j)$. The mass arriving at j is $N_M \sum_k \bar{F}_{kj} = w_j n_j$ by symmetry, so sizes and mover shares are constant across periods; arriving movers also have type distribution D and hence the same mean as those departing, so the mean-composition balance condition of Assumption 1 holds exactly. (iv) *Stayer types.* Stayers at j draw α_i from any distribution with mean $(\bar{\alpha}_j - w_j \bar{m})/(1 - w_j)$, so the location composition equals $\bar{\alpha}_j$ in both periods. (v) *Outcomes.* Location effects are the given ψ_j, and ε_{it} is i.i.d. mean zero, independent of types and assignments, so exogenous mobility holds. The induced directed flow is $F_{jk} = \bar{F}_{jk}$, whose symmetrization is $\bar{F}$. □

Three remarks. First, the construction loads all compositional heterogeneity on stayers and makes movers homogeneous on average; this is enough because every estimand in the paper—the transition moments, the cross-sectional moments, and hence θ_∞, the pairs $(\beta_\Delta, \lambda_\Delta)$ and (β_M, λ_M), and the benchmark shares—depends on the assignment process only through $(\psi, \bar{\alpha}, \bar{F})$. Second, richer constructions keep movers sorted rather than homogeneous, and dispense with stayers entirely. Assigning a mover travelling between j and k a type centred on the symmetric edge potential $\varphi_j + \varphi_k$, with φ solving $(I + W)\varphi = \bar{\alpha}$ for W the row-normalised symmetrised flow, makes every location's movers—leavers and arrivers alike—average its composition $\bar{\alpha}_j$ in both periods, with no stayers required; the solve is exact whenever no connected component of the flow is bipartite—a bipartite component being precisely what places -1 in the spectrum of W and makes $I + W$ singular. This restriction is empirically vacuous for the connected flows used here: a connected component is bipartite only if its locations split into two classes with every move crossing between them and none within—equivalently, the component contains no odd cycle—a configuration ruled out by any three of its locations exchanging workers in a triangle, and thus absent from realized mover networks. A directional balancing, $\sum_k \Phi_{jk} = \sum_k \Phi_{kj} = \deg_j \bar{\alpha}_j$ (the divergence-free condition of Section S2), drops the non-bipartiteness requirement but is itself feasible only on sufficiently connected supports, not on every flow with a cycle. The homogeneous construction of the lemma remains the universal witness precisely because stayers can absorb any composition; that is why the lemma itself needs no connectivity or parity condition at all. In each, a mover matches every location's *mean* composition in both periods but, carrying one type across two firms, adds dispersion: movers share stayers' location means, not their full distribution. The simulations of Section VI use this edge-potential economy. Third, by Proposition 1 any directed flow F with symmetrization $\bar{F}$—including heavily history-dependent ones—induces the same estimand, so the lemma's symmetric choice of F is itself without loss of generality.

S5 Descriptive Statistics for the Estimation Sample

Table 3 compares the estimation sample of Section VII—full-time male workers at firms above the 50th percentile of total worker-period observations—and the denser 75th-percentile universe of Supplemental Appendix S7 with the broader population of all full-time male stayers and one-time movers, over 2009–2016. The estimation samples are modestly older, somewhat higher-paid, and sit at substantially larger firms, but are otherwise similar.

Table 3: Descriptive Statistics: Estimation Sample versus Full Population

	All stayers & movers (full population)	Estimation sample (above median)	Sensitivity (above 75th pct.)
Worker–year observations	10,621,264	5,309,689	2,654,108
Age (years)	46.7 (9.6)	47.8 (9.4)	48.0 (9.4)
Annual gross earnings (euros)	52,491 (50,060)	55,310 (41,251)	57,838 (43,483)
Hourly wage (euros)	25.5 (24.5)	26.8 (20.0)	28.0 (21.0)
Firm size (employees)	780 (811)	1,416 (667)	1,895 (259)
One-time mover (share of worker-years)	—	0.16	0.12

Notes. CBS linked employer–employee data, 2009–2016. The full population is all full-time male stayers and one-time movers; the estimation sample restricts to firms above the 50th percentile of total worker-period observations (Section VII), and the sensitivity column to firms above the 75th percentile (Supplemental Appendix S7). All entries are worker-year means, with standard deviations in parentheses for continuous variables. Hourly wage is gross annual earnings spread over a 52-week, 40-hour year, adjusted for the contracted full-time factor. "One-time mover" is the share of worker-years contributed by workers whose single employer change is observed within the sample panel (a worker-*level* share of 13.9 percent in the estimation sample); it is not reported for the full population. Firm size is the employer's total headcount from register size-intervals.

S6 Robustness: Unadjusted Log Hourly Wage

The paper's outcome is the log gross hourly wage residualized on age and calendar-year indicators in a preliminary step. This section repeats the full analysis on the *unadjusted* log hourly wage, on the identical sample. Every conclusion is unchanged: the affine benchmark is rejected ($\widehat{\lambda}_M = 2.84$, CI $[2.49, 3.27]$), both isotropy restrictions reject, and the joint Wald is 99.1 (χ^2_2 $p = 3.0 \times 10^{-22}$).

Magnitudes shift moderately relative to the adjusted outcome—the mover coefficient rises to 0.167 and the transition projection to 0.184—with no sign or significance changes. Figure 2 shows the event study: pre-event coefficients from $r = -6$ onward are at most 0.014 in absolute value, while the post-event drift is more pronounced than in the adjusted specification, consistent with the age and year indicators absorbing life-cycle and calendar wage growth; the adjusted outcome, on which the paper's objects are defined, remains the main specification.

Table 4: Mover-Design Diagnostics, Unadjusted Log Hourly Wage

	Estimate	SE	95% CI
Panel A. Mover-design coefficient and cross-sectional benchmarks			
Mover regression $\widehat{\theta}$	0.167	0.005	[0.156, 0.177]
Transition projection $f(\widehat{\beta}_\Delta, \widehat{\lambda}_\Delta)$	0.184	0.006	[0.173, 0.197]
θ_{iso} (midpoint)	0.258	0.010	[0.239, 0.277]
θ_{var}	0.263	0.020	[0.226, 0.304]
θ_{sd}	0.513	0.019	[0.475, 0.551]
S_ψ^{eq}	0.253	0.017	[0.219, 0.287]
Panel B. Sorting geometry: cross-section versus transitions			
Cross-sectional slope $\widehat{\beta}_M$	−0.018	0.060	[−0.123, 0.107]
Transition slope $\widehat{\beta}_\Delta$	−0.161	0.051	[−0.248, −0.056]
Cross-sectional residual intensity $\widehat{\lambda}_M$	2.840	0.194	[2.489, 3.271]
Transition residual intensity $\widehat{\lambda}_\Delta$	3.849	0.247	[3.408, 4.391]
Panel C. Isotropy test			
Slope restriction $\widehat{g}_\beta$	−0.143	0.038	[−0.221, −0.068]
Residual restriction $\widehat{g}_Q$ ($\times 10^4$)	2.450	0.406	[1.647, 3.271]
Joint Wald $\mathcal{W}$ (χ^2_2 $p = 3.0 \times 10^{-22}$)	99.139	—	—

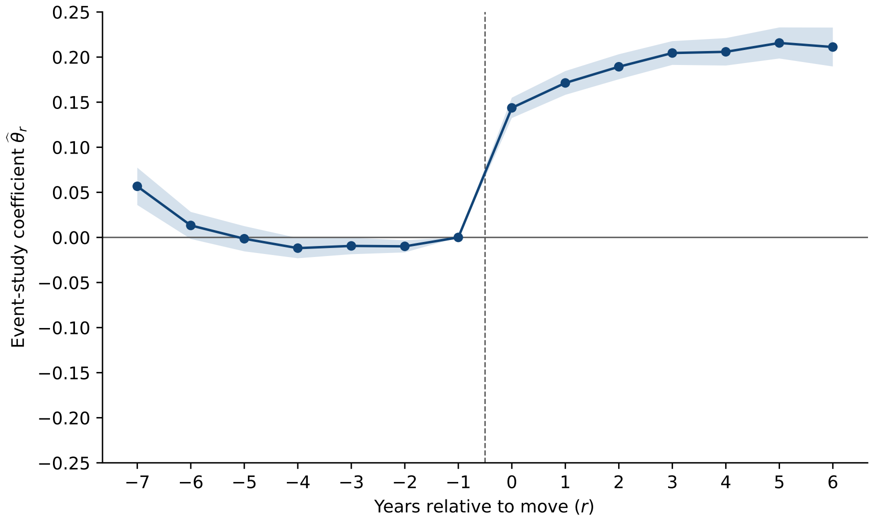
Notes. Identical sample, implementation, and inference as Table 2 of the paper; the only change is that the outcome is the raw log gross hourly wage. Both reported joint-test p-values are interior results; the bootstrap value is below 0.001.

S7 Sensitivity: Firms above the 75th Percentile

This section repeats the main analysis—adjusted log hourly wage, identical implementation and inference—on the denser universe of firms above the 75th percentile of total worker-period observations: 2,654,108 worker-years, 416,343 workers, 43,728 one-time movers, and 220 firms with a mean firm-year size of 1,604 workers. The diagnostic core retains 176 firms and 12,682 movers, or 29.0 percent of the 43,728 movers; both split connected components retain every mover, so attrition again reflects the complete-window endpoint requirement rather than graph disconnection. This is the regime in which the fixed-J, large-location asymptotic frame is least contestable.

Table 5 reports the results. The sorting geometry differs from the median-cut sample: among the largest firms the cross-sectional slope is positive and significant ($\widehat{\beta}_M = 0.83$, CI $[0.44, 1.32]$)—higher-premium firms employ higher-composition workers—and the transition slope is commensurate ($\widehat{\beta}_\Delta = 0.63$), so the slope restriction does not reject ($\widehat{g}_\beta = -0.19$, bootstrap $p = 0.36$). The residual

Figure 2: Event Study, Unadjusted Log Hourly Wage



Notes. As in Figure 1 of the paper, with the unadjusted log hourly wage as the outcome.

coordinate is decisive instead: residual composition is large in the cross-section ($\widehat{\lambda}_M = 5.43$) and transition routes oversample it by roughly seventy percent ($\widehat{\lambda}_\Delta = 9.24$), so the residual restriction rejects ($p < 0.001$) and the joint Wald is 15.3 (χ^2_2 $p = 4.7 \times 10^{-4}$; bootstrap $p < 0.001$). The two coordinates of the diagnostic therefore separate exactly as intended: routes reproduce the cross-sectional slope but not the residual geometry. With $\widehat{\beta}_M > 0$ the benchmark ordering of Corollary 1 reverses relative to the median-cut sample ($\theta_{\text{var}} \le \theta_{\text{iso}} \le S_\psi^{\text{eq}}$), and the mover coefficient ($\widehat{\theta} = 0.135$) lies below θ_{iso}, S_ψ^{eq}, and θ_{sd}, but above $\theta_{\text{var}} = 0.114$.

S8 Recovery Without Stationarity

The paper's recovery and inference results, and their empirical use, maintain Assumption 1. This final section records a separate deviation. It uses one common full-window location attribute and centered moments of the directed mover flow; the undirected-flow reduction is not carried over.

Fix a common calendar window $\mathcal{T}$, and assume every retained location-year cell in that window has positive population mass. For each location define the equally year-weighted population attribute

$$\bar{Y}_j := \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \mathbb{E}[Y_{it} \mid j(i,t) = j,\ i \text{ eligible at } t], \qquad \bar{\alpha}_j := \bar{Y}_j - \psi_j.$$

For the canonical mover regression, its full-sample leave-one-out counterpart first averages eligible workers within location-year and then averages those location-year means equally over $t \in \mathcal{T}$:

$$\overline{Y}_j^{(-i)} := \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \overline{Y}_{j,t}^{(-i)}, \qquad \delta_i^{(-i)} := \overline{Y}_{d(i)}^{(-i)} - \overline{Y}_{o(i)}^{(-i)}.$$

Table 5: Mover-Design Diagnostics, Firms above the 75th Percentile

	Estimate	SE	95% CI
Panel A. Mover-design coefficient and cross-sectional benchmarks			
Mover regression $\widehat{\theta}$	0.135	0.006	[0.123, 0.147]
Transition projection $f(\widehat{\beta}_\Delta, \widehat{\lambda}_\Delta)$	0.137	0.007	[0.124, 0.150]
θ_{iso} (midpoint)	0.208	0.024	[0.162, 0.257]
θ_{var}	0.114	0.011	[0.094, 0.134]
θ_{sd}	0.338	0.016	[0.306, 0.366]
S_ψ^{eq}	0.303	0.046	[0.217, 0.397]
Panel B. Sorting geometry: cross-section versus transitions			
Cross-sectional slope $\widehat{\beta}_M$	0.829	0.222	[0.439, 1.318]
Transition slope $\widehat{\beta}_\Delta$	0.635	0.089	[0.472, 0.829]
Cross-sectional residual intensity $\widehat{\lambda}_M$	5.434	0.907	[3.716, 7.232]
Transition residual intensity $\widehat{\lambda}_\Delta$	9.242	0.630	[8.151, 10.510]
Panel C. Isotropy test			
Slope restriction $\widehat{g}_\beta$	−0.194	0.214	[−0.653, 0.186]
Residual restriction $\widehat{g}_Q$ ($\times 10^4$)	1.132	0.296	[0.603, 1.741]
Joint Wald $\mathcal{W}$ (χ^2_2 $p = 4.7 \times 10^{-4}$)	15.308	—	—

Notes. As in Table 2 of the paper, restricting to firms above the 75th percentile of total worker-period observations. N = 2,654,108 worker-years, 416,343 workers, 372,615 stayers and 43,728 one-time movers, 220 firms; the diagnostic core retains 176 firms and 12,682 movers, all with complete canonical leave-one-out gaps. The mover regression, the transition moments, and the cross-sectional benchmarks are all computed on this common core sample (n = 12,682). SEs and percentile CIs from the worker-level multiplier bootstrap with 1,000 draws; the SE for $\widehat{\theta}$ is the conventional OLS SE, and its CI the corresponding analytic normal interval. These conventional quantities treat the generated leave-one-out location averages as fixed and do not incorporate their sampling variation under the paper's asymptotics. The midpoint and projection identities hold in the estimates as reported. The reported χ^2_2 and bootstrap joint-test p-values are interior results; the bootstrap value is below 0.001.

The cross-split diagnostic estimates the same population attribute with the split-specific common full-window means of Supplemental Appendix S3. The mover regression includes an intercept, $\Delta Y_i = c + \theta\, \delta_i^{(-i)} + u_i$, and all transition moments are centered over movers. In the large-cell limit its slope is therefore

$$\theta_\infty = \frac{\mathrm{Cov}_F(\Delta\psi_i, \Delta\bar{Y}_i)}{\mathrm{Var}_F(\Delta\bar{Y}_i)}.$$

Let F_{jk} be the directed mover flow, normalized to sum to one, $\bar{F}_{jk} := (F_{jk} + F_{kj})/2$, and

$$m_F(f) := \sum_{j,k} F_{jk}(f_k - f_j).$$

For any location-level f, g, direct expansion gives

$$\mathrm{Cov}_F(\Delta f, \Delta g) = \sum_{j<k} 2\bar{F}_{jk}(f_j - f_k)(g_j - g_k) - m_F(f)m_F(g). \tag{S21}$$

Thus the raw second-moment term is determined by $\bar{F}$, but the centered covariance generally retains the direction of F through the mean product. Under Assumption 1, $m_F(f) = 0$ for every f, which is exactly what restores Proposition 1.

Weight cross-sectional moments by mover incidence $\mu_j := \{\sum_k F_{jk} + \sum_k F_{kj}\}/2$, and form transition moments as centered moments under F. Define the cross-sectional and transition projection pairs from those moments exactly as in the main text.

Theorem S3 (Recovery against a common full-window cross-section)**.** *Maintain the AKM model, exogenous mobility, and the large-location limit, but not Assumption 1. Suppose*

$$\mathrm{Var}_\mu(\psi_j) > 0, \qquad \mathrm{Var}_F(\Delta\psi_i) > 0, \qquad \mathrm{Var}_\mu(\bar{Y}_j) > 0, \qquad \mathrm{Var}_F(\Delta\bar{Y}_i) > 0.$$

Then:

(i) $\theta_\infty = f(\beta_\Delta, \lambda_\Delta)$, *with* $f(\beta, \lambda) = (1+\beta)/\{(1+\beta)^2 + \lambda\}$.

(ii) *The transition and cross-sectional projection pairs coincide if and only if the centered directed-flow moments reproduce the full-window cross-sectional sorting projection. When they coincide,*

$$\theta_\infty = \theta_{\mathrm{iso}} = \tfrac{1}{2}\big(\theta_{\mathrm{var}} + S_\psi^{\mathrm{eq}}\big),$$

with all cross-sectional objects evaluated on the μ-weighted full-window cross-section.

(iii) *Under the remaining fixed-J, stable-core, moment, interior, and nonsingular-covariance conditions of Proposition 2 and Supplemental Appendix S3, the same fixed-dimensional smooth-map argument yields the worker multiplier inference for the centered restriction vector.*

Proof. The common full-window regressor converges to $\Delta\bar{Y}_i = \Delta\psi_i + \Delta\bar{\alpha}_i$, while residual exogeneity gives $\mathrm{Cov}_F(\Delta\varepsilon_i, \Delta\bar{Y}_i) = 0$. Dividing the centered numerator and denominator by $\mathrm{Var}_F(\Delta\psi_i)$ gives (i). Part (ii) is the same ratio algebra as Theorem 1 and Corollary 1, now using the centered directed-flow moments in (S21). For (iii), the centered moments and the common full-window location means remain smooth functions of finitely many worker averages, so the influence expansion and multiplier argument apply under the stated conditions. □

Dropping stationarity removes balance restrictions and therefore weakly enlarges the admissible set of directed configurations. It does not preserve the undirected-flow reduction: the mean correction in (S21) is the precise additional object.