

Technical Supplement to FEVC: Leave-Out Bias-Corrected Variance Decompositions in Stata

Johannes Schmieder

October 2026

Abstract

This supplement gives the assumptions, derivations, numerical procedures and source-bound evidence supporting the accompanying FEVC article. Exact leave-out and centering identities are distinguished from finite-probe approximations and experimental component inference. The evidence includes failures, unavailable intervals and numerical sensitivity. The source notes and original experimental receipts are preserved in the local replication archive.

Contents

S1 Framework, notation and populations	3
S1.1 Model and notation guide	3
S1.2 Frequency compression and target mass	4
S1.3 Controls and Frisch–Waugh–Lovell reduction	4
S1.4 Retained population	5
S2 Exact leave-out estimation and centering	5
S2.1 Deletion blocks and identification	6
S2.2 Why the block correction removes plug-in bias	6
S2.3 Observation-deletion setup	7
S2.4 Centering and an exact correction	7
S2.5 Worker–firm matches	9
S2.6 Arbitrary correlated deletion blocks	10
S3 Randomized computation and numerical uncertainty	14
S3.1 Normalized leverage estimates and finite-probe adjustment	14
S3.2 Inverse adjustment, controls and blocks	15
S3.3 Physical directions and target contractions	16
S3.4 Randomized Corrected centering	17
S3.5 Numerical covariance from all point-estimation probes	17

S4 Fixed-effect projections	19
S4.1 Projection covariance by deletion block	19
S4.2 Normalization and supported routes	20
S5 Experimental component inference	20
S5.1 Independent units for movers and stayers	21
S5.2 The structured variance calculation	22
S5.3 High rank and one dominant mode	24
S5.4 Exact observation inference and cross-target covariance	26
S5.5 Mean inference: fixed-origin reference and omitted randomness	26
S5.6 Exact observation-deletion algebra	27
S5.7 High-rank covariance and uncentered noise probes	27
S5.8 One leading mode and its remainder	28
S5.9 Frequency weights and declared matches	29
S5.10 What estimating the mean omits	30
S5.11 JLA and numerical uncertainty	32
S5.12 Supported requests and interpretation	33
S5.13 Mover-only public Veneto example	34
S5.14 Pooled public Veneto component inference	36
S6 Statistical validation	37
S6.1 Fixed-design bias experiment	37
S6.2 Random-projection approximation	38
S6.3 Projection experiment	40
S6.4 A numerical illustration	41
S6.5 Pooled structured component inference	42
S6.6 Observation and mover-only calibration by variance estimator	43
S6.7 Numerical sensitivity	46
S6.8 Observation and mover-only Mean diagnostics	49
S6.9 Observation and mover-only Mean assessment	50
S6.10 Observation and mover-only diagnostic cells	52
S7 Computational evidence and reproducibility	53
S7.1 Timing, memory, and numerical settings	53
S7.2 Synthetic graph classes	53
S7.3 Agreement between randomized estimates	54
S7.4 Source-only Mata computation	54
S7.5 Supporting Veneto results	54
S7.6 Local match-inference size ladder	57
S7.7 Reproducing documents and locating evidence	58
S8 Implementation comparisons	59

S1 Framework, notation and populations

This supplement develops the estimators and computational approximations used by Schmieder (2026), the accompanying October 2026 article. Sections S2–S5 separate exact expectation identities, randomized calculations and experimental sampling inference. Sections S6–S8 report evidence by method and population. The replication records bind each experiment to its software source, inputs and numerical settings.

S1.1 Model and notation guide

Following Kline et al. (2020, pp. 1861–1862, equation (1)), on the physically expanded retained sample let

$$y_i = x_i' \beta + \varepsilon_i, \quad i = 1, \dots, n, \quad S_{xx} = X'X, \quad \hat{\beta} = S_{xx}^{-1} X'y, \quad \theta_t = \beta' A_t \beta. \quad (\text{S1.1})$$

Here X has k linearly independent columns after fixed-effect normalization, and A_t is a known symmetric $k \times k$ target matrix. Set

$$P = X S_{xx}^{-1} X', \quad M = I_n - P, \quad B_t = X S_{xx}^{-1} A_t S_{xx}^{-1} X', \quad e = My. \quad (\text{S1.2})$$

For observation deletion define $d_{ti} = B_{t,ii}/M_{ii}$ and the KSS leave-out kernel

$$C_t = B_t - \frac{1}{2} \{ \text{diag}(d_t) M + M \text{diag}(d_t) \}, \quad C_{t,ii} = 0. \quad (\text{S1.3})$$

This is equation (8) of Kline et al. (2020, p. 1873); its off-diagonal entries are $C_{t,ij} = B_{t,ij} - M_{ij}(d_{ti} + d_{tj})/2$. The uncentered leave-out estimate is $y' C_t y$. The subscript t always identifies a target when more than one target is under discussion.

Table S1.1: Notation and conditioning conventions

Symbol	Meaning
i, n, k	Physical observation, physical sample size, identified coefficient dimension.
$w(i), j(i)$	Worker and firm of physical observation i .
s, N_s, f_s	Stored row, number of stored rows, positive integer frequency.
g, G, m_g	Deletion block, number of blocks, physical block size.
t	Component label: α , ψ , $\alpha\psi$, or total T .
b, R, T_p	Random direction, leverage-probe count, target-probe count.
X_s, W_f	Stored-row design and $\text{diag}(f_s)$; X without a qualifier is physical.
a_s, K_a	Complete stored-row target mass and its centered mass matrix.
Z, γ	Regression controls and their coefficients.
J, N_B	Block-indicator matrix and $\text{diag}(m_g)$.
$\mathcal{D}$	Fixed design, retained population, weights, deletion partition and loadings.
$\bar{y}, c, c_0$	Observed mean, fixed centering constant, genuinely nonrandom reference mean.

Expectations of statistical identities condition on $\mathcal{D}$; when an exact identity needs only X , this is a special case. Selection, weights or projection covariates constructed from outcomes require their conditional mean-zero and dependence assumptions to be justified in the retained population. Numerical expectations instead hold the complete observed dataset fixed and average over independent computational directions. Neither conditioning on a fitted offset nor treating

an observed mean as fixed makes it nonrandom under repeated sampling. Spectral eigenvectors and eigenvalues refer to the KSS decomposition of $S_{xx}^{-1/2} A_t S_{xx}^{-1/2}$; Q is reserved for an orthonormal spectral eigenvector matrix and D for its eigenvalue matrix.

S1.2 Frequency compression and target mass

A stored row s represents f_s physical copies with equal recorded outcome and regressors. With $n = \sum_{s=1}^{N_s} f_s$ and $W_f = \text{diag}(f_s)$,

$$S_{xx} = X'_s W_f X_s, \quad X'y = X'_s W_f y_s, \quad \tilde{X}_s = W_f^{1/2} X_s, \quad \tilde{y}_s = W_f^{1/2} y_s. \quad (\text{S1.4})$$

This is an algebraic compression of recorded rows. Statistical independence is an assumption about physical error units, not something created by duplicating a stored value. Observation deletion removes one physical copy; deleting a stored row instead removes all its copies. Whole-match deletion removes every copy in the declared match.

Let $a_s > 0$ be the complete target mass of row s , $\rho_s = a_s / \sum_v a_v$, and $K_a = \text{diag}(\rho) - \rho\rho'$. Each physical copy receives mass a_s/f_s . Without `targetweight()`, $a_s = f_s$. Specifying target mass changes the estimand, without changing S_{xx} or the regression fit. Write the stored-row design as $[X_{\text{worker}} \ X_{\text{firm}} \ Z_s]$ and let $X_\alpha = [X_{\text{worker}} \ 0 \ 0]$ and $X_\psi = [0 \ X_{\text{firm}} \ 0]$. In identified coefficient coordinates,

$$\begin{aligned} A_\alpha &= X'_\alpha K_a X_\alpha, & A_\psi &= X'_\psi K_a X_\psi, \\ A_{\alpha\psi} &= \frac{1}{2}(X'_\alpha K_a X_\psi + X'_\psi K_a X_\alpha), & A_T &= A_\alpha + A_\psi + 2A_{\alpha\psi}. \end{aligned} \quad (\text{S1.5})$$

The covariance target contains the factor $1/2$; the total contains twice that covariance. Controls have zero target rows and columns. These centered targets are invariant to admissible location normalizations.

S1.3 Controls and Frisch–Waugh–Lovell reduction

For joint controls, write $X = [X_{\text{FE}} \ Z]$, where $X_{\text{FE}} = [X_{\text{worker}} \ X_{\text{firm}}]$. If $Z'Z$ is nonsingular, the two-way coefficient block solves

$$X'_{\text{FE}} M_Z X_{\text{FE}} \hat{\delta} = X'_{\text{FE}} M_Z y, \quad M_Z = I - Z(Z'Z)^{-1}Z'. \quad (\text{S1.6})$$

The target matrices have zero control rows and columns, but the block projection P_g still uses the full joint design. Consequently, deleting a block changes the control coefficient and the two-way coefficients together. This is the general KSS linear-model cross-fit: the coefficient estimated without g is a function only of observations outside g .

Under fixed-offset controls, the full model is estimated once and the command forms $y^* = y - Z\hat{\gamma}$. The subsequent design is X_{FE} , not $[X_{\text{FE}} \ Z]$. Equation (S2.5) then describes deletion from the two-way fit to y^* while holding $Z\hat{\gamma}$ fixed. This is why joint and fixed-offset controls can give different finite-sample corrections even when they share the same full-sample fitted values. If $\hat{\gamma}$ was estimated on that same full sample, y^* at and outside g contains a random quantity influenced by g . Equation (S2.10) does not then follow from block independence alone. Same-sample fixed offset is therefore an approximation requiring a separate justification. The scalable component implementation and its omitted nuisance uncertainty are described in section S5.

S1.4 Retained population

The match-deletion sample is constructed from complete cases for the requested outcome, identifiers, controls, and weights. The command selects a largest connected component and repeatedly removes workers with only one retained deletion unit, worker articulation vertices, and deletion blocks whose removal disconnects the graph. Default pair-defined matches count firms; explicit `deletionid()` counts declared blocks, including parallel blocks within one worker–firm coordinate. These remain separate edges. The sample is recomputed until these restrictions stop changing it. A largest component is chosen by number of firms, then physical observation mass; a tie on both criteria withholds a result.

Worker-articulation pruning is stronger than requiring connectedness after each match deletion. For example, suppose worker w is observed at firms A, B, C, D , worker u at A, B , and worker v at C, D . Removing w disconnects the network, although deleting any single worker–firm match leaves it connected.

Initial mover/stayer eligibility is frozen from complete-case histories. Observation deletion and default matches count estimation firms; explicit match deletion IDs count declared units. Thus a worker at one estimation firm with two declared blocks can enter the mover graph. If retained after graph and rank checks, the worker’s blocks receive whole-block deletion under either `stayers()` choice. Multiple blocks do not guarantee final retention. With `stayers(both)`, the command adds original one-unit stayers attached to retained mover firms if they have at least two physical observations. Their rows enter the regression jointly with movers, but their deletion units are individual physical observations. Workers made into stayers by pruning are excluded. Under `stayers(movers)`, the population consists only of the retained movers. Target weights are normalized on the resulting sample, and the full regression design must satisfy the deleted-rank conditions in Section S2.1.

Structured match component inference uses the same retained population and mixed deletion partition: mover units and individual physical stayer observations. It assumes independence across all these units while allowing dependence within mover units. Its variance model has separate coefficients for the two unit types, fitted jointly on the combined design (Section S5). Estimated controls are held fixed in this calculation. Observation-deletion inference instead uses individual physical observations throughout its requested mover/stayer population.

S2 Exact leave-out estimation and centering

This section states expectation identities for exact inverse actions. They do not establish unbiasedness of a finite-probe randomized calculation or coverage of an interval. The three command conventions are None (the uncentered KSS proxy), Mean (subtract the retained physical-observation mean from its outcome factor), and Corrected (remove that mean-induced bias). Mean is the current default. Corrected is a point-estimation option; the covariance identities below are mathematical results, not additional public projection or component-inference capabilities.

S2.1 Deletion blocks and identification

For the deletion identities, let X and y denote the expanded physical design and outcome, and X_s, y_s the stored rows. The full-sample compression in (S1.4) gives

$$X'X = X'_s W_f X_s, \quad X'y = X'_s W_f y_s, \quad (\text{S2.1})$$

with target mass a_s/f_s per physical copy. The deletion unit must also be specified: deleting one weighted stored row removes all f_s copies. In the block formulas below, X_g, y_g therefore contain physical rows. A weighted block is interchangeable only when the deletion removes every copy of its stored rows, as in whole-match deletion. Work in any identified coordinate system for which

$$S_{xx} = X'X \quad (\text{S2.2})$$

is nonsingular. Partition rows into deletion blocks g , write X_g and y_g for a block, and define

$$P_g = X_g S_{xx}^{-1} X'_g, \quad M_g = I - P_g. \quad (\text{S2.3})$$

Let $S_{xx,-g} = S_{xx} - X'_g X_g$ and let $\hat{\beta}_{-g}$ be the least-squares coefficient after deleting block g . Provided $S_{xx,-g}$ is nonsingular, the standard block-deletion identity gives

$$y_g - X_g \hat{\beta}_{-g} = M_g^{-1} (y_g - X_g \hat{\beta}) \quad (\text{S2.4})$$

$$= M_g^{-1} \hat{e}_g. \quad (\text{S2.5})$$

The determinant lemma also gives

$$\det(S_{xx,-g}) = \det(S_{xx}) \det(M_g). \quad (\text{S2.6})$$

Thus the deleted coefficient problem is identified if and only if the block residual maker is nonsingular in the identified coordinate representation.

S2.2 Why the block correction removes plug-in bias

Suppose

$$y = X\beta + \varepsilon, \quad \mathbb{E}(\varepsilon_g | \mathcal{D}) = 0, \quad \mathbb{E}(\varepsilon_g \varepsilon'_h | \mathcal{D}) = 0 \quad (g \neq h), \quad (\text{S2.7})$$

while the covariance matrix Ω_g inside a block is unrestricted. For a symmetric target $A = A_t$, let

$$B_{t,g} = X_g S_{xx}^{-1} A S_{xx}^{-1} X'_g. \quad (\text{S2.8})$$

The conditional expectation of the plug-in quadratic form is

$$\mathbb{E}(\hat{\beta}' A \hat{\beta} | \mathcal{D}) = \beta' A \beta + \sum_g \text{tr}(B_{t,g} \Omega_g). \quad (\text{S2.9})$$

Because $\hat{\beta}_{-g}$ depends only on errors outside block g , (S2.7) implies

$$\mathbb{E}\{y'_g B_{t,g} (y_g - X_g \hat{\beta}_{-g}) | \mathcal{D}\} = \text{tr}(B_{t,g} \Omega_g). \quad (\text{S2.10})$$

Summing (S2.10) and using (S2.5) yields

$$\hat{b}_A = \sum_g y'_g B_{t,g} M_g^{-1} \hat{e}_g. \quad (\text{S2.11})$$

Combining this correction with (S2.9) gives

$$\mathbb{E}(\hat{\theta}_t^{\text{KSS}} | \mathcal{D}) = \beta' A \beta. \quad (\text{S2.12})$$

This fixed-design result requires conditional mean zero, cross-block uncorrelatedness, and deleted-sample identification. The selection of blocks determines which error covariances enter the correction.

S2.3 Observation-deletion setup

For this subsection $r_i = e_i/M_{ii} = y_i - x_i' \hat{\beta}_{-i}$, $B_{ii} = B_{t,ii}$ for a fixed target $A = A_t$, and $\bar{y} = n^{-1} \sum_i y_i$. The errors have finite second moments and are uncorrelated across observations. The uncentered proxy $y_i r_i$ has expectation σ_i^2 .

S2.4 Centering and an exact correction

S2.4.1 The bias introduced by the sample mean

Define the centered leave-out variance estimate

$$\hat{\sigma}_{i,c}^2 = (y_i - \bar{y})r_i. \quad (\text{S2.13})$$

The superscript 2 denotes the variance being estimated: neither this estimate nor its KSS counterpart is necessarily nonnegative. To collect the observations, write

$$P = X S_{xx}^{-1} X', \quad M = I_n - P, \quad d_i = 1 - P_{ii} > 0, \quad \Delta_0 = \text{diag}(d_i), \quad e = My. \quad (\text{S2.14})$$

Here $e_i = d_i r_i$ is the ordinary OLS residual. Bold symbols $\boldsymbol{\sigma}^2$ and $\hat{\boldsymbol{\sigma}}_c^2$ collect the true and centered estimated error variances.

Since $\mathbb{E}[r_i | \mathcal{D}] = 0$, the bias from subtracting $\bar{y}$ is its covariance with the leave-out residual:

$$\mathbb{E}[\hat{\sigma}_{i,c}^2 | \mathcal{D}] - \sigma_i^2 = -\text{Cov}(\bar{y}, r_i | \mathcal{D}) = -\frac{1}{n d_i} \sum_j M_{ij} \sigma_j^2. \quad (\text{S2.15})$$

The bias can have either sign. It is zero under homoskedasticity when X spans a constant, because $M\mathbf{1} = 0$. The factor $1/n$ alone does not establish that it is negligible: small d_i can magnify it. Multiplying by d_i and collecting observations gives the useful identity

$$\boxed{\mathbb{E}[\Delta_0 \hat{\boldsymbol{\sigma}}_c^2 | \mathcal{D}] = \left(\Delta_0 - \frac{M}{n} \right) \boldsymbol{\sigma}^2.} \quad (\text{S2.16})$$

This identity does not require an intercept.

S2.4.2 Unbiasedness and existence

Proposition S2.1 (Exact centered correction). *Suppose $d_i > 0$ for every i and $n - k \geq 2$. The system*

$$\boxed{\left(\Delta_0 - \frac{M}{n} \right) \tilde{\boldsymbol{\sigma}}^2 = \Delta_0 \hat{\boldsymbol{\sigma}}_c^2} \quad (\text{S2.17})$$

has a unique solution with $\mathbb{E}[\tilde{\sigma}_i^2 | \mathcal{D}] = \sigma_i^2$. Consequently,

$$\tilde{\theta} = \hat{\theta}_{\text{PI}} - \sum_i B_{ii} \tilde{\sigma}_i^2 \quad (\text{S2.18})$$

is unbiased for θ . The system is singular when $n - k = 1$.

Proof. The matrix $\Delta_0^{-1/2} M \Delta_0^{-1/2}$ is positive semidefinite, has unit diagonal, trace n , and rank $n - k$. If the rank is at least two, its largest eigenvalue is strictly below n . Therefore

$$\Delta_0^{-1/2} \left(\Delta_0 - \frac{M}{n} \right) \Delta_0^{-1/2} = I_n - \frac{\Delta_0^{-1/2} M \Delta_0^{-1/2}}{n} \succ 0. \quad (\text{S2.19})$$

At rank one, its only positive eigenvalue is n , making the system singular. In the nonsingular case, taking expectations in (S2.17) and using (S2.16) proves unbiasedness of the corrected variances. Equation (S2.9) then proves the component result. $\square$

The correction uses only a known design matrix and observed centered variance estimates. It does not require a preliminary estimate of the unknown variances. Existence does not guarantee good conditioning: an eigenvalue close to the singular boundary can amplify numerical error and sampling noise. A singular full system also need not rule out correction of a particular scalar target.

If $Xb_0 = \mathbf{1}$, both factors in (S2.13) are invariant to $y \mapsto y + c\mathbf{1}$, and so is $\tilde{\sigma}^2$. The complete component estimate is invariant when $Ab_0 = 0$, as for appropriately centered fixed-effect variance components. Other quadratic targets can change with the outcome level; their unbiasedness still holds.

As in KSS's Remark 1, the same corrected variances estimate the covariance matrix of OLS coefficients:

$$\tilde{V}[\hat{\beta}] = S_{xx}^{-1} \left(\sum_i x_i x_i' \tilde{\sigma}_i^2 \right) S_{xx}^{-1}. \quad (\text{S2.20})$$

For any fixed ℓ , $\ell' \tilde{V}[\hat{\beta}] \ell$ is unbiased for $\text{Var}(\ell' \hat{\beta} \mid \mathcal{D})$. Thus one corrected vector serves all quadratic components and linear-contrast variances. These expectation results do not imply positive semidefiniteness, minimum mean squared error (MSE), or valid small-sample normal-Wald inference. Nor do KSS's asymptotic inference results automatically apply to the modified estimator.

S2.4.3 One additional weighted regression

A stronger leverage condition gives a convenient computation:

$$nd_i > 1 \quad \text{for every } i, \quad \tau_i = \frac{nd_i}{nd_i - 1}, \quad T = \text{diag}(\tau_i). \quad (\text{S2.21})$$

Regress the centered variance estimates on the original design with these weights, and add a scaled residual from this auxiliary regression:

$$\hat{\gamma} = (X' T X)^{-1} X' T \hat{\sigma}_c^2. \quad (\text{S2.22})$$

$$\boxed{\tilde{\sigma}_i^2 = \hat{\sigma}_{i,c}^2 + \frac{\hat{\sigma}_{i,c}^2 - x_i' \hat{\gamma}}{nd_i - 1}.} \quad (\text{S2.23})$$

These are computational weights, not frequency counts. The dependent variable is the centered variance estimate, rather than the original outcome.

To verify the formula, write $\tilde{\sigma}^2 = T \hat{\sigma}_c^2 - (T - I_n) X \hat{\gamma}$. The weighted normal equations imply $X' \tilde{\sigma}^2 = S_{xx} \hat{\gamma}$, hence $P \tilde{\sigma}^2 = X \hat{\gamma}$. Multiplication of (S2.23) by $nd_i - 1$ then gives $(n \Delta_0 - I_n + P) \tilde{\sigma}^2 = n \Delta_0 \hat{\sigma}_c^2$, which is (S2.17). No independence between the auxiliary fit and the original fit is needed.

Conditional on the original residuals and leverages, the extra work is one weighted regression and linear passes over the observations. It requires no new leverage diagonal or separate correction fit for each target. The condition $nd_i > 1$ permits this positive-weight implementation; it is stronger than the existence condition in Proposition S2.1. When it fails, the symmetric system (S2.17) can still be solved if $n - k \geq 2$.

S2.5 Worker–firm matches

Partition the n physical observations into G disjoint matches of sizes m_g , with $\sum_g m_g = n$. Errors may be arbitrarily correlated within matches but are uncorrelated across matches. Suppose every regressor is constant within a match:

$$X_g = \mathbf{1}_g x'_g, \quad Y_g = \frac{\mathbf{1}'_g y_g}{m_g} = x'_g \beta + \eta_g, \quad \sigma_g^2 = \text{Var}(\eta_g \mid \mathcal{D}) = \frac{\mathbf{1}'_g \Omega_g \mathbf{1}_g}{m_g^2}. \quad (\text{S2.24})$$

Here Ω_g is the within-match error covariance matrix. Worker and firm indicators meet the constant-regressor condition. The original OLS fit is the weighted regression of match averages Y_g on x_g , with weights m_g .

Let X_c collect the rows x'_g , and define

$$N_B = \text{diag}(m_g), \quad S_{xx} = X'_c N_B X_c, \quad P_w = N_B^{1/2} X_c S_{xx}^{-1} X'_c N_B^{1/2}, \quad M_w = I_G - P_w. \quad (\text{S2.25})$$

The match leverage complement is $d_g = (M_w)_{gg} = 1 - m_g x'_g S_{xx}^{-1} x_g$, with $d_g > 0$ required for leave-match-out identification. Write $\Delta_c = \text{diag}(d_g)$. Retain the physical-observation mean and form

$$\bar{y} = \frac{1}{n} \sum_g m_g Y_g, \quad r_g = \frac{Y_g - x'_g \hat{\beta}}{d_g} = Y_g - x'_g \hat{\beta}_{-g}, \quad \hat{\sigma}_{g,c}^2 = (Y_g - \bar{y}) r_g. \quad (\text{S2.26})$$

In this section, r_g is the scalar held-out residual for the match average and σ_g^2 is its error variance.

The symmetric counterpart of (S2.16) is

$$\mathbb{E}[N_B \Delta_c \hat{\sigma}_c^2 \mid \mathcal{D}] = \left(\Delta_c - \frac{N_B^{1/2} M_w N_B^{1/2}}{n} \right) N_B \sigma^2. \quad (\text{S2.27})$$

Indeed $\mathbb{E}[Y_g r_g \mid \mathcal{D}] = \sigma_g^2$, and $\text{Cov}(\bar{y}, r_g \mid \mathcal{D}) = \sum_h m_h \sigma_h^2 (\mathbf{1}\{g = h\} - m_h x'_g S_{xx}^{-1} x_h) / (nd_g)$. Thus the match correction is obtained by replacing the left side of (S2.27) by its observed value and solving for $N_B \tilde{\sigma}^2$.

Proposition S2.2 (One weighted match regression). *If $nd_g > m_g$ for every match, define*

$$\tau_g = \frac{nd_g}{nd_g - m_g}, \quad \hat{\gamma}_c = \left(\sum_g m_g \tau_g x_g x'_g \right)^{-1} \sum_g m_g \tau_g x_g (m_g \hat{\sigma}_{g,c}^2). \quad (\text{S2.28})$$

Then the corrected estimate

$$\boxed{\tilde{\sigma}_g^2 = \hat{\sigma}_{g,c}^2 + \frac{m_g \hat{\sigma}_{g,c}^2 - x'_g \hat{\gamma}_c}{nd_g - m_g}} \quad (\text{S2.29})$$

is unbiased for σ_g^2 . The auxiliary regression has dependent variable $m_g \hat{\sigma}_{g,c}^2$ and weight $m_g \tau_g$.

Proof. Put $T_c = \text{diag}(\tau_g)$ and $t = N_B \tilde{\sigma}^2$. The formula gives $t = T_c N_B \hat{\sigma}_c^2 - (T_c - I_G) X_c \hat{\gamma}_c$. The normal equations imply $X_c' N_B t = S_{xx} \hat{\gamma}_c$. Multiplying (S2.29) by $m_g(nd_g - m_g)$, collecting matches, and substituting the normal equations gives

$$(n\Delta_c - N_B)t + N_B X_c S_{xx}^{-1} X_c' N_B t = n N_B \Delta_c \hat{\sigma}_c^2.$$

Since $N_B^{1/2} M_w N_B^{1/2} = N_B - N_B X_c S_{xx}^{-1} X_c' N_B$, this is the system in (S2.27). Its unique solution is unbiased by that expectation identity. Positive weights ensure nonsingularity, as also follows from the rank argument below. $\square$

With every $d_g > 0$, the match correction exists if and only if $G - k \geq 2$. To see this, normalize its symmetric matrix by $\Delta_c^{-1/2}$ on both sides. The subtracted matrix

$$\frac{1}{n} \Delta_c^{-1/2} N_B^{1/2} M_w N_B^{1/2} \Delta_c^{-1/2} \quad (\text{S2.30})$$

is positive semidefinite, has rank $G - k$, and has trace $\sum_g m_g/n = 1$. The eigenvalue argument in Proposition S2.1 applies. The condition concerns residual degrees of freedom across matches; repeated observations within a match do not add independent match-level information. For a connected worker–firm graph without other regressors, $G - k$ is the number of independent cycles. Again, $nd_g > m_g$ is a stronger condition for positive regression weights, not the existence condition.

The corrected match variances enter the targets with squared counts:

$$\begin{aligned} \tilde{\theta} &= \hat{\theta}_{\text{PI}} - \sum_g m_g^2 B_{gg} \tilde{\sigma}_g^2, & B_{gg} &= x_g' S_{xx}^{-1} A S_{xx}^{-1} x_g, \\ \tilde{V}_\ell &= \sum_g m_g^2 (\ell' S_{xx}^{-1} x_g)^2 \tilde{\sigma}_g^2. \end{aligned} \quad (\text{S2.31})$$

For equal match sizes m , $n = mG$, and the calculation reduces to the observation formula applied to match averages with sample size G . For unequal sizes, replacing n by G is incorrect. The mean, dependent variable, regression weights, and target weights must retain their counts.

Time-varying controls invalidate the constant-regressor assumption. Partialling them out once does not make the residualized rows constant within matches. Deleting a whole worker history while retaining that worker’s unrestricted effect also violates deleted-design identification; that case requires a different identified parameterization or transformation.

S2.6 Arbitrary correlated deletion blocks

The match calculation in Section S2.5 uses constant regressors within deletion units. This section allows regressors to vary within units and recovers the full covariance blocks. Using both orientations of the centered product reduces the unresolved bias to one scalar per block. A directional leverage condition then permits one additional regression with rank-one block weight adjustments.

S2.6.1 Block notation and the source of bias

Partition the n physical observations into G disjoint deletion units $\mathcal{I}_g$, of sizes m_g , so $\sum_g m_g = n$. Let $\mathbf{1}_g$ denote the m_g -vector of ones. Retain the original X, S_{xx}, P, M , but replace the diagonal covariance assumption by

$$\Omega = \text{Var}(\varepsilon \mid \mathcal{D}) = \text{blockdiag}(\Omega_1, \dots, \Omega_G). \quad (\text{S2.32})$$

Errors in different deletion units are uncorrelated; within each unit Ω_g is arbitrary and finite. Independence across units is sufficient but not necessary for the expectation calculations. All regressors, including any controls, belong to X . Assume every X_{-g} has full column rank. Then

$$\Delta_g = M_{gg} = I_{m_g} - P_{gg} \succ 0, \quad \mathcal{D}_B = \text{blockdiag}(\Delta_g), \quad \mathcal{R} = \mathcal{D}_B^{-1}M. \quad (\text{S2.33})$$

For completeness, the full and deleted normal equations imply $\hat{\beta} - \hat{\beta}_{-g} = S_{xx}^{-1}X'_g r_g$, where $r_g = y_g - X_g \hat{\beta}_{-g}$. Thus $(My)_g = r_g - P_{gg}r_g = \Delta_g r_g$, proving $r = \mathcal{R}y$. In particular, $\mathcal{R}_{gg} = I_{m_g}$ and $\mathcal{R}X = 0$. For a square matrix B , write $\text{sym}(B) = (B + B')/2$. Let $z = y - \bar{y}\mathbf{1}$ and define

$$\hat{\Omega}_g^M = \text{sym}(z_g r'_g), \quad u_g = \Omega_g \mathbf{1}_g, \quad u = (u'_1, \dots, u'_G)'. \quad (\text{S2.34})$$

Here u_g is a vector of covariance row sums, not a vector of marginal variances. Since $\Omega \mathbf{1} = u$,

$$c_g := \text{Cov}(r_g, \bar{y} \mid \mathcal{D}) = \frac{1}{n}(\mathcal{R}u)_g, \quad \mathbb{E}[\hat{\Omega}_g^M \mid \mathcal{D}] = \Omega_g - \text{sym}(\mathbf{1}_g c'_g). \quad (\text{S2.35})$$

Thus only n nuisance quantities are needed to remove the centering bias, even though the blocks together can contain many more covariance entries.

S2.6.2 Use both product orientations to isolate the bias

The useful simplification is to use the two orientations of the centered product before discarding their difference. Define

$$t_g = z_g(\mathbf{1}'_g r_g), \quad v_g = r_g(\mathbf{1}'_g z_g), \quad \xi_g = \frac{t_g - v_g}{m_g}, \quad \kappa_g = \mathbf{1}'_g c_g. \quad (\text{S2.36})$$

The vectors t_g, v_g, ξ_g are observed; κ_g is unknown. These within-block calculations do not change the center in the KSS proxy: it remains $z = y - \bar{y}\mathbf{1}$ throughout. From $\mathbb{E}[z_g r'_g] = \Omega_g - \mathbf{1}_g c'_g$,

$$\mathbb{E}[t_g] = u_g - \mathbf{1}_g \kappa_g, \quad \mathbb{E}[v_g] = u_g - m_g c_g, \quad \mathbb{E}[\xi_g] = c_g - \frac{\kappa_g}{m_g} \mathbf{1}_g. \quad (\text{S2.37})$$

We suppress conditioning on X in this subsection to lighten notation. Also $\mathbf{1}'_g \xi_g = 0$ identically. Thus ξ_g already estimates the within-block contrast part of the centering bias without bias. Let

$$\hat{\Omega}_g^{M,0} = \hat{\Omega}_g^M + \text{sym}(\mathbf{1}_g \xi'_g). \quad (\text{S2.38})$$

Equations (S2.35) and (S2.37) give

$$\mathbb{E}[\hat{\Omega}_g^{M,0}] = \Omega_g - \frac{\kappa_g}{m_g} \mathbf{1}_g \mathbf{1}'_g, \quad \hat{\Omega}_g^{M,0} \mathbf{1}_g = t_g. \quad (\text{S2.39})$$

The second identity follows from $\hat{\Omega}_g^M \mathbf{1}_g = (t_g + v_g)/2$ and $m_g \xi_g = t_g - v_g$. Only one scalar bias per block remains.

Let J be the $n \times G$ matrix of block indicators: column g is one on $\mathcal{I}_g$ and zero elsewhere. Write $N_B = J'J = \text{diag}(m_g)$, and stack t_g, ξ_g, c_g into t, ξ, c . Define the positive directional residual leverage and the G -dimensional system

$$a_g = \frac{\mathbf{1}'_g \Delta_g \mathbf{1}_g}{m_g} > 0, \quad (\text{S2.40})$$

$$\mathcal{K} = n \text{diag}(a_g) - J' M J, \quad b = J' M t - n J' \mathcal{D}_B \xi. \quad (\text{S2.41})$$

All entries of $\mathcal{K}$ depend only on the design and deletion units.

Proposition S2.3 (Exact general-block correction). *If $\mathcal{K}$ is invertible, set*

$$\hat{\kappa} = \mathcal{K}^{-1}b, \quad \hat{c}_g = \xi_g + \frac{\hat{\kappa}_g}{m_g} \mathbf{1}_g, \quad \boxed{\tilde{\Omega}_g = \hat{\Omega}_g^{\text{M},0} + \frac{\hat{\kappa}_g}{m_g} \mathbf{1}_g \mathbf{1}_g'} \quad (\text{S2.42})$$

Then $\mathbb{E}[\hat{\kappa} \mid \mathcal{D}] = \kappa$, $\mathbb{E}[\hat{c}_g \mid \mathcal{D}] = c_g$, and $\mathbb{E}[\tilde{\Omega}_g \mid \mathcal{D}] = \Omega_g$. If X contains an intercept, all these observed statistics are invariant to a common shift of the outcomes.

Proof. In stacked notation, (S2.37) says $\mathbb{E}[t] = u - J\kappa$ and $\mathbb{E}[\xi] = c - JN_B^{-1}\kappa$. Equation (S2.35) implies $n\mathcal{D}_B c = Mu$. Therefore

$$\mathbb{E}[b] = J'M(u - J\kappa) - nJ'\mathcal{D}_B(c - JN_B^{-1}\kappa) \quad (\text{S2.43})$$

$$= \{nJ'\mathcal{D}_B JN_B^{-1} - J'MJ\}\kappa = \mathcal{K}\kappa. \quad (\text{S2.44})$$

Applying the fixed inverse proves the first assertion. The other expectation claims follow from (S2.37) and (S2.39). With an intercept, $\mathcal{R}\mathbf{1} = 0$, so both z and r are unchanged by $y \mapsto y + d\mathbf{1}$. Every subsequent statistic is a fixed function of these quantities. Only finite second moments are required because all estimators displayed here are fixed quadratic forms in y . $\square$

For a contrast, put $s = XS_{xx}^{-1}\ell$. For a quadratic component with symmetric fixed A , put $B_g = X_g S_{xx}^{-1} A S_{xx}^{-1} X_g'$. Then

$$\tilde{V}_\ell = \sum_g \left\{ (s'_g z_g)(s'_g r_g) + (s'_g \mathbf{1}_g)(s'_g \hat{c}_g) \right\}, \quad (\text{S2.45})$$

$$\tilde{\theta} = \hat{\beta}' A \hat{\beta} - \sum_g \left\{ z'_g B_g r_g + \mathbf{1}'_g B_g \hat{c}_g \right\} \quad (\text{S2.46})$$

are unbiased for $\text{Var}(\ell' \hat{\beta} \mid \mathcal{D})$ and $\beta' A \beta$. These contractions avoid forming the covariance matrices. The same $\hat{c}$ serves every target. The whole component estimator is shift invariant when $Ab_0 = 0$, where $Xb_0 = \mathbf{1}$; its covariance correction is shift invariant regardless. Unbiasedness does not imply positive semidefiniteness, minimum variance, or finite sampling variance under only second-moment assumptions.

S2.6.3 Invertibility

Let $T_B = n \text{diag}(a_g) \succ 0$. The normalized system is

$$T_B^{-1/2} \mathcal{K} T_B^{-1/2} = I_G - T_B^{-1/2} J' M J T_B^{-1/2}. \quad (\text{S2.47})$$

The subtracted matrix is positive semidefinite, has rank $\text{rank}(MJ)$, and has trace

$$\sum_g \frac{(J'MJ)_{gg}}{na_g} = \sum_g \frac{\mathbf{1}'_g \Delta_g \mathbf{1}_g}{na_g} = \sum_g \frac{m_g}{n} = 1. \quad (\text{S2.48})$$

It has at least one positive eigenvalue because every $a_g > 0$. If its rank is one, that eigenvalue is exactly one. If its rank is at least two, each positive eigenvalue is strictly less than one. Hence

$$\boxed{\mathcal{K} \succ 0 \iff \text{rank}(MJ) \geq 2.} \quad (\text{S2.49})$$

Equivalently $\text{rank}([X \ J]) - k \geq 2$: adding deletion-unit indicators to the design must create at least two additional independent directions. This is the exact condition for the full correction, not a numerical leverage cutoff. Nearly collinear residualized block indicators can still make the inverse poorly conditioned.

Using only v_g instead would give the valid moment equation $\mathbb{E}[v] = (I - \text{blockdiag}(m_g I) \mathcal{R}/n)u$. That simpler-looking inverse can be singular even when (S2.49) holds. The difference of the two orientations retains information that this single moment discards. The present construction is therefore preferable for general deletion units.

S2.6.4 One additional regression with rank-one block weights

Define $\delta_g = na_g - m_g$. If every $\delta_g > 0$, set

$$W_g = I_{m_g} + \frac{\mathbf{1}_g \mathbf{1}_g'}{\delta_g}, \quad W = \text{blockdiag}(W_g), \quad (\text{S2.50})$$

$$p_g^* = t_g - \mathbf{1}_g \frac{\mathbf{1}_g' \Delta_g \xi_g}{a_g}, \quad q_g^* = \mathbf{1}_g' t_g - n \mathbf{1}_g' \Delta_g \xi_g. \quad (\text{S2.51})$$

The vector p^* is an auxiliary dependent variable. Fit it on the original full design using the block weights W :

$$\hat{\gamma} = (X'WX)^{-1}X'Wp^*, \quad \boxed{\hat{\kappa}_g = \frac{q_g^* - (X_g' \mathbf{1}_g)' \hat{\gamma}}{\delta_g}}. \quad (\text{S2.52})$$

This is one extra regression shared by all targets. The weights equal one on within-block contrasts and $na_g/(na_g - m_g)$ on the normalized block-mean direction. It retains time-varying regressors inside blocks.

Here is a direct proof of equivalence to (S2.41). Put $F = X'J$, $\Delta = \text{diag}(\delta_g)$ and $q^* = (q_g^*)_g$. Then

$$\mathcal{K} = \Delta + F' S_{xx}^{-1} F, \quad b = q^* - F' S_{xx}^{-1} X' t, \quad (\text{S2.53})$$

$$X'WX = S_{xx} + F \Delta^{-1} F', \quad X'Wp^* = X' t + F \Delta^{-1} q^*. \quad (\text{S2.54})$$

The regression equation and recovery formula imply $S_{xx} \hat{\gamma} = X' t + F \hat{\kappa}$. Substituting this identity into $\Delta \hat{\kappa} = q^* - F' \hat{\gamma}$ yields $\mathcal{K} \hat{\kappa} = b$. Because $\Delta \succ 0$, $\mathcal{K} \succ 0$ and the solution is unique. No stochastic independence between this auxiliary fit and the original estimate is needed.

This construction is cheap to apply after the exact leave-block quantities are available. The only required extra local design vector is $\Delta_g \mathbf{1}_g$; it supplies both a_g and $\mathbf{1}_g' \Delta_g \xi_g$. With that vector cached, forming t, v, ξ, p^*, q^* and applying W cost $O(n)$. In particular,

$$W_g v_g = v_g + \mathbf{1}_g (\mathbf{1}_g' v_g) / \delta_g. \quad (\text{S2.55})$$

There is no need to store dense new weight matrices or factor a new matrix for each deletion unit. Matrix products with $X'WX$ use multiplication by X , group sums, and multiplication by X' . Obtaining $\Delta_g \mathbf{1}_g$ may still be costly if the original implementation does not expose exact block leverage information; that cost is not removed by the algebra. A new regression factorization or preconditioner may also be needed. This is not merely a new right-hand side for the original factorization.

If $\rho = \max_g m_g/(na_g) < 1$, then $I \preceq W \preceq (1 - \rho)^{-1}I$, and consequently $S_{xx} \preceq X'WX \preceq (1 - \rho)^{-1}S_{xx}$. This gives a useful preconditioning bound when the directional leverage condition is comfortably satisfied.

If a δ_g is zero or negative, the positive-weight formula should not be used. This is a failure of that convenient implementation, not of (S2.49). When the rank condition holds, solve the positive definite system (S2.47) instead. For a G -vector v ,

$$\mathcal{K}v = \Delta v + F'S_{xx}^{-1}Fv. \quad (\text{S2.56})$$

Thus every matrix application requires group aggregation and one application of the original regression solve. No dense $n \times n$ or $G \times G$ matrix is required. Iteration counts depend on conditioning, so this route need not cost only one fit. An intercept-only design with block sizes $(2, 1, 1)$ illustrates the distinction: one δ_g is zero, but $\text{rank}(MJ) = 2$ and the correction is well defined.

S2.6.5 Singular systems and the boundary of the result

Full recovery requires $\text{rank}(MJ) \geq 2$, even when $n - k \geq 2$. For a particular target set $w_g = \mathbf{1}'_g B_{t,g} \mathbf{1}_g / m_g$. If $\mathcal{K}\lambda = w$ has a solution, the scalar correction $\sum_g \text{tr}(B_{t,g} \hat{\Omega}_g^{M,0}) + \lambda'b$ is unbiased for the plug-in bias, where $\hat{\Omega}_g^{M,0}$ denotes the preliminary block proxy in (S2.38). This follows from $\mathbb{E}b = \mathcal{K}\kappa$ and symmetry of $\mathcal{K}$. It is a target-specific mathematical identity, not a claim that the command implements a generalized-inverse fallback. When regressors are constant within every match, residualized block indicators span the collapsed residual space, so the rank condition becomes $G - k \geq 2$. With varying controls within matches one must use the full block formulas; the scalar match shortcut is not justified by matching worker and firm IDs.

S3 Randomized computation and numerical uncertainty

This section conditions on the complete observed data and randomizes only computational directions. `fevc` uses the normalized stochastic leverage approximation and finite-probe adjustment of Kline et al. (2021) (October 2026 revision), building on the JLA method in Kline et al. (2020). We summarize the derivation and describe its implementation with controls, deletion blocks and frequency weights, followed by randomized centering and numerical MCSE. The adjustment removes a leading ratio-expansion term; finite-probe JLA remains an approximation to exact leave-out.

S3.1 Normalized leverage estimates and finite-probe adjustment

First take an exact orthogonal projector P and $M = I - P$. For a fixed row i let $h = P_{ii}$, $m = M_{ii} = 1 - h$, and $s_i = \sum_j P_{ij}^2 M_{ij}^2$. For independent Rademacher directions q_b , $b = 1, \dots, R$, put

$$U_b = (Pq_b)_i^2, \quad V_b = (Mq_b)_i^2, \quad \bar{p} = R^{-1} \sum_b U_b, \quad \bar{m} = R^{-1} \sum_b V_b, \quad \hat{h} = \frac{\bar{p}}{\bar{p} + \bar{m}}, \quad \hat{m} = \frac{\bar{m}}{\bar{p} + \bar{m}}. \quad (\text{S3.1})$$

The same direction enters U_b and V_b . Define empirical raw moments $a = R^{-1} \sum_b U_b^2$, $b_2 = R^{-1} \sum_b V_b^2$, and $c_2 = R^{-1} \sum_b U_b V_b$. These are fourth moments of projected directions, not squared estimated

leverages. The implemented leading bias and variance are

$$\hat{b}_i = \{\hat{m}a - \hat{h}b_2 + (\hat{m} - \hat{h})c_2\}/R, \quad (\text{S3.2})$$

$$\hat{v}_i = \{\hat{m}^2a + \hat{h}^2b_2 - 2\hat{h}\hat{m}c_2\}/R. \quad (\text{S3.3})$$

The raw moments are not divided by further powers of $\bar{p} + \bar{m}$.

For fixed vectors u, v , independence and $q_j^2 = 1$ imply

$$\mathbb{E}[(u'q)^2(v'q)^2] = (u'u)(v'v) + 2(u'v)^2 - 2\sum_j u_j^2 v_j^2. \quad (\text{S3.4})$$

The three pairings count the all-equal coordinate three times, whereas its fourth moment is one; subtracting twice its contribution proves the formula. Using $P^2 = P$, $M^2 = M$ and $PM = 0$ gives

$$\mathbb{E}U_b^2 = h^2 + 4h^2m - 2s_i, \quad \mathbb{E}V_b^2 = m^2 + 4hm^2 - 2s_i, \quad \mathbb{E}U_bV_b = hm - 2s_i. \quad (\text{S3.5})$$

Write $a_0 = \bar{p} - h$, $b_0 = \bar{m} - m$, $d_0 = a_0 + b_0$ and $L = hb_0 - ma_0$. The exact ratio identity is

$$\hat{m} - m = L - Ld_0 + \frac{Ld_0^2}{1 + d_0}, \quad -Ld_0 = ma_0^2 + (m - h)a_0b_0 - hb_0^2. \quad (\text{S3.6})$$

Since $(Pq_b)_i + (Mq_b)_i = q_{bi}$, their squared values sum to at least $1/2$, so $1 + d_0 \geq 1/2$. The moment formulas give $R\mathbb{E}a_0^2 = 4h^2m - 2s_i$, $R\mathbb{E}b_0^2 = 4hm^2 - 2s_i$ and $R\mathbb{E}a_0b_0 = -2s_i$. Hence, at a fixed exact design as $R \rightarrow \infty$,

$$\begin{aligned} \mathbb{E}\hat{m} - m &= 4(h - m)s_i/R + O(R^{-2}), \\ \text{Var}(\hat{m}) &= 2\{2h^2m^2 - (m - h)^2s_i\}/R + O(R^{-2}). \end{aligned} \quad (\text{S3.7})$$

Substituting the empirical raw moments and normalized shares into the leading bias and variance expressions gives (S3.2)–(S3.3); see also Kline et al. (2021, pp. 2–6).

S3.2 Inverse adjustment, controls and blocks

Without separately handled controls the adjusted scalar deleted residual is

$$z_i^{\text{adj}} = e_i F_i, \quad F_i = \hat{m}^{-1} + \hat{b}_i \hat{m}^{-2} - \hat{v}_i \hat{m}^{-3}. \quad (\text{S3.8})$$

Expanding the reciprocal around the true residual share shows why the leading bias enters with a plus sign and the leading variance with a minus sign. This expansion requires separation from a zero inverse denominator. The bound $\bar{p} + \bar{m} \geq 1/2$ does not imply $\hat{m} > 0$ on every draw. True identified deletion does not guarantee that every finite sketch has a usable inverse, nor does withholding an unusable sketch imply unbiasedness conditional on acceptance. No uniform approximation near the deletion boundary is asserted. Solver error is another approximation, outside this expansion.

At the pinned implementation, generic JLA certifies control rank using within-worker-firm-cell scatter and directly factored deleted scatter. This is a sufficient computational screen, not a necessary identification condition. For example, in the complete three-worker/three-firm grid with two observations per cell, the control $z_{wj} = wj$ has zero within-cell scatter although the full design and every match-deleted design have rank six. The screen can return `UNVERIFIED_DELETION_RANK`

for this identified design, including under fixed-offset control treatment. The exact route instead checks the relevant deleted design matrices.

For exact low-dimensional control geometry, write P_0 for the fixed-effects projection, $U_Z = (I - P_0)Z$, and $S_Z = U_Z'U_Z$ on identified control columns. Then $P = P_0 + U_Z S_Z^{-1} U_Z'$. Sketch P_0 and $I - P_0$ in (S3.1); if $c_i^Z = (U_Z S_Z^{-1} U_Z')_{ii}$, replace the inverse denominator in (S3.8) by $\ell_i = \hat{m}_i - c_i^Z$. Keep the same $\hat{b}_i, \hat{v}_i$. Thus $F_i = \ell_i^{-1} + \hat{b}_i \ell_i^{-2} - \hat{v}_i \ell_i^{-3}$.

In a stored, frequency-transformed match, let v_g have entries $\sqrt{f_s/F_g}$, $F_g = \sum_{s \in g} f_s$, and let $\Pi_{Z,g}$ be the exact control-projection block. Put

$$K_g = (I - \Pi_{Z,g} - \hat{h}_g v_g v_g')^{-1}, \quad a_g = K_g e_g^w, \quad w_g = K_g v_g, \quad t_g = v_g' a_g, \quad k_g = v_g' w_g. \quad (\text{S3.9})$$

The implemented block adjustment is

$$z_g^{\text{adj}} = a_g + (\hat{b}_g - \hat{v}_g k_g) w_g t_g. \quad (\text{S3.10})$$

Here $\hat{b}_g$ is the leading bias of the residual share, so its negative is the leverage bias. Differentiating the inverse with respect to leverage gives $\partial a_g / \partial h = w_g t_g$ and $\partial^2 a_g / \partial h^2 = 2k_g w_g t_g$, proving the signs and factor in (S3.10). The full differential used for MCSE is

$$dz_g = w_g t_g \{ (1 + 2\hat{b}_g k_g - 3\hat{v}_g k_g^2) d\hat{h}_g + d\hat{b}_g - k_g d\hat{v}_g \}. \quad (\text{S3.11})$$

For observations the analogous differential is $dF_i = (-\ell_i^{-2} - 2\hat{b}_i \ell_i^{-3} + 3\hat{v}_i \ell_i^{-4}) d\hat{m}_i + \ell_i^{-2} d\hat{b}_i - \ell_i^{-3} d\hat{v}_i$. The derivatives of both moment corrections are included, holding the probe count fixed. These scalar and block formulas do not replace the rank checks.

S3.3 Physical directions and target contractions

Frequency compression must reproduce directions on physical copies. Here P_0 is the FE-only projector; exact residualized-control geometry is added separately as in section S3.2. If $S_{sb} = \sum_{a=1}^{f_s} q_{sab}$, a projected value is common across copies, but residual projected directions are $q_{sab} - (P_0 q_b)_s$. Their second and fourth moments must retain the physical-copy sums. Averaging nonlinear inverse adjustments across the copies is distinct from applying an inverse adjustment to averaged moments. For whole matches, contract the physical direction as $F_g^{-1/2} \sum_{s \in g} S_{sb}$ and the fitted direction as $F_g^{-1/2} \sum_{s \in g} f_s (P_0 q_b)_s$ before forming match moments.

For target mass a_s and $A_+ = \sum_s a_s$, an independent target direction is $d_s = \zeta_s - (a_s/A_+) \sum_v \zeta_v$ with $\zeta_s = \sqrt{a_s/(f_s A_+)} S_s$. Then $\mathbb{E} dd' = K_a$. The full-model target prediction columns are

$$p_\alpha = X_s S_{xx}^{-1} X_\alpha' d, \quad p_\psi = X_s S_{xx}^{-1} X_\psi' d. \quad (\text{S3.12})$$

The inverse includes the joint controls, even though the target loadings have zero control columns. These are predictions from the complete target solutions, not isolated worker or firm coefficient blocks. For observation deletion the three target corrections from one direction are

$$\sum_s f_s y_s^* z_s^{\text{adj}} (p_{\alpha,s}^2, p_{\psi,s}^2, p_{\alpha,s} p_{\psi,s})', \quad (\text{S3.13})$$

where y_s^* is the working outcome with the requested centering. The frequency representation in this expression denotes the average of the copy-specific adjustment where required. For matches, set $A_{\alpha,g} = \sum_{s \in g} f_s y_s^* p_{\alpha,s}$ and similarly $A_{\psi,g}$. The coefficient column multiplying the frequency-transformed adjusted residual at stored row s is $\sqrt{f_s} (A_{\alpha,g} p_{\alpha,s}, A_{\psi,g} p_{\psi,s}, (A_{\alpha,g} p_{\psi,s} + A_{\psi,g} p_{\alpha,s})/2)'$. Whole deletion units are contracted before numerical covariance outer products.

S3.4 Randomized Corrected centering

The following map defines the increment at the documentation source, in physical coordinates. For each leverage pool $\mathcal{P}$ (full, first half, second half), form the normalized FE leverage shares from that pool only. For an observation put $\hat{\Delta}_{i,\mathcal{P}} = 1 - \hat{h}_{i,\mathcal{P}} - c_i^Z$; for a match put $\hat{\Delta}_{g,\mathcal{P}} = I - \Pi_{Z,g} - \hat{h}_{g,\mathcal{P}}v_gv_g'$ in the equivalent frequency-transformed coordinates of (S3.9). Map back to physical copies for the formulas below. Use $r_{g,\mathcal{P}} = \hat{\Delta}_{g,\mathcal{P}}^{-1}e_g$, without applying the scalar finite-probe adjustment again. With $z = y - \bar{y}\mathbf{1}$, form $t_{g,\mathcal{P}}$ and $\xi_{g,\mathcal{P}}$ by (S2.36) using this pool's residual. Let $\hat{\mathcal{D}}_{B,\mathcal{P}}$ be the block diagonal of its local makers and define

$$\hat{a}_{g,\mathcal{P}} = \mathbf{1}_g' \hat{\Delta}_{g,\mathcal{P}} \mathbf{1}_g / m_g, \quad (\text{S3.14})$$

$$\hat{\mathcal{K}}_{\mathcal{P}} = n \text{diag}(\hat{a}_{g,\mathcal{P}}) - J' M J, \quad (\text{S3.15})$$

$$\hat{b}_{\mathcal{P}} = J' M t_{\mathcal{P}} - n J' \hat{\mathcal{D}}_{B,\mathcal{P}} \xi_{\mathcal{P}}, \quad (\text{S3.16})$$

$$\hat{c}_{g,\mathcal{P}} = \xi_{g,\mathcal{P}} + \mathbf{1}_g (\hat{\mathcal{K}}_{\mathcal{P}}^{-1} \hat{b}_{\mathcal{P}})_g / m_g. \quad (\text{S3.17})$$

Here the global $M = I - X S_{xx}^{-1} X'$ uses ordinary model solves, not a second sketch of its full matrix. The local makers and their inverses are randomized; control geometry and frequencies are held fixed. Stored-row implementations evaluate the physical sums using frequency weights; an observation unit is one physical copy, whereas a match unit contains all its copies. Eligible stayer observations remain separate units.

For the original independent target pool, let $\hat{B}_{t,gg}$ be its average target outer product: $p_{\alpha,g} p'_{\alpha,g}$ for worker variance, $p_{\psi,g} p'_{\psi,g}$ for firm variance, and the symmetrized cross product for covariance, using (S3.12). The signed change to the point estimate from one leverage pool is

$$\Delta_{\mathcal{P},t} = - \sum_g \mathbf{1}_g' \hat{B}_{t,gg} \hat{c}_{g,\mathcal{P}}. \quad (\text{S3.18})$$

The opposite sign is added to the estimated plug-in bias before subtracting that bias. This fixes the sign convention relative to Mean. The same target pool is used for all three maps. Corrected JLA uses

$$\Delta_{\text{JLA},t} = 2\Delta_{R,t} - \frac{1}{2}(\Delta_{R/2,1,t} + \Delta_{R/2,2,t}), \quad \hat{\theta}_t^{\text{Corrected,JLA}} = \hat{\theta}_t^{\text{Mean,JLA}} + \Delta_{\text{JLA},t}. \quad (\text{S3.19})$$

The public calculation requires an even $R \geq 4$ and splits logical leverage indices into equal halves. Each pool must pass the local maker, auxiliary-system and complete-equation residual checks; failure of a half is not handled by dropping it. Writing $\hat{\delta}_{g,\mathcal{P}} = n\hat{a}_{g,\mathcal{P}} - m_g$, the weighted solve handles units with $\hat{\delta}_{g,\mathcal{P}} > 0.1n\hat{a}_{g,\mathcal{P}}$ directly. It handles at most 64 remaining units through a separately checked positive-definite Schur system; exceeding that resource limit withholds the calculation. Thus a nonpositive auxiliary weight is not by itself the full acceptance rule. These are numerical support conditions, not the exact rank equivalences of section S2 applied to a sketch.

The split calculation cancels a leading R^{-1} term when a regular expansion exists. It does not establish that expansion near singular local or global makers, make Corrected JLA exactly unbiased, or supply an unbiasedness claim conditional on all three pools passing. Solver error remains separate. The exact centering theorems in section S2 therefore do not extend to arbitrary finite probe counts.

S3.5 Numerical covariance from all point-estimation probes

This calculation holds all observed inputs fixed, including the outcome, retained sample, weights, and deterministic preparation. Its randomness comes only from the point-estimation probes. Write

$\hat{\theta} = (\hat{V}_\alpha, \hat{V}_\psi, \hat{C}_{\alpha\psi})'$ for the three primitive corrected components. Let S denote the entire leverage sketch and $\mathcal{T}$ the independent target sketch. Whenever the unfiltered randomized calculation is defined and has finite second moments, the law of total covariance gives

$$\Sigma_{\text{MC}} = \mathbb{E}_S\{\text{Cov}_{\mathcal{T}}(\hat{\theta} \mid S)\} + \text{Cov}_S\{\mathbb{E}_{\mathcal{T}}(\hat{\theta} \mid S)\}. \quad (\text{S3.20})$$

The conditional target-probe diagnostic estimates $\text{Cov}_{\mathcal{T}}(\hat{\theta} \mid S)$ at the realized S . The all-probe diagnostic also estimates the leverage contribution using a local linearization.

More precisely, let U_b collect sufficient-statistic contributions from leverage direction b , $\hat{u} = R^{-1} \sum_{b=1}^R U_b$, and $z = h_R(\hat{u})$ be the adjusted deleted-residual vector. The function h_R includes the implemented finite-projection adjustment, with R held fixed when differentiating. Target direction b supplies a linear operator L_b , so its correction is $c_b = L_b z$ and $\hat{\theta} = \hat{\theta}_{\text{plugin}} - \bar{c}$. For $T_p \geq 2$ target directions, the conditional covariance estimate is

$$\hat{\Sigma}_{T|S} = \frac{1}{T_p(T_p - 1)} \sum_{b=1}^{T_p} (c_b - \bar{c})(c_b - \bar{c})'. \quad (\text{S3.21})$$

Split the target directions into two disjoint folds A, B by their logical indices, and let $\bar{L}_A, \bar{L}_B$ be their separate means. With $J_R = \partial h_R(u) / \partial u'|_{u=\hat{u}}$, define $\psi_b^A = -\bar{L}_A J_R (U_b - \hat{u})$ and $\psi_b^B = -\bar{L}_B J_R (U_b - \hat{u})$. Center each set of scores over leverage directions to obtain a_b, d_b . For $R \geq 2$, the raw estimate is

$$\hat{\Sigma}_{\text{MC,raw}} = \hat{\Sigma}_{T|S} + \frac{1}{2R(R-1)} \sum_{b=1}^R (a_b d_b' + d_b a_b'). \quad (\text{S3.22})$$

Independent target folds avoid counting their interaction with leverage randomization twice. Each score sums all deletion units before taking cross-products, including both movers and stayers when requested. Replay uses the original leverage directions, not new independent draws. The public probe count sets $R = T_p = \text{probes}()$.

For an affine h_R with fixed derivative, the cross-fold construction is unbiased for the complete-run covariance under independent stages and finite moments. The actual deletion adjustment is nonlinear, so (S3.22) is a local approximation. Calibration can deteriorate near singular deletion or at small probe counts. It does not estimate finite-probe bias, solver error, or sampling uncertainty, and does not by itself justify numerical confidence intervals about exact KSS.

The four-component covariance is obtained by applying the linear map whose rows are $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, and $(1, 1, 2)$ to the primitive covariance. In particular, the MCSE of the total includes cross-component covariances; marginal MCSEs cannot simply be added. The public matrices `e(mcse_cov_raw)` and `e(mcse_cov)` contain the raw and usable four-by-four numerical covariances. A materially negative eigenvalue or a failed derivative or replay calculation withholds usable MCSEs while preserving valid point output and any computed raw covariance. A tolerance-level adjustment of negative eigenvalues is recorded separately. Neither numerical covariance is posted as sampling `e(V)`.

For an affine $h_R(u) = h_0 + J_R u$, conditional expectation over the independent target folds replaces their cross-product by the product of population target operators. The sample covariance of the leverage directions is unbiased for their one-direction covariance; division by R gives the covariance of their mean. The conditional target term already includes target-leverage interaction, so using an independent-fold cross-product in the leverage term counts that interaction once. This proof

requires finite moments and a deterministic derivative. Substituting a derivative at the observed sketch makes the implemented calculation a local approximation.

For Corrected JLA, the reported MCSE is Mean's MCSE using the same probes. It excludes the variance of the added centering increment and its covariance with the Mean estimate. It is therefore neither a complete MCSE for Corrected nor guaranteed conservative; the omitted uncertainty need not be negligible.

S4 Fixed-effect projections

S4.1 Projection covariance by deletion block

Let R_p collect the projection covariates, including the intercept, let W_p contain their positive projection weights, and let $E\beta$ select the worker or firm effect attached to each projection row. The population projection and its loading matrix are

$$\pi = L'\beta, \quad L' = (R_p'W_pR_p)^{-1}R_p'W_pE, \quad (\text{S4.1})$$

provided $R_p'W_pR_p$ is nonsingular. The projection rows may use stored observations with their corresponding frequency or target weights. Set $S = XS_{xx}^{-1}L$, with block rows S_g and row i equal to ℓ'_i . For stored rows the corresponding loading is $\ell'_s = x'_sS_{xx}^{-1}L$. Independence across deletion blocks and the exclusion of block g from $\hat{\beta}_{-g}$ imply

$$\mathbb{E}\{y_g(y_g - X_g\hat{\beta}_{-g})' \mid \mathcal{D}\} = \Omega_g. \quad (\text{S4.2})$$

Set $z_g = y_g$ for `centering(none)` or $z_g = y_g - \bar{y}\mathbf{1}_g$ for the default `centering(mean)`, where $\bar{y} = \sum_s f_s y_s / \sum_s f_s$ is the retained working-outcome mean. Target and projection weights do not enter this mean. The symmetric proxy is

$$\hat{\Omega}_g = \frac{1}{2}(z_g\hat{e}'_{g,-g} + \hat{e}_{g,-g}z'_g). \quad (\text{S4.3})$$

Substitution gives

$$\hat{V}_\pi = \sum_g S'_g \hat{\Omega}_g S_g = L'S_{xx}^{-1} \left(\sum_g X'_g \hat{\Omega}_g X_g \right) S_{xx}^{-1} L. \quad (\text{S4.4})$$

For one-observation blocks this reduces to $z_i \hat{e}_{i,-i} \ell'_i$. Mean centering leaves the fit, deleted residuals, and loadings unchanged, but estimating $\bar{y}$ can invalidate (S4.2) under heteroskedasticity. More explicitly, for $c_g = \text{Cov}(\hat{e}_{g,-g}, \bar{y} \mid \mathcal{D})$,

$$\mathbb{E}(\hat{V}_\pi^{\text{Mean}} \mid \mathcal{D}) - V_\pi = - \sum_g S'_g \text{sym}(\mathbf{1}_g c'_g) S_g. \quad (\text{S4.5})$$

This follows by contracting the centered block-product identity in (S2.35). For a reported contrast, variance consistency must control this discrepancy as well as covariance-estimation variability. Corrected centering is not implemented for projections. Under fixed-offset controls, centering follows subtraction of the fitted control contribution.

The implementation evaluates (S4.4) without expanding positive integer frequencies. For a mover match g , define

$$a_g = \sum_{s \in g} f_s \ell_s z_s, \quad b_g = \sum_{s \in g} f_s \ell_s \hat{e}_{s,-g}, \quad (\text{S4.6})$$

where ℓ'_s is the stored-row loading. Its covariance contribution is $\frac{1}{2}(a_g b'_g + b_g a'_g)$. An eligible stayer row contributes $f_s z_s \hat{e}_{s,-s} \ell'_s$. Here $-s$ means deletion of one physical copy, not the entire stored row. To make the distinction explicit, let $h_s = x'_s S_{xx}^{-1} x_s$ and $\hat{e}_s = y_s - x'_s \hat{\beta}$ on stored rows. The two residuals are

$$\hat{e}_{s,-\text{one copy}} = \frac{\hat{e}_s}{1 - h_s}, \quad \hat{e}_{s,-\text{all copies}} = \frac{\hat{e}_s}{1 - f_s h_s}, \quad (\text{S4.7})$$

whenever the respective deleted designs remain identified. Observation deletion uses the first residual f_s times. A declared match uses the full block residual $M_g^{-1} \hat{e}_g$. Thus the compressed covariance statistics preserve both the weights and the requested deletion rule.

The full-sample residual covariance replaces $\hat{\Omega}_g$ in (S4.4) by $\hat{e}_g \hat{e}'_g$, retaining the same blocks and loadings. It uses full-sample residuals without the leave-out noise correction. It is a residual-based comparison, not a regression treating the estimated firm effects as error-free observations. Its standard errors need not always be smaller than the KSS standard errors.

Symmetrization does not ensure a positive-semidefinite covariance estimate. `fevc` rejects a projection request with a nonpositive estimated variance or a materially indefinite covariance. A small adjustment near zero is permitted within the numerical tolerance and reported in the diagnostics. For a contrast $c'\pi$, the estimation error is $\sum_g c' S'_g \varepsilon_g$. Normal intervals require a limit approximation for this sum and consistent estimation of its variance, beyond the finite-sample covariance identity above.

S4.2 Normalization and supported routes

If R_p includes an intercept, a location renormalization of the selected fixed effects changes that intercept and leaves the nonconstant projection slopes unchanged. Specify the fixed-effect normalization when reporting the intercept. Changing projection weights changes the projection estimand and loadings; it does not redefine the working-outcome mean.

The default projection route is exact Mata, limited by `exact_limit()`. The sparse route requires explicit `backend(rust)`, `rng(counter_v1)` and `algorithm(jla)`. Set the preconditioner explicitly to `diagonal` or `cmg`. The engine must be generic: an explicit `engine(generic)`, `engine(auto)` or an omitted engine is admitted, while an automatic preconditioner is not. Both support observation or match deletion, joint controls and the requested mover/stayer partition. Mean and None are supported; Corrected is unavailable. Projection coefficients do not change with outcome centering, but their leave-out covariance does. The four-component numerical MCSE does not quantify uncertainty in projection covariance.

S5 Experimental component inference

The calculations in this section are opt-in and experimental. Current Mean inference holds the observed retained mean fixed inside the variance and reference-distribution calculations. This approximation omits terms derived below. None uses the uncentered working outcome. Corrected component inference is unavailable. A fixed estimated control offset introduces a separate omission; neither approximation is justified merely by a large sample or few controls. Point estimation remains the default and posts no sampling $\mathbf{e}(\mathbf{V})$.

The first formulas use None, or a genuinely fixed outcome origin. For Mean, replace the outcome in its point/influence contractions as specified later in this section; keep residuals and noise

probes uncentered. KSS develop inference for leave-out quadratic forms under both diffuse and concentrated spectra. The calculations below combine their reference distributions with practical variance estimators. The structured variance fit does not implement KSS’s unrestricted sample-split variance products (Kline et al., 2020, sections 5 and 7).

Products of ordinary leave-out error-variance estimates are generally biased. Sections 5.2 and 6.2 of Kline et al. (2020) use independent split-sample predictions for these products, with conservative adjustments where that construction is unavailable. In the first-differenced two-period worker–firm model, identification after any two deletions suffices for the two predictions to exist; identification after any three deletions suffices to avoid the remaining conservative variance-product substitutions. `fevc` instead uses fitted error variances, so those conservative-coverage results do not by themselves justify its intervals.

Sections S5.1–S5.4 give the None/fixed-origin construction. Sections S5.5–S5.11 revisit its kernels to distinguish a fixed center from the observed mean and derive the omitted terms. Sections S5.12–S5.13 give supported requests and examples; Section S6 reports evidence separately by procedure and population.

S5.1 Independent units for movers and stayers

After holding the control offset fixed, let $y_i^* = y_i - z_i' \hat{\gamma}$. The two-way regressor x_g is constant within a declared match. With frequency masses f_i , define

$$F_g = \sum_{i \in g} f_i, \quad y_g^c = \frac{\sum_{i \in g} f_i y_i^*}{\sqrt{F_g}}, \quad x_g^c = \sqrt{F_g} x_g. \quad (\text{S5.1})$$

For the mover population these sufficient rows preserve the two-way normal equations: $\sum_i f_i x_i x_i' = \sum_g x_g^c x_g^{c'}$ and $\sum_i f_i x_i y_i^* = \sum_g x_g^c y_g^c$. Deleting one collapsed row deletes the entire original match. The collapse is exact for a given offset. Target mass still defines A_t separately from regression mass.

For the pooled population, let $\mathcal{G}_M$ index retained mover deletion units and $\mathcal{I}_S$ index retained physical stayer observations. A mover has multiple original deletion units, possibly at a common worker–firm coordinate; a stayer has one. Original stayers enter only if they have at least two physical observations at a retained mover firm. Sample pruning does not reclassify a mover as a stayer. Stack the transformed rows as

$$(x_a^c, y_a^c) = \begin{cases} (\sqrt{F_g} x_g, \sum_{i \in g} f_i y_i^* / \sqrt{F_g}), & a = g \in \mathcal{G}_M, \\ (x_i, y_i^*), & a = i \in \mathcal{I}_S. \end{cases} \quad (\text{S5.2})$$

Each frequency copy of a stayer row is a separate physical observation. The combined $X_c' X_c$ and $X_c' y_c$ equal the physical-sample normal equations. All four A_t use the pooled target weights, including between-group differences in mean effects. The calculation is one pooled decomposition, not an average of two group decompositions.

For a known correct offset, let u_g contain the stored-row averages of the physical errors within each frequency group, with covariance Γ_g and frequency vector f_g . The scalar error has variance

$$\tau_g^2 = \text{Var}(f_g' u_g / \sqrt{F_g}) = f_g' \Gamma_g f_g / F_g. \quad (\text{S5.3})$$

No restriction on covariance inside g is needed for this identity. Independence across declared matches is needed for the scalar inference model. Frequency mass does not create additional

independent matches, and the independence assumption also applies to distinct matches of the same worker. For pooled inference, $\tau_a^2 = \sigma_i^2$ on a physical stayer row. Every mover aggregate and every stayer observation is assumed mutually independent. Thus arbitrary within-mover-unit dependence is allowed, whereas serial correlation among a stayer's observations is excluded. Unequal variances are allowed subject to the structured model below.

When controls are estimated on the same sample, however,

$$y_i^* = x_i' \beta + u_i + z_i'(\gamma - \hat{\gamma}). \quad (\text{S5.4})$$

The last term depends on outcomes across matches. Its induced covariance is absent from (S5.3). Conditioning on the fitted offset does not automatically restore mean zero or independence. `fevc` makes no joint-nuisance, delta-method, or second-stage correction here.

S5.2 The structured variance calculation

Use X_c and y_c for the independent-unit design and outcome in (S5.2), or its mover-only restriction, and write $S_{xx} = X_c' X_c$, $P = X_c S_{xx}^{-1} X_c'$, $M = I - P$, and $B_t = X_c S_{xx}^{-1} A_t S_{xx}^{-1} X_c'$. For exact inverse actions, the corrected quadratic form has symmetric zero-diagonal kernel

$$C_t = B_t - \frac{1}{2} \{ \text{diag}(d_t) M + M \text{diag}(d_t) \}, \quad d_{ta} = B_{t,aa} / M_{aa}. \quad (\text{S5.5})$$

The sparse JLA implementation approximates the required diagonal and inverse actions and reports numerical diagnostics separately.

Let $\mu_c = \mathbb{E}(y_c \mid \mathcal{D})$ and $V = \text{diag}(\tau_a^2)$. With independent units and suitable finite moments, the exact-kernel covariance is

$$\text{Cov}(\hat{\theta}_t, \hat{\theta}_s \mid \mathcal{D}) = 4(C_t \mu_c)' V (C_s \mu_c) + 2 \text{tr}(V C_t V C_s). \quad (\text{S5.6})$$

Zero diagonal removes the own-error square terms. `fevc` substitutes realized influences $C_t y_c$ and a fitted positive variance vector $\hat{V}$, using the corresponding minus-trace adjustment,

$$\hat{\Omega}_{ts} = 4(C_t y_c)' \hat{V} (C_s y_c) - 2 \text{tr}(\hat{V} C_t \hat{V} C_s). \quad (\text{S5.7})$$

The minus sign removes quadratic noise already present in the realized influences. It is not a claim that the population trace term is negative. The trace term is computed from common Gaussian draws $z^{(b)} = \hat{V}^{1/2} \xi^{(b)}$, $\xi^{(b)} \sim N(0, I)$. The sample covariance of the three scalars $z^{(b)'} C_t z^{(b)}$ estimates $2 \text{tr}(\hat{V} C_t \hat{V} C_s)$. Sharing a draw across targets estimates the off-diagonal covariances as well as the variances. `fevc` streams these scalar moments without storing a unit-by-probe matrix.

Let T append the row $(1, 1, 2)$ to the 3×3 identity. The posted four-component covariance is $T \hat{\Omega} T'$, so its null direction is $(-1, -1, -2, 1)'$. This singularity is required by the adding-up identity. By contrast, a nonpositive primitive variance or a materially negative eigenvalue of the primitive covariance withholds the joint covariance, not necessarily every individual interval. The structured calculation allows only a roundoff-scale correction, bounded by 10^{-8} times the largest absolute primitive diagonal. Such a covariance failure concerns the realized variance estimate; it is distinct from failure to solve the regression equations or accurately recover a mode.

The structured model uses a common residual-moment fit across targets. For match deletion, residuals use the transformed outcome and fixed-effects projection above: aggregate mover outcomes

first, then square residuals; stayer observations remain separate. For observation deletion, use the individual outcome and the full-model projection, including jointly estimated controls. Write $e = My$ in the appropriate independent-unit representation. Let Z_v contain the selected outcome-free variance predictors and suppose the variance vector is $Z_v\eta$. Then, with $\circ$ denoting elementwise multiplication,

$$\mathbb{E}[Z'_v(e \circ e) \mid \mathcal{D}] = K\eta, \quad K = Z'_v(M \circ M)Z_v. \quad (\text{S5.8})$$

Identifying the variance coefficients requires K to be nonsingular. It is the Gram matrix of $M \text{diag}(z_j)M$, where z_j is column j of Z_v , under the Frobenius inner product. Thus no nonzero coefficient vector δ may satisfy $M \text{diag}(Z_v\delta)M = 0$; full column rank of Z_v alone is insufficient. This identification condition is distinct from the numerical rank checks and consistency requirements for the fitted variances. The residual-moment adjustment matters because squared fitted residuals mix the variances of different units. Regressing them directly on Z_v generally does not undo that mixing. The raw leave-out products used for point correction and q1 recentering are unchanged.

For pooled match inference, the variance design has separate mover and stayer coefficient blocks. Within each type, transform leverage and the three primitive target diagonals to normalized midranks $2(\text{rank} - 1/2)/n_{\text{type}} - 1$, assigning equal ranks to ties. Movers also use regression mass. **structured_common** includes an intercept, linear terms, squares and pairwise interactions: 21 candidate columns for movers and 15 for stayers. **structured_leverage** has an intercept, leverage rank and its square for each type. These are different variance specifications, not data-selected alternatives.

The implementation fits the mover and stayer variance coefficients in one stacked residual-moment system. In the supported fixed-offset match geometry, the exact residual maker and Gram separate by type: stayer firm indicators lie in the span of stayer worker indicators. The finite-probe Gram can have nonzero cross-type entries from simulation noise. The common fitted variance vector enters one pooled component-covariance calculation, whose target and effect loadings can couple the two populations. Joint-control observation designs need not have this residual separation.

Before ranking pooled leverage and target-diagonal features, the calculation rounds each nonzero feature vector to the binary grid $2^{\lfloor \log_2 \max_a |v_a| \rfloor - 40}$, with half ties away from zero. This outcome-free convention preserves design ties across solver roundoff. For fixed-offset stayers, the firm and worker-firm covariance diagonal features are structural zeros, because a stayer outcome shock is absorbed by that worker's effect; these features are set to zero before ranking. Regression-mass ranks are exact. Deterministic, outcome-free removal of redundant columns preserves the original model span to a scaled reconstruction tolerance of 10^{-12} . It is not model selection or a ridge fallback. Each type must supply at least five independent units per active term; the fit applies the 10^{-10} scaled rank tolerance.

Mover-only Rust match inference uses the 21-column candidate basis on its single type. Structured observation inference uses the joint-control projection and 15 candidate columns; its leverage-only alternative has three. These single-population native bases use normalized midranks directly. The explicit structured Mata match route uses the rounded-feature convention.

Rust approximates the small matrix K without constructing M . For numerical draws $\xi_r \sim N(0, I)$, put $a_r = Z'_v[(M\xi_r) \circ (M\xi_r)]$ and $\bar{a} = R_G^{-1} \sum_r a_r$. The Gaussian quadratic-form identity gives

$$K = \frac{1}{2} \text{Var}_\xi(a_r), \quad \widehat{K} = \frac{1}{2(R_G - 1)} \sum_{r=1}^{R_G} (a_r - \bar{a})(a_r - \bar{a})'. \quad (\text{S5.9})$$

The implementation forms $M\xi_r = \xi_r - P\xi_r$ using projection solves that meet the residual tolerance, streams the centered covariance, and solves $\widehat{K}\widehat{\eta} = Z'_v(e \circ e)$. This direct residual calculation can be compared with a subtractive representation that removes a term using approximate leverages; that subtraction could make the estimated Gram matrix indefinite. There is no ridge, eigenvalue clipping, exact-matrix fallback or automatic probe increase. Rank and inverse-residual checks still apply.

Exact structured Mata evaluates the same Gram using coefficient-space contractions. With $H = S_{xx}^{-1}$ and z_a denoting column a of Z_v ,

$$K_{ab} = \sum_i (1 - 2P_{ii})z_{ia}z_{ib} + \text{tr}\{HX'_c \text{diag}(z_a)X_cHX'_c \text{diag}(z_b)X_c\}. \quad (\text{S5.10})$$

This follows by expanding $(I - P) \circ (I - P)$; it forms no observation-square matrix. Exact refers to the design solves and Gram. Both backends still use common Gaussian draws for covariance traces and simulation for q1 critical values, subject to the same structured-model interpretation.

For Rust the default is $R_G = 2,048$, controlled by `inferencegramprobes()`. Accepted integers range from 512 to 2,147,483,647, subject to count limits and the selected memory policy. An explicit budget warns by default; only `memorycheck(error)` rejects an over-budget forecast. Actual allocation failures remain errors, and the inference memory bounds remain conservative. Values below 2,048 trade precision for speed and are not recommended for reported inference. This budget is separate from the default 200 JLA point probes, 1,000 covariance probes and critical-value simulations. Changing it does not change the other numerical streams. The finite-probe inverse and estimated variance predictors do not inherit an exact- K unbiasedness claim.

Predictions below $10^{-8} \text{median}_g[e_g^2/(1 - P_{gg})]$ are raised to that floor, using the available leverage estimates. A nonpositive residual scale, an entirely floored fit, inadequate support or a failed numerical accuracy check withholds inference. The resulting positive vector defines $\widehat{V}$. Diagnostics report the Gram probe count, conditioning, residual checks and positivity adjustments; there are no cross-validation penalties or folds in this fitter. Flooring is a safeguard, not evidence that the variance specification is adequate. Neither the residual-moment adjustment nor agreement with the leverage-only sensitivity model detects variance drivers omitted from both models or implements KSS's unrestricted independent sample splitting.

S5.3 High rank and one dominant mode

KSS distinguish diffuse spectral sequences from sequences with finitely many important modes. For target t , let $A_t a_1 = \lambda_1 S_{xx} a_1$, $a'_1 S_{xx} a_1 = 1$, and $v_1 = X_c a_1$. Order the generalized eigenvalues by absolute magnitude. The leading spectral share and the leading share within the remainder are, respectively,

$$\frac{\lambda_1^2}{\sum_j \lambda_j^2}, \quad \frac{\lambda_2^2}{\sum_{j \geq 2} \lambda_j^2}. \quad (\text{S5.11})$$

The second denominator excludes the leading mode: a small full-spectrum remainder can still be concentrated. These are properties of B_t , whose nonzero eigenvalues equal those of $S_{xx}^{-1/2} A_t S_{xx}^{-1/2}$; they do not use error-variance weights. The largest unit weight $\max_a v_{1a}^2$ measures concentration of the leading eigenvector across independent units.

The reported numerical diagnostics replace $\sum_j \lambda_j^2 = \text{tr}(B_t^2)$ by $T_t = \max\{\widehat{\text{tr}}(B_t^2), \widehat{\lambda}_1^2 + \widehat{\lambda}_2^2\}$. They are $\widehat{\lambda}_1^2/T_t$ and $\widehat{\lambda}_2^2/(T_t - \widehat{\lambda}_1^2)$. A trace estimate more than six of its own numerical standard errors below the two-mode sum withholds the diagnostic. These standard errors concern the spectral-trace

calculation. Thus these finite-probe ratios, and their averages in the calibration results, differ from exact spectral shares. The command also reports the largest influence share,

$$\frac{\max_a \{\hat{\tau}_a^2 g_{ta}^2\}}{\sum_a \hat{\tau}_a^2 g_{ta}^2}, \quad g_t = C_t y_c, \quad (\text{S5.12})$$

using its numerical influence estimates and positive fitted unit variances. For Mean, use the centered influence defined in (S5.25) and the mixed independent-unit coordinates in Section S5.9. The total influence is the worker influence plus the firm influence plus twice the covariance influence. None of these diagnostics automatically selects q .

For $q = 1$, the leading score and its raw leave-out recenter are

$$\hat{b}_1 = v_1' y_c, \quad \hat{V}_{b_1}^{LO} = \sum_g v_{1g}^2 y_g^c \hat{e}_{g,-g}^c, \quad \hat{r} = \hat{\theta}_t - \lambda_1 (\hat{b}_1^2 - \hat{V}_{b_1}^{LO}). \quad (\text{S5.13})$$

The implementation checks this remainder against the direct rank-one-subtracted kernel. The raw recenter must not be replaced by a positive fitted variance. Fitted variances instead supply the covariance of $(\hat{b}_1, \hat{r})$, denoted $\hat{S} = \begin{pmatrix} V_b & C_{br} \\ C_{br} & V_r \end{pmatrix}$. Specifically, put $B_t^R = B_t - \lambda_1 v_1 v_1'$ and form C_t^R from (S5.5) using B_t^R and its diagonal. With $h_R = C_t^R y_c$, the entries are

$$V_b = v_1' \hat{V} v_1, \quad C_{br} = 2v_1' \hat{V} h_R, \quad (\text{S5.14})$$

$$V_r = 4h_R' \hat{V} h_R - 2 \text{tr}(\hat{V} C_t^R \hat{V} C_t^R). \quad (\text{S5.15})$$

The remainder trace uses the same Gaussian draws as the full covariance. **fevc** requires $V_b, V_r > 0$ and $1 - C_{br}^2/(V_b V_r) > 10^{-8}$ for the Rust q1 attachment. No ridge or positivity floor is applied to these two-dimensional covariance estimates. The relevant maximal curvature is

$$\kappa = \frac{2|\lambda_1|V_b}{\sqrt{V_r - C_{br}^2/V_b}}. \quad (\text{S5.16})$$

The numerator contains a variance, not a standard deviation. This agrees with the published KSS expression (Kline et al., 2020, p. 1878); it is invariant to rescaling the outcome. **fevc** uses a dimensionless covariance check and maps the calibrated ellipse through $(b, r) \mapsto \lambda_1 b^2 + r$ (Andrews and Mikusheva, 2016).

For confidence level $1 - \alpha$, let $c_{1-\alpha}(\kappa)$ be the $(1 - \alpha)$ quantile of

$$D_\kappa = \sqrt{Z_1^2 + (|Z_2| + 1/\kappa)^2} - 1/\kappa, \quad Z_1, Z_2 \stackrel{\text{iid}}{\sim} N(0, 1). \quad (\text{S5.17})$$

Its zero-curvature limit is $|Z_2|$ and its infinite-curvature limit is $\sqrt{Z_1^2 + Z_2^2}$. The public command uses $\max(100,000, 100R)$ critical draws per eligible target, where $R = \text{inferencesimulations}()$ is the covariance-probe count. It selects the ordered draw at rank $\lceil (1 - \alpha)N_{\text{crit}} \rceil$. Critical-value draws have a separate random-number domain from covariance probes. The endpoints are the minimum and maximum of $\lambda_1 b^2 + r$ over

$$\left[\begin{pmatrix} b \\ r \end{pmatrix} - \begin{pmatrix} \hat{b}_1 \\ \hat{r} \end{pmatrix} \right]' \hat{S}^{-1} \left[\begin{pmatrix} b \\ r \end{pmatrix} - \begin{pmatrix} \hat{b}_1 \\ \hat{r} \end{pmatrix} \right] \leq c_{1-\alpha}(\kappa)^2. \quad (\text{S5.18})$$

The implementation parametrizes the ellipse boundary by angle, searches an 8,192-point global grid, and refines its extrema by bounded search. This is the image of the calibrated joint ellipse, not a Wald interval centered on $\hat{\theta}_t$.

Under KSS’s joint-limit conditions and consistent covariance estimation, the resulting interval has asymptotic coverage at least equal to its nominal level, uniformly over the leading-score nuisance mean (Kline et al., 2020, Theorem 3 and Lemma 7). This guarantee is not inherited merely by substituting a structured fit. One-mode inference requires a diffuse remainder; several important modes, weak or null designs, and misspecified variance models can fall outside the claim. A target-specific q1 covariance or mode that fails the stated numerical requirements produces missing endpoints, while shared solver or variance-fit failures stop the attachment. Rejection of the full joint covariance alone does not discard valid target-specific intervals. Ordinary Wald columns are not substitute q1 intervals.

Applying these reference laws to the feasible calculation requires errors from nuisance estimation and numerical approximation to vanish on the sampling scale. For highrank, the difference from the reference estimator, divided by its sampling standard deviation, must be $o_p(1)$, and the fitted variance must be ratio-consistent; joint inference requires the corresponding covariance consistency on the nonsingular primitive-target space. For q1, the score and remainder errors must be negligible on their joint covariance scale, with consistent covariance, mode and critical-value calculations. A uniform coverage claim also requires these approximations to hold uniformly over the leading-score nuisance mean. These conditions supplement the spectral and moment assumptions. They are not established by fixed probe budgets or successful numerical checks. Sections S5.10 and S5.11 describe the omitted-mean and numerical terms; estimated control offsets require a separate negligibility argument.

S5.4 Exact observation inference and cross-target covariance

The exact Mata observation-deletion smoother uses a different variance approximation. For a target A , it smooths $y_i \hat{e}_{i,-i}$ on leverage and the diagonal of $XS_{xx}^{-1}AS_{xx}^{-1}X'$, using rank bins and a weighted local-linear tricube fit. The default `inferencebins()` is 1,000. Materially negative fitted variances withhold the request; only tiny negative roundoff is set to zero. Let $\mathcal{V}(A)$ denote the scalar variance estimate obtained from that target’s influence weights and Gaussian simulations under its fitted variance vector. For each pair of primitive targets, Mata refits the smoother for their sum and sets

$$\hat{\Omega}_{ts}^{\text{Mata}} = \frac{1}{2}\{\mathcal{V}(A_t + A_s) - \mathcal{V}(A_t) - \mathcal{V}(A_s)\}. \quad (\text{S5.19})$$

All fits use the same standardized Gaussian simulation draws. The resulting 3×3 covariance is checked for positive semidefiniteness, then mapped by T to four targets. Thus `lincom` has a coherent joint covariance and reproduces the total identity. Coherence does not make the separate smoothers a common variance model or establish the validity of every Wald contrast. This exact observation family is distinct from the sparse residual-moment procedure and its estimator-specific results below.

S5.5 Mean inference: fixed-origin reference and omitted randomness

Write u for the working outcome, d for the direction of a common physical outcome shift ($d = \mathbf{1}$ for unit observation rows), and $z = u - cd$. For the following fixed-origin formulas c is genuinely nonrandom. For the implemented approximation it is set equal to the observed retained physical mean and then held fixed inside the inference calculation.

S5.6 Exact observation-deletion algebra

Condition on a fixed, identified design X and write $u = \mu + \varepsilon$, $\mu = X\beta$, $\mathbb{E}\varepsilon = 0$. Let errors be independent with $V = \text{diag}(\sigma_i^2)$. All inverses below are taken on the identified coefficient space. Put

$$S_{xx} = X'X, \quad P = XS_{xx}^{-1}X', \quad M = I - P, \quad B_t = XS_{xx}^{-1}A_tS_{xx}^{-1}X', \quad (\text{S5.20})$$

where A_t is the symmetric coefficient-space matrix for target $\theta_t = \beta'A_t\beta$. Define

$$d_{ti} = \frac{B_{t,ii}}{M_{ii}}, \quad C_t = B_t - \frac{1}{2}\{\text{diag}(d_t)M + M\text{diag}(d_t)\}. \quad (\text{S5.21})$$

Existence requires the usual admissible deletion makers. Then $\text{diag}(C_t) = 0$, $e = Mu = M\varepsilon$, and the deleted residual is $\hat{e}_{i,-i} = e_i/M_{ii}$. This is the KSS leave-out kernel; see (Kline et al., 2020, p. 1862, equation (1); p. 1873, equation (8)).

For FEVC's variance and covariance targets, a common outcome shift is in the model span and leaves the plug-in component unchanged:

$$M\mathbf{1} = 0, \quad B_t\mathbf{1} = 0. \quad (\text{S5.22})$$

Consequently $Mz = Mu$, and for any genuinely fixed scalar c ,

$$\begin{aligned} \hat{\theta}_t(c) &= u'B_tu - \sum_i B_{t,ii}(u_i - c)\hat{e}_{i,-i} \\ &= z'B_tz - z'\text{diag}(d_t)Mz = z'C_tz. \end{aligned} \quad (\text{S5.23})$$

The centered proxy $s_i(c) = (u_i - c)\hat{e}_{i,-i}$ satisfies $\mathbb{E}s_i(c) = \sigma_i^2$: $\mathbb{E}\hat{e}_{i,-i} = 0$ and $\mathbb{E}[u_i\hat{e}_{i,-i}] = \sigma_i^2$. Thus $\mathbb{E}\hat{\theta}_t(c) = \theta_t$ under the fixed- c model. This argument is a change of outcome origin in the KSS identity, not a claim that replacing c by a sample statistic preserves exact unbiasedness.

Two identities are useful for implementation and for the omitted term:

$$C_t\mathbf{1} = -\frac{1}{2}Md_t, \quad \mathbf{1}'C_t\mathbf{1} = 0. \quad (\text{S5.24})$$

In general $C_t\mathbf{1}$ is not zero. Therefore centering does change the leave-out estimator even though it leaves the plug-in component and fitted residuals unchanged. The changed influence is

$$g_t(c) = C_tz = C_tu + \frac{c}{2}Md_t. \quad (\text{S5.25})$$

Computations may retain the original fitted coefficients because $B_tu = B_tz$; the outcome multiplying the diagonal correction must nevertheless be z . No new model fit is needed.

S5.7 High-rank covariance and uncentered noise probes

Set $\mu_c = \mu - c\mathbf{1}$. For fixed c , independent mean-zero errors, and finite fourth moments, the zero diagonal gives

$$\Omega_{ts}(c) = 4(C_t\mu_c)'V(C_s\mu_c) + 2\text{tr}(VC_tVC_s). \quad (\text{S5.26})$$

No Gaussian assumption is needed for this identity: diagonal terms that would involve third moments or fourth cumulants vanish. Gaussian error probes are a convenient way to compute the trace, not an assertion that observed errors must be Gaussian. With V known,

$$\hat{\Omega}_{ts}(c; V) = 4g_t(c)'Vg_s(c) - 2\text{tr}(VC_tVC_s) \quad (\text{S5.27})$$

is unbiased for (S5.26). The minus sign removes the quadratic noise already contained in the realized influence. Substituting an estimated variance model is an additional approximation and can yield invalid scalar or joint covariance estimates; the existing failure/status rules remain.

For independent numerical draws $\eta^{(r)} \sim N(0, \hat{V})$,

$$q_t^{(r)} = \eta^{(r)'} C_t \eta^{(r)}, \quad \text{Cov}(q_t^{(r)}, q_s^{(r)} \mid \hat{V}) = 2 \text{tr}(\hat{V} C_t \hat{V} C_s). \quad (\text{S5.28})$$

The original kernel acts on these zero-mean errors. Subtracting c would add a spurious signal term. Subtracting a draw-specific mean would replace C_t by a different kernel and approximate a different sampling model. This rule also applies to Gram, spectral, and critical-value simulation where those domains are already defined by the existing implementation.

Two retained variance constructions. Structured inference in Rust and exact Mata match calculations uses one positive common variance vector across all component targets. The residual-moment fitter models variances in an outcome-free basis Z_v , with $\mathbb{E}(e \circ e) = (M \circ M) Z_v \gamma$. It estimates the small residual Gram through direct covariance probes in Rust or coefficient contractions in Mata, then solves the moment equations. Since $M(u - c\mathbf{1}) = Mu$, its residual response, Gram operator, and design-only basis are unchanged by a constant shift. Both `structured_common` and `structured_leverage` retain their registered support, conditioning, positivity, and failure rules. Agreement between these models is not a test of omitted variance drivers.

The exact Mata observation method instead smooths the raw proxy $s_i(c)$ using leverage and a target-specific diagonal, with its existing binning and local-linear rules. It refits each primitive and each polarized target. If F_t denotes the scalar calculation with target t 's fit, its cross covariance is $(F_{t+s} - F_t - F_s)/2$. Centering therefore enters every such fit and influence. The structured common-vector formula is not substituted for this family. Neither family implements the paper's unrestricted variance-product construction (Kline et al., 2020, pp. 1874–1875).

For a valid primitive covariance, the fourth row and column follow from the fixed map

$$(\theta_\alpha, \theta_\psi, \theta_{\alpha\psi}) \mapsto (\theta_\alpha, \theta_\psi, \theta_{\alpha\psi}, \theta_\alpha + \theta_\psi + 2\theta_{\alpha\psi}).$$

Structured target availability and joint-matrix availability remain distinct. The high-rank Gaussian reference requires diffuse spectrum and influence; a requested option is not evidence for these conditions (Kline et al., 2020, p. 1873, Theorem 2).

S5.8 One leading mode and its remainder

For a particular target, let (λ, v) be its largest-absolute-value nonzero observation-space eigenpair, $B_t v = \lambda v$, $v'v = 1$. Equivalently it is the generalized eigenmode of (A_t, S_{xx}) . Since $B_t \mathbf{1} = 0$, $v' \mathbf{1} = 0$ and the leading score

$$\hat{b} = v'z = v'u \quad (\text{S5.29})$$

is algebraically invariant to centering. The raw recenter is not:

$$L(c) = \sum_i v_i^2 z_i \hat{e}_{i,-i}. \quad (\text{S5.30})$$

It is essential to retain this realized leave-out product separately from the positive modeled leading-score variance $\hat{V}_b = v' \hat{V} v$. Replacing $L(c)$ by $\hat{V}_b$ changes the remainder center while leaving its influence and trace tied to the original remainder kernel.

Define

$$B_R = B_t - \lambda v v', \quad d_{R,i} = \frac{B_{t,ii} - \lambda v_i^2}{M_{ii}}, \quad (\text{S5.31})$$

$$C_R = B_R - \frac{1}{2} \{ \text{diag}(d_R) M + M \text{diag}(d_R) \}. \quad (\text{S5.32})$$

The exact identity and remainder influence are

$$\hat{R}(c) = \hat{\theta}_t(c) - \lambda \{ \hat{b}^2 - L(c) \} = z' C_R z, \quad g_R = C_R z. \quad (\text{S5.33})$$

Production independently evaluates both sides and retains its scale-relative identity gate. At fixed c , $\mathbb{E}\hat{R} = R_0 = \theta_t - \lambda b_0^2$, where $b_0 = v' \mu$. The one-mode reference approximates $(\hat{b}, \hat{R})$ jointly by a Gaussian vector, under the additional mode-weight and diffuse-remainder conditions in (Kline et al., 2020, p. 1876, Assumption 2 and Theorem 3).

For a common variance fit, its covariance estimate is

$$\hat{\Sigma} = \begin{pmatrix} v' \hat{V} v & 2v' \hat{V} g_R \\ 2v' \hat{V} g_R & 4g_R' \hat{V} g_R - 2 \text{tr}(\hat{V} C_R \hat{V} C_R) \end{pmatrix}. \quad (\text{S5.34})$$

The exact Mata observation smoother uses its target-specific variance fit in this expression, with the centered raw proxy. Structured match inference uses the common variance fit on both backends. The corresponding population remainder variance has the plus trace sign, as in (S5.26). A second large mode, concentrated mode weights, or concentrated remainder influence can invalidate the one-mode reference even if the interval computation succeeds.

Write the entries as V_b , C_{bR} , and V_R and require the existing positivity and determinant checks. The maximal curvature is

$$\kappa = \frac{2|\lambda|V_b}{\sqrt{V_R - C_{bR}^2/V_b}}. \quad (\text{S5.35})$$

The numerator uses a variance, not a standard deviation. If $\rho_{1-\alpha}(\kappa)$ is the existing critical radius, form the ellipse

$$\mathcal{E} = \{ (b, R) : [(\hat{b}, \hat{R})' - (b, R)']' \hat{\Sigma}^{-1} [(\hat{b}, \hat{R})' - (b, R)'] \leq \rho_{1-\alpha}(\kappa)^2 \}. \quad (\text{S5.36})$$

The interval endpoints are the minimum and maximum of $\lambda b^2 + R$ over $\mathcal{E}$. Centering changes the observed center and covariance, not the deterministic image calculation or critical-value construction. The radius is based on the Andrews–Mikusheva curvature bound, not a same-design empirical quantile of shortest distances to a parabola (Andrews and Mikusheva, 2016, pp. 1256–1257). With the required asymptotic conditions and consistent studentization, KSS establish coverage of at least the nominal level (Kline et al., 2020, p. 1878, Lemma 7); this does not guarantee exact size, or establish those conditions for FEVC’s fixed-observed-mean approximation. Structured q1 retains target-local statuses and does not post a Gaussian $\mathbf{e}(V)$ or silently replace a failed q1 interval by a q0 interval.

S5.9 Frequency weights and declared matches

The supported match component route first freezes the nuisance offset. Every row in a declared match g then has the same FE row x_g . Let $F_g = \sum_{i \in g} f_i$ and define the mover aggregates

$$U_g = \frac{\sum_{i \in g} f_i u_i}{\sqrt{F_g}}, \quad X_g^c = \sqrt{F_g} x_g, \quad d_g = \sqrt{F_g}. \quad (\text{S5.37})$$

For each physical stayer observation append $U_i = u_i$, $X_i^c = x_i$ and $d_i = 1$. Let $N = \sum_{g \in \mathcal{G}_M} F_g + |\mathcal{I}_S| = d'd$. The physical-frequency mean and centered mover aggregate are

$$c = \frac{d'U}{N}, \quad Z_g = U_g - cd_g = \sqrt{F_g}(\bar{u}_g - c). \quad (\text{S5.38})$$

They are not obtained by subtracting the unweighted mean of U_g . For a stayer row the centered outcome is $Z_i = u_i - c$. Thus the mean is computed on the retained physical population before the mixed transformation. Target weights continue to determine A_t separately from regression mass. Distinct declared deletion IDs at a common coefficient cell remain distinct independent units.

For completeness, represent the original stored rows in square-root-frequency coordinates and let $w_{gi} = \sqrt{f_i/F_g}$. Within a match, $w'_g w_g = 1$, $\tilde{X}_g = w_g X_g^c$, and

$$P_{gg} = h_g w_g w'_g, \quad B_{t,gg} = b_{tg} w_g w'_g, \quad w'_g M_{gg}^{-1} \tilde{e}_g = \frac{w'_g \tilde{e}_g}{1 - h_g}. \quad (\text{S5.39})$$

Hence the centered physical-block correction reduces exactly to

$$\tilde{z}'_g B_{t,gg} M_{gg}^{-1} \tilde{e}_g = b_{tg} Z_g \hat{E}_{g,-g}. \quad (\text{S5.40})$$

Within-match contrasts orthogonal to w_g vanish from the target. Applying (S5.21) to the collapsed design gives the ordinary scalar kernel on the mixed independent units, with $Md = B_t d = 0$. Equations (S5.23)–(S5.34) therefore apply with $\mathbf{1}$ replaced by d , u by U , and z by Z . The q1 raw recenter sums $v_g^2 Z_g \hat{E}_{g,-g}$ over mover units and $v_i^2 Z_i \hat{e}_{i,-i}$ over physical stayer observations. Mover terms use products of match contractions rather than sums of rowwise products.

If Γ_g is the covariance of the unscaled stored-row error vector within match g , the aggregate variance is

$$\tau_g^2 = \frac{f'_g \Gamma_g f_g}{F_g}, \quad V = \text{diag}(\tau_g^2). \quad (\text{S5.41})$$

In literal physical-copy coordinates the same aggregate has variance $\mathbf{1}' \Gamma_g^{\text{phys}} \mathbf{1} / F_g$, where Γ_g^{phys} is the covariance on those expanded coordinates. Use one independent error probe per mover unit and per physical stayer observation. Frequencies create separate stayer observations but do not create additional independent mover matches. Whole-cluster deletion permits dependence within a match under independence across matches (Kline et al., 2020, p. 1863, Remark 3); it does not permit unrestricted correlation across different matches of a worker. If the nuisance offset was estimated on the same sample, the omitted offset uncertainty is a separate approximation that mean centering does not repair.

S5.10 What estimating the mean omits

The following derivation covers unit observations, mover aggregates and the mixed mover–stayer representation. Let d be the constant-shift direction, $a = d/(d'd)$, $c_0 = a'\mu$, $\bar{u} = a'u$, and $\delta = \bar{u} - c_0 = a'\varepsilon$. For observations $d = \mathbf{1}$; for matches use (S5.37). Put $h_t = \text{diag}(d_t)d$. The exact kernel satisfies

$$C_t d = -\frac{1}{2} M h_t, \quad d' C_t d = 0, \quad (\text{S5.42})$$

so

$$\hat{\theta}_t(\bar{u}) - \hat{\theta}_t(c_0) = \delta S_t, \quad S_t = h'_t M \varepsilon. \quad (\text{S5.43})$$

Both factors have mean zero. Their product need not have mean zero. In particular the exact estimated-mean bias is

$$b_t^{\text{mean}} = \mathbb{E}(\delta S_t) = a'VMh_t. \quad (\text{S5.44})$$

For unit observations this is $n^{-1}\mathbf{1}'VMd_t$. It vanishes under homoskedastic independent errors, but generally not under heteroskedasticity. There is no uniform claim that ignoring the product is conservative.

A sufficient target-specific condition for the fixed- c_0 approximation is

$$\delta S_t = o_p(s_t), \quad s_t^2 = \text{Var}\{\hat{\theta}_t(c_0)\}. \quad (\text{S5.45})$$

Root- n orders for both factors would give an $O_p(n^{-1})$ product, which is negligible if $s_t \asymp n^{-1/2}$; more generally it suffices that $n^{-1} = o(s_t)$. An upper bound $s_t = O(n^{-1/2})$ alone is insufficient. These orders are not automatic with many effects, large leverage, concentrated match masses, or a nearly degenerate target. For q1 the leading score is unchanged because $v'd = 0$. The relevant joint Mahalanobis scale can be smaller than the marginal remainder SD: using population covariance entries, a sufficient additional condition is

$$\delta S_R = o_p\left(\sqrt{V_R - C_{bR}^2/V_b}\right), \quad S_R = (\text{diag}(d_R)d)'M\varepsilon. \quad (\text{S5.46})$$

A large stored-row count and a small observed point-estimate difference do not verify (S5.45). The feasible covariance fit must also be stable to replacing c_0 by $\bar{u}$, particularly for the exact target-specific smoother.

S5.10.1 Exact Gaussian diagnostics

Let $J_c = I - da'$. The actual sample-mean estimator is

$$\hat{\theta}_t(\bar{u}) = u'K_t u, \quad K_t = J_c' C_t J_c = C_t + \frac{1}{2}\{ah_t' M + Mh_t a'\}. \quad (\text{S5.47})$$

Its diagonal is $K_{t,ii} = a_i(Mh_t)_i$, generally nonzero. Under Gaussian errors the exact moments are

$$\begin{aligned} \mathbb{E}\hat{\theta}_t(\bar{u}) &= \theta_t + \text{tr}(K_t V), \\ \text{Cov}\{\hat{\theta}_t(\bar{u}), \hat{\theta}_s(\bar{u})\} &= 4(K_t \mu)' V (K_s \mu) + 2 \text{tr}(V K_t V K_s). \end{aligned} \quad (\text{S5.48})$$

These dense formulas provide reference values for checking the implementation. Because $M\mu = 0$ and $Md = 0$, $K_t \mu = C_t(\mu - c_0 d)$. Thus mean estimation leaves the population linear influence unchanged and modifies only the quadratic-noise kernel in this maintained model. Applying (S5.47) to C_R gives the exact q1 remainder moments. The Gaussian leading/remainder covariance is $2v'VK_R\mu$; it equals the fixed- c_0 population covariance. With true V substituted, the expectation of the implemented frozen-mean covariance formula can also be checked exactly:

$$\mathbb{E}\hat{\Omega}_{ts}(\bar{u}; V) = 4g_{0t}' V g_{0s} + 4 \text{tr}(C_t V C_s J_c V J_c') - 2 \text{tr}(V C_t V C_s), \quad (\text{S5.49})$$

where $g_{0t} = C_t(\mu - c_0 d)$. Comparing this expression with (S5.48) isolates the omitted-mean effect before variance fitting.

For a direct scale diagnostic define $A_0 = a'Va$, $J_t = h_t' M V M h_t$, and $D_t = a'VMh_t$. Gaussian moment factoring then gives

$$\mathbb{E}[(\delta S_t)^2] = A_0 J_t + 2D_t^2, \quad \text{Var}(\delta S_t) = A_0 J_t + D_t^2. \quad (\text{S5.50})$$

The bias-to-SD and RMS-discrepancy-to-SD ratios therefore can be checked without repeated outcomes. They should be reported for each target and eligible q1 remainder, together with support and spectral diagnostics.

For non-Gaussian errors, the nonzero diagonal of K_t brings third moments and fourth cumulants back into the actual-Mean covariance. For independent errors, if $g_t = K_t\mu$, $m_{3i} = \mathbb{E}\varepsilon_i^3$, and $\kappa_{4i} = \mathbb{E}\varepsilon_i^4 - 3\sigma_i^4$, the terms added to the Gaussian covariance in (S5.48) are

$$2 \sum_i m_{3i}(g_{ti}K_{s,ii} + g_{si}K_{t,ii}) + \sum_i \kappa_{4i}K_{t,ii}K_{s,ii}. \quad (\text{S5.51})$$

For the diagnostic with independent standardized t_8 errors, symmetry sets the third moments to zero and the excess kurtosis is 1.5. Its actual-Mean covariance therefore adds $1.5 \sum_i \sigma_i^4 K_{t,ii}K_{s,ii}$ to the Gaussian reference. Equation (S5.44) remains exact. At fixed c_0 , the zero-diagonal covariance (S5.26) remains exact, but the estimator is not thereby Gaussian. These distinctions prevent a Gaussian moment comparison from being read as general sampling-coverage evidence.

When the diagnostic estimates a nuisance offset on each outcome draw, the dense point oracle applied to that draw's transformed outcome still checks the estimator's algebra. A diagonal-Gaussian moment oracle for a known offset is only a reference for that experiment: same-sample $\hat{\gamma}$ makes the transformed error dependent on the outcomes and generally correlated across units. The unconditional law must include that estimation step. Also, empirical q1 joint-covariance comparisons computed only for available fits are conditional on success; they cannot replace the separately reported unconditional attempt and availability counts.

S5.11 JLA and numerical uncertainty

The exact kernels above define the statistical reference and independent test oracles. Native JLA retains its finite-projection point formula, leverage/target probe domains, maker corrections, and solver contracts. An exact zero-diagonal identity must not be asserted for an arbitrary finite randomized approximation merely because it holds for the underlying exact operator. The implementation centers the working outcome after fitting, then passes it to the existing component attachment; match compression uses (S5.38). Residual-moment fitting retains the unchanged residuals.

Numerical diagnostics must distinguish point JLA, residual Gram, covariance trace, eigenmode, and critical-value error from sampling error. The reported point MCSE does not include all of these quantities and is not an econometric standard error. A repeated-seed/probe-budget experiment at fixed data tests numerical stability, while a sampling experiment redraws outcomes and recomputes the retained observed mean on every replication. It holds that newly observed mean fixed only inside that replication's inference calculation.

Checking the one-mode remainder with inexact solves. The remainder identity also depends on the residual of the original fit. To isolate numerical error, write $z = X\hat{\beta}_c + e$ and let $\Delta_R = \text{diag}(d_R)$ contain the remainder correction ratios. Define

$$q = A_t\hat{\beta}_c + \frac{1}{2}X'\Delta_R z, \quad r_0 = X'z - S_{xx}\hat{\beta}_c = X'e, \quad r_1 = q - S_{xx}h, \quad (\text{S5.52})$$

where h is the numerical solution of the combined influence system. After cancellation of the common leading-mode terms, the difference between the directly evaluated remainder $\hat{R}_{\text{dir}}$ and its

raw leave-out decomposition $\hat{R}_{\text{raw}}$ obeys

$$\begin{aligned}\hat{R}_{\text{dir}} - \hat{R}_{\text{raw}} &= z'Xh - \hat{\beta}'_c q \\ &= (S_{xx}\hat{\beta}_c + r_0)'h - \hat{\beta}'_c(S_{xx}h + r_1) = r_0'h - \hat{\beta}'_c r_1.\end{aligned}\tag{S5.53}$$

Symmetry of S_{xx} gives the final equality; neither residual is assumed zero. The numerical certificate subtracts this independently evaluated signed contribution before applying its existing discrepancy threshold, while retaining the raw absolute discrepancy in the diagnostic. It changes no point, influence, covariance or public solver tolerance. A discrepancy not explained by these signed terms fails the check, including the tested inconsistent recenter perturbations. The identity does not certify a deletion maker or recenter used consistently on both sides: the cancellation also holds for a common incorrect diagonal correction matrix. Those inputs require separate verification. This is a consistency check for accepted inexact solves, not a coverage result or a bound on all finite-probe error.

S5.12 Supported requests and interpretation

Table S5.1 distinguishes the statistical method from the backend. All component requests are opt-in, support Mean or None, and reject Corrected centering. The structured match routes accept the default `stayers(both)` or explicit `stayers(movers)`, positive integer frequency weights, `targetweight()` and `deletionid()`. The variance model must be explicitly selected as `structured_common` or `structured_leverage`.

Table S5.1: Component-inference methods and populations

Method	Calculation	Population and controls	Weights
Structured match	Exact Mata or Rust/JLA	Movers and eligible stayers, or movers only; fixed offsets	Frequencies and target weights
Structured observation	Rust/JLA	Both groups, or movers only; joint controls	No frequency-weight declaration
Observation smoother	Exact Mata	Observation deletion; joint or fixed-offset controls	Unit frequencies

For example, with original deletion units identified by `match`, the small-design structured route is

```
fevc y controls, worker(worker) firm(firm) ///
    deletion(match) deletionid(match) stayers(both) ///
    nuisance(fixedoffset) backend(mata) algorithm(exact) ///
    inference(highrank) inferencemodel(structured_common)
```

For a larger design use

```
fevc y controls, worker(worker) firm(firm) ///
    deletion(match) deletionid(match) stayers(both) ///
    nuisance(fixedoffset) backend(rust) rng(counter_v1) ///
    algorithm(jla) engine(generic) preconditioner(cmg) ///
    inference(highrank) inferencemodel(structured_common)
```

Here `controls` stands for the application’s control variables and can be omitted. `deletionid()` can be omitted for worker–firm pairs. The Rust route also accepts explicit `preconditioner(diagonal)`. Use `inference(q1)` when the one-mode assumptions are appropriate; there is no automatic reference selection. Native observation inference requires explicit `deletion(observation)`, `backend(rust)`, `rng(counter_v1)`, `algorithm(jla)`, a supported explicit preconditioner and structured model, and joint controls. Its engine can be omitted, `auto`, or `generic`. Even an all-one frequency-weight declaration is rejected on that observation route. Omitting `inferencemodel()` selects the exact Mata observation smoother.

Both structured backends use the covariance-draw count set by `inferencesimulations()`. The Rust Gram-probe count is set by `inferencegramprobes()`. Mata’s coefficient-contraction Gram uses no probes. The observation smoother uses `inferencebins()`. Exact Mata remains subject to its coefficient-dimension and memory limits; there is no structured Mata/JLA attachment. Structured component inference cannot share a call with `project()`; the exact Mata observation method can combine the two.

The mixed partition is recorded in three counts:

- `e(inference_mover_units)`: mover units;
- `e(inference_stayer_units)`: physical stayer observations;
- `e(inference_independent_units)`: their sum.

Highrank posts `e(V)` only when its joint covariance and all individual checks pass. Structured `q1` posts no substitute Gaussian `e(V)`. A missing interval, invalid covariance or failed mode is not a zero standard error. Report the variance model, centering, deletion population, numerical budgets and interval availability. Sampling terms from estimating the control offset and observed mean are omitted; component inference does not supply a joint-control match correction. Section S6 reports evidence separately for the distinct procedures and populations.

S5.13 Mover-only public Veneto example

This example uses None centering, a mover-only target and the explicitly fixed numerical settings below. Section S5.14 considers the pooled target with Mean centering.

The example uses the same 71,614-row subsample as the article’s decomposition: workers with regional code 1 observed in both years. Select it from the full public dataset and define the match identifier before requesting inference:

```
import delimited using kss_veneto.csv, ///
varnames(nonames) clear
rename (v1 v2 v3 v4 v5) (worker firm year region log_wage)
keep if region == 1
bysort worker (year): keep if _N == 2
egen long match = group(worker firm)
```

Component inference uses the mover population and fixed-offset convention:

```

fevc log_wage, worker(worker) firm(firm) ///
    deletion(match) deletionid(match) stayers(movers) ///
    nuisance(fixedoffset) backend(rust) rng(counter_v1) ///
    algorithm(jla) engine(generic) preconditioner(cmg) ///
    probes(200) batch(16) seed(8675309) ///
    tolerance(1e-10) maxiter(10000) memory_gib(16) ///
    memorycheck(error) ///
    centering(none) inference(q1) inferencemodel(structured_common) ///
    inferencegramprobes(2048) inferencesimulations(1000) ///
    inferenceeseed(8675309)

```

Replacing `inference(q1)` by `inference(highrank)` requests $q=0$. These calls use the same mover sample and point-estimation settings. Since no controls are supplied, there is no estimated control offset whose uncertainty is omitted. The structured variance model, cross-match independence, and q -specific concentration conditions still matter.

With the direct residual Gram and 2,048 probes, both requests complete on 12,828 retained observations and return all four intervals. The corresponding point-only call gives the same estimates and sample. These mover matches each contain one observation in the teaching extract; the synthetic size ladder supplies the separate repeated-match example. The corrected total variance is 0.06376, with a Wald standard error of 0.00107. The covariance interval includes zero under both references; the firm-variance $q1$ interval is wider and asymmetric around the estimate (table S5.2).

Table S5.2: Approximate component inference on mover matches

Target	Estimate	Wald SE	95% Wald interval	$q=1$ interval
Worker variance	0.04484	0.00213	[0.04066, 0.04902]	[0.04160, 0.04949]
Firm variance	0.01458	0.00146	[0.01172, 0.01743]	[0.01241, 0.01837]
Worker–firm covariance	0.00217	0.00136	[-0.00049, 0.00483]	[-0.00121, 0.00389]
Total variance	0.06376	0.00107	[0.06166, 0.06585]	[0.06166, 0.06586]

Notes: Direct residual-moment results at inference seed 8675309. The point estimates refer to movers, a different population from the 56,044-observation combined decomposition in the article. Wald columns use the $q0$ call; the $q1$ column uses the separate one-mode reference. Both use 200 point probes, 2,048 Gram probes and 1,000 covariance probes; $q1$ also uses 100,000 critical-value draws per target. This is a public implementation example, not a coverage experiment.

When a $q0$ request succeeds with an admissible joint covariance, ordinary Stata postestimation can be used:

```

lincom worker_variance + firm_variance + ///
    2*worker_firm_covariance
estat diagnostics

```

The $q1$ request does not post a Gaussian covariance. Ordinary `lincom` uses the covariance from an admissible $q0$ call; it does not form a $q1$ interval for an arbitrary new target. The researcher should inspect target-specific withholding and justify the variance model and dependence partition before interpreting the intervals.

Holding the point seed fixed, six preselected inference seeds return all 48 requested intervals. Across seeds, the coefficients of variation of the Wald standard errors are 5.17, 4.55, 6.47, and 1.04 percent for worker, firm, covariance, and total targets. The largest $q1$ endpoint range is 12.18

percent of the corresponding median interval width; at most 0.234 percent of fitted unit variances reach the positivity floor. These calculations measure numerical sensitivity on one outcome with point randomization held fixed. They neither establish sampling coverage nor justify selecting a seed or reference after seeing the intervals. The experiments below report the conditioning failures and numerical sensitivity under the different variance estimators.

S5.14 Pooled public Veneto component inference

The pooled illustration uses the same 71,614-row subsample and Mean convention as the article’s decomposition. Import the full public dataset with `asdouble`, rename its first five columns as in the article, and retain workers with regional code 1 observed in both years. The full registered highrank request is

```
set processors 4
fevc log_wage, worker(worker) firm(firm) ///
  deletion(match) stayers(both) nuisance(fixedoffset) ///
  backend(rust) rng(counter_v1) algorithm(jla) ///
  engine(generic) preconditioner(cmg) centering(mean) ///
  probes(200) seed(8675309) nativethreads(4) ///
  tolerance(1e-10) maxiter(10000) ///
  memory_gib(16) memorycheck(error) ///
  inference(highrank) inferencemodel(structured_common) ///
  inferencegramprobes(2048) inferencesimulations(1000) ///
  inference_seed(8675309) level(95)
```

The point-only comparison omits the inference options. Replacing `inference(highrank)` by `inference(q1)` gives the second reference calculation. The settings, order and comparisons were registered before execution; the model and numerical budgets were not selected from these interval results.

All three calls retain exactly the same 56,044 physical observations, 28,022 workers and 1,684 firms, with matching hashes of the retained rows. The inference partition has 12,828 mover units and 43,216 physical stayer observations. Each mover match in this extract contains one observation; repeated mover units are covered separately by the synthetic assessment. All four point estimates agree across the point-only and inference calls. No controls are included, eliminating estimated-offset uncertainty in this example. Mean estimation uncertainty remains omitted.

Table S5.3: Pooled component inference in the public Veneto example

Target	Estimate	Highrank SE	95% highrank interval	95% q1 interval
Worker variance	0.08514	0.00120	[0.08278, 0.08749]	[0.08273, 0.08760]
Firm variance	0.01024	0.00103	[0.00822, 0.01226]	[0.00819, 0.01235]
Worker–firm covariance	0.00468	0.00077	[0.00318, 0.00618]	[0.00310, 0.00623]
Variance of sum	0.10473	0.00041	[0.10394, 0.10553]	[0.10393, 0.10553]

Notes: The targets describe movers and eligible stayers together, with equal physical-observation target mass. Both requests use Mean, `structured_common`, 200 point probes, 2,048 Gram probes and 1,000 covariance simulations; q1 uses 100,000 critical-value draws per target. The highrank covariance is admissible and posted as $\mathbf{e}(\mathbf{V})$. The q1 call returns all four target intervals and posts no Gaussian covariance. Intervals rely on the respective spectral conditions, the variance model, independence across mover units and individual stayer observations, and negligible omitted-mean terms. This is not a coverage experiment.

The two reference calculations give similar intervals on this fixed outcome (Table S5.3). Highrank gives $[0.08278, 0.08749]$ for worker variance and $[0.00822, 0.01226]$ for firm variance. Both covariance intervals exclude zero under their maintained assumptions. The total-component standard error is smaller than either variance-component standard error because the joint calculation accounts for covariance among the component estimates.

Table S5.4: Spectral and influence diagnostics for the pooled Veneto targets

Target	Leading share	Remainder share	Largest mode weight	Largest influence share
Worker variance	0.044433	0.043833	0.269162	0.021719
Firm variance	0.046599	0.046075	0.266211	0.012187
Worker–firm covariance	0.048567	0.048106	0.268108	0.024057
Variance of sum	0.000034	0.000034	0.000372	0.001127

Notes: Leading and remainder shares use (S5.11); the largest mode weight is $\max_a v_{1a}^2$. The largest influence share is the variance-weighted concentration ratio in (S5.12). Shares use finite numerical budgets and are diagnostics, not automatic acceptance thresholds for a reference distribution.

The leading shares are below 0.05, but the largest squared leading-mode weight is about 0.27 for each primitive target. Similar highrank and q1 intervals therefore do not establish diffuse mode weights or justify a one-mode asymptotic approximation. Researchers must assess those conditions and the stayer-independence assumption separately. The input and package hashes, complete command logs, statuses, returned matrices and retained-sample hashes are recorded under `replication/experiments/pooled_inference_20261010/`.

S6 Statistical validation

The experiments below answer distinct questions, with separate sample definitions, failures and denominators. Algebra checks concern identities; fixed-data seed comparisons concern numerical error; coverage experiments redraw outcomes. Passing one layer does not establish another. The replication records bind each experiment to its source and input hashes, registration, raw results and exclusions.

S6.1 Fixed-design bias experiment

The experiment uses six workers, four firms, and four periods, with 24 observations and 17 declared matches. The worker–firm allocation is shown in table S6.1; a match consists of all observations for a worker–firm pair. Here w indexes workers, j firms, and τ periods. The signal is

$$y_{w\tau} = \alpha_w + \psi_{j(w,\tau)} + 0.3(\tau - 1.5) - 0.21\{\tau = 2\} + \varepsilon_{w\tau}, \quad (\text{S6.1})$$

with coefficient vectors

$$\begin{aligned} (\alpha_0, \dots, \alpha_5) &= (-0.8, -0.45, -0.1, 0.25, 0.65, 1.05), \\ (\psi_0, \dots, \psi_3) &= (0.35, -0.4, 0.55, 0). \end{aligned}$$

The true values are the quadratic targets evaluated at these coefficients and the retained sample selected from the zero-error outcome.

Table S6.1: Worker–firm allocation in the fixed-design experiments

Worker	$\tau = 0$	$\tau = 1$	$\tau = 2$	$\tau = 3$
0	0	0	1	1
1	0	2	2	1
2	1	2	3	3
3	2	3	0	0
4	3	1	1	2
5	3	3	2	0

Notes: Entries are firm identifiers. The bias and random-projection experiments use this allocation, with different outcome constructions.

The independent-error specification draws normal errors with standard deviation $0.35 + 0.08j + 0.04\tau$ and uses observation deletion. The match-error specification sets

$$\varepsilon_{w\tau} = 0.6\eta_{g(w,\tau)} + 0.25\nu_{w\tau}, \quad \eta_g, \nu_{w\tau} \stackrel{\text{iid}}{\sim} N(0, 1), \quad (\text{S6.2})$$

and uses match deletion. Each match shares η_g and is independent of other matches. Both regressions include the two controls in their deleted fits. There are 5,000 replications per specification, with no failed calls. The reported Monte Carlo standard errors use the variation across outcome replications.

S6.2 Random-projection approximation

The approximation exercise uses the same allocation but fixes the outcome at $y_{w\tau} = 1.5 + 0.3w - 0.2j(w, \tau) + u_{w\tau}$, without controls. For workers $w = 0, \dots, 5$, the four-period vectors u_w are, respectively, $(.2, -.1, .1, -.2)$, $(-.2, .3, -.1, .1)$, $(.1, -.2, .2, -.1)$, $(-.1, .2, -.2, .1)$, $(.3, -.2, .1, -.2)$, and $(-.2, .1, .2, -.1)$. The balanced design uses unit frequencies. The concentrated design assigns frequency 12 to worker 0’s two observations at firm 0, changing both the regression information and default target weights.

The reference is the deterministic exact Mata match-deletion calculation, checked against an independent dense calculation that expands the frequency weights and refits the regression after each literal match deletion. The randomized calls use generic Mata JLA, diagonal preconditioning, automatic batching, solver tolerance 10^{-12} , at most 10,000 iterations, and a strict 16-GiB direct-allocation budget. Each of the six probe counts in table S6.2 uses 1,000 independently seeded calls for each design. Separate Rust calls use Counter-V1 at 200 and 1,000 probes, also with 1,000 seeds per design. Equal seed numbers across the two backends do not imply equal random projections. Both stages use the stated probe count; all calls use the uncentered estimator, request `mcse(all)`, and hold the outcome fixed.

For each component, the table compares the standard deviation across complete point estimates with the square root of their mean raw estimated numerical variance (“All”). The conditional column (“Cond.”) reports the root mean square of the conditional target-probe diagnostic. These are estimates of one run’s numerical dispersion, without division by the square root of the number of runs, and are themselves subject to Monte Carlo error. Bias is the mean randomized estimate minus exact KSS. The replication records also report RMSE, mean absolute error, and the 95th percentile of absolute error for every component and cell.

Availability is counted over all planned runs. Finite raw covariances enter the variance average even when a materially indefinite matrix withholds the usable MCSE; averaging only available

diagnostics would select on the estimated covariance. Failure counts and any nonfinite raw estimates are retained separately. These fixed-data experiments assess the local approximation on the stated graphs and do not establish sampling coverage or uniform accuracy near singular deletion.

At 25 probes, usable MCSEs are withheld in one of the 1,000 balanced Mata runs and ten of the 1,000 concentrated Mata runs because the raw covariance is materially indefinite. Their point estimates remain valid. All other Mata diagnostics and all Rust diagnostics in the grid are available. There are no point-call failures in the 16,000-run synthetic experiment.

Table S6.2: Complete-run dispersion and numerical MCSE

Design	Target	Probes	Bias	Seed SD	All	Cond.	Usable
<i>Mata</i>							
Balanced	Worker	25	-0.00066	0.01154	0.01167	0.00851	999/1000
	Firm	25	-0.00018	0.00625	0.00620	0.00390	
	Covariance	25	-0.00025	0.00886	0.00896	0.00889	
	Sum	25	-0.00134	0.01949	0.01985	0.01558	
Balanced	Worker	50	-0.00014	0.00819	0.00829	0.00597	1000/1000
	Firm	50	-0.00007	0.00427	0.00435	0.00271	
	Covariance	50	0.00013	0.00599	0.00628	0.00624	
	Sum	50	0.00004	0.01439	0.01402	0.01086	
Balanced	Worker	100	-0.00027	0.00585	0.00577	0.00412	1000/1000
	Firm	100	-0.00003	0.00298	0.00303	0.00187	
	Covariance	100	0.00011	0.00430	0.00439	0.00436	
	Sum	100	-0.00008	0.00964	0.00982	0.00758	
Balanced	Worker	200	-0.00006	0.00414	0.00405	0.00289	1000/1000
	Firm	200	0.00008	0.00206	0.00213	0.00131	
	Covariance	200	0.00016	0.00311	0.00310	0.00308	
	Sum	200	0.00034	0.00716	0.00694	0.00536	
Balanced	Worker	500	0.00005	0.00262	0.00256	0.00183	1000/1000
	Firm	500	-0.00002	0.00130	0.00134	0.00082	
	Covariance	500	-0.00001	0.00199	0.00196	0.00195	
	Sum	500	0.00000	0.00446	0.00437	0.00337	
Balanced	Worker	1000	-0.00004	0.00184	0.00181	0.00130	1000/1000
	Firm	1000	-0.00002	0.00093	0.00095	0.00058	
	Covariance	1000	-0.00001	0.00142	0.00139	0.00138	
	Sum	1000	-0.00008	0.00295	0.00309	0.00239	
Concentrated	Worker	25	0.00034	0.02178	0.02115	0.01320	990/1000
	Firm	25	0.00027	0.00626	0.00630	0.00262	
	Covariance	25	0.00046	0.01374	0.01344	0.01161	
	Sum	25	0.00154	0.04648	0.04539	0.03044	
Concentrated	Worker	50	-0.00062	0.01527	0.01557	0.00921	1000/1000
	Firm	50	-0.00017	0.00455	0.00460	0.00167	
	Covariance	50	-0.00004	0.00991	0.00965	0.00820	
	Sum	50	-0.00086	0.03339	0.03343	0.02147	
Concentrated	Worker	100	-0.00004	0.01089	0.01099	0.00625	1000/1000
	Firm	100	-0.00007	0.00309	0.00326	0.00111	
	Covariance	100	0.00023	0.00710	0.00678	0.00574	
	Sum	100	0.00034	0.02397	0.02361	0.01487	
Concentrated	Worker	200	0.00048	0.00794	0.00773	0.00438	1000/1000
	Firm	200	0.00018	0.00230	0.00233	0.00077	
	Covariance	200	0.00010	0.00488	0.00475	0.00402	
	Sum	200	0.00086	0.01711	0.01658	0.01034	

Table S6.2 (continued)

Design	Target	Probes	Bias	Seed SD	All	Cond.	Usable
Concentrated	Worker	500	0.00015	0.00492	0.00491	0.00276	1000/1000
	Firm	500	0.00006	0.00149	0.00147	0.00047	
	Covariance	500	0.00011	0.00295	0.00300	0.00253	
	Sum	500	0.00042	0.01053	0.01052	0.00651	
Concentrated	Worker	1000	0.00014	0.00350	0.00348	0.00195	1000/1000
	Firm	1000	0.00004	0.00103	0.00104	0.00033	
	Covariance	1000	0.00002	0.00215	0.00213	0.00179	
	Sum	1000	0.00022	0.00748	0.00745	0.00461	
<i>Rust</i>							
Balanced	Worker	200	-0.00002	0.00398	0.00407	0.00291	1000/1000
	Firm	200	0.00004	0.00211	0.00213	0.00131	
	Covariance	200	0.00011	0.00298	0.00312	0.00309	
	Sum	200	0.00024	0.00665	0.00694	0.00536	
Balanced	Worker	1000	0.00003	0.00175	0.00181	0.00129	1000/1000
	Firm	1000	0.00005	0.00092	0.00095	0.00058	
	Covariance	1000	-0.00005	0.00143	0.00138	0.00137	
	Sum	1000	-0.00002	0.00308	0.00309	0.00238	
Concentrated	Worker	200	0.00018	0.00773	0.00778	0.00440	1000/1000
	Firm	200	-0.00003	0.00230	0.00232	0.00076	
	Covariance	200	0.00019	0.00460	0.00474	0.00400	
	Sum	200	0.00053	0.01626	0.01664	0.01036	
Concentrated	Worker	1000	-0.00007	0.00346	0.00348	0.00195	1000/1000
	Firm	1000	-0.00004	0.00104	0.00104	0.00033	
	Covariance	1000	0.00006	0.00210	0.00213	0.00179	
	Sum	1000	0.00001	0.00743	0.00746	0.00461	

The article’s public Veneto decomposition is also repeated at 200 and 1,000 probes with eight seeds specified before execution, using the displayed Rust call with the default mean centering. This is an application sensitivity check; eight runs do not give a precise calibration test. The source-only Mata example is checked separately at its displayed seed. Numerical MCSE does not quantify the sampling uncertainty of the decomposition or the numerical uncertainty of the separate projection covariance.

For firm variance at 200 probes, the square root of the mean estimated numerical variance is 0.0000897, and the eight corrected estimates have a standard deviation of 0.0000883 and range from 0.01014 to 0.01038. At 1,000 probes, the corresponding values are 0.0000388 and 0.0000162, with estimates from 0.01027 to 0.01032. With only eight runs, the smaller dispersion at 1,000 probes is imprecisely estimated. All 16 Rust diagnostics and the Mata diagnostic are available. Numerical variation is therefore small relative to the bias correction of 0.00958 even at 200 probes.

S6.3 Projection experiment

The projection experiment uses an independent dense Python calculation with exact deleted fits. It has 40 workers ($w = 0, \dots, 39$), 20 periods ($\tau = 0, \dots, 19$), and 20 firms. The allocation $j(w, \tau) = (w + \lfloor \tau/2 \rfloor) \bmod 20$ gives 800 observations in 400 two-observation matches. Firm effects are projected on an intercept and $\sin(w/5) + \cos(j(w, \tau)/3) + \tau/20$, with unit weights and no controls. The signal has $\alpha_w = 0.03w$, $\psi_j = 0.03(40 + j)$ for $j = 0, \dots, 18$, and $\psi_{19} = 0$. Errors are jointly normal within each match and independent across matches, with within-match correlation 0.45 and standard deviation $0.4 + 0.0002i$ for observation $i = 20w + \tau$ (zero-based in this design).

The experiment draws 3,000 outcome replications using seed 20260902 and evaluates the block covariance formula with exact linear algebra. It measures the statistical behavior of that formula; comparisons between the Stata command's Mata and Rust calculations use separate fixed-input examples. The true covariance of the projection intercept and slope is

$$\begin{pmatrix} 0.01042258 & -0.00090719 \\ -0.00090719 & 0.00032228 \end{pmatrix}.$$

Across 3,000 replications, its mean estimate is

$$\begin{pmatrix} 0.01032362 & -0.00089505 \\ -0.00089505 & 0.00031931 \end{pmatrix}.$$

The 95 percent slope interval covers in 0.9400 of replications, with Monte Carlo standard error 0.00434. One replication has a nonpositive estimated slope variance and is counted as noncoverage.

A separate fixed-input Stata fixture compares exact literal expansion, exact stored-row aggregation, and generic Rust/JLA. It crosses observation and match deletion, positive frequencies, joint controls, eligible stayers, and equivalent normalization shifts. Counter-V1 seeds 20260901–20260908 use 600 and 1,200 probes, batch size 8, and tolerance 10^{-12} . The repeated-seed means of covariance cells, standard errors and endpoints are compared with the exact results using the deterministic floor $10^{-8} \max(1, |a|, |b|)$ plus six empirical numerical MCSEs. Seven 600-probe streams return a valid projection covariance; the eighth is withheld by the positive-semidefinite check. All eight 1,200-probe streams return valid results. The accepted comparisons pass the numerical criterion and complete-system residual checks. This fixture is separate from the 3,000-outcome experiment above.

S6.4 A numerical illustration

Consider seven worker–firm observations, indexed by the pairs

$$(2, 2), (1, 1), (0, 2), (2, 1), (1, 0), (0, 1), (2, 0). \quad (\text{S6.3})$$

Use worker indicators for workers 0, 1, 2 and firm indicators for firms 0, 1, grounding firm 2. Thus $k = 5$, $n - k = 2$, and each observation can be deleted. Set $\boldsymbol{\sigma}^2 = (1, 1, 1, 1, 1, 4)'$ and $\boldsymbol{\ell} = (0, 0, 0, -1/2, 0)'$. This loading is the slope from projecting the three firm effects on a constant and the firm indices 0, 1, 2.

For $V_\ell = \text{Var}(\ell' \hat{\beta} \mid \mathcal{D})$, the variance estimates use weights $(x_i' S_{xx}^{-1} \ell)^2$. Here $V_\ell = 2/3$. The uncentered and full-mean-centered estimates have exact expectations

$$\frac{2}{3}, \quad \frac{53}{84}, \quad (\text{S6.4})$$

respectively. Their relative biases are 0 and -5.36% . The proposed correction restores expectation $2/3$.

For the precision comparison, take independent Gaussian errors with the specified variances and set

$$\mu_i = 10 + g_i - \bar{g}, \quad g_i = 0.5 \sin(1.3 \text{ worker}_i) + 0.3 \cos(1.7 \text{ firm}_i), \quad (\text{S6.5})$$

where $\bar{g}$ averages g_i over the seven observations. This mean vector lies in the column space of X . Table S6.3 gives exact Gaussian moments, not simulated point estimates.

Table S6.3: Estimating a true contrast variance of 2/3

Method	Expectation	Bias (%)	Sampling SD	MSE
Uncentered KSS	0.666667	0.00	3.14775	9.90833
Centered	0.630952	-5.36	0.77655	0.60430
Corrected centered	0.666667	0.00	0.88418	0.78178

Centering delivers a large MSE reduction at this outcome location. The unbiased correction retains most of it, although the biased full-mean version has lower MSE in this case. With the common level changed from 10 to zero, raw and corrected-full-mean MSEs are instead 0.69543 and 0.78178: the correction does not uniformly dominate the original KSS estimator.

S6.5 Pooled structured component inference

The pooled assessment evaluates the common structured variance calculation on exact Mata and Rust/JLA, separately for highrank and q1. The design has a 12×12 mover grid with unequal integer mover-unit masses and 36 attached stayers with histories of two, four or eight physical observations. Stayers receive either 25% or 75% of total target mass. Their errors are independent across physical observations. Positive quadratic variance functions in within-type leverage rank differ between movers and stayers and lie in the specified model spans on the exact design. The q1 designs concentrate stayer target mass on one worker. The full geometry, variance functions and outcome-free checks are registered with the replication files.

The eight backend/reference/mass cells each use 1,000 outcomes and two paired arms: the observed retained Mean and a fixed population-mean origin. This gives 16,000 calls and 64,000 target attempts. All four highrank targets are primary; worker, firm and total variance are primary for q1. The q1 covariance target is a predeclared diagnostic because its remainder has another concentrated mode. Every primary interval is available, and the registered descriptive screens pass (Table S6.4). These are bounded synthetic results, not general coverage qualification.

Table S6.4: Pooled structured inference: primary synthetic designs

Calculation	Reference	Stayer mass (%)	Availability (%)	Coverage range (%)
Exact Mata	Highrank	25	100	93.9–95.3
Exact Mata	Highrank	75	100	93.3–94.9
Exact Mata	q1	25	100	94.8–96.7
Exact Mata	q1	75	100	96.0–97.1
Rust/JLA	Highrank	25	100	96.0–96.7
Rust/JLA	Highrank	75	100	93.6–95.7
Rust/JLA	q1	25	100	94.2–97.0
Rust/JLA	q1	75	100	95.8–98.7

Notes: Nominal coverage is 95%. Each row summarizes 1,000 outcomes in the observed-Mean arm. Coverage ranges span the four highrank primary targets or the three q1 primary targets; q1 covariance is excluded from the primary set before outcomes. Availability and coverage use all attempts as the denominator. Rust uses 200 point probes and 2,048 Gram probes; both backends use 1,000 covariance simulations and the same q1 critical-draw floor. Backend-specific outcome streams make cross-backend coverage comparisons unpaired. No targets are selected after seeing coverage.

The screens require at least 95% availability and assess bias, lower coverage, highrank RMS standard errors and the difference between the paired centering arms. Their broad tolerances include a 90% coverage floor with Monte Carlo allowances; q1 has no upper-coverage penalty. Passing them does not establish exact nominal size. Across both arms, 32 q1 covariance intervals are withheld: eight in Mata at 25% stayer mass, six in Mata at 75%, and eighteen in Rust at 75%. There are no failures affecting all targets in a call. The largest primary coverage reduction from estimating the mean is 0.1 percentage point, and the design-based RMS omitted-mean discrepancy is at most 0.000421 target sampling standard deviations. These small discrepancies concern the registered designs, not arbitrary samples.

Stress cases and numerical sensitivity. Nine diagnostic cells use 100 paired outcomes each: 1,800 calls and 7,200 target attempts, of which 446 lack intervals. All calls complete, but successful execution is not a coverage criterion. In the observed-Mean arm, q1 null worker and total targets have 44% availability and 42% all-attempt coverage. A deliberately multiple-mode q1 worker target has 77% availability and 74% coverage. Weak highrank firm and covariance targets have 91% and 90% availability and 86% and 83% coverage, respectively. Under severe variance-model misspecification, the q1 covariance diagnostic has 93% availability and 82% coverage. The complete inventory also retains short histories, heavy tails, estimated offsets and leverage-model sensitivity. These diagnostic cells have no primary coverage acceptance threshold.

A separate fixed-outcome grid contains 96 calls with four outcomes, three numerical seeds and four budget settings. All target intervals are returned. Increasing point probes from 200 to 800 changes the more affected interval endpoint by a median 2.35% and maximum 23.00% of the baseline interval width across primary targets. Raising Gram probes from 2,048 to 8,192 gives 1.40% and 9.15%; raising covariance simulations from 1,000 to 4,000 gives 0.058% and 0.276%. These comparisons measure sensitivity to individual budgets, not convergence or sampling coverage. Neither seeds nor budgets were retuned after inspecting these results. All failures, source identities, registrations and machine-readable summaries accompany the assessment.

S6.6 Observation and mover-only calibration by variance estimator

The experiments below evaluate three distinct variance estimators. Results for one procedure are not pooled with those for another. We begin with the direct residual-Gram estimator used by the command, then report the cross-fitted and subtractive alternatives. All unavailable intervals remain in the stated counts.

Direct residual-Gram estimator. The direct residual-Gram experiment uses 2,048 Gram probes and 200 point probes. Across eight primary cells and 200 saved outcomes per cell, all 1,600 calls complete with unchanged point estimates, but three cells fail their prespecified calibration criteria. A separate exact-Gram calculation on those 600 saved outcomes resolves the match one-mode firm case and the observation dominant-mode firm case. Their 2,048-probe RMS standard errors are 9.3 and 5.9 percent below their exact-Gram counterparts. For observation deletion with jointly estimated controls, the difference is about 1.0 percent, and the exact-Gram SD/RMS-SE ratio remains 1.151. This last comparison concerns joint observation controls, not omitted fixed-offset uncertainty. Exact Gram calculation is diagnostic only, not a production fallback or an alternative variance estimator.

A separate comparison uses 70 fixed saved-outcome replays, including severe misspecification, null, weak and multiple-mode diagnostics. All 280 target attempts are accounted for: 277 intervals are computed and three are withheld by target-specific checks, with no shared fitting failures. The primary replays agree with the saved 2,048-probe calculation, and four observation/match q0/q1 cases agree with the public Stata command. These fixed-outcome comparisons measure agreement between implementations; they do not estimate coverage.

Cross-fitted variance estimator. Table S6.5 reports primary correct-model results for the cross-fitted match variance estimator. The two independent confirmation campaigns each use 400 matches, 14 designs, and 2,500 outcome replications per design. The q0 campaign computes 59,996 of 60,000 primary target intervals; the q1 campaign computes 44,995 of 45,000. The remaining designs diagnose misspecification, controls, weak/null signals, multiple modes, and solver behavior. Their attempts remain in the full inventory of 140,000 per campaign.

The within-match designs vary physical covariance while keeping the scalar aggregate variance in the specified model. Independent, common-shock, and serial errors thus exercise different dependence shapes, not different assumptions about cross-match independence. The primary common-model variance is affine in normalized match-mass rank; the leverage sensitivity design uses its specified leverage model. These are fixed-offset designs without re-estimated controls. The one-mode designs have mean leading shares 0.907–0.935 and remainder shares 0.304–0.408 for their eligible targets. The covariance target has a remainder share around 0.94 and is excluded from the q1 coverage claim. These are averages of the finite-probe ratios in (S5.11), using the numerical denominators described above. The diagnostics and coverage results do not demonstrate a diffuse remainder along an asymptotic sequence.

Table S6.5: Cross-fitted variance estimator: primary correct-model designs

Design	Returned	W	F	C	T	MCSE range
<i>Panel A: diffuse designs, $q = 0$ (10,000 attempts per row)</i>						
Equal mass, independent	10,000	94.92	94.92	95.44	95.00	0.42–0.44
Unequal mass, independent	10,000	94.60	93.88	95.20	94.72	0.43–0.48
Unequal mass, common shock	9,996	94.36	95.04	95.24	94.44	0.43–0.46
Unequal mass, serial	10,000	94.60	94.16	94.84	94.84	0.44–0.47
Equal aggregate variance	10,000	94.48	95.56	94.88	95.92	0.40–0.46
Leverage-model sensitivity	10,000	94.20	95.16	94.52	94.92	0.43–0.47
<i>Panel B: one-mode designs, $q = 1$ (7,500 eligible attempts per row)</i>						
Equal mass, independent	7,499	95.48	96.40	—	96.20	0.37–0.42
Unequal mass, independent	7,497	95.20	95.40	—	95.72	0.40–0.43
Unequal mass, common shock	7,500	95.20	95.28	—	95.32	0.42–0.43
Unequal mass, serial	7,499	95.12	94.92	—	95.56	0.41–0.44
Equal aggregate variance	7,500	94.24	95.72	—	95.92	0.40–0.47
Leverage-model sensitivity	7,500	95.76	96.44	—	96.76	0.35–0.40

Notes: Coverage and Monte Carlo SE are in percentage points. W , F , C , and T denote worker variance, firm variance, their covariance, and total variance. Each target has 2,500 attempts. Returned counts pool the displayed targets; coverage divides by returned intervals for each target. MCSE is $\sqrt{\widehat{p}(1-\widehat{p})/n_{\text{returned}}}$. Dashes mark covariance targets excluded from q1 calibration before the run. These results use the cross-fitted variance estimator, not the direct residual-moment fitter. All displayed rows meet the prespecified availability, coverage, bias, and standard-error criteria described below. Counts for diagnostic designs and all failures affecting multiple targets are reported in the text.

The correct-model calibration criteria require at least 98 percent availability for each target, absolute coverage error no greater than $\max(0.03, 3 \text{ MCSE})$, empirical-SD/mean-SE ratios in $[0.82, 1.18]$, and bias within four MCSEs. Passing these tolerances is evidence for the tested designs, not exact 95 percent calibration. Both campaigns use 256 estimator probes, 512 covariance probes, 128 spectrum probes, and 256 spectral iterations, with fixed outcome-free folds and independent confirmation seeds. The q1 scientific intervals use an independent deterministic critical-value reference; production simulated critical values were checked separately. These tests are distinct from the public command’s larger critical-draw budget.

Across all designs, q0 returns 138,988 intervals and withholds 1,012 target attempts: 256 through variance-fit positivity failures and 756 through shared covariance failures. The q1 campaign returns 119,564 intervals, with 208 target-specific and 20,228 shared failures; 20,000 of the latter are weak/null diagnostic attempts. Each shared failure counts against all four targets. Severe omitted variance drivers reduce total coverage to 80.44 percent in the q0 campaign and 84.00 percent in the q1 campaign. Neither successful fits nor the passing primary rows erase these limitations. Exact seeds, excluded targets, failure inventories, and source identities accompany the machine-readable replication records.

Estimated controls remain a separate limitation. In a paired 400-match diagnostic with two controls and 1,000 replications, total-component q1 coverage is 95.1 percent with a known offset and fitted match variances, 75.9 percent with an estimated offset and fitted variances, and 76.1 percent when the estimated-offset calculation receives the true original match variances. Better fitting those scalar variances does not recover the covariance induced by estimating controls. This design is not a prediction of coverage in every application, but it rules out justifying the approximation by the number of controls alone.

The cross-fitted Rust observation experiment uses joint controls, unit frequencies and movers; it does not assess the stayer population. An observation-q1 confirmation fails one prespecified standard-error-ratio criterion: 1.101204 against a maximum of 1.10, with 93.52 percent coverage in that firm-target design. The reported ratio exceeds that criterion. These results concern the cross-fitted observation procedure used in that experiment.

Subtractive variance estimator. The 512-probe subtractive variance estimator has a separate assessment. It reuses 400 saved outcomes in each of 28 match cells, comparing the cross-fitted and subtractive fitters at identical numerical settings: 22,400 calls with 200 point probes and 1,000 covariance probes. All four targets and failed attempts are retained. An additional 160 calls replay four saved outcomes in each of 20 observation cells under both interfaces; this small slice compares numerical results without estimating observation coverage. Point estimates agree across fitters, and the subtractive fitter has no failures affecting all targets.

In the 24 eligible correct-model q0 target rows, subtractive-fitter coverage ranges from 93.00 to 96.75 percent; in the 22 eligible q1 rows, it ranges from 92.50 to 97.25 percent. All intervals in those rows are available. The empirical-SD/mean-SE ranges are 0.934–1.067 and 0.948–1.057, respectively. These rows meet both the prespecified lower-coverage requirement and the two-sided calibration criteria. The outcomes were reused from the experiments that informed the estimator; this comparison is not an independent confirmation of nominal coverage.

The subtractive fit retains the substantive limits. Severe variance-model misspecification gives total-component coverage of 81.75 percent for q0 and 88.25 percent for q1. With estimated offsets in the controls diagnostic, the corresponding rates are 66.25 and 79.00 percent, despite complete

computational availability. Weak/null designs can still yield unavailable or miscalibrated intervals. The replication records retain all 22,560 calls and 90,240 target attempts, including exclusions, diagnostics, and the software identities needed to reproduce them.

For perspective, reaggregating the subtractive fitter’s 400 saved draws in the observation-controls cell gives firm SD/RMS-SE ratios of 1.150 in the first half, 0.949 in the second, and 1.054 overall. Corresponding coverage is 92.0, 95.5 and 93.75 percent. These are subtractive-fitter results and do not supply a 400-draw assessment of the adopted direct implementation. They show why a small calibration sample should not be treated as a precise population statement; the three prespecified calibration failures remain.

These results support treating component inference as an experimental, model-based extension. Compare a few inference seeds chosen in advance and report material sensitivity; more Gram probes can reduce numerical error, but neither a stable answer nor a preferred seed or q establishes coverage.

Centering convention. The preceding component-inference campaigns used None. Their results are not a validation of the later Mean default.

S6.7 Numerical sensitivity

The point-estimation probe experiment varies complete-call seeds and compares against exact targets. The projection fixtures additionally vary probe counts and seeds for a complete covariance matrix, including the withheld 600-probe run. These checks show why an accepted regression residual and a small numerical MCSE for the point estimates do not by themselves establish accuracy of every reported inferential quantity. The current default includes both leverage and target probes for the four main point estimates; it does not provide numerical MCSEs for projection covariances, component sampling covariances, or interval endpoints. The experiments below retain their original source identities and numerical diagnostics.

A separate fixed-data experiment repeats q_0 and q_1 on the public Veneto example and the 61,440-row synthetic dataset from the size ladder. The retained samples contain 12,828 and 30,720 matches, respectively. Veneto has no controls; the synthetic specification retains its two estimated controls and fixed-offset convention. Outcomes, controls, target definitions, solver tolerances, and the estimator implementation remain unchanged.

The estimator block crosses five seeds with 200 and 800 point probes, holding the inference seed and 1,000 covariance probes fixed. The inference block crosses five inference seeds with 1,000 and 4,000 covariance probes, holding the estimator seed and 800 point probes fixed. A shared setting appears in both blocks for each dataset and q procedure, giving 76 unique calls rather than 80. The q_1 critical-draw count is 100,000 or 400,000 per eligible target. The public spectrum budget stays at 128 probes and iterations. Inference seeds also change the variance-regression folds and spectral calculations, so this block is not a pure critical-value simulation check. The numerical settings were specified before execution.

The experiment returns 299 of 304 target intervals in 27.6 minutes, with at most two concurrent eight-processor Stata/MP calls on the local workstation. Inaccurately recovered spectral modes withhold one synthetic q_0 call and the firm interval in the corresponding q_1 call; the other three q_1 targets remain available. There are no timeouts, memory stops, or missing attempts. All returned calls retain the same observations within each dataset; where both calls return, q_0/q_1 point estimates agree.

Table S6.6: Cross-fitted variance estimator: numerical sensitivity across five seeds

Data	Mode	Seed varied	Budget	SE range (%)	Endpoint range (% of width)	Intervals
Veneto	$q = 0$	Estimator	200	42.4	75.3	20/20
Veneto	$q = 0$	Estimator	800	21.9	39.0	20/20
Veneto	$q = 0$	Inference	1,000	8.1	4.0	20/20
Veneto	$q = 0$	Inference	4,000	4.8	2.4	20/20
Veneto	$q = 1$	Estimator	200	42.4	49.6	20/20
Veneto	$q = 1$	Estimator	800	21.9	20.2	20/20
Veneto	$q = 1$	Inference	1,000	8.1	6.3	20/20
Veneto	$q = 1$	Inference	4,000	4.8	6.6	20/20
Synthetic	$q = 0$	Estimator	200	0.2	2.2	20/20
Synthetic	$q = 0$	Estimator	800	0.1	1.1	20/20
Synthetic	$q = 0$	Inference	1,000	1.3	0.6	20/20
Synthetic	$q = 0$	Inference	4,000	0.4	0.2	16/20
Synthetic	$q = 1$	Estimator	200	0.2	2.2	20/20
Synthetic	$q = 1$	Estimator	800	0.1	1.1	20/20
Synthetic	$q = 1$	Inference	1,000	1.3	0.7	20/20
Synthetic	$q = 1$	Inference	4,000	0.4	0.4	19/20

Notes: These results use the cross-fitted variance model, not the direct residual-moment fitter. Budget means point probes in the estimator block and covariance probes in the inference block. For each target, SE range is the maximum minus minimum SE divided by its median. Endpoint range is the larger of the lower- and upper-endpoint ranges divided by median interval width. Each displayed percentage is the maximum across the four targets, not necessarily the same target in both columns. q1 uses its AM endpoints; its SE column describes the accompanying high-rank SE. Interval counts include four targets at five seeds. Four shared calls appear in both blocks; they are counted once in the 76-call inventory. These are numerical sensitivity measures, not coverage rates or estimates of sampling error.

Veneto is sensitive to estimator randomization (table S6.6). The largest endpoint range across targets falls from 75.3 to 39.0 percent of interval width for q0 and from 49.6 to 20.2 percent for q1 when point probes increase from 200 to 800. Point-estimate ranges reach 2.14 reported SEs at 200 probes and 1.10 at 800. For the worker–firm covariance, whether the interval excludes zero changes across estimator seeds for both procedures at 200 probes and for q0 at 800. With the estimator setting fixed, inference-seed endpoint ranges are smaller: at most 4.0 percent for q0 and 6.6 percent for q1 across the two covariance budgets. Increasing that budget does not uniformly reduce q1 variation.

The synthetic example is more stable. Its largest estimator-seed endpoint range is 2.2 percent at 200 probes and 1.1 percent at 800 for either procedure; inference-seed ranges are below one percent among returned intervals. This contrast does not establish that larger datasets are easier. It shows why numerical precision should be checked on the application rather than inferred from a successful return or a fixed probe count. In particular, the cross-fitted example’s 200-probe intervals illustrate the workflow but should not be treated as numerically settled empirical results.

For Veneto, an additional experiment repeats the same five estimator seeds at 3,200 and 12,800 point probes, holding the inference seed, 1,000 covariance probes, and all other settings fixed. It adds 20 calls and reuses the 20 lower-budget Veneto estimator-block calls in table S6.7.

Table S6.7: Cross-fitted variance estimator: Veneto sensitivity by probe count

Point probes	Estimate range (SE units)		Endpoint range (% of width)		Intervals	
			$q = 0$	$q = 1$	$q = 0$	$q = 1$
200	2.14	75.3	49.6	20/20	20/20	20/20
800	1.10	39.0	20.2	20/20	20/20	20/20
3,200	0.49	12.9	12.8	20/20	20/20	20/20
12,800	0.28	7.4	7.4	20/20	20/20	20/20

Notes: These results use the cross-fitted variance model. Each range is the maximum across the four targets at five estimator seeds. Estimate range is divided by the median high-rank SE; endpoint range uses the definition in table S6.6. q1 uses AM endpoints. The 200- and 800-probe rows use the same calls as the sensitivity grid; the 3,200- and 12,800-probe rows extend the probe range. All other settings are held fixed.

All 20 higher-probe calls return all 80 intervals in 33.4 minutes. At 12,800 probes, the largest endpoint range is 7.4 percent of interval width for either procedure, the largest point-estimate range is 0.28 high-rank SEs, and the largest relative SE range is 3.9 percent. The largest paired endpoint change from 3,200 to 12,800 probes is 8.8 percent of higher-budget width for q0 and 7.2 percent for q1. No target changes zero exclusion across seeds at either higher budget. Nevertheless, all q0 covariance intervals exclude zero whereas all q1 covariance intervals include it: reducing numerical variation does not make the two references interchangeable. Median command times rise from about 91 seconds at 3,200 probes to about 308 seconds at 12,800, per q procedure. The higher budget improves precision on this example; it is not a universally sufficient setting.

The higher-budget results are not exact references, and five seeds do not bound the tail of numerical error. Independent deterministic calculations of the critical value and ellipse image check these numerical steps separately. These numerical exercises do not test the omitted-offset or structured- variance assumptions.

S6.8 Observation and mover-only Mean diagnostics

The diagnostic uses four fixed outcomes, three numerical seeds (410081, 410089, and 410099), and four budget settings. All 48 paired tasks, 96 calls, and 384 target rows pass engineering checks; the outcome-file hashes are identical across numerical settings. The baseline uses 200 point probes, 2,048 Gram probes, and 1,000 covariance simulations. Each alternative increases only one budget, respectively to 800, 8,192, or 4,000. For Mean, Table S6.8 reports the largest absolute change across seeds and targets when increasing point probes. Each difference is divided by that target’s dense exact-estimator sampling SD, not by its numerical MCSE. The highrank maxima occur for the

Table S6.8: Fixed-data sensitivity to 200 versus 800 point probes

Route and reference	Point change / SD	Endpoint change / SD
Observation, highrank	0.131	0.151
Match, highrank	1.090	1.150
Observation, q1	0.011	0.333
Match, q1	0.018	0.782

covariance target: seed 410081 for matches and seed 410099 for observations. The q1 endpoint maxima occur for firm variance, at seeds 410089 and 410081, respectively. Raising only the Gram budget changes an endpoint by at most 0.166 SD; raising only the covariance budget changes one by at most 0.015 SD. Neither changes the point estimate. These are differences between two finite budgets, not estimates of absolute approximation error. Four outcomes cannot characterize sampling coverage or provide a general numerical-error bound. No budget, seed, reference, or acceptance rule was retuned after seeing these results. The dense moments describe the exact estimator; finite-probe JLA error is a separate contribution and is material in some of these examples. The companion `numerical-summary.json` records the aggregate results.

Numerical identity checks. The signed two-solve identity in section S5.11 accounts for both the influence solve’s residual and the original fit’s residual under Mean and None. Numerical checks include injected inconsistencies and pass their registered criteria. This consistency check does not explain or change the separate sampling assessment’s failures.

Independent derivation check. The companion `verify_moments.py` uses a product three-node Gauss–Hermite rule on seven independent Gaussian units (2,187 integration nodes per design). It imports no production FEVC operators. The rule integrates the relevant polynomials, of degree at most four, exactly apart from floating-point roundoff. Unit-observation and unequal-mass match geometries both pass the 10^{-10} absolute tolerance: the largest error is 1.8×10^{-14} . Checks include the raw-proxy point identity, estimated-mean product, kernel expansion, unchanged population influence, actual-Mean expectation and covariance, expectation of the frozen-mean covariance formula, discrepancy second moment, and q1 score/recenter/remainder cross covariance. It also verifies the signed two-solve certificate with deliberately inexact numerical solutions and rejects an injected recenter error. The receipt is `quadrature.json`. These tiny generic linear-model examples establish neither a many-effects regime nor coverage of the feasible smoothing and JLA procedures.

Outcome-free assessment geometry. The archived check using two simulated outcomes covers all 12 route/reference/size cells of the new registration: equal worker and firm counts $N_W = N_F \in \{12, 20\}$, observation samples of 432 and 1,201 independent rows, and match samples of 144 and 400 independent aggregates. The companion `oracle-summary.json` records every target’s exact Gaussian moments and realized support/spectrum diagnostics. Across highrank primary targets, the largest squared-eigenvalue share is 0.2143; the q1 primary targets have leading shares 0.9140–0.9594 and remainder shares at most 0.1359. Covariance is a predeclared q1 diagnostic because another concentrated mode remains. These finite shares describe the fixtures, not a universal cutoff. The omitted-mean RMS discrepancy divided by actual target SD is at most 0.00131 for highrank primary targets and 0.00103 for q1 primary targets; for the latter its conditional-remainder-scale ratio is at most 0.000493. These are analytic outcome-free values, not coverage estimates.

The identified coefficient dimension is $k = N_W + N_F - 1$, hence 23 or 39, not the size index 12 or 20 in the original harness filenames. Workers w and firms j are indexed from zero in this design. Observation cells repeat each worker–firm pair $2 + \{(13w + 7j + 1) \bmod 3\}$ times. Match cells store two rows per pair, with frequencies $1 + \{(w + 2j) \bmod 3\}$ and $1 + \{(2w + j) \bmod 2\}$. The complete pair target mass is $0.85 + \{(19w + 11j + 3) \bmod 31\}/100$, divided equally among its stored rows; q1 multiplies the mass of pair $(0, 0)$ by 1,000. The physical signal is $3 + \alpha_w + \psi_j$, with

$$\alpha_w = 2 \sin\{0.73(w + 1)\} + w/N_W, \quad \psi_j = -1.4 \cos\{1.07(j + 1)\}.$$

Errors in the independent-unit geometry are Gaussian, with variance $0.7 + 0.3r_i + 0.15r_i^2$, where $r_i = (\text{rank}(P_{ii}) - 1/2)/n_u$ uses average ranks for ties and n_u is the number of independent units. Observation units are physical rows; match units are square-root-mass-scaled aggregates as in section S5.9. Within a match the primary design replicates its one aggregate shock across physical copies. This is a particular within-match covariance design, not an assumption of independent copies.

The DGP variance is a positive quadratic in ranks of *exact* leverage, which lies in the ideal leverage basis. Native fitting uses ranks of JLA-estimated design quantities. The DGP need not lie exactly in that realized finite-probe basis. Feasible inference therefore includes this covariate-approximation effect as well as variance-fit, Gram, and covariance simulation error. A claim of exact finite-probe native specification would be stronger than this design establishes.

S6.9 Observation and mover-only Mean assessment

The registered primary assessment contains 12 cells with 2,000 paired outcome draws per cell: four exact observation-smoother cells and eight structured Rust observation/mover-match cells. The complete union has 24,000 paired outcomes, 48,000 command calls, and 192,000 target rows. All planned command calls and target records are accounted for, including failed calls and unavailable intervals. No fixture, target classification, threshold, seed, or estimator was retuned after assessment outcomes. Highrank includes all four targets. The q1 primary set is worker variance, firm variance, and total variance; covariance is a predeclared multimode diagnostic, retained in the output but outside primary screens.

Exact observation smoother: interval availability. The primary assessment fails 22 registered arm/target availability screens, all on the exact route: both highrank sizes and the smaller q1 cell. These are scientific screen failures, not malformed runs. Across the exact assessment, 9,419

command calls fail atomically and 38,090 target rows lack intervals: 19,644 report negative inference variance, 18,032 report a non-positive-semidefinite inference covariance, and 414 report target status 5. Table S6.9 retains every attempted draw. Mean interval availability is only 11.4% and 18.8% in the highrank cells and 32.4–35.45% for the smaller q1 primary targets. The larger q1 cell has 98.0–98.15% availability and passes the broad descriptive screens. The fixed- c_0 arm has the same broad pattern; this comparison alone does not identify the cause of any failure.

Table S6.9: Exact observation smoother: counts per target out of 2,000 attempts

Reference, N_W	Arm	Posted points	Posted SEs	Intervals
Highrank, 12	Mean	228	228	228
	Fixed c_0	228	228	228
Highrank, 20	Mean	376	376	376
	Fixed c_0	377	377	377
q1, 12	Mean	722	648–709	648–709
	Fixed c_0	722	652–710	652–710
q1, 20	Mean	1,964	1,960–1,963	1,960–1,963
	Fixed c_0	1,964	1,960–1,963	1,960–1,963

Ranges run across the eligible primary targets. Posted SEs are the harness’s finite highrank-SE diagnostic, including on q1 requests; they are not a replacement for q1 joint covariance or its curved interval. Target-level counts, failures, and both coverage denominators are retained in `primary-summary.json`.

Conditional on an interval, Mean coverage ranges from 96.93–97.37% in highrank $N_W = 12$, 94.41–95.48% in highrank $N_W = 20$, 95.83–97.60% in q1 $N_W = 12$, and 95.36–96.23% in q1 $N_W = 20$. Counting unavailable intervals as not covered gives only 11.05–11.10%, 17.75–17.95%, 31.45–34.60%, and 93.45–94.45%, respectively. Conditional coverage cannot rescue the failed availability screens. Reported bias and empirical SD also use finite posted points only: atomic failures suppress all points, so these quantities are conditional on point posting. The ratio of RMS available SE to posted-point SD is not unconditional calibration.

The dense Gaussian covariance of the three primitive targets is positive definite in all four exact cells. In highrank, its raw condition numbers are about 1.59×10^4 and 3.81×10^4 , largely reflecting the small covariance-target scale; the corresponding correlation-matrix condition numbers are only 2.48 and 3.24. The PSD gate operates before adding the redundant total-variance row. Thus neither the total identity nor a nearly singular standardized population contrast explains the observed failures. Distinct primitive/polarized local-linear variance fits and finite simulated trace subtraction are plausible sources of instability, but the recorded failure labels do not establish that mechanism. No repair or change of acceptance rule was made from these outcomes.

Structured observation and mover-match inference. Every target in the eight native primary cells posts a point, finite SE, and interval in all 2,000 attempts in each arm. No registered primary screen fails. Table S6.10 reports the coverage ranges, with conditional and all-attempt denominators identical. Mean highrank coverage ranges from 93.60–95.85%; q1 primary coverage ranges from 95.05–97.15%. The highrank RMS-SE/empirical-SD ratio ranges from 0.989 to 1.040. The q1 covariance diagnostic has Mean coverage of 91.75–94.00%, and remains outside the primary set as registered. These observations support only the bounded descriptive comparison.

Table S6.10: Structured observation and mover-match coverage, percent; 2,000 intervals per target

Deletion	Reference	N_W	Mean	Fixed c_0
Observation	Highrank	12	94.90–95.75	94.90–95.80
Observation	Highrank	20	95.05–95.85	95.05–95.85
Match	Highrank	12	93.60–95.00	93.60–94.95
Match	Highrank	20	94.45–95.35	94.45–95.40
Observation	q1	12	95.05–96.40	95.05–96.40
Observation	q1	20	95.05–97.15	95.05–97.15
Match	q1	12	96.25–97.10	96.25–97.10
Match	q1	20	95.90–97.15	95.90–97.15

The registered descriptive screens require at least 95% availability; absolute bias no larger than 0.10 oracle SD plus three bias MCSEs; conditional coverage plus three nominal-95% binomial MCSEs of at least 90%; a highrank RMS-SE/empirical-SD ratio in $[0.8, 1.2]$; and an all-attempt Mean coverage drop relative to fixed c_0 no larger than 0.02 plus three paired MCSEs. The exact Gaussian omitted-mean RMS/actual-SD screen is 0.10. There is no upper-coverage penalty for q1. These broad descriptive checks are not a new public coverage standard. A coverage estimate near 95% from 2,000 independent draws has binomial MCSE about 0.49 percentage points. Across eligible primary targets the largest absolute paired Mean-versus-fixed- c_0 difference in all-attempt coverage is 0.15 percentage points. Small paired differences and the small analytic omitted-mean terms do not imply that variance fitting, reference approximation, or numerical error is negligible in other designs.

S6.10 Observation and mover-only diagnostic cells

The separate diagnostic profile completes 10 cells with 100 paired draws each: 2,000 calls and 8,000 target rows. All calls and target records are accounted for; 2,264 target rows lack intervals. The raw aggregation label **SCREENS_PASS** means that no primary scientific screens apply to this profile; it is not a scientific pass. Exact highrank availability is 14% under the homoskedastic diagnostic and 28% under weak signal; exact q1 has no available interval under the null. Under concentrated match mass, native highrank has no covariance interval (status 6), while its other three targets are available on every draw. In these four diagnostic cells, all 1,600 paired target attempts agree between Mean and fixed c_0 on availability and return/status/failure codes. This documents a shared limitation on these draws, not a causal diagnosis.

The remaining diagnostics retain misspecified variance, multimode, non-Gaussian-error, leverage-basis, and estimated-offset cases. With only 100 draws per cell, they expose examples rather than establish robust coverage. Their Gaussian dense-moment references have the t_8 and estimated-offset qualifications derived above, and finite-probe JLA variation is a separate effect. q1 covariance comparisons use successful intervals only. The companion summaries retain all attempts, both arms, failures, and the applicable conditioning.

The replication records also retain the two-outcome engineering **SCREEN_FAILURE** and invalidated pilot acceptance labels; these are not coverage evidence. Observation-q1 calibration and individual-inference campaign failures retain their status. This assessment does not replace them with a confirmation claim.

S7 Computational evidence and reproducibility

S7.1 Timing, memory, and numerical settings

The synthetic and private Veneto comparisons use identical retained inputs within each package pair. Setup-inclusive time includes package setup and fresh Matlab pool creation after common import and validation; estimator-only time excludes that setup. Time points are medians of three measured calls, and time-ratio points are medians of three same-host paired ratios. The packages run sequentially, with one warm-up per package and rotated execution order. Hardware is fixed within each pair and thread sweep, but may differ across size cells and sweeps.

The synthetic grids contain 97 distinct configurations and 102 evaluated grid cells. Four configurations are measured separately in the size and thread grids; the near-degree-three 400,000-row, 28-thread measurement appears in both views without double-counting its pairs. The 606 measured calls and 202 warm-ups give 808 reported synthetic calls. All pass the output checks; the maximum complete-system `fevc` residual is 9.9921×10^{-6} , below the 10^{-5} acceptance limit. Matlab’s logged PCG residuals are at most 9.5×10^{-11} . These solver-specific diagnostics are not directly interchangeable accuracy measures. Calls that fail during process launch or pool creation, or exceed scheduler deadlines or external memory limits, are excluded from timing summaries; the replication records enumerate all attempts and exclusions.

The memory figures use peak process-tree proportional set size (PSS) over the setup-inclusive interval: the largest sum of current process PSS across sampling sweeps, rather than a sum of individual process peaks. Linux `smaps_rollup` supplies KiB, divided by 2^{20} for GiB. PSS counts private resident pages fully and apportions shared pages across mappings. It includes loaded data and application overhead. Summed resident set size (RSS) instead counts each resident mapping in full. Tables S7.2 and S7.3 report median peak summed RSS across three calls; these measurements are not converted to PSS.

PSS measurements cover 64 distinct pairs: 54 synthetic match-deletion cells and five thread counts for each Veneto sample. All pass the output and residual checks. One memory pair per cell does not supply repetition uncertainty. Three synthetic 400,000-row, 28-thread measurements appear in both the size and thread panels.

Nominal sampling intervals are 0.05 seconds for synthetic `fevc` at up to 400,000 observations and one second otherwise. Calls have at least 11 complete setup-inclusive sweeps; the largest gap between complete sweeps is 4.00 seconds. Sequential process reads are not atomic and may miss short-lived peaks. Observer work, external page sharing, page residency, and allocator behavior also affect the measured footprint. No application swap was observed. PSS is observed memory, not a forecast, reservation, or exact minimum RAM.

Software versions, source and input hashes, hardware, numerical settings, resource limits, and per-call outcomes are recorded by experiment in the replication materials. The private Veneto records contain only aggregate timing, memory, and numerical diagnostics.

S7.2 Synthetic graph classes

Every synthetic worker has distinct firm matches, with one observation per match, and every firm has degree 100. The degree-five and mixed-degree designs have $N/5$ workers and $N/100$ firms; the degree-four design has $N/4$ workers. Mixed degrees run from two to eight with mean five,

approximately equal worker counts at each degree, and balanced integer rounding. The near-degree-three class has $\lfloor N/3 \rfloor$ workers: $N - 3\lfloor N/3 \rfloor$ workers have degree four and all others degree three. This negligible adjustment affects one or two workers at each size and preserves exact round observation counts and firm degree 100.

The segmented and bottleneck designs split firms into two equal markets. Exactly one percent of workers cross between them; within each worker-degree stratum, paired workers exchange half their matches, rounded down. The crossing share and firm degrees remain fixed as the sample grows. The well-mixed designs randomly pair worker and firm slots while preserving degrees and excluding duplicate worker–firm matches.

S7.3 Agreement between randomized estimates

Table S7.1 summarizes differences between corrected point estimates in all 303 measured pairs of the six-class benchmark. Each comparison uses the same target and identical input data. The packages use different random projections and finite-projection adjustments, so these differences do not measure error relative to exact KSS or establish equal approximation accuracy.

Table S7.1: Numerical differences in the six-class benchmark

Corrected target	Median absolute gap	Maximum absolute gap
Worker variance	0.000063	0.000501
Firm variance	0.000013	0.000123
Covariance	0.000010	0.000083
Total variance	0.000076	0.000448

Notes: Absolute differences between paired `fevc` and KSS Matlab corrected estimates after multiplying Matlab’s raw targets by $(N - 1)/N$. All 303 measured pairs are included; both packages use 200 probes. Differences are descriptive, not error relative to exact KSS.

S7.4 Source-only Mata computation

A synthetic computation experiment uses 100, 1,000, 5,000, and 10,000 workers, each with six stored rows in three matches. It includes integer frequency weights and separate target weights, with no controls and movers only. Both generic and compressed Mata JLA calculations use 200 probes, batch size 16, diagonal preconditioning, tolerance 10^{-10} , and at most 20,000 iterations. On an Apple M2 Ultra with 192 GiB of memory and Stata 18 using four processors, each configuration has five measured calls after one same-process warmup. Time runs from command invocation to return. At 60,000 stored rows, the median times are 21.531 seconds for the generic calculation and 17.808 seconds for the compressed calculation. All measured calls complete, and the four corrected components agree exactly across the two representations within each sample size and repetition. These results describe source-only computation on this repeated-match design; the do-file and replication records supply its full construction and settings.

S7.5 Supporting Veneto results

The private benchmarks use two pooled samples with movers and stayers. Both use uniform physical-row targets and 200 probes. Table S7.2 reports both deletion modes for the 1999/2001

sample; table S7.3 reports match deletion for the 1996–2001 sample. Sample preparation, deletion methods, and execution hosts differ, so their comparison does not isolate the effect of panel length. The comparator is the KSS Matlab package, distinct from the Econometrica replication code.

S7.5.1 1999/2001 sample

The two-period sample contains 1,319,994 adjusted observations, including stayers: 22 more rows than the published pooled sample of 1,319,972 (Kline et al., 2020, table I). It is not an exact reconstruction of that sample. Observation deletion follows the main two-period KSS application; the KSS longer-panel appendix also compares match and worker deletion (Kline et al., 2020, section 9 and appendix A).

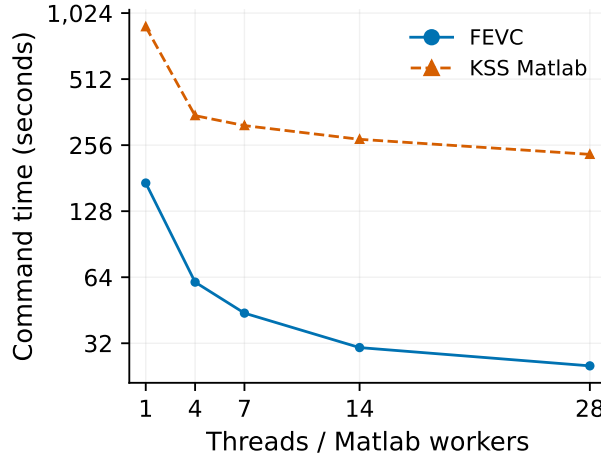
Both modes use the adjusted outcome without additional controls. Observation deletion removes individual rows; match deletion uses the match/stayer-hybrid correction. Each mode runs on one host at one, four, seven, 14, and 28 threads, with one warm-up and three measured calls per package and thread count. All 80 reported calls pass the output checks. The maximum absolute corrected-target gap across 30 measured pairs is 0.000247 after normalization; it does not establish equal finite-projection precision. Figure S7.1 shows the observation-deletion results separately from the combined display in the accompanying article.

Table S7.2: 1999/2001 Veneto timing and summed RSS by thread count

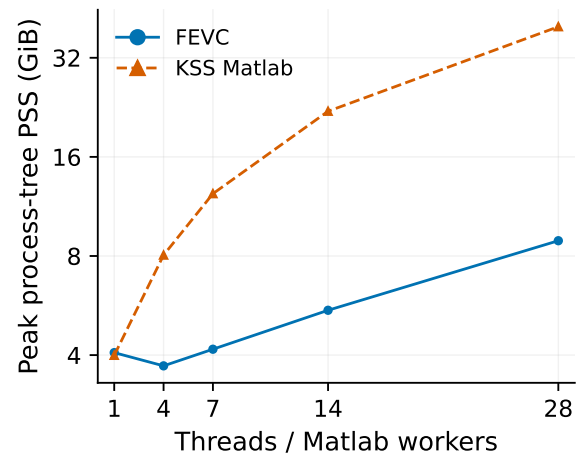
Threads	FEVC (s)	Matlab (s)	Speedup	Est. ratio	FEVC GiB	Matlab GiB
<i>Observation deletion</i>						
1	172.0	887.2	5.16	0.209	4.07	4.47
4	60.8	349.0	5.68	0.217	3.61	9.98
7	43.9	314.4	7.16	0.216	3.75	15.64
14	30.6	272.2	8.92	0.199	5.29	28.35
28	25.2	232.4	9.11	0.178	8.60	52.53
<i>Match deletion</i>						
1	181.5	888.9	4.89	0.211	4.27	4.28
4	69.6	312.6	4.49	0.224	4.18	9.70
7	51.1	214.4	4.13	0.216	3.84	15.18
14	36.8	158.9	4.33	0.195	5.63	27.13
28	31.3	197.7	6.30	0.176	8.80	51.28

Notes: 1,319,994 adjusted observations, including stayers; 200 probes and three measured pairs per row. Times are separate package medians, including setup and fresh Matlab pool creation. Speedup is the reciprocal of the median paired **fevc**-to-Matlab time ratio. The estimator-only ratio excludes setup; values below one favor **fevc**. Memory is median peak setup-inclusive summed process-tree RSS, not the PSS in the figures. Each deletion mode’s sweep uses one host; the modes use different hosts. Stata uses $\min(T, 4)$ processors.

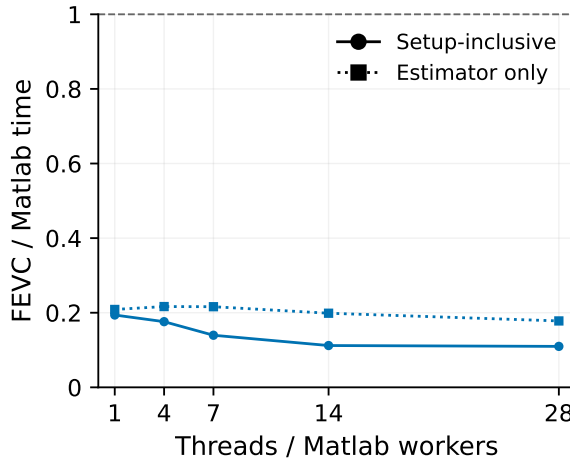
Figure S7.1: Observation deletion on the 1999/2001 Veneto sample



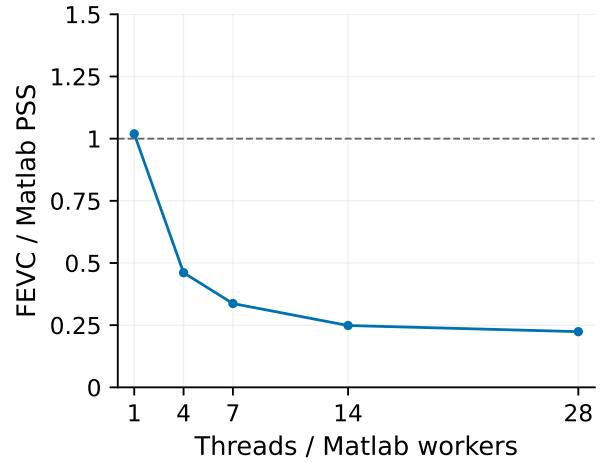
(a) Setup-inclusive time



(b) Peak process-tree PSS



(c) Paired time ratios



(d) Paired PSS ratios

Notes: 1,319,994 adjusted observations, movers and stayers, and 200 probes; observation deletion throughout. Time panels show three-run medians and paired ratios. Command time includes setup and fresh pool creation; dotted timing ratios exclude setup. PSS covers the same interval and apportions shared pages, but can miss brief peaks. Upper vertical axes use powers-of-two log ticks; lower axes are linear with parity at one. Stata uses $\min(T, 4)$ processors, native `fevc` T threads and Matlab T single-threaded workers. Each package pair uses the same host and input.

S7.5.2 1996–2001 sample

The six-year sample contains 5,163,428 adjusted observations, including stayers, and uses the match/stayer-hybrid correction. Its construction and differences from the specification in KSS table A.I are described in the accompanying article. The sweep over five thread counts has 40 reported calls, including warm-ups, and 15 measured package pairs; all calls pass the output checks. The maximum absolute corrected-target gap is 0.005622 after normalization. As in the synthetic comparison, this describes agreement at the specified settings rather than error relative to exact KSS.

Table S7.3: 1996–2001 Veneto panel: timing and summed RSS

Threads	FEVC (s)	Matlab (s)	Speedup	Est. ratio	FEVC GiB	Matlab GiB
1	551.9	1585.5	2.88	0.318	7.63	5.61
4	217.9	549.9	2.53	0.324	7.68	11.95
7	163.8	396.8	2.39	0.324	7.81	18.59
14	121.5	321.3	2.65	0.289	11.03	33.10
28	104.5	278.8	2.67	0.264	16.48	61.34

Notes: 5,163,428 adjusted observations, including stayers; 200 probes and three measured pairs per row, all on one host. Times are separate package medians; speedups and estimator ratios use paired times. The match/stayer-hybrid specification differs from an exact Table A.I replication. Memory is sampled peak setup-inclusive summed RSS, not the PSS in the figures; shared pages can be counted repeatedly and brief peaks missed. Stata uses $\min(T, 4)$ processors, native `fevc` T threads and Matlab T single-threaded workers.

Pooled component calculation. Structured match inference operates on the number of retained mover units plus physical stayer observations. Rust applies its covariance and spectral operators through sparse solves; exact Mata contracts the variance-model Gram in coefficient space. Both require additional covariance simulations beyond point estimation. Their numerical budgets are distinct from the point-probe budget. The public pooled illustration in Section S5.14 records complete command settings and outcomes; it is not a timing or memory benchmark.

S7.6 Local match-inference size ladder

The direct residual-Gram timing experiment uses fixed datasets with 7,680, 30,720, and 122,880 matches, with two physical observations in each. No outcomes are regenerated. Workers have three distinct firm matches; the long-range degree-three topology has 40 worker layers per firm. Independent normal worker and firm effects have standard deviations 1 and 0.6. Two independent normal controls enter with coefficients 0.5 and -0.3 . Errors combine an independent match shock of variance one and row noise of variance 0.25, giving within-match correlation 0.8. Regression and target masses are uniform. The controls are estimated, then held fixed: this is a computational test of the approximation, not a test that omitted nuisance uncertainty is negligible.

Point-only, high-rank, and q1 calls use Rust, Counter-V1, generic JLA, CMG, 200 probes, batch 16, seed 8675309, solver tolerance 10^{-10} , a maximum of 10,000 iterations, and a 16 GiB direct-allocation budget, enforced by the strict-budget policy (`memorycheck(error)`). Inference uses `structured_common` with the common residual-moment fitter, 2,048 direct residual Gram probes, 1,000 covariance probes, and 128 spectrum probes. The spectral iteration budgets are 512 for q0 and 128 for q1. A q1 target uses 100,000 critical-value draws; the inference seed is 8675309. Dataset

seed 20260905 is addressed by size. Each fixed dataset is reused for three timing repetitions, with call order rotated across sizes and repetitions. Stata/MP 19 runs in fresh processes with eight active processors on macOS ARM64 and an Apple M2 Ultra.

The 27-call experiment completes in 38.23 minutes. All calls finish within the 600-second process limit; none reaches the 24 GiB measured process-tree stop. Of 72 target-interval attempts, 69 return intervals. The three withheld intervals are the q1 covariance target at the smallest size. These are timing repetitions on fixed outcomes, not coverage replications.

A separate 36-call experiment uses the cross-fitted variance estimator. With that fitter and 128 spectral iterations for q0, three smallest-size q0 calls failed to recover the spectral modes to the required tolerance, and six inference calls at 983,040 observations timed out. The direct residual-Gram experiment does not include the million-observation size. Both experiments retain their failures and preselected settings.

Table S7.4: Runtime and interval availability in the local match-inference experiment

Rows	Matches	Call	Complete	Seconds	GiB	Timeouts	Intervals
15,360	7,680	Point	3/3	0.62	0.21	0	—
15,360	7,680	q=0	3/3	6.43	0.22	0	12/12
15,360	7,680	q=1	3/3	4.27	0.22	0	9/12
61,440	30,720	Point	3/3	3.76	0.32	0	—
61,440	30,720	q=0	3/3	67.36	0.34	0	12/12
61,440	30,720	q=1	3/3	37.62	0.35	0	12/12
245,760	122,880	Point	3/3	20.54	0.95	0	—
245,760	122,880	q=0	3/3	397.33	1.05	0	12/12
245,760	122,880	q=1	3/3	216.33	1.06	0	12/12

Notes: Complete counts successful command returns; interval availability counts returned intervals out of four targets per planned call. Time and GiB entries are medians among completed calls only. Missing timing entries are not zero. Command time excludes process startup and import; the ten-minute timeout applies to the whole process. RSS is absolute process-tree memory sampled every 0.1 seconds during the command, without baseline subtraction. It differs from the direct allocation budget. Failed and timed-out calls remain in the counts.

Returned calls retain the same observations and point estimates across inference settings, satisfy the adding-up identities, and meet the solver residual tolerances. Per-call results include covariance, spectral, and variance-model diagnostics and account for every attempted interval. Because outcomes are fixed, interval availability in this experiment describes computation rather than sampling coverage.

S7.7 Reproducing documents and locating evidence

The replication repository separates document builds, reconstruction from saved receipts and estimator execution. From its root, `make paper` compiles the supplement first and then the article; `make supplement` builds the standalone supplement; `make combined` concatenates their PDFs without renumbering. `make check` verifies active evidence bindings, displayed arithmetic and document/tooling contracts using existing data. It does not rerun estimators. LaTeX/latexmk and Poppler are required for documents; Python packages are pinned in `requirements/dev.txt`. The combined-copy helper also requires `pypdf`; set `FEVC_PDF_PYTHON` to a Python installation providing it. Delivered files are under `output/pdf/`.

`replication/publication_manifest.json` lists active source, exhibit, generator and evidence

hashes. `SOURCE_PROVENANCE.md` maps experiment families to their source identities; raw receipts are under `replication/results/raw/`, and protocols under `replication/experiments/` and `replication/benchmarks/`. The October 9 public Mean example is `public_mean_20261009` at source `13a98492`; September 30 numerical-MCSE evidence uses `f9d3f2e2`; the performance archive uses snapshot `b14496bc`. The Mean-inference assessment is bound to `b9f80ce9`, with its source-specific compatibility record. None of these identities is replaced by the documentation pin.

Reconstruct a historical result only with its recorded source, inputs, retained sample, seed, numerical budgets and environment. Each protocol identifies any required public download or licensed runtime. Restricted Veneto microdata and licensed third-party code are not distributed; their authorized evidence consists of aggregate receipts. The preserved comparator manifest and archive index give the exact source boundaries. Missing external dependencies remain explicit, rather than being replaced by a later source version.

S8 Implementation comparisons

The KSS Matlab package, `VarianceComponentsHDFE.jl`, `LeaveOutKSS`, and `PyTwoWay` provide related estimators (Saggio, 2024; High Dimensional Economics Lab, 2024; Moghani, 2026; Lamadon and Oppenheimer, 2022). Table S8.1 summarizes their main interfaces. Comparisons require matching retained observations, target weights, deletion blocks, the treatment of controls, and outcome centering. Identical variable names or raw inputs alone do not ensure that two outputs estimate the same quantity.

Table S8.1: Main-interface choices in KSS implementations

Interface	Deletion	Controls	Weights/targets	Additional inference
<code>fevc</code>	Observation or declared match	Joint or fixed offset	Frequency and separate target weights	Worker/firm projections; experimental components
KSS Matlab	Observation or pair-defined match	Fixed offset	Observation mass	Firm projections
Julia	Observation or pair-defined match	Fixed offset	No separate frequency/target arguments	Firm linear combinations
R	Observation or pair-defined match	Fixed offset	No separate frequency/target arguments	Firm projections
Python	Observation HE path with package cleaning	In model design	Observation weights; programmable targets	Selected coefficients

`fevc` permits declared deletion blocks nested within the estimation pair and separates integer regression frequencies from target mass. Its default joint-control calculation re-estimates the control coefficients after deletion, following the general KSS linear model. The principal Matlab, Julia, and R workflows instead partial controls out once. The additional KSS Matlab routine `leave_out_COMPLETE.m` supports joint controls for observation deletion and uses fixed offsets for match deletion.

The main Matlab routine centers the working outcome after transforming match averages by the square roots of match sizes. `fevc` instead uses the mean over retained physical observations. Thus

the default `centering(mean)` does not imply identical centering for unequal match sizes.

The KSS Matlab main interface provides firm-effect projection inference but does not expose variance-component inference; additional routines remain in the package. `fevc`'s component procedures are experimental and use the additional assumptions in section S5. The Econometrica replication code is a separate KSS implementation and is identified as such in the replication documentation.

The comparison describes the archived source snapshots in the publication manifest and comparator audit, not an assertion about current upstream heads. The KSS Econometrica archive, maintained Matlab main routine and additional Matlab routines are separate comparison objects. Differences in numerical adjustments must be distinguished from differences in statistical targets.

Pinned semantic comparison. The comparator inventory records KSS Matlab `LeaveOutTwoWay` commit `8b957ffe` (October 17, 2024); Julia `VarianceComponentsHDFE.jl` version 0.1.0 at `fa3d66ea`; R `LeaveOutKSS` version 0.1.0, archive SHA-256 beginning `f70360b6`; and Python `PyTwoWay` version 0.3.21 at `7208bda3`. Full hashes are in `replication/manifests/comparators.json`. Julia, R and Python entries are source/interface audits, not executed performance comparisons. The current FEVC row is updated to the documentation pin while historical experiment rows retain their original versions. A short commit prefix identifies the full hash in that inventory; it does not establish behavior of a later release.

References

- Andrews, I. and A. Mikusheva (2016). A geometric approach to nonlinear econometric models. *Econometrica* 84(3), 1249–1264.
- High Dimensional Economics Lab (2024). VarianceComponentsHDFE.jl. Software repository.
- Kline, P., R. Saggio, and M. Sølvesten (2020). Leave-out estimation of variance components. *Econometrica* 88(5), 1859–1898.
- Kline, P., R. Saggio, and M. Sølvesten (2021). Improved stochastic approximation of regression leverages for bias correction of variance components. Technical note, January 21; revised October 2026.
- Lamadon, T. and A. A. Oppenheimer (2022). PyTwoWay: Two-way models in Python. Software repository.
- Moghani, V. (2026). *LeaveOutKSS: Leave-Out Variance Component Estimation for Two-Way Fixed Effects Models*. Comprehensive R Archive Network.
- Saggio, R. (2024). LeaveOutTwoWay: Bias-corrected estimates of variance components in two fixed-effects models. Software repository.
- Schmieder, J. (2026). FEVC: Leave-out bias-corrected variance decompositions in Stata. Working paper, Boston University. October. Article accompanying this technical supplement.