

Response to Dan's Feedback

COMPREHENSIVE RESPONSE TO DAN'S FEEDBACK

HOW I'VE ADDRESSED EACH POINT

COMMENT 1: Introduction - State the actual problem, avoid broad characterizations

Dan's Feedback: > “The introduction doesn't state the actual problem that motivates your intervention, and has these broad lines about individual vs. group-level studies that don't help you advance your argument and yet may well trigger reviewers who disagree with these sweeping characterizations.”

What I've Done: Addressed in Introduction - Cut broad characterizations about individual vs group-level studies - Tried to state the problem clearly and directly: interaction terms are unaccounted for in LDA, leading to misallocated votes - Tried to get to the point faster

COMMENT 2: Clarify what's doing the work - interaction terms

Dan's Feedback: > “Clarify what's really doing the work here—why does your approach differ from prior approaches? My sense is that it's the interaction terms, no? Give readers some intuition for what the interactions represent.”

What I've Done: Addressed in Section 2 (Lines 165-199) - Added explicit substantive explanation BEFORE math: what interaction terms represent - Concrete example: Hispanic voters both growing AND shifting Republican creates multiplicative effect - Explained why LDA misses these: “By omitting these terms, the LDA implicitly assumes they are zero or trivially small” - Forward reference: “However, as I demonstrate in Section 4, these interaction terms can represent millions of votes” - Simplified Taylor series explanation to plain language

about linear approximations - Made it clear: RCVD captures interactions through sequential calculation, LDA doesn't

COMMENT 3: Avoid math jargon, lead with intuition

Dan's Feedback: > "You go very quickly to the math terms (derivatives, eigenvalues, Taylor series approximation, etc.)... Why not just say 'linear transformation' instead of eigenvalue? Lead with intuition, not math terms. The math doesn't speak for itself."

What I've Done: Addressed throughout Sections 2, 3, 4 - **Section 2:** Removed ALL "derivative" language, changed to "Linear Difference Approach" - **Section 2:** "eigenvalue" → "scaling factor" (line 453 in Non-Uniqueness) - **Section 2:** Added substantive explanations before all math - **Section 3:** Removed "derivative-based calculation" language - **Section 3:** Created simple worked example with INTUITION first, math second - Global replacement: DBA (derivative-based approach) → LDA (Linear Difference Approach) - Every technical concept now has plain-language explanation first

COMMENT 4: Establish functions are continuous/differentiable OR avoid derivative language

Dan's Feedback: > "You talk about derivatives, but I didn't see whether they are in fact smooth, differentiable, etc. That has to be established first before you talk about their derivatives. Do Grimmer et al. refer to their method as derivative-based? If not, maybe emphasize change rather than derivatives."

What I've Done: Addressed throughout - Removed all derivative language - Changed from "derivative-based approach" to "Linear Difference Approach" (LDA) - No longer claiming to take formal derivatives - Focus on differences/changes, not derivatives - This sidesteps the continuity/differentiability question entirely

COMMENT 5: Clarify role of Volume

Dan's Feedback: > “What exactly is the role of volume? You have a section ‘The Problem is Not Volume’ showing errors aren’t from volume.”

What U’ve Done: Addressed in two ways: 1. **Section 2:** Clarified Volume’s role as component that ensures votes add up correctly 2. **Section 3:** MOVED “The Problem is Not Volume” to appendix with footnote - Footnote explains: Even with no Volume change, LDA still has 1.7M residual - This proves residual is from interaction terms, not missing Volume - Following Dan’s “fight on one front” advice - not being defensive in main text

COMMENT 6: Walk through calculations, provide worked example

Dan's Feedback: > “You use both methods but only present the output, so it’s hard to build intuition about why the two techniques differ. Is it possible to work through the calculations and provide more of an explanation for their differences, instead of simply saying ‘I did the math, here’s what they look like.’ Maybe derive a quantity of interest and explain how you calculated it under each approach?”

What I’ve Done: Addressed in Section 3 (NEW subsection, lines 242-371) - **Created brand new simple worked example** at start of Section 3 - Hypothetical data: 2 groups, 100→120 voters, all whole numbers - **Step-by-step LDA calculation** showing how it gets -0.9 vote residual - **Step-by-step RCVD calculation** showing how it gets 0.0 residual (exactly) - Readers can verify with a calculator - Shows EXACTLY where the approaches differ

COMMENT 7: Order of calculation - clarify assumptions

Dan's Feedback: > “The fact that the calculation depends on the order makes sense... are these two things linked? If so, I’d be super clear about the connections.”

What I’ve Done: Addressed in Section 3, Non-Uniqueness subsection (lines 438-456) - Removed “state-space is a ball” jargon - Removed “eigenvalue” language → “proportional scaling factor” - **Two clear substantive justifications:** 1. Rate first ensures continuity with existing literature (validates past work) 2. Volume last ensures it acts as clean proportional scaling - Concrete example: “if we shifted Composition first...the Volume effect would be confounded” - Attempted to take on a less defensive tone, more matter-of-fact

COMMENT 8: Figures - walk readers through, make take-home points clear

Dan's Feedback: > “Figures can be better than tables, but only when they clearly illustrate what you want us to take away. Walk the reader through the figures in more detail. What key take-home points do you want the reader to leave with? Consider redesigning the figure to clarify that.”

What I've Done: Addressed in Section 4 - **Reviewed all 6 figures** - I feel like each serves clear purpose: - Fig 1: Quantifies LDA error (365k votes) - Fig 2: Shows error relative to effects (700k error vs 1.5M composition) - Fig 3: KEY - ANES vs PEW comparison (4M vs 600k error) - Fig 4: 2016-2020 baseline - Fig 5: CORE FINDING - 2020-2024 reversal - Fig 6: 2016-2024 synthesis - Each figure caption clearly explains what to see - Text walks through what each figure shows - No redundant figures - **I'd like more feedback here if I still need to cut!**

COMMENT 9: ANES vs PEW analysis doesn't help main argument

Dan's Feedback: > “The analyses of Pew vs. ANES for 2020 don't help you advance your main argument. If Pew is closer to correct, why not just use Pew? The message I took is that the survey source is what really matters—not presumably the message you want to foreground.”

What I've Done: Addressed in Section 4 - Streamlined by ~50% - Cut from ~22 lines to ~11 lines - Shortened methodological justification significantly - Got to the Hispanic voter finding faster - Kept key point: PEW shows Rate effect, ANES shows Composition - Shortened strategic implications - Attempted to be more focused, less digressive

COMMENT 10: Show the method matters substantively

Dan's Feedback: > “If you want to garner significant attention, you'll need to show that the choice of method matters. Is it possible to find a point on which your method and the Grimmer et al. approach reach different substantive conclusions?”

What I've Done: Addressed throughout - **Core take-home:** LDA misallocates millions of votes due to interaction terms; RCVD fixes this - **Section 3:** Shows 1.4M residual in real 2016-2020 data - **Section 4:** Shows different interpretation of 2024 (Rate vs Composition decoupling) - **Streamlined** to focus on this one main point - Attempted to cut defensive asides and tangential arguments - Everything subordinated to: “RCVD gives accurate decomposition, LDA doesn't”

SUMMARY OF MAJOR CHANGES:

Section 2: Literature Review & RCVD Introduction - Removed derivative language throughout - Added substantive explanation of interaction terms - Changed “eigenvalue” to “scaling factor” - Explained RCVD intuition before math

Section 3: Proof of Concept - NEW: Simple worked example - Streamlined Two-Group Simulation (~20% shorter) - Moved “Problem is Not Volume” to appendix - Rewrote Non-Uniqueness subsection (clearer, less jargon) - All DBA → LDA globally

Section 4: Empirical Application - Streamlined ANES vs PEW by ~50% (Dan’s specific request) - Tightened all subsections (~30% shorter overall) - Attempted to remove overly dramatic language - All figures reviewed and justified

Conclusion - Broke up overly long sentences - Softened overly strong claims - Attempted to be more confident, less defensive

Overall Paper - Global DBA → LDA replacement (all 10+ instances) - Removed ALL derivative/eigenvalue jargon - Lead with intuition, not math - ~25-30% shorter overall - “Fight on one front” - focused on interaction term problem