

Day 10 Presenter Notes

Shahryar Minhas

Purpose of the Day

Day 9 separated broad actor tendencies from discrete relational roles. Day 10 asks what we gain when recurring relationship patterns are represented with continuous geometry. The class begins with a latent distance model because the map is intuitive, puts that map under pressure with a complementary-role network, derives the factor model from an ordinary regression interaction, makes the factor idea concrete through a movie preference matrix and its singular value decomposition, and then applies a dynamic rank-two AME to changing ICEWS state-pair relationships.

The central distinction is between broad partner counts and exact ordered pairs. An additive effect can tell us that a state appears in high-volume relationships with many partners. A multiplicative term can tell us that one particular ordered pair is much more or less likely than those broad patterns and the measured characteristics would suggest.

The final application uses 13 annual directed networks with 462 observed relationship changes. Rank two improves complete-pair prediction and reproduces much

more year-to-year movement than the dynamic SRM. It also misses reciprocity in several years and four of twelve transition intervals. The class therefore gets a model that is substantively useful without pretending that a good application must fit every feature perfectly.

Files and Setup

Open the Day 10 deck and the rendered walkthrough before class. The deck supplies the route; the walkthrough supplies the explanations, code, output, and caveats. Each section below names the walkthrough anchor and the chunks used there.

Run the **setup** chunk after restarting R. The downloaded folder contains the data and saved fits used by the walkthrough. Expensive calls are wrapped in **cache_fit()**, so unchanged model code loads the saved result instead of re-running estimation.

Do not delete a cache during class unless a refit is intentional. The Gade latent-distance fit and the dynamic AME with 100 bootstrap refits are not useful places to make the room wait.

Keep this sentence in view: distances, fitted probabilities, and the fitted multiplicative surface are meaningful model quantities; the displayed orientation of a map or factor axis is not.

Opening: Where Geometry Enters

Walkthrough anchor: **#sec-where**

Suggested opening: “Day 9 showed that some organizations

work with many partners and that cooperation can center on a small tactical core. Today we ask what those summaries still miss. When a state appears in many high-volume relationships, do all of its relationships move together, or can one directed state pair have its own history?”

Put the general linear predictor on the screen:

$$\text{eta_ij} = \text{x_ij}' \text{beta} + \text{a_i} + \text{b_j} + \text{latent relationship_ij}$$

Walk through the jobs in plain language. The measured variables describe things we observed about a pair. The additive actor terms represent actors that take part in many relationships across many partners. The final term asks whether this particular pair fits together more or less strongly than those broad tendencies predict.

Ask: “If I know that two states appear in many high-volume relationships, have I learned whether they cross the threshold with each other?”

Expected answer: No. Broad state involvement is not the same as the fit of one ordered pair.

Likely question: “Did we already fit AME on Day 9?”

Answer: “No. Day 9 deliberately stopped with the random-effects SRM and blockmodels. Day 10 is where we build the multiplicative part and connect it to AME.”

Transition: “Before we choose a geometry, let’s be clear about what an unmeasured relationship pattern could represent.”

What the Missing Pattern Could Represent

Walkthrough anchor: #sec-political-question

Use the two applications to keep the discussion concrete. For the Syrian organizations, ask whether tactical cooperation was scattered or concentrated among organizations with overlapping partner lists. For ICEWS, ask whether states became more involved across many high-volume relationships or whether particular directed relationships changed in their own ways.

Keep four possible uses separate:

1. Proximity: actors may occupy similar unmeasured positions, making a relationship more likely.
2. Complementary roles: two actors may recur together because their relationship profiles fit one another.
3. Dependence adjustment: the latent term may represent connected residual structure while a measured association is the focus.
4. Measurement: the fitted relational surface may itself be the quantity we want to describe.

Suggested language: “A map that predicts ties well is not automatically a map of ideology. A factor that separates countries is not automatically a discovered security dimension. The model estimates a recurring relationship pattern. Naming the process behind that pattern takes outside evidence.”

Ask: “What information would we want before calling a fitted pattern ideology, threat, or institutional alignment?”

Listen for historical records, treaty institutions, text, geog-

raphy, expert coding, or validation on data that did not generate the same factor fit.

Transition: “Those uses lead to different model choices, so let’s put the main options next to one another.”

The Applied Model Compass

Walkthrough anchor: `#sec-model-compass`

For the additive SRM, say: “This model asks whether some actors appear across many relationships, regardless of partner. It cannot give one particular bilateral relationship its own history.”

For the latent distance model, say: “This model asks whether cooperation is concentrated among organizations with overlapping partner lists. Organizations with similar overall cooperation patterns are placed near one another.”

For the factor or eigenmodel, say: “This model asks whether similar roles and complementary roles both matter. It can represent actors that connect to the same partners even when they do not connect to one another.”

For AME, say: “This model separates broad state involvement from the history of a particular directed pair, while also estimating the measured associations and connection between the two directions.”

Likely question: “Is the most flexible model automatically the best one?”

Answer: “No. More flexibility can improve training fit while making the result unstable or worse on held-out data. We choose complexity using the relationship question, compu-

tation, fit checks, and a clearly defined validation target.”

Transition: “Let’s build the distance model before discussing its estimator.”

How the Distance Model Is Estimated

Walkthrough anchor: `#sec-estimation-roadmap`

Use this section only after the distance claim, equation, and likelihood have been explained.

Suggested language: “We have already said what the model assumes: nearby pairs receive higher tie probabilities. Estimation now runs that story backward. The network goes in, and the model looks for coefficients and positions that make all of the ties and non-ties plausible together.”

Walk through the sequence: choose the tie model and dimension, initialize coefficients and positions, update coefficients conditional on positions, update positions conditional on coefficients, align equivalent orientations, and inspect sampling plus reproduced network features.

Emphasize that dimension is a restriction chosen before estimation. A two-dimensional picture is readable, but readability does not make two the true dimension.

State the target before the update sequence. The likelihood rewards observed ties with high fitted probabilities and observed non-ties with low fitted probabilities. The posterior combines that likelihood with priors on the coefficients, positions, and position variance.

Make the distinction explicit: the posterior says what is being learned, while MCMC says how the computer explores

it. The factor and AME model get their own estimation explanation after their substantive and regression intuition.

Transition: “MCMC is one way to carry uncertainty through that joint estimation. Let’s make its mechanics concrete.”

How MCMC Represents Uncertainty in the Distance Model

Walkthrough anchor: `#sec-latent-mcmc`

Use the seating-chart analogy. We observe who interacts but not the arrangement or coefficient values. The sampler begins with a provisional account of both, then updates one part while holding the other at its current value. Each complete state is weighted by both its fit to the network and the prior structure.

Suggested language: “One completed round gives us one plausible state of the model. A sequence of those states is a chain. The chain is not collecting new networks. It is revisiting plausible explanations of the same observed network.”

Explain burn-in as the early part discarded because it may still reflect the starting state. Explain thinning as a storage decision, not a cure for poor exploration. Explain effective sample size as how much independent Monte Carlo information remains after accounting for correlation between saved draws. Explain R-hat as a check on whether separately started chains reached the same posterior region.

Make the latent-model complication explicit. A distance map can rotate or reflect without changing any fitted distance. A factorization can also change orientation while

preserving its fitted pair surface. Raw coordinate traces can therefore be misleading; diagnostics should include invariant quantities.

Likely question: “If the effective sample size is low, should we thin more?”

Answer: “No. Thinning can save storage, but it cannot create information. We need longer exploration, better initialization, a more stable parameterization, or a closer look at weak identification.”

Likely question: “Does a longer chain improve a causal claim?”

Answer: “No. It can reduce Monte Carlo error. It cannot repair measurement, create comparison cases, or separate a measured predictor from a correlated omitted process.”

Transition: “Now we can see the first geometry in its simplest form: nearby positions raise the chance of a tie.”

The Distance Model

Walkthrough anchor: **#sec-idea**

Chunks: **sim-space**, **sim-density**

Put the model on the screen:

$\text{logit } \Pr(Y_{ij} = 1) = \alpha + x_{ij}' \beta - \|u_i - u_j\|$

Read it one term at a time. The intercept sets a baseline. Measured pair characteristics move the score. The distance between the latent positions is subtracted, so moving farther apart lowers the fitted tie probability.

Before running **sim-space**, ask students where they expect

the ties to concentrate. Then run the chunk and connect one visibly close pair and one visibly distant pair to the probability rule.

Run **sim-density**. In this seeded simulation, pairs below the median latent distance tie at about 0.52, while pairs above it tie at about 0.09.

Suggested language: “The tie is still a random draw. Some close pairs do not connect and some distant pairs do. Distance changes the probability; it does not make the network deterministic.”

Explain the triangle inequality without turning it into a closure coefficient. If actor *i* is very close to both *j* and *k*, *j* and *k* cannot be arbitrarily far apart. Because tie probability falls with distance, the geometry tends to produce clustered and transitive patterns.

Likely question: “Are the coordinates observed traits?”

Answer: “No. We created them in this simulation so we could watch the model run forward. In an application, the model estimates positions from each actor’s complete relationship pattern.”

Likely question: “Does one triangle show that closure occurred?”

Answer: “No. The geometry makes clustered configurations more likely, but one observed triangle does not identify the process that created it.”

Transition: “The map is not a decoration added after the regression. It is part of the probability calculation for every pair.”

Reading the Distance Linear Predictor

Walkthrough location: `#sec-idea`, under “The Model, Written Down”

Suggested language: “For any pair, begin with the baseline, add measured explanations, add actor activity terms if the specification contains them, subtract the fitted distance, and transform the result into a probability.”

Be precise about broad activity. The basic undirected distance model can generate some degree variation because a centrally placed actor may sit close to many others. It does not give every actor a separate free activity parameter. `rsociality()` or AME additive effects are needed when broad actor participation remains important.

Likely question: “Does being in the center of the map mean an actor is powerful?”

Answer: “No. It means the fitted geometry places that actor near many others under this model. Power needs its own definition and outside evidence.”

Transition: “Let’s fit this model to a cooperation network where the tie has a concrete meaning.”

Syrian Armed-Organization Cooperation

Walkthrough anchor: `#sec-cameo`

Chunks: `load-data`, `fit-2d`, `mcmc-2d`, `plot-2d`, `summary-2d`, `cameo-transitivity`

Set up the case before running code. The data come from Gade et al. (2019) and cover July 2012 through June

2015. The 31 nodes are Syrian armed organizations. An undirected tie means that a pair took part in at least one recorded joint operation during the full period.

State the measurement limits immediately. This is a binary whole-period teaching reanalysis, not a replication of the article's square-root count model. A tie records tactical cooperation, not ideological agreement, friendship, or a permanent coalition. When an operation involved more than two organizations, the projection records every pair, so one operation can mechanically create several ties and a triangle.

State the question before the fit: "Was tactical cooperation scattered across unrelated pairs, or concentrated among organizations with overlapping sets of partners?" Also state the limited purpose: this map is here to make distance-model reasoning visible. It should not be sold as a new explanation of the Syrian conflict.

Run `load-data`. The network has 31 organizations, 86 cooperative pairs, density 0.185, and weak transitivity 0.357.

Suggested language: "The weak transitivity value is a descriptive feature of this binary network. It is not evidence that one organization's partnership caused another partnership."

Run or load `fit-2d`. Point out that `d = 2` makes a readable map but imposes a two-dimensional Euclidean representation. The intercept and positions are learned jointly. There are no measured covariates and no separate organization-activity effect in this deliberately simple fit.

Display `mcmc-2d`. Read the traces as computational checks.

The plot focuses on the intercept and latent-position variance rather than raw coordinates because an equivalent map can rotate or reflect while preserving every fitted distance.

Display **plot-2d**. Use a disciplined reading routine:

1. Point out that Al-Nusrah Front and Ahrar al-Sham Islamic Movement sit in the same tightly connected part of the map. They recorded an operation with one another, and each also recorded operations with 18 of the same other organizations.
2. Find one distant pair and describe it as receiving a lower fitted probability.
3. Refuse to name the horizontal or vertical axis.
4. Check outside histories before assigning a substantive label to proximity.

Display **summary-2d**. The posterior mean intercept is about 1.00. Explain that this is the fitted log-odds for a pair at zero latent distance, not the overall cooperation probability and not the effect of moving one unit along a displayed axis.

Run **cameo-transitivity**. An independent-tie baseline at the observed density is about 0.185. Simulated networks from the distance fit average 0.287 on weak transitivity, compared with 0.357 in the observed network; the simulation interval is about 0.174 to 0.390.

Suggested language: “The fitted geometry moves the replicated networks toward the clustering we observe, and the observed value lies inside the simulated range. That means the model can reproduce much of this summary. It does not show that closure caused the joint operations.”

Then say: “This mostly reinforces the tactical-core result

from Day 9. It does not give us a substantively new explanation, so we use it to understand the distance model and move on.”

Likely question: “If two organizations appear close, were they allies?”

Answer: “Not necessarily. They receive a higher fitted probability of at least one recorded joint operation in this projection. Tactical cooperation can coexist with rivalry, ideological disagreement, or changes over time.”

Likely question: “Why not add covariates right away?”

Answer: “We could in a fuller application. We keep this first fit simple so the assumptions and limitations of distance geometry remain visible.”

Transition: “This map is useful when proximity is a plausible relationship story. Now we will give it a pattern that proximity handles poorly.”

Where the Distance Map Struggles

Walkthrough anchor: `#sec-foil`

Chunks: `sim-disassort`, `disassort-density`, `fit-ldm-dis`, `fit-ldm-dis-long`, `kmeans-ldm`, `hrt-check`

Set up the simulation as complementary roles. Ties are uncommon within each role type and common across the two types. The pattern could resemble buyers and sellers, patrons and clients, or two specialized roles that need one another.

Run `sim-disassort` and ask: “How can one Euclidean map

keep every preferred cross-role partner close while keeping same-role pairs far apart?”

Let students identify the contradiction. Bringing the groups together helps the cross-role ties but also raises probabilities within each group. Pulling same-role actors apart makes it hard to keep all their cross-role partners nearby.

Run **disassort-density**. The seeded network has within-role density about 0.079 and across-role density about 0.868.

Run or load **fit-ldm-dis**. The two-dimensional distance model predicts about 0.461 within roles and 0.470 across roles. It matches the overall density, about 0.465, while nearly erasing the defining contrast.

Suggested language: “This is a useful warning about averages. A model can match the network’s average tie rate and still miss the relationship pattern we care about.”

The **fit-ldm-dis-long** branch shows that extending this particular run does not recover the contrast. Do not claim that this proves every distance model at every dimension must fail. The conclusion is narrower: this small Euclidean representation is inefficient for a sharply disassortative role pattern.

The **kmeans-ldm** result recovers the simulated roles at 55%, close to chance for this balanced example. The model-based clustering extension in **hrt-check** also does not rescue this specification. Treat both as illustrations, not universal tests.

Likely question: “Could more dimensions or another geometry fit this pattern?”

Answer: “Possibly. More dimensions or a different manifold can represent more patterns, but that adds complexity and

changes the assumptions. Our result is about the familiar two-dimensional Euclidean map.”

Transition: “A distance map is still a map built under choices about dimension and geometry. We will keep that caution visible, then replace closeness with compatibility.”

Optional Walkthrough Box: Identification and Procrustes Alignment

Walkthrough anchor: `#sec-procrustes`

Chunks: `rotate-demo`, `distances-identical`, `procrustes-fn`, `procrustes-recover`

This material is no longer a deck stop. It appears in one collapsed optional-details box in the walkthrough. Leave it closed during the main presentation unless a student asks why two latent maps can face different directions or how independently estimated configurations are compared.

State the identification result directly. The distance likelihood depends on pairwise distances. Translating, rotating, or reflecting the whole configuration leaves every distance and fitted probability unchanged.

Run `rotate-demo`. Ask: “Which map is correct?” The answer is that both are equivalent displays of the same fitted relationship structure.

Run `distances-identical`. The largest difference between the two complete distance matrices is about $8.9\text{e-}16$, which is numerical zero.

Give the permission list. Students may discuss close and distant pairs, compare distances, describe stable clusters cau-

tiously, and calculate fitted probabilities. They may not call the horizontal axis ideology, interpret left versus right, or assign meaning to north and south from the map alone.

Explain Procrustes in one sentence: “It translates, rotates, and, when needed, reflects one configuration so that it lines up with a reference as closely as possible.”

Run `procrustes-fn` and `procrustes-recover` if the room benefits from the mechanics. In this controlled example, mean squared disagreement is about 1.015 before alignment and about $5e-32$ afterward.

Emphasize that this version does not rescale the map. Stretching it would change pairwise distances and therefore change the fitted model.

Likely question: “Does a residual after aligning two independently estimated maps prove that one chain failed?”

Answer: “No. It can reflect posterior uncertainty, Monte Carlo error, weak identification, local modes, or genuine differences in the fitted configurations. We read it with invariant quantities and standard diagnostics.”

Likely question: “Can I align annual maps and interpret movement?”

Answer: “Only if the model actually estimates time-varying positions, aligns them, and carries uncertainty through the comparison. The ICEWS application later does estimate changing profiles, so we will inspect their inner products and fitted probabilities rather than naming the axes.”

Optional point: The optimal Procrustes rotation itself uses an SVD of the cross-product between centered configurations. This gives a natural preview of the next section.

If you open the box, close with: “The displayed orientation is arbitrary, so we interpret distances and fitted probabilities. Now we return to the main argument and replace closeness with compatibility.”

Two Kinds of Similarity

Walkthrough anchor: **#sec-factor**

Chunk: **two-similarities**

Begin by distinguishing homophily from stochastic equivalence.

Homophily means actors with similar attributes or positions are more likely to connect to one another. Distance geometry represents this naturally.

Stochastic equivalence means actors play similar roles because they connect to the same partners. Two actors can be stochastically equivalent even if they never connect directly.

Run **two-similarities**. In the first panel, similar actors connect within a group. In the second, A and B occupy the same role because both connect to actors 1, 2, and 3, even though A and B do not connect to one another.

Suggested language: “Homophily says similar actors connect to each other. Stochastic equivalence says similar actors connect to the same others.”

Connect back to Day 9. A blockmodel represents stochastic equivalence with discrete roles. A factor model represents it with continuous profiles.

Likely question: “Does stochastic equivalence mean the ac-

tors are substantively similar?”

Answer: “It means their relationship profiles are similar under the model. Whether they share an ideology, institution, strategy, or something else requires outside evidence.”

Transition: “Before we use a matrix decomposition, let’s build the factor model from the regression language we already know.”

Building the Factor Model From a Regression

Walkthrough anchor: #sec-regression-bridge

Deck slides: Start With the Regression We Already Know; The SRM Adds Broad Source and Target Differences; A Factor Is an Interaction With Unknown Inputs; One Source Can Fit Two Targets Differently; Rank Is the Number of Latent Interactions

Begin with the ordinary dyadic regression:

$$g(E[Y_{ij}]) = \eta_{ij} = \beta_0 + x_{ij}' \beta$$

Suggested language: “Nothing about the outcome or link function changes when we move to a factor model. We still have a relationship outcome, a linear predictor, an intercept, and measured covariates.”

Add the source and target terms from Day 9:

$$\eta_{ij} = \beta_0 + x_{ij}' \beta + a_i + b_j$$

Say exactly what the additive terms can and cannot do. The source effect moves every relationship beginning with actor i . The target effect moves every relationship aimed at actor

j. They can explain broadly high rows and columns, but not why one particular source-target combination is high while another combination involving the same source is low.

Now introduce a measured interaction. Let c_i be a measured source characteristic such as military capability and r_j a measured target characteristic such as vulnerability. State that their separate main effects are already included in x_{ij} :

$$\eta_{ij} = \beta_0 + x_{ij}'\beta + a_i + b_j + \delta c_i r_j$$

Suggested language: “This is ordinary regression. We multiply two observed characteristics, put the product in the design matrix, and estimate its coefficient. The same source characteristic can matter differently across targets because it is multiplied by the target characteristic.”

Then replace the observed inputs with latent inputs:

$$\eta_{ij} = \beta_0 + x_{ij}'\beta + a_i + b_j + \delta u_i v_j$$

Suggested language: “The multiplication is unchanged. What changes is that u and v are not columns in our dataset. The model has to learn them from repeated patterns across the whole relationship matrix.”

Work through the numerical example slowly. With $u_A = 2$, $u_B = -1$, $v_C = 1.5$, $v_D = -2$, and $\delta = 1$, the contributions are A to C = 3, A to D = -4, B to C = -1.5, and B to D = 2.

Ask: “Does A have a generally positive factor effect?” The answer is no. A’s contribution is positive with C and negative with D. The factor term describes pair-specific fit, not general source activity.

Move from rank one to rank R:

$$\eta_{ij} = \beta_0 + x_{ij}' \beta + a_i + b_j + \sum_r \delta_r u_{ir} v_{jr}$$

Suggested language: “Rank is simply the number of latent interactions. Rank one gives us one hidden source-target pairing. Rank two gives us two. The model adds their contributions to obtain one score for the relationship.”

Explain the directed parameterization. We can write the weights as a diagonal D in $u_i' D v_j$, or absorb them into the scale of U or V and write $u_i' v_j$. These are equivalent directed factorizations. In the symmetric eigenmodel, retaining the diagonal weights makes their positive and negative signs useful for distinguishing assortative and disassortative components.

Connect this to estimation. The probability model must learn β , a , b , U , V , reciprocity, and variance quantities jointly. In a probit AME fit, the sampler first represents each observed zero or one with a latent continuous score on the appropriate side of zero. Conditional on those scores, it updates the regression and additive terms, then U given V , then V given U , then the variance and reciprocity quantities.

If V were known, estimating U would look like a regression problem. If U were known, estimating V would look like a regression problem. ALS alternates between those two jobs and updates the other parameters. MCMC samples the unknown blocks conditional on their current values. The target for the MCMC fit is the joint posterior, not the squared reconstruction error from one SVD.

Likely question: “Is this just an interaction model?”

Answer: “Algebraically, yes, it is a small set of multiplicative interactions. Statistically, it is harder because both sides of every interaction are latent and estimated from the same relationship matrix.”

Likely question: “Can I call the first factor ideology?”

Answer: “Not from the network alone. A latent direction can mix several processes, and its orientation is not uniquely identified. Give it a substantive name only after comparing it with outside information.”

Transition: “Now the movie matrix will show how repeated row-column patterns can reveal those unknown profiles.”

Movie Preferences and the SVD

Walkthrough anchor: `#sec-svd-movies`

Chunks: `movie-preferences`, `movie-svd`, `movie-svd-reconstruct`, `movie-svd-profiles`, `movie-cell-products`

Run `movie-preferences`. Rows are viewers, columns are movies, and the cell is a recorded score from 0 to 5. A zero is an observed zero in this toy example, not a missing rating.

Ask students to describe the repeated patterns before showing any algebra. Mike, Cindy, Hyerin, and Emily differ mostly in the strength of a similar science-fiction preference profile. Juan favors the two romantic films. Cassy and Max mix the patterns.

Suggested language: “We could describe all 35 cells separately, but that misses the repetition. The SVD asks whether a small number of shared row and column patterns can reconstruct almost the whole matrix.”

Write:

$$M = U D V'$$

Read the pieces by their jobs. Rows of U describe how viewers load on shared directions. Rows of V describe how movies load on those directions. Diagonal entries of D order the directions by how much squared matrix magnitude they reconstruct.

Then write the rank- R approximation:

$$M_R = U_R D_R V_R'$$

Split the singular values evenly across the two sides:

$$P = U_R D_R^{(1/2)}$$

$$Q = V_R D_R^{(1/2)}$$

$$M_R = P Q'$$

$$\hat{m}_{ij} = p_i' q_j$$

Suggested language: “A viewer does not receive one universal ‘likes movies’ score, and a movie does not receive one universal ‘good movie’ score. The fitted cell depends on how the viewer profile and movie profile line up, direction by direction.”

Run `movie-svd`. The singular values are approximately 12.481, 9.509, and 1.346, with the final two at numerical zero. The first direction retains about 62.8% of squared matrix magnitude, and the first two together retain 99.3%. The rank-two reconstruction RMSE is 0.227.

Be precise about language. Because the matrix was not centered, call 99.3% the share of squared matrix magnitude reconstructed, not variance explained. Do not claim that

two psychological traits generated the ratings.

Run movie-svd-reconstruction. Ask students to compare the observed and reconstructed cells. The point is not perfect recovery. The point is that two repeated row-column patterns reproduce nearly all 35 scores.

Run movie-svd-profiles. Explain that the plotted vectors use balanced coordinates P and Q. The orientation may rotate or flip, so the axes do not arrive named “science fiction” and “romance.” We recognize genre patterns because the movie labels provide outside information.

Run movie-cell-products. Work through one fitted cell slowly:

Emily and Star Wars: $4.828 + 0.142 = 4.970$, compared with an observed score of 5.

Juan and Casablanca: $0.081 + 4.835 = 4.917$, compared with an observed score of 5.

Cassy and Alien: $1.130 + 0.162 = 1.292$, compared with an observed score of 2.

Suggested language: “The multiplication is literal. Emily’s first coordinate raises the Star Wars reconstruction because it lines up with the corresponding movie coordinate. The same viewer coordinate can do little for another movie because every contribution is a product of both sides.”

Likely question: “Did the SVD discover the movie genres?”

Answer: “No. It found directions that reconstruct the matrix under squared-error loss. We can connect those directions to genres only because we already know what the movie labels mean.”

Likely question: “Why can the reconstruction miss a cell?”

Answer: “Rank two forces 35 cells to share only two repeated patterns. The residual is the part that those two directions do not reconstruct.”

Likely question: “What if zero means the movie was not rated?”

Answer: “Then ordinary SVD would wrongly treat missingness as a recorded dislike. We would need a method designed for incomplete matrices or a likelihood fitted only to observed entries.”

Transition: “Now carry the same row-column multiplication into a directed relationship matrix.”

From Movie Profiles to Network Profiles

Walkthrough anchors: **#sec-svd-network** and the “SVD Is the Scaffold, Not the Network Estimator” subsection

Write the directed approximation:

z_{ij} approximately equals $u_i' v_j = \sum_r u_{ir} v_{jr}$

Suggested language: “The same actors appear on both sides of a directed network, but the row and column jobs differ. Row i describes actor i as a source. Column j describes actor j as a target. The product asks whether this source profile and this target profile line up.”

Separate additive and multiplicative jobs. A source effect says that actor i tends to direct many relationships toward many targets. A target effect says that actor j appears on the other side of many relationships. The product says that

this particular source-target combination recurs more or less often than those broad tendencies predict.

Use the table's examples only to explain the algebra. A donor-by-recipient aid matrix, a country-by-resolution voting matrix, and a source-by-target event matrix all have row and column profiles with different actions behind them.

Then state why SVD is only the scaffold:

1. SVD minimizes squared reconstruction error; a latent network model uses a likelihood suited to the outcome.
2. SVD treats every supplied cell as observed; a network model can define missing dyads, structural zeros, and the risk set.
3. SVD has no measured predictors, actor effects, or reciprocity; AME can include them.
4. SVD gives a point decomposition; MCMC or bootstrap refits can quantify uncertainty when they behave well.
5. SVD decomposes the matrix handed to it; AME estimates the low-rank surface jointly with the other model pieces.

Suggested language: “We do not run `svd()` once on a binary adjacency matrix and call that an AME fit. SVD gives us the low-rank intuition. The statistical model puts that product on the right probability scale and estimates it with the other terms.”

Transition: “The factor model keeps the product but places it inside the relationship model.”

The Multiplicative Term and Eigenvalue Signs

Walkthrough anchors: `#sec-factor` and `#sec-eigen`

Chunks: `fit-eigen`, `kmeans-factor`, `ldm-home`, `eigen-plot`

For a directed network, write:

$$u_i' v_j = u_{i1} v_{j1} + \dots + u_{iR} v_{jR}$$

For the symmetric eigenmodel, write:

$$u_i' \text{Lambda } u_j = \sum_r \text{lambda}_r u_{ir} u_{jr}$$

Explain the moving parts separately. The vectors describe continuous relationship profiles. The weight `lambda_r` tells us whether matching or opposing coordinates produce a positive pair contribution on that direction.

When `lambda_r` is positive, same-sign coordinates produce a positive contribution. This supports assortative, distance-like structure.

When `lambda_r` is negative, opposite-sign coordinates produce a positive contribution. This supports disassortative or complementary-role structure.

Run or display `fit-eigen`. In the distance-generated network, the posterior-mean weights are about +33.7 and +3.7, with the dominant weight positive. In the disassortative network, the weights are about -51.4 and -2.7, with the dominant weight negative.

Suggested language: “A negative eigenvalue does not mean a negative relationship. We must multiply the weight by both actor coordinates. With a negative weight, opposite-signed coordinates can create a positive pair contribution.”

Run `kmeans-factor` and `ldm-home`. On the distance-generated network, the factor model and distance model both recover the coarse simulated grouping at 95%. On the disassortative network, the factor model recovers the two roles at 100%, while the two-dimensional distance positions gave 55%.

Display `eigen-plot`. Keep the claim narrow. This is an illustration consistent with Hoff's containment result, not proof that a rank-two factor model beats every distance model on every network.

Likely question: "Is a factor the same as a block?"

Answer: "No. A block assigns or mixes actors among discrete roles. A factor gives each actor a continuous profile. Both can represent recurring roles, but they impose different structure and produce different summaries."

Likely question: "Does rank two mean two groups?"

Answer: "No. Rank two means two continuous profile directions. It does not imply two clusters or two automatically named concepts."

Transition: "We now put the factor together with the additive terms and let all three relationship pieces move across the ICEWS panel."

From Broad State Involvement to Particular Directed Relationships

Walkthrough anchor: `#sec-fits`

For the directed ICEWS application, write:

$$Y_{ijt} = \alpha_t + x_{ijt}' \beta + a_{it} + b_{jt} + u_{it}'v_{jt} + e_{ijt}$$

$$Y_{ijt} = 1(Y_{ijt} > 0)$$

Translate every term into the relationship being studied. The year intercept lets the overall threshold rate change. The measured pair variables describe Polity gap and same region. The additive effects describe broad source-side and target-side involvement. The multiplicative surface describes which ordered pairs fit together after those pieces enter.

Source state i receives u_{it} , target state j receives v_{jt} , and the product $u_{it}'v_{jt}$ represents source-target compatibility in year t . The changing surface is identified more directly than the separate raw factor matrices.

Likely question: “What makes the rank-zero model different from the dyadic probit?”

Answer: “Rank zero adds changing source and target effects, so broadly involved states can differ from less involved states. Rank two then adds changing partner-specific structure.”

Transition: “With the model pieces separated, let’s return to the ICEWS data and verify that the relationship changes enough to estimate movement.”

CHANGING STATE-PAIR RELATIONSHIPS WITH DYNAMIC AME

Walkthrough anchor: #sec-fits

Deck slide: Same Data, A More Specific Question

Set up the bridge from Day 9 before showing the new equation. Say: “Yesterday we tracked whether ICEWS

coded events from or toward each state across many partners. Those summaries could not tell us which bilateral relationships were driving the change. Today we let each directed state pair have its own path after accounting for those broad state patterns.”

Ask: “If I know that Iran and Syria each appear in many high-volume relationships, do I yet know whether Iran toward Syria crosses the threshold?”

Expected answer: No. Broad involvement is not the same thing as the fit of one ordered pair.

State the question in one clean sentence: “After accounting for states that appear with many partners, does the chance of a high-volume relationship change for particular ordered pairs?”

Keep the outcome wording precise. A one means that ICEWS contains more than 20 coded material-conflict events from one state toward another in a year. It does not mean that the source initiated a war, that the target suffered more harm, or that the relationship was more severe. It is a high coded event-volume relationship.

Transition: “Now that the relationship and question are clear, we can give every part of the model one job.”

THE DYNAMIC AME EQUATION

Walkthrough anchor: #sec-fits

Deck slide: The Dynamic AME Keeps Five Jobs Separate

Write the equation slowly:

$$Y_{ijt} = \alpha_t + x'_{ijt} \beta + a_{it} + b_{jt} + u'_{it} v_{jt} + e_{ijt}$$

$$Y_{ijt} = 1(Y^*_{ijt} > 0)$$

Read from left to right. Alpha_t is the baseline for the whole network in year t. The x-beta term contains Polity gap and same region. The a_it term adjusts for source states that appear with many partners in year t. The b_jt term does the same for target states. The u-v inner product adjusts this exact ordered pair beyond those broad partner counts.

Suggested language: “The broad paths tell us whether ICEWS records events from or toward a state across many partners. The pair-specific part tells us whether one directed state relationship has a different history than those broad patterns would suggest.”

Likely question: “Are u and v positions?”

Answer: “They are source and target profiles whose inner product creates a fitted relationship adjustment. A plot can make them look like positions, but the model uses the inner products. The axes can rotate or reflect, so we do not name them as if the model discovered two literal traits.”

Likely question: “Is this already AME?”

Answer: “Yes. Rank-two ALS with additive and multiplicative effects is a point-estimated AME model. Moving to MCMC would change how we represent uncertainty, not create a different model family.”

Transition: “Before fitting movement, we need to establish that the observed network actually moves.”

RETURN TO THE ICEWS PANEL

Walkthrough anchor: #sec-icews-data

Chunks: icews-dynamic-data, icews-dynamic-change

Deck slides: Same Data, A More Specific Question; There Is Enough Movement to Ask

Run icews-dynamic-data. The panel contains 18 states, 13 years from 2002 through 2014, and 3,978 ordered state-pair-years. About 21.8 percent of the observations cross the threshold.

Explain the case selection. These are the 18 highest-volume ICEWS source states among states with complete Polity and GDP coverage across the panel. The selection is purposive because the classroom example needs enough positive relationships and enough change. It is not a probability sample, and the event rate should not be generalized to the international system.

Run icews-dynamic-change. Across adjacent years, 244 ordered relationships move from zero to one and 218 move from one to zero. The total is 462 changes.

Suggested language: “A dynamic model cannot manufacture information about movement. Here we have hundreds of observed changes, so time-varying effects have something to learn.”

Likely question: “Why use a threshold of 20?”

Answer: “It creates a binary high-volume relationship with enough ones and enough changes for this class example. It is a measurement choice, not a universal definition of conflict. In a research project, I would check nearby thresholds and the count outcome itself.”

Likely question: “Why keep direction?”

Answer: “Iran toward Syria and Syria toward Iran can have different coded histories. A directed model keeps that difference instead of averaging it away.”

Transition: “Now we turn one long pair-year table into the exact aligned matrices the estimator needs.”

BUILD THE MATRICES WITH NETIFY

Walkthrough anchor: `#sec-icews-netify`

Chunk: `icews-dynamic-netify`

Deck slide: `netify` Keeps the Panel Aligned

Run the chunk and point to the arguments rather than reading every line. `actor1` is source, `actor2` is target, `time` is year, `symmetric` is `FALSE`, and `weight` is the binary threshold outcome.

Explain `missing_to_zero = FALSE`. The source data already contain every ordered pair in every year. Every off-diagonal zero is therefore observed. We do not need the software to invent zeros for absent rows.

Point out `validate_netify()`. It checks actor labels, time slices, direction, missingness, and agreement among the object’s representations. The `stopifnot()` line makes the walk-through stop immediately if a coherence check fails.

Then explain `to_lame()`. It returns 13 outcome matrices and 13 predictor arrays with the same state order. This prevents a very dangerous error in longitudinal relational models: placing one state’s outcome in another state’s row because the ordering changed.

Checkpoint: 13 outcome matrices, 18 actors in each matrix, and 2 predictors.

Transition: “We have aligned inputs. Next we decide which parts are allowed to change.”

WHAT IS DYNAMIC

Walkthrough anchor: #sec-dynamic-pieces

Deck slide: Broad Involvement Is Not Partner Fit

Separate the three moving parts.

α_t lets the baseline event rate change for the entire network.

a_{it} and b_{jt} let each state’s broad source-side and target-side involvement change.

u_{it} and v_{jt} let the source-target compatibility surface change.

Explain $\text{dynamic_beta} = \text{“intercept”}$. Only the intercept moves. The Polity-gap and same-region coefficients remain pooled across the 13 years. This keeps the applied question manageable and avoids estimating a noisy annual coefficient for every predictor.

Explain $\text{dynamic_beta_kind} = \text{“rw1”}$. A first-order random-walk penalty discourages sharp back-and-forth changes in the yearly intercept. It allows sustained drift without pulling the path toward one fixed long-run mean.

Important wording: Do not call ρ_{ab} , ρ_{uv} , or ρ_{beta} estimated persistence. In dynamic ALS they are fixed smoothing values derived from the prior settings. They tell the optimizer how strongly adjacent years should borrow information.

Suggested language: “Smoothing is partial pooling across neighboring years. It keeps us from treating 2008 and 2009

as unrelated studies, while still allowing a path to move when the observations consistently support that movement.”

Transition: “The estimator now cycles through these pieces instead of trying to solve everything in one step.”

HOW DYNAMIC ALS WORKS

Walkthrough anchor: `#sec-dynamic-als`

Deck slides: What Dynamic ALS Is Trying to Do and What Dynamic ALS Minimizes

Begin with the two jobs. The fit should reconstruct which state-pair-years cross the event threshold, and it should keep neighboring years connected unless the observations support a sustained movement.

Show the simplified objective. The first term is weighted squared error between a binary-response working score and the fitted predictor. The other terms penalize large year-to-year changes in the actor paths, factor paths, and intercept. The lambda values say how expensive those jumps are.

Walk students through the update sequence.

First, begin with provisional fitted probabilities for every ordered pair-year.

Second, turn those probabilities into a working response and weights.

Third, update the yearly baseline and the two measured coefficients while holding the actor and factor paths fixed.

Fourth, update the broad source and target paths.

Fifth, update the changing source and target factor profiles, using low-rank matrix updates to revise the pair-specific sur-

face.

Sixth, recalculate the probabilities and repeat until the objective and fitted surface barely change.

Suggested language: “It is a coordinated back-and-forth. Each step solves an easier problem while treating the other pieces as temporarily known. The cycle continues until the pieces agree.”

Distinguish ALS from MCMC. For a binary model, ALS solves a sequence of weighted least-squares approximations and returns one penalized point estimate. It does not directly maximize the exact Bernoulli likelihood in one step. With rank greater than zero, the problem is nonconvex, so multiple starts matter. MCMC would return draws from a posterior distribution. The parametric bootstrap later repeatedly simulates and refits the ALS model to quantify one form of model-based uncertainty.

Likely question: “Does fast mean approximate or bad?”

Answer: “Fast means we are optimizing for a point estimate rather than sampling a posterior. We still have to check convergence, multiple starts, held-out prediction, and goodness of fit. Speed does not remove those responsibilities.”

Transition: “Let’s look at the actual call and the evidence that it found a stable solution.”

FIT THE DYNAMIC SRM AND DYNAMIC AME

Walkthrough anchor: #sec-dynamic-fit

Chunks: icews-dynamic-fit, icews-dynamic-fit-call

Deck slide: The Actual lame() Call

Load the cached fits. Do not run the build script during

class. The downloaded folder contains the dynamic rank-zero fit and the dynamic rank-two fit with 100 bootstrap refits.

Read the key arguments. `R = 2` adds two multiplicative dimensions. `dynamic_ab = TRUE` turns on changing additive effects. `dynamic_uv = TRUE` turns on changing factor profiles. `dynamic_beta = "intercept"` allows the baseline to move. `method = "als"` selects penalized point estimation. `als_stability = "validation"` runs four jittered starts. `bootstrap = 100` requests 100 parametric refits.

Both main fits converged. The rank-two fit took 48 iterations. All four stability starts converged. Their fitted probability surfaces correlate at essentially 1.00, and the largest surface RMSE is below 0.0001.

Suggested language: "Different starts reached the same relationship surface. That does not prove the model is correct, but it is strong evidence that our displayed point estimate is not a random local solution."

All 100 bootstrap refits succeeded. Explain that the bootstrap simulates a panel from the fitted model, refits the same model, and records how estimates vary. It is conditional on the fitted model, rank, smoothing setup, outcome threshold, and case selection.

Transition: "We will read the model in the same order as its equation: broad state paths first, then exact partner profiles."

READ THE ADDITIVE PATHS WITH AB_PLOT

Walkthrough anchor: `#sec-dynamic-ab`

Chunk: `icews-dynamic-ab-plot`

Deck slide: Broad State Paths Come First

Display the source and target panels for the United States, Syria, Iran, and the Russian Federation. The plot uses `lame::ab_plot()`, which reads the model's dynamic additive-effect arrays directly.

Start with the source-side plot. A higher line means the model needs a larger broad source adjustment for that state after the year baseline, measured predictors, and multiplicative surface enter.

Then read the target-side plot. A higher line means the model needs a larger broad target adjustment.

Do not say that a higher source effect means a state started more conflicts. Do not say that a higher target effect means a state was a greater victim. The outcome is high coded event volume, and the additive scores pool across all partners.

Suggested language: "These paths tell us whether events are coded from or toward a state across many different partners. They do not identify which bilateral relationship accounts for the pattern. That is the next question."

Likely question: "Where are the uncertainty intervals?"

Answer: "The bootstrap refits contain uncertainty for the additive paths. The main display stays readable by showing the point paths. If we report one state's path substantively, we should add its bootstrap interval and remember that the interval is conditional on the model and measurement choices."

Transition: "Now we move from how often a state appears broadly to which exact partners fit."

READ THE CHANGING FACTOR PROFILES WITH UV_PLOT

Walkthrough anchor: #sec-dynamic-uv

Chunk: icews-dynamic-uv-plot

Deck slide: Then Ask Which Directed Relationships Changed

The code first calls `align_directed_uv()` and then `uv_plot()`. The helper stacks the source and target profiles, applies one common rotation, and verifies that every fitted source-target inner product remains unchanged. This keeps arbitrary changes in pose from masquerading as movement without changing the fitted relationship surface.

Explain source and target profiles. A source point and target point that line up contribute a positive inner product. Points facing in opposing directions contribute a negative inner product. The exact contribution is $u_{it}'v_{jt}$.

Do not name either axis. The axes can rotate, reflect, rescale, or mix while preserving the fitted surface. The estimand is the matrix of inner products and the resulting probability, not one raw coordinate column.

Suggested language: “The plot helps us inspect the surface. The model is not claiming that a state literally moved east or west in a geopolitical space.”

Likely question: “Why align if the axes do not matter?”

Answer: “Alignment removes arbitrary changes in pose so we can see whether the invariant relationship pattern changed. It makes the picture readable without making the axes substantive.”

Transition: “Before telling a story from two dimensions, we need to show that rank two earned its complexity.”

CHOOSE THE RANK

Walkthrough anchor: #sec-dynamic-rank

Chunk: icews-dynamic-rank-cv

Deck slide: Rank Two Is a Practical Choice

Define the held-out unit carefully. Every unordered pair is assigned to a fold. When a pair is held out, both directions and all 13 years are hidden together. This prevents the model from seeing Iran-Syria in another year or Syria-Iran in reverse while scoring that pair.

Explain the metrics briefly. ROC AUC asks whether positive relationships tend to receive higher scores than zero relationships. Precision-recall AUC focuses on how well positives are concentrated near the top and is useful with an imbalanced outcome. Brier score is the average squared probability error. Log loss strongly penalizes confident wrong probabilities.

Mean results: rank zero has ROC AUC 0.825, precision-recall AUC 0.676, and Brier 0.118. Rank one has 0.837, 0.694, and 0.113. Rank two has 0.837, 0.695, and 0.113. Ranks three and four do not provide a consistent gain.

Suggested language: “Rank one and rank two are essentially tied on hidden pair histories. I keep rank two because it also improves the transition check and gives us a two-dimensional profile display we can inspect. That is a balance among prediction, temporal reproduction, and interpretability, not proof that reality has exactly two dimensions.”

Transition: “Now we can turn the abstract inner products into four ordered relationship histories.”

FOLLOW SPECIFIC STATE-PAIR PATHS

Walkthrough anchor: #sec-dynamic-pairs

Chunk: icews-dynamic-pair-paths

Deck slide: Direction and Timing Matter

Explain the symbols before interpreting the lines. The line is the estimated probability of crossing the threshold. A filled point means the observed relationship crossed the threshold in that year. An open point means it did not.

Iran toward Syria remains below the threshold through 2011 and crosses it in 2012, 2013, and 2014. Its fitted probability rises from about 2 percent in 2002 to 92 percent in 2012. The 2012 bootstrap interval is roughly 57 to 98 percent.

Syria toward Iran crosses the threshold in 2011 and stays above it through 2014. Its fitted probability rises from about 1 percent in 2002 to 69 percent in 2011. The 2011 bootstrap interval is roughly 27 to 83 percent. The model gives the two directions separate paths because the observed directed relationships differ.

The United States toward Syria crosses the threshold in nearly every year after 2002 and receives a consistently high fitted probability. Syria toward the United States is more irregular early and becomes consistently above the threshold later.

Suggested language: “The rise involving Iran and Syria is not simply a uniform increase across every relationship involving either state. Iran toward Syria and Syria toward

Iran follow different paths.”

Keep the causal boundary visible. The paths can be connected to dated events in a research project, but the fitted movement does not identify why the coded event relationship changed.

Likely question: “Can we say Iran and Syria became closer?”

Answer: “Only in a carefully qualified relational sense. Their fitted high-volume event relationship became more compatible with the observed network pattern. We should not turn that into a claim about general diplomatic closeness.”

Transition: “A compelling pair story still needs a whole-network check.”

CHECK ANNUAL GOODNESS OF FIT

Walkthrough anchor: #sec-dynamic-gof

Chunks: icews-dynamic-gof-summary, icews-dynamic-transition-plot

Deck slides: The Annual Shape Is Mostly Right; Rank Two Gets Much Closer Over Time

Explain the simulation check. The code holds the fitted ALS solution fixed, simulates 500 complete panels, and compares each simulated panel with the observed one. It does not refit the model inside every GOF simulation and does not integrate over parameter uncertainty.

The rank-two model covers density, source-rate variation, and target-rate variation in all 13 years. It covers cyclic and transitive triadic dependence in 12 of 13 years. It covers reciprocity in only 6 of 13 years.

Translate those statistics. Density asks whether the total amount of above-threshold relationship volume is right. Source- and target-rate variation ask whether involvement is as uneven across states as observed. Reciprocity asks whether opposite directions line up as often as observed. The triadic checks ask whether three-state relationship patterns are reproduced.

Suggested language: “The annual networks mostly have the right volume, unevenness, and three-state pattern. Reciprocity is the main annual weakness.”

Now show the transition plot. Black points are observed numbers of ordered pairs that changed threshold status. The line is the simulated median, and the ribbon contains the middle 95 percent of simulated counts.

Rank zero covers only 2 of 12 observed transitions. Rank two covers 8 of 12. Rank two closes much of the temporal gap, but four transitions remain outside the interval.

Suggested language: “The dynamic factors make the simulated movie much more realistic. They do not make it perfect. The model still understates the persistence or timing of some relationships.”

Likely question: “If annual GOF is good, why can transition GOF fail?”

Answer: “Two movies can contain similar still frames but arrange the changes differently. Annual checks inspect each network separately. The transition check inspects how one frame turns into the next.”

Transition: “We now have enough evidence to state the result without overselling it.”

WHAT THE CODED EVENT RECORDS SHOW

Walkthrough anchor: #sec-dynamic-after

Deck slide: What the Iran-Syria Paths Show

Suggested full interpretation: “Across these 18 states, the model separates a general rise in how often a state appears in above-threshold relationships from a rise concentrated in one directed pair. The estimated chance for Iran toward Syria rises from about 2 percent in 2002 to 92 percent in 2012. Syria toward Iran rises on a different timetable and reaches about 69 percent in 2011. These changes occur around the escalation of the Syrian conflict, but the event records and model do not establish why they occurred or measure battlefield severity. The richer model reproduces much more of the observed year-to-year change than the broad state-level model, although it still misses reciprocity in several years and four annual transitions.”

Give students the post-estimation checklist.

First, check convergence and multiple starts.

Second, define and inspect the held-out prediction target.

Third, inspect annual structure and transitions.

Fourth, interpret additive paths separately from partner-specific inner products.

Fifth, use `ab_plot()`, `uv_plot()`, and `latent_positions()` as entry points, then return to fitted inner products and probabilities.

Sixth, connect interesting paths to historical evidence and alternative outcome definitions.

Transition: “The dynamic example shows what AME can

represent well. It still does not solve omitted-variable bias by itself.”

Published Illustration: Conflict in Nigeria

Walkthrough location: the collapsed Nigeria illustration following `#sec-gof`

State that this is a published application from Dorff, Gallop, and Minhas (2020), not the model just fit in class.

The authors study directed ACLED battle relationships among 37 Nigerian armed organizations from 2000 through 2016. They ask: “Who fights whom, when, and how did the wider conflict system change after Boko Haram entered it?”

Explain the AME pieces in actions. Measured variables include prior attacks on civilians, geographic spread, government involvement, election years, conflict in neighboring countries, and the period after Boko Haram’s 2009 uprising. Additive source and target effects represent organizations that initiate or receive battles broadly. Residual reciprocity represents retaliation. Multiplicative factors represent recurring opponent profiles.

Use the Boko Haram, MASSOB, and MEND result to explain stochastic equivalence. These organizations operated in different places and pursued different projects, but all repeatedly fought Nigerian military and police forces. The model placed them in similar relational roles because of whom they fought, not because they formed one group or shared one ideology.

The published substantive comparisons are:

1. Moving from no attacks on civilians to 17 attacks multiplies the estimated risk of initiating a battle by about 1.58 and the risk of being targeted by about 1.55, with the paper's comparison values held fixed.
2. The estimated directed-pair battle risk is about 2.82 times as large after Boko Haram's 2009 uprising, including pairs that did not contain Boko Haram.
3. Under the paper's validation procedure, AME reaches ROC AUC 0.92 and precision-recall AUC 0.33. A regression with the same measured predictors plus a lagged outcome reaches 0.82 and 0.26, while a measured-predictor regression reaches 0.79 and 0.15.

Suggested language: "The model shows that violence was organized as a system. Some organizations initiated conflict broadly, some were targeted broadly, retaliation connected directions, and organizations could occupy similar roles by fighting the same state forces."

State the boundaries. The post-2009 comparison is not a randomized effect of Boko Haram's entry. The civilian-targeting results are conditional associations. ACLED captures coded reports, not every unreported event or each battle's severity. The evidence comes from one conflict and one validation design.

Likely question: "Why is precision-recall AUC important here?"

Answer: "Battles are rare among all possible directed organization pairs. Precision-recall performance focuses on how successfully the model concentrates actual battles among its highest-ranked predictions."

Transition: “Both applications improve our description of relational dependence, but neither makes a correlated omitted process disappear. That is the last major boundary for Day 10.”

What Latent Models Did Not Buy Us

Walkthrough anchor: `#sec-notbuy`

Begin with the central sentence: “A latent term can represent residual dependence without automatically controlling confounding.”

Use the notation:

$$Y^* = \beta X + \gamma W + e$$

$$W = \alpha X + z$$

Substitute:

$$Y^* = (\beta + \gamma \alpha)X + \gamma z + e$$

Explain each symbol. X is the measured predictor. W is an omitted relational pattern. β is the association we would like to separate. γ says how W enters the outcome. α describes how W moves with X . The fitted coefficient can combine β with γ times α .

Suggested language: “If the omitted relationship pattern rises where X rises and that omitted pattern also raises the outcome, a simple coefficient can give X credit for both. If the omitted pattern moves in the opposite direction, the coefficient can be pushed downward. If the two are unrelated after the other controls, this particular covariance channel does not move the slope, although dependence and uncer-

tainty problems can remain.”

Connect directly to the random-effects and AME context. The additive SRM treats actor effects as random quantities with a relationship to X specified by the model. AME adds a low-rank random relationship surface. Neither model automatically proves that those latent quantities are independent of the measured predictor.

Use the ICEWS setup. Suppose unmeasured international news visibility raises coded event volume and is also associated with regime profile. The random source, target, and factor effects can represent part of that leftover pattern, but the same event records may still be unable to separate the Polity-gap association from the portion carried by visibility. The model’s ability to fit the relational pattern does not identify the direct effect of regime difference.

Explain Minhas et al. (2022) in basic terms. When the omitted low-rank structure is independent of X , AME can absorb the patterned residual and recover the measured association much better than an independence model. When the omitted structure and X are correlated, the same network may not contain enough information to decide how much shared pattern belongs to X and how much belongs to the latent surface.

Likely question: “Can we regress the estimated factors on X to test exogeneity?”

Answer: “That describes overlap in the fitted representation. It cannot test whether the true unobserved process was independent of X , especially because the factors were learned from the same outcome.”

Likely question: “Does a stable coefficient across ranks prove exogeneity?”

Answer: “No. It can be stably biased. Movement across ranks shows sensitivity, not which specification is causal.”

Likely question: “Does lagging X solve the problem?”

Answer: “No. Lagging can improve temporal ordering, but it does not establish strict or sequential exogeneity when ties and attributes can affect one another.”

State the practical response. Measure important pre-treatment confounders, use fixed effects or correlated random-effects strategies when the needed within-actor variation exists, exploit design when available, and report sensitivity across defensible relational specifications.

Transition: “We can now separate three questions that are too often blurred: did the algorithm run, does the model reproduce or predict the network, and does the design support the substantive claim?”

After the Model Runs

Walkthrough anchor: **#sec-after-fit**

Give students a four-part reporting routine:

1. Translate measured associations into meaningful probabilities or outcome contrasts, state what is held fixed, and report uncertainty only when the inferential computation supports it.
2. Interpret invariant relationship quantities such as fitted probabilities, pairwise distances, inner products, or complete surfaces, not arbitrary axes.

3. State what the latent term could contain, including institutions, strategy, history, geography, measurement, and shared exposure.
4. Report whether the model earned the interpretation through computation, simulation checks, a clearly defined held-out target, rank sensitivity, and visible failures.

Use this Gade summary: “Al-Nusrah Front and Ahrar al-Sham Islamic Movement sit in the same tightly connected part of the fitted map. They recorded an operation with one another, and each also recorded operations with 18 of the same other organizations. Simulated networks from the model reproduce much of the observed clustering. This mostly reinforces the tactical-core result from Day 9. It does not explain why the organizations cooperated or give substantive meaning to the displayed axes.”

Use this ICEWS summary: “Across 18 states from 2002 through 2014, the model separates a general rise in how often a state appears in above-threshold relationships from a rise concentrated in one directed pair. Iran toward Syria rises from about 2 percent in 2002 to 92 percent in 2012, while Syria toward Iran rises on a different timetable and reaches about 69 percent in 2011. The richer model reproduces much more of the observed year-to-year movement than the broad state-level model, but reciprocity and four transition intervals remain weaknesses. The event threshold measures coded volume rather than conflict initiation or severity.”

Use this Nigeria summary: “Dorff, Gallop, and Minhas (2020) show that directed additive and multiplicative

structure can separate broad initiators, broad targets, retaliation, and recurring opponent profiles. Their AME model improves out-of-sample battle prediction, but its post-2009 and civilian-targeting comparisons remain conditional associations.”

Ask: “If a paper reports only an AME coefficient table, what is missing?”

Listen for the relationship surface, pair-level results, probability scale, chain diagnostics, prediction target, posterior or simulation fit, rank sensitivity, outcome definition, and causal limits.

Transition: “The last step is practice. The exercises ask students to move from model output to claims they could defend in a paper.”

Exercises and Discussion

Walkthrough anchor: **#sec-exercises**

For exercise 1, have students find the closest and farthest organizations in the Gade map, check whether the close pair has a recorded tie, and write one sentence that treats distance as a whole-network summary rather than proof of a shared ideology.

For exercise 2, compare Iran toward Syria with Syria toward Iran. Require the first threshold-crossing year in each direction, a description of the fitted probability paths, and one sentence explaining why direction should not be collapsed.

For exercise 3, compare ranks zero through four. Require students to say that both directions and all 13 years for

complete unordered state pairs were held out and that the exercise tests unseen relationships among known states.

For exercise 4, use both `gof_coverage` and `transition_plot_data`. The dynamic rank-two model covers annual density and source- and target-rate variation in all 13 years, both triadic summaries in 12 of 13 years, reciprocity in 6 of 13 years, and adjacent-year change counts in 8 of 12 transitions. Require students to identify reciprocity as the largest annual weakness and explain why annual fit does not guarantee transition fit.

For exercise 5, fit the Gade distance model under a second seed and align the maps. Require students to distinguish arbitrary pose from remaining uncertainty or optimization differences.

For exercise 6, have students define a validation target before interpreting a factor fit on their own network.

Likely discussion prompt: “Which failure is more serious for the claim you want to make: weak held-out prediction, poor reproduction of a specific network statistic, or nonconvergent coefficient inference?”

Expected answer: It depends on the claim. A predictive claim needs held-out performance, a structural descriptive claim needs the relevant reproduction check, and an inferential coefficient claim needs valid sampling plus defensible identification.

Closing and Bridge to Day 11

Walkthrough anchor: `#sec-missing`

Suggested closing: “AME represents higher-order structure through shared latent profiles. It does not estimate a closure coefficient, and reproducing transitivity does not identify closure as the generating process.”

Day 11’s ERGM asks a different question by placing chosen graph statistics inside a joint probability model for the whole network. AME uses latent actor and pair terms in the link-scale mean. ERGM uses change in explicit graph statistics when a tie is toggled, conditional on the rest of the graph.

End with the substantive distinction: “Broad actor involvement, recurring pair compatibility, and explicit graph configurations are different kinds of structure. A useful model tells us which one it represents, what relationship question it answers, and where the evidence stops.”

Final Reminders

Do not call the Gade tie an alliance. It is at least one recorded joint operation in a whole-period binary projection.

Do not interpret horizontal or vertical directions as named concepts. Use distances, fitted probabilities, and the invariant multiplicative surface.

Do not treat rank as a number of groups. Rank controls the dimension of a continuous low-rank surface.

Do not call ALS coefficients posterior estimates or attach intervals to them.

Do not treat complete-pair validation as future forecasting. It tests unseen ordered relationship histories among the same known states within 2002 through 2014.

Do not treat a good prediction score as a causal result. The random effects and multiplicative surface adjust modeled dependence but do not establish exogeneity.

Do not read `ab_plot()` paths as conflict initiation or victimization. They are broad adjustments for a high coded event-volume outcome.

Do not name raw `uv_plot()` axes. Align the paths, then interpret inner products and fitted probabilities.

Do not hide the temporal evidence. Rank two improves the transition check substantially, but reciprocity and four transition intervals remain weak.

Do not call the ALS smoothing values estimated persistence parameters. They are fixed tuning values derived from the prior settings.

End with what the model reveals about relationships, not with the package name.