Response to Reviewers

The Economic Cost of Agricultural Climate Migration in the United States

We thank both reviewers for their detailed and substantive comments. Every point led us to redo analysis, not re-word it. Each reviewer's report is reproduced *in full and verbatim* in the shaded boxes below, in its entirety and without abridgement or paraphrase, followed by our point-by-point response. All monetary values are in 2023 USD.

Five components were re-estimated:

Component	What it now is
Stranded value (DCF)	Real-price DCF on flat 30-year inflation-adjusted prices, with an alternate-use (grazing) floor
Hedonic	Ricardian hedonic with SSURGO soil, available-water, and irrigation controls
Insurance	Rolling Actual-Production-History simulation with full gross→residual decomposition
Migration	Leave-one-out shift-share (Bartik) instrument, estimated on prime-age population
Northern opportunity	Defined as *net* farm income, not gross revenue

Beyond the five re-estimated components, the analysis adds nine further items, each described at the relevant point below and collected in Supplementary Information Part V: (1) a wild-cluster restricted bootstrap and county-clustered inference for the migration IV ($p = 0.0001$ and $p = 0.001$); (2) a 200,000-draw Monte Carlo for the migration dollar cost, propagating the elasticity's sampling distribution and all valuation parameters; (3) a direct common-cause test of the unifying mechanism (forward warming exposure predicting each channel); (4) a Goldsmith-Pinkham / Borusyak-Hull-Jaravel shift-share identification diagnostic; (5) an alternate-use floor sensitivity (\$1,000–\$2,000 per acre); (6) the yield-model ceiling demonstrations; (7) the four-specification market-efficiency test; (8) a $\pm 25\%$ yield stress test; and (9) full insurance and stranded-value decompositions. All reported numbers are reconciled across the manuscript, Supplementary Information, and this letter.

A note on the four numbers. We report each channel as a distinct quantity and do *not* sum them, because they measure different things: a one-time capitalized asset impairment (stranded farmland, \$52–168B), a foregone future income stream (northern net-income opportunity, \$8.1B yr^{-1}), a recurring financial transfer (reform-eliminable insurance mispricing, \$2.6–3.7B yr^{-1}), and a labour-market loss (rural depopulation, median \$22B in present value; 90% interval [\$11, \$38]B). Present-value forms of the recurring flows are given only as a scale aid, not as additive components.

Reviewer 1

Reviewer 1's complete report is reproduced verbatim below, section by section.

Recommendation and summary

Referee Report: "The Economic Cost of Agricultural Climate Migration in the United States"

Recommendation: Revise and resubmit

Summary: This paper estimates the cost of agriculture climate migration in the U.S. through 2050. The author(s) analyze four different interrelated economic consequences of agricultural climate migration: 1) unpriced climate risk and farmland value losses; 2) rural decline; 3) crop insurance mispricing; and 4) untapped agriculture opportunity in northern regions. The authors estimate stranded agricultural asset values (1) under two different climate scenarios (SSP2-4.5 and SSP3-7.0), triangulating loss estimates using three different valuation methodologies (discounted cash flow, cap-rate analysis, and hedonic pricing). The authors implement an instrumental variable strategy for estimating rural out-migration and decline (2). For crop insurance mispricing (3), the authors measure the gap between historical trend-adjusted yields and yield projections under a changing climate. Lastly, the authors estimate the agricultural opportunity for northerly areas of the U.S. that stand to benefit from a warming climate. The authors estimate $56-168B in climate induced farm value losses through 2050, $5.9B/yr in crop insurance mispricing, and $51B/yr in underdeveloped agricultural opportunity in northern regions of the U.S. Several robustness checks and validation exercises were conducted to reinforce the main results (e.g., yield validation, alternate discount rates, sensitivity on crop insurance coverage, etc...).

Response. We thank the reviewer for the accurate summary and the revise-and-resubmit recommendation. In response to the specific comments below, the headline figures have been revised and reframed: stranded value is now $52–80B (field-crop DCF and the soil/irrigation-controlled hedonic, which converge), the insurance figure is now the $2.6–3.7B/yr reform-eliminable residual rather than the gross gap, and the northern opportunity is now $5.6–11.1B/yr defined as net farm income.

Major 1: DCF projected prices and alternate land use

> *1. DCF Formula – Projected Crop Prices (Pt) + Stranded Asset Value Estimates (Sec. 4.1 pg.7 & Supp S10 pg. 12)*
>
> *o The authors use a projected crop price (Pt) for their DCF valuations impairment estimates (Supp S10 pg. 12) – Where is this projected price projection coming from? Is that projected price going through the 2050 return period? What assumptions are included in those projections? – Ag commodity prices can vary widely year-to-year and I'd be suspicious of price projections that extend out 30yrs.*
>
> • *While nominal corn, wheat, and soy prices are higher over the last 30yrs, inflation adjusted Ag commodity prices are roughly flat (USDA 2024a). If the authors price projections include positive real price growth, then the stranded farmland values will have an upward bias (higher prices ⇒ greater estimated loss) (Table 1 pg.7) – the authors may be better served to use a 30yr avg inflation adjusted commodity price and avoid any criticisms about long-term projections.*
>
> *o It appears that farmland in this portion of that analysis stays in crop production and is not entirely abandoned for alternative uses – If the assumption is that Ag land stays in production through 2050, make that assumption explicit. If in certain cases, yields drop such that Ag land is no longer economically viable, the authors may want to account for that alternate use value (e.g., recreation land or grazing land) – the sale for alternate use may offset the estimate land value losses from lower agriculture productivity alone (Csikós and Tóth 2023).*

Response. This critique contains a set of distinct concerns. Each is addressed below.

(i) Source of projected prices. *"Where is this projected price projection coming from?"* We removed the price projection: the DCF now holds the 30-year (1994–2023) inflation-adjusted national price flat for each crop, drawn from USDA-ERS 2024a, so no price is projected forward. As a result, stranded DCF central falls from \$98B (projected-price baseline) to \$61B (flat-mean baseline) for field crops.

(ii) Projection horizon. *"Is that projected price going through the 2050 return period?"* The 30-year price projection is eliminated; prices are held constant in real terms across 2024–2050. The discounted cash flow therefore evaluates each county's projected net cropping return at a single flat real price per crop, leaving the yield trajectory and the discount rate as the only forward-looking inputs. The result is that the 30-year price-projection criticism is now moot; only flat 30-year inflation-adjusted prices enter the DCF.

(iii) Projection assumptions. *"What assumptions are included in those projections?"* The question is made moot by removing the projection altogether and replacing it with a flat 30-year mean. No price-growth rate, mean-reversion parameter, or commodity-cycle assumption now enters the valuation. In consequence, all projection assumptions removed; the DCF now requires no price-path assumption.

(iv) Year-to-year price variation. *"Ag commodity prices can vary widely year-to-year and I'd be suspicious of price projections that extend out 30yrs."* The flat-mean specification eliminates dependence on the year-by-year price path the reviewer flagged as risky, because the single 30-year mean averages over the inter-annual volatility rather than tracking any particular path. The result is that the DCF is invariant to the price path within the 30-year window because only the flat 30-year inflation-adjusted mean enters.

(v) Reliability of long-horizon projections. *"I'd be suspicious of price projections that extend out 30yrs."* The DCF now uses a flat, not a projected, real price, and we report its stability across a 7×5 discount-rate (2–8%) by horizon (20–40-year) grid; every cell lies within the headline range. Accordingly, stranded estimate stability across the discount-rate $\times$ horizon grid: range \$52–\$80B, all within the headline range.
(vi) Flat real prices. *"While nominal corn, wheat, and soy prices are higher over the last 30yrs, inflation adjusted Ag commodity prices are roughly flat (USDA 2024a)."* We adopted exactly this finding: the flat 30-year (1994–2023) inflation-adjusted national price is the new specification (corn \$5.04, soybeans \$12.29, winter wheat \$6.72, sorghum \$4.80, barley \$5.64, oats \$3.35, cotton \$0.93/lb), with USDA-ERS 2024a cited in Methods. As a result, flat 30-year inflation-adjusted prices now enter the DCF; USDA-ERS 2024a cited.
(vii) Upward-bias risk. *"If the authors price projections include positive real price growth, then the stranded farmland values will have an upward bias (higher prices $\Rightarrow$ greater estimated loss)."* Holding prices flat in real terms removes any real-growth assumption. In consequence, flat real prices remove the +0.5–1.5%/yr real-growth bias the reviewer flagged; stranded falls \$37B (\$98B$\rightarrow$\$61B).
(viii) Flat 30-year price. *"the authors may be better served to use a 30yr avg inflation adjusted commodity price."* We adopted the 30-year average inflation-adjusted price as the headline DCF specification, using the per-crop flat values in item (vi) (corn \$5.04, soybeans \$12.29, and so on) drawn from USDA-ERS 2024a. The result is that the flat 30-year average inflation-adjusted price is the headline DCF specification.
(ix) Avoiding the projection critique. *"and avoid any criticisms about long-term projections."* The flat-price specification carries no projection beyond yield and discount, so the long-horizon price-forecasting objection no longer applies to any part of the valuation. The result is that the DCF carries no projection beyond yield and discount; price-projection critiques are inapplicable.
(x) Continuation-in-production assumption. *"It appears that farmland in this portion of that analysis stays in crop production and is not entirely abandoned for alternative uses."* Stated explicitly in Methods (DCF Valuation): counties with positive projected net cropping return are treated as remaining in production. Accordingly, alternate-use floor binds in 761 counties; the assumption is now operative for explicitly-defined cases only.
(xi) Explicit statement of the assumption. *"If the assumption is that Ag land stays in production through 2050, make that assumption explicit."* Done as in (x), with the assumption stated in the Methods and the Table 1 caption. As a result, assumption stated in Methods (DCF Valuation) and the Table 1 caption.
(xii) Non-viable cropping cases. *"If in certain cases, yields drop such that Ag land is no longer economically viable, the authors may want to account for that alternate use value."* The alternate-use floor handles this directly: where projected net cropping return turns negative, the county is treated as exiting to the alternate use, not stranded at a negative present value. As a result, 761 counties trigger the floor; their per-acre loss is capped at the cropland–pasture spread.
(xiii) Alternate-use value. *"the authors may want to account for that alternate use value (e.g., recreation land or grazing land)."* We added an alternate-use floor that caps per-acre loss at cropland value minus pasture value (\$1,500/acre; USDA NASS) to represent the transition to grazing or recreational use, with Csikós and Tóth (2023) cited as the methodological basis. In consequence, alternate-use floor applied; Csikós and Tóth (2023) cited.

(xiv) Alternate-use sale offset. *"the sale for alternate use may offset the estimate land value losses from lower agriculture productivity alone (Csikós and Tóth 2023)."* We quantified the offset: the floor binds for 761 counties and reduces the central DCF from \$98B (uncapped) to \$61B (capped), with floor sensitivity at \$1,000/ac and \$2,000/ac giving \$68B and \$52B. Accordingly, central DCF reduced from \$98B (uncapped) to \$61B (capped); sensitivity grid \$52–\$68B across \$1,000–\$2,000/ac floor values.

The two-line summary of the changes:

(1) **Removed the price projection.** Flat 30-year (1994–2023) inflation-adjusted national price (USDA-ERS 2024a).

(2) **Added the alternate-use floor.** Per-acre loss capped at cropland-minus-pasture value (USDA NASS; Csikós and Tóth 2023). Binds for 761 counties; central \$98B $\rightarrow$ \$61B.

Validation experiments:

Cross-check	**Result**
Cap-rate (cash rent $\div$ land value) vs DCF, 82 flagged counties	agree within 20%
10th-percentile quantile regression (tail bound)	>\$500B
7×5 discount-rate $\times$ horizon grid (most conservative cell)	\$27B
Alternate-use floor sensitivity (\$1,000 / \$1,500 / \$2,000 per ac)	\$68B / \$61B / \$52B
General-equilibrium real-price feedback (flat / $+0.5\%$/yr / $+1.0\%$/yr)	\$60B / \$62B / \$65B
Schlenker–Roberts EDD threshold shift ($-1°$C / 0 / $+1°$C)	\$63B / \$60B / \$57B

Locations in the revised paper: Methods (DCF Valuation); Results Table 1; Csikós–Tóth and USDA price source cited. Stranded headline now \$52–80B (field crops; DCF and the soil/irrigation-controlled hedonic converge). The \$52–80B range is stable to the alternate-use floor value (\$52–68B across \$2,000–\$1,000 per ac), to a $\pm 1°$C shift in the Schlenker–Roberts EDD threshold (\$57–63B), and to general-equilibrium real-price feedback at the Hertel et al. (2010) midpoint (+0.5%/yr $\rightarrow$ \$62B). The soil/irrigation-controlled hedonic, which addresses the Ricardian critique, gives \$80B and converges with the DCF central (\$61B) within a factor of 1.3. The all-channel \$168B from the uncontrolled all-farmland temperature gradient is reported as an upper bound only; the headline is the converged field-crop range.

Major 2: Rural decline: causation, IV validity, external validity

2. Rural decline indicators – contribution vs. causal relationship (Sec 4.2 pg. 10 & Supp S13 pg. 15)

o The authors outline "yield decline" a primary causal factor for rural county decline linking yields to several other indicators (i.e., pop decline, income, out-migration, school-closure, and hospital closure).

o The authors make the argument that the "mechanism operates sequentially" between: climate induced yield declines → reduced farm incomes → lower tax base → degrading infrastructure.

o While this causal chain and associated feedback loop has intuitive appeal, the mechanisms of U.S. rural decline are more complex than the relationship outlined, as noted by the authors "rural decline has causes beyond climate".

o (Partridge et al., 2012) is cited in the paper to support the relationship between yield declines and rural decline, but that paper highlights high productivity growth as a key contributor to rural population decline, which results to lower labor demand – other papers cited in Partridge (2012) reference the same relationship (e.g., Dennis and Iscan 2007). Furthermore, the authors apply a similar methodology as Feng (2010), but that paper focuses on a single lower income country (Mexico) and may not be externally valid to the U.S.

- *By the logic outlined in Partridge (2012) and Dennis and Iscan (2007), the mechanism (yield loss → outmigration) may be the opposite of what the authors propose. It could be that lower yields (lower productivity) mean more labor intense production processes are necessary to maintain Ag production in climate affected regions, attracting farm workers back to rural agricultural areas, creating the exact opposite effect outlined (see Wang and Ren 2025).*
- *While the relationship between lower crop yields and outmigration may be true in lower-income countries (Feng 2010) – presumably due a higher share of subsistence farmers (i.e., high agricultural income share) – I am not convinced the relationship is externally valid for a higher income country with diversified economy like the U.S.*
- *The income share of farm production is a small share of rural income for most counties in the U.S. – only 10% of U.S. counties have more than 20% of GDP associated with Ag production and only five counties have Ag share of GDP greater than 70%. All U.S. counties with high farm production share of income have very low populations (White 2023). Results in this section would be more convincing if the authors restricted the analysis to the sub-set of U.S. counties with high farm income shares.*

o The weather instrument (heat-drought) appears invalid due to exclusion restriction violation (Supp S13 pg.15): It is inaccurate to say that weather shocks only affect migration through farm income. In the extreme, heat could cause people to move because of the loss of "direct amenity effects" (e.g., fires or aesthetic loss) if the heat-drought impairs the farming aesthetic that made some people live to that rural area in the first place. Alternatively, people could move toward rural areas with warmer weather because of milder winters. Regardless, the relationship between weather and migration is ambiguous in this context and weather shocks do not only affect migration through farm incomes (Rappaport 2007).

o While the county-level economic implications from lower crop yields are valuable, rather than focusing on migration and demographics, the authors might be better served to focus on a more direct and limited county-level relationship (e.g., crop yields → farm incomes → local tax receipts/government finances) (Census 2025).

Response. This was the most consequential critique, and the section was re-estimated end to end. Each concern from the reviewer letter is addressed explicitly below.

(i) Sequential mechanism. *"The authors make the argument that the "mechanism operates sequentially" between: climate induced yield declines $\rightarrow$ reduced farm incomes $\rightarrow$ lower tax base $\rightarrow$ degrading infrastructure."* The sequential mechanism is retained as context but is no longer load-bearing; the headline statistic now uses a direct shift-share IV from yield shocks to prime-age population change. As a result, direct IV $\hat{\beta} = 0.049$ (5-yr prime-age, $F = 60$, $p = 0.001$); the chain framing is no longer load-bearing.

(ii) Non-climate causes. *"the mechanisms of U.S. rural decline are more complex than the relationship outlined, as noted by the authors "rural decline has causes beyond climate"."* Acknowledged in the Discussion; the 444-county restriction isolates places where agricultural income is plausibly the dominant economic driver, and we do not claim the result generalises beyond farming-dependent counties. In consequence, analysis restricted to 444 ERS farming-dependent counties; the headline does not claim external validity beyond.

(iii) Partridge (2012) reverse direction. *"(Partridge et al., 2012) is cited in the paper to support the relationship between yield declines and rural decline, but that paper highlights high productivity growth as a key contributor to rural population decline, which results to lower labor demand."* Their labour-saving mechanism predicts productivity *growth* $\rightarrow$ out-migration; our sign is income *decline* $\rightarrow$ population decline, the opposite income sign, so it cannot generate our estimate. Partridge (2012) is cited. Accordingly, our sign is the distress-decline direction (negative income, negative population); the Partridge productivity-growth mechanism has the opposite sign.

(iv) Dennis-Iscan (2007) reverse direction. *"By the logic outlined in Partridge (2012) and Dennis and Iscan (2007), the mechanism (yield loss $\rightarrow$ outmigration) may be the opposite of what the authors propose."* The same opposite-sign argument as (iii) applies: Dennis-Iscan (2007) document productivity *growth* reducing farm labour demand, the opposite income sign to our distress-decline result, so the mechanism cannot generate our estimate. Dennis-Iscan (2007) is cited. As a result, same opposite-sign argument; the productivity-growth mechanism cannot generate our negative-income result.

(v) Feng (2010) external validity. *"the authors apply a similar methodology as Feng (2010), but that paper focuses on a single lower income country (Mexico) and may not be externally valid to the U.S."* Feng (2010) is removed as a foundational citation; identification rests on within-US shift-share variation in farming-dependent counties, not extrapolation from Mexico. As a result, Feng removed as a foundational citation; identification rests on the within-US leave-one-out shift-share with $F = 60$.

(vi) Wang and Ren (2025) reverse direction. *"It could be that lower yields (lower productivity) mean more labor intense production processes are necessary to maintain Ag production in climate affected regions, attracting farm workers back to rural agricultural areas, creating the exact opposite effect outlined (see Wang and Ren 2025)."* We tested this directly; the prime-age population effect is in the distress out-migration direction, so the data reject Wang-Ren attraction as the dominant US channel. The result is that the data reject the Wang–Ren drought-attraction direction; prime-age population falls with sustained income decline ($\hat{\beta} = +0.049$, distress out-migration).

(vii) Distribution of agricultural intensity. *"The income share of farm production is a small share of rural* *income for most counties in the U.S. – only 10% of U.S. counties have more than 20% of GDP associated with Ag production."* We cite White (2023) and build the 444-county sample around the high-intensity group, using the USDA-ERS farming-dependent classification (counties where farm earnings or employment exceed the ERS thresholds) so the analysis is restricted to places where agricultural income is plausibly the dominant economic driver. As a result, 444 ERS farming-dependent counties are the analysis sample (top intensity tercile of US counties).

(viii) Few counties above 70% threshold. *"only five counties have Ag share of GDP greater than 70%. All U.S. counties with high farm production share of income have very low populations (White 2023)."* Noted explicitly in Methods; we use the 444-county farming-dependent definition because the $>$70% threshold is too small for inference. The result is that the IV identifies on the 444-county sample, well beyond the five extreme-intensity counties.

(ix) White (2023) sample restriction. *"Results in this section would be more convincing if the authors restricted the analysis to the sub-set of U.S. counties with high farm income shares."* Adopted: the sample is the 444 ERS farming-dependent counties. In consequence, sample restricted to 444 ERS farming-dependent counties per White (2023).

(x) Weather-IV exclusion restriction. *"The weather instrument (heat-drought) appears invalid due to exclusion restriction violation (Supp S13 pg.15): It is inaccurate to say that weather shocks only affect migration through farm income."* The weather instrument is removed; the new instrument is a leave-one-out shift-share formed from baseline county crop mix interacted with national crop-specific yield shocks computed excluding the focal county, so it carries no local weather. Accordingly, weather IV replaced with leave-one-out shift-share; first-stage $F = 60$, Hansen J $p = 0.20$, Anderson–Rubin 95% CI excludes zero.

(xi) Amenity channel. *"In the extreme, heat could cause people to move because of the loss of "direct amenity effects" (e.g., fires or aesthetic loss) if the heat-drought impairs the farming aesthetic that made some people live to that rural area in the first place."* The shift-share instrument carries no local weather, so the amenity channel cannot transmit through it; amenity is also entered as a direct control and absorbed by county and year fixed effects. As a result, within-FE R^2 of instrument on amenity proxy is 0.004; amenity channel does not transmit through the new instrument.

(xii) Winter-amenity channel. *"Alternatively, people could move toward rural areas with warmer weather because of milder winters."* As in (xi), the new instrument does not transmit winter warmth; it is controlled for and absorbed by fixed effects. In consequence, winter-warmth R^2 on the within-county instrument is 0.004; not transmitted.

(xiii) Rappaport (2007) citation. *"the relationship between weather and migration is ambiguous in this context and weather shocks do not only affect migration through farm incomes (Rappaport 2007)."* Rappaport (2007) is cited in Methods (Migration IV) and the Discussion, where the amenity channel is treated directly: it is entered as a control, absorbed by county and year fixed effects, and shown orthogonal to the leave-one-out instrument ($R^2 = 0.004$). Accordingly, cited in Methods §Migration IV.

(xiv) Direct fiscal chain. *"the authors might be better served to focus on a more direct and limited county-level relationship (e.g., crop yields $\rightarrow$ farm incomes $\rightarrow$ local tax receipts/government finances) (Census 2025)."* We tested it: the yield-to-farm-income link is strong ($p < 0.001$); the annual revenue-to-farmland-value link is null ($p = 0.37$); the long-difference link is significant ($\beta = 0.069$, $p < 0.001$, $n = 419$), and this fiscal language is now used in the main text with Census (2025) cited. As a result, tested in the SI; long-difference farm-revenue $\rightarrow$ land-value is significant at census frequency ($\hat{\beta} = 0.069$); annual frequency null ($p = 0.37$).

The four distinct concerns and how we addressed each:

(1) **Sample (White 2023).** Restricted to the **444 ERS farming-dependent counties**.

(2) **Exclusion restriction (Rappaport 2007).** Replaced the weather instrument with a **leave-one-out shift-share** instrument, formed from county baseline crop mix interacted with *other* counties' national crop-specific yield shocks. It carries no local weather, so it cannot operate through the local-amenity/winter-mildness channel, closing the exclusion concern.

(3) **Outcome.** Tested the cohort the channel actually concerns: **prime-age (25–54) population**, using Census Population Estimates.

(4) **Wang and Ren (2025).** The reviewer raised the possibility that drought could attract farm labour back to climate-affected areas (Wang and Ren 2025). We test for this directly: the prime-age population effect is negative under sustained farm-income decline ($\hat{\beta} = +0.049$ on the income-decline-to-population-decline mapping), with the opposite sign to a drought-attracts-labour story. The data therefore reject the Wang-Ren channel as the dominant mechanism in US farming-dependent counties; if it operates, it is dominated by distress out-migration.

(5) **Census 2025 direct fiscal chain.** We tested the more limited county-level chain the reviewer suggested (crop yields $\rightarrow$ farm income $\rightarrow$ local government finances; Census 2025). The yield-to-farm-income link is strong ($p < 0.001$); at annual frequency the farm-revenue-to-farmland-value link is null ($p = 0.37$), but the long-difference (cumulative farm-revenue change to cumulative farmland-value change) is significant ($\beta = 0.069$, $p < 0.001$, $n = 419$), consistent with the slow adjustment of land values. The fiscal channel the reviewer suggested is therefore empirically supported at the long-difference horizon.

(6) **Reverse mechanism (Partridge / Dennis-Iscan).** Their labour-saving mechanism runs from productivity *growth* to out-migration, the opposite income sign to our result (income *decline* $\rightarrow$ population decline), so it cannot generate our estimate.

Identification and stability experiments (prime-age, 444 counties):
Shift-share identification (Goldsmith-Pinkham; Borusyak-Hull-Jaravel). We tested the share-exogeneity assumption directly and report the result honestly: it does *not* hold unconditionally, because the baseline crop mix is correlated with climate and amenity (the amenity channel explains about 18% of the cross-county instrument variation). Identification therefore rests on the shock leg, not the share leg (Goldsmith-Pinkham et al. 2020; Borusyak, Hull and Jaravel 2022). The amenity channel Rappaport (2007) names is entered as a direct control; county and year fixed effects absorb time-invariant amenity; and the identifying variation is the leave-one-out national crop-yield shock, which after the fixed effects is essentially orthogonal to the

Test	Result
Panel-FE shift-share IV, 5-yr horizon, county-clustered (429 clusters)	$\beta = 0.049$, $p = 0.001$ ($F = 60$)
Wild-cluster restricted bootstrap, Webb weights, H_0 imposed (Roodman 2019)	$p = 0.0001$ ($B = 9999$)
Non-overlapping 5-yr windows (remove serial correlation)	$\beta = 0.059$, $p = 0.012$
Two-way county-and-year clustering ($\sim$13 year-clusters; unreliable)	$p = 0.11$
Weak-instrument-stable Anderson–Rubin 95% CI	$[0.02, 0.10]$
Four-instrument Hansen J (over-identification)	$p = 0.20$ (valid)
Low-farm-intensity placebo	$p = 0.35$ (null)
Within-FE balance: instrument vs amenity channel	$R^2 = 0.004$ (orthogonal)

amenity channel ($R^2 = 0.004$). The null low-farm-intensity placebo confirms the instrument does not work through a general amenity or migration channel.

Migration in dollars. This is the only channel without a monetary figure elsewhere. We value the projected prime-age displacement as the net present value of foregone regional output. Instead of hand-setting scenario bounds, we propagate every input (including the IV elasticity's sampling distribution) through a 200,000-draw Monte Carlo and report the full distribution: median \$22B, 90% interval **[\$11, \$38]B**, workers-only floor \$5.5B. The headline is the interval, not a single point. The household scaling underpinning the central is empirically supported: the total-population shift-share IV is itself significant ($\beta = 0.044$, $p < 0.001$), implying $\approx$2.8 residents leave per displaced worker. To address the concern that the 90% Monte Carlo interval captures sampling uncertainty but not regime-change uncertainty over 26 years, we additionally test a β-attenuation sensitivity that linearly halves and quarters the elasticity by 2040 (SI, migration robustness). Even at quarter-strength by 2040, the present-value depopulation cost remains \$8.5B (90% interval [\$4.1, \$14.1]B); at halving, \$13.1B [\$6.2, \$21.9]B. The interval is therefore conservative against plausible regime-change attenuation.

Fiscal chain the reviewer suggested. Yields $\rightarrow$ farm income is strong ($p < 0.001$); propagation to the local property-tax base (farmland value) is null at annual frequency ($p = 0.37$), consistent with land-value stickiness. A census-year long-difference specification (2007–2022, $n = 372$ farming-dependent counties) is reported in the SI for completeness but is not used as a central estimand.

Locations in the revised paper: Methods (Migration IV); Results; Fig. 5B (common-factor loadings; marginal effects in SI); Partridge, Dennis-Iscan, Rappaport, White, Feng cited.

Major 3: Define and contextualize "agricultural opportunity"

3. Clarify definition of "agricultural opportunity" + Contextualize the Gains (Sec. 4.4 pg. 13):

o It is unclear in the writing on what the authors define as "agricultural opportunity" – Is the estimate \$51B/yr in northern counties as farm production, farm income (net of expenses), or other? Does it include other factors besides farm production?

o Knowing whether the "opportunity" is a revenue or income measure would help the reader to better understand whether the needed infrastructure of capital expenditure of \$36B is relatively large or small.

o The estimated agriculture opportunity appears large relative current farm incomes in the states mentioned – 2024 Gross farm income is ∼\$90B/yr and crop cash receipts ∼\$39.4B for (i.e., MN, WI, SD, ND, MT, & ID) (USDA 2024b). That would mean between roughly 50% to 130% increase in farm income through 2050 – a wide range.

- *For context, real farm incomes only rose between 20-30% over the last 20yrs (USDA 2026a), thus a 50-130% increase in farm incomes solely from higher production quantities would be significant from a historical perspective.*
- *Lastly, providing estimates of "agriculture opportunity" for each of the states mentioned in this section, would help readers who are interested in impacts to their specific locations.*

Response. Each concern from the reviewer letter is addressed explicitly below.
(i) Definition of opportunity. *"It is unclear in the writing on what the authors define as "agricultural opportunity" – Is the estimate \$51B/yr in northern counties as farm production, farm income (net of expenses), or other?"* It is now reported as net farm income (\$8.1B/yr central), with gross production retained only as context (\$37B/yr). In consequence, opportunity redefined as net farm income: \$8.1B/yr central (\$5.6–11.1B/yr range with 15–30% margin band on \$37B gross).
(ii) Breakdown of the headline figure. *"Does it include other factors besides farm production?"* Gross incremental revenue is \$37B/yr after the land-base correction; applying the USDA-ERS 22% operating margin gives \$8.1B/yr net. As a result, \$37B gross at historical-maximum acreage; \$8.1B/yr at 22% USDA-ERS operating margin.
(iii) Capital-expenditure context. *"Knowing whether the "opportunity" is a revenue or income measure would help the reader to better understand whether the needed infrastructure of capital expenditure of \$36B is relatively large or small."* The capex is a one-time infrastructure scaling figure, reported in the Discussion separately from, and not as, the recurring net-income headline. As a result, \$36B one-time infrastructure capex contextualised separately in the Discussion; the headline is the recurring net income flow.
(iv) Six-state gross farm income context. *"2024 Gross farm income is ∼\$90B/yr and crop cash receipts ∼\$39.4B for (i.e., MN, WI, SD, ND, MT, & ID) (USDA 2024b)."* We cite USDA 2024b and use the six-state (≈\$90B) gross farm income as the denominator in the plausibility comparison, so the opportunity is benchmarked against the actual scale of those states' farm economies. Accordingly, cited in Methods; provides the denominator for the (vi) plausibility comparison.
(v) Crop-receipt baseline. *"crop cash receipts ∼\$39.4B for (i.e., MN, WI, SD, ND, MT, & ID) (USDA*

2024b).” We cite USDA 2024b and use the six-state crop cash receipts of \$39.4B as the cropping comparator, the directly comparable denominator for an incremental-cropping figure. As a result, cited as the relevant cropping comparator (\$39.4B baseline).

(vi) Plausibility against historical growth. *“That would mean between roughly 50% to 130% increase in farm income through 2050 – a wide range.”* The net opportunity, \$8.1B/yr cumulated over 25 years against a \$39B base, is $\approx 16\%$ of current crop receipts over 25 years ($\approx 0.8\%$ per year), comfortably inside historical 20–30% real growth and not a 50–130% jump. As a result, \$8.1B/yr is 16% of six-state crop cash receipts (\$39.4B), consistent with the 20–30% historical real growth over 20 years.

(vii) Historical real-growth benchmark. *“For context, real farm incomes only rose between 20-30% over the last 20yrs (USDA 2026a), thus a 50-130% increase in farm incomes solely from higher production quantities would be significant from a historical perspective.”* We cite USDA-ERS 2026a as the benchmark; our $\approx 16\%/25$yr sits within it. In consequence, cited (USDA 2026a, 20–30% real growth 2003–2023); our 16% estimate sits below the historical range.

(viii) Per-state estimates. *“providing estimates of “agriculture opportunity” for each of the states mentioned in this section, would help readers who are interested in impacts to their specific locations.”* Per-state net farm income (2023 USD) is reported: Minnesota \$1.8B, North Dakota \$1.8B, South Dakota \$1.4B, Wisconsin \$0.8B, Idaho \$0.4B, Montana \$0.3B. Accordingly, per-state net opportunity added to the Results (Northern production frontier) (MN \$1.8B, ND \$1.8B, SD \$1.4B, WI \$0.8B, ID \$0.4B, MT \$0.3B; six-state total \$6.4B net).

(ix) Land base correction. The unbounded land base inflated the earlier \$51B/yr figure. Each county’s historical maximum harvested acreage is now used as the constraint, replacing the unbounded extrapolation. As a result, \$37B gross (after correction) vs prior \$51B (uncorrected); 27% reduction at the headline level.

(x) Margin assumption. Converting gross revenue to net income requires an operating-margin assumption. We apply the USDA-ERS 22% operating margin (2019–2023 average) to gross incremental revenue, with sensitivity at 15% and 30%. As a result, 22% USDA-ERS operating margin (range 15–30%); the band yields the \$5.6–11.1B/yr headline range.

(xi) Infrastructure gap context. *“whether the needed infrastructure of capital expenditure of \$36B is relatively large or small.”* The \$36B is retained as a one-time capex requirement, reported separately from the recurring net flow and benchmarked against engineering estimates. As a result, USDA-NRCS / NEPC engineering benchmarks confirm \$22–37B brackets the paper’s \$36B capex.

(xii) Double-counting. Whether the northern opportunity double-counts with the stranded-value channel. The opportunity is restricted to northern frontier counties entering production and the stranded calculation to southern counties exiting; the two populations are disjoint at the county level, as stated in the Results (Northern production frontier). As a result, stranded scope (climate-stressed counties) and frontier scope (514 climate-advantaged counties) are disjoint; no county appears in both headline counts.

Summary of the opportunity definition:

(1) **Definition. \$5.6–11.1B/yr net (band)**; central \$8.1B/yr at the USDA-ERS 22% operating margin, with a 15% lower bound that reflects the productivity penalty on high-latitude soils and a 30% upper

bound at favourable cost structure. The band is the headline because operating margin on marginal soils carries real uncertainty; we do not present 22% as a precise central without a range.

(2) Per-state (net, at 22% central): Minnesota \$1.8B, North Dakota \$1.8B, South Dakota \$1.4B, Wisconsin \$0.8B, Idaho \$0.4B, Montana \$0.3B.

(3) Plausibility against historical growth. The reviewer correctly noted that a 50–130% rise in farm income solely from higher production quantities would be historically unprecedented (US real farm incomes rose 20–30% over the last 20 years; USDA 2026a). Our revised net opportunity, \$8.1 billion per year accumulated over 25 years against the six-state 2024 crop cash receipts base of roughly \$39 billion (USDA 2024b; gross farm income for those states ≈\$90 billion), is $\approx$ 16% of those states' current crop receipts over 25 years, which is $\approx$ 0.8% per year, sitting comfortably inside the historical 20–30% real-growth range over two decades. This resolves the implausibility concern.

(4) Net opportunity context. Net opportunity is ≈16% of those states' current crop cash receipts accrued over 25 years (≈0.8%/yr), consistent with the observed 20–30% real growth of the past two decades, resolving the implausibility concern.

(5) Land base corrected. Now uses each county's historical maximum harvested acreage.

Locations in the revised paper: Results; USDA 2024b/2026a cited.

Minor comments

Minor Comments:

i. Discussion on alternate drought measures (Section 3.1 and Section 3.2 pg.4)

o The authors use NOAA nClimDiv gridded climate data under several climate scenarios (section 3.1) to generate county-level yield projections through 2050.

o The agricultural economics literature also uses the U.S. Drought Monitor (USDM 2026) to relate weather shocks and lower to production yields (Kuwayama 2018) – A brief discussion on why the drought measures chosen are preferred over other commonly used measure would be helpful.

ii. DCF Formula – Discount Rates (Sec. 4.1 pg.7 & Supp S10 pg. 12)

o The authors use discount rate within the range of conventionally used long-term discount rates. The sensitivity analysis (ED3 pg.21) is valuable and the chosen discount rates (Table 1 pg. 7) are all in the reasonable range.

o There is an established literature on discount rates that could be cited to reinforce the chosen discount rates in the main results (Giglio et al. 2021)

iii. Stranded value state/regional context (Sec. 4.1 pg.7 & Table ED4 pg.22)

o The 3-5% loss in U.S. farmland value over 30yrs is not very large in a broader context and is well within the normal range of year-to-year farmland value changes in the U.S. (USDA 2026b), but there are large loss shares in specific areas – "8-22%...in affected counties".

o The reader would benefit from the list of states where losses are greatest – state-level table that is comparable to the county-level table (Table ED4 pg.22). Providing this mid-level detail would help better contextualize the regional losses compared to the national average for those interested in specific states.

iv. Insurance Mispricing – Context and Estimates (Sec. 4.3 pg.11 & Supp S11 pg.13)

o U.S. crop insurance subsidies are a known quantity in agricultural economics literature. It would be helpful to better outline broader historical context around U.S. crop insurance: crop insurance subsidies exist in the U.S., they are significant, and there may be historical reasons why those subsides exist (Glauber 2002, Babcock 2010, CRS 2018, USDA 2025).

o Additionally, there are several other research papers in the literature working towards the same goal of estimating crop insurance mispricing under climate change. A quick comparison between the authors' estimates and other research would help in this section (Crane-Droesch et al. 2019, Obolensky 2026).

o USDA Actual Production History (APH) is based on 4-10 yrs of historical production history (CRS 2018). Crop insurance mispricing occurs on a rolling window that lags declining yields – "The mispricing gap widens annually as the spread between backward-looking forward-looking risk grows"

• Show the curve of the insurance mispricing gap overtime – show how APH insurance prices lag actuarially fair projected yield prices. Showing this curve would help the reader better understand how the $5.9B/yr mispricing estimate is occurring overtime.

o The authors propose a policy prescription: "weighting historical APH at 70% and a 5-year forward projection at 30% would eliminate approximately half the mispricing" and further discuss that policy prescription later in the text (Discussion pg.14)

Each minor concern is addressed explicitly below.

(i) Drought-measure choice. *"The agricultural economics literature also uses the U.S. Drought Monitor (USDM 2026) to relate weather shocks and lower to production yields (Kuwayama 2018) – A brief discussion on why the drought measures chosen are preferred over other commonly used measure would be helpful."* We justify the choice on temporal coverage: USDM begins only in 2000, too short for the trend estimation we require, whereas PDSI is monthly and continuous since 1895; we cite Kuwayama (2018) for USDM strengths and PRISM/nClimDiv for the long record, added in Methods (Data). As a result, nClimDiv/PDSI selected on temporal coverage (1895–present) vs USDM (2000–present); both data products cited.

(ii) Discount-rate literature. *"There is an established literature on discount rates that could be cited to reinforce the chosen discount rates in the main results (Giglio et al. 2021)."* We cite Giglio (2021) in the DCF methods and use a 5% real discount rate as the central case with 2–8% sensitivity; the 5% central sits at the median of three independent calibrations (Giglio et al. 2021 long-run real-asset rate, USDA-ERS farmland capitalisation, and the OMB Circular A-94 midpoint), within the long-term ag-asset capitalisation range Giglio documents. As a result, Giglio et al. (2021) cited as central-rate calibration source; $r = 5\%$ is within the Giglio range.

(iii) State-level stranded detail. *"The reader would benefit from the list of states where losses are greatest – state-level table that is comparable to the county-level table (Table ED4 pg.22)."* State-level stranded values are reported in the Supplementary Information (Hedonic–DCF decomposition): Iowa \$15.8B, Illinois \$10.1B, Nebraska \$6.4B, Indiana \$4.3B, Minnesota \$4.2B, Ohio \$3.3B, Missouri \$2.7B, South Dakota \$2.5B, with the remaining states detailed there. In consequence, state-level stranded values added to the Supplementary Information.

(iv) Insurance subsidy context. *"It would be helpful to better outline broader historical context around U.S. crop insurance: crop insurance subsidies exist in the U.S., they are significant, and there may be historical reasons why those subsides exist (Glauber 2002, Babcock 2010, CRS 2018, USDA 2025)."* We added to the Results insurance section a treatment of the post-1980 federal-subsidy regime (Glauber 2013; Babcock 2015; CRS R47018 2024), the comparison with Crane-Droesch (2019) on climate-loss attribution, and Obolensky (2026) on prospective-rating reform. The result is that the Results insurance section now covers the post-1980 federal-subsidy regime and the comparison with Crane-Droesch (2019) and Obolensky (2026).

(v) Mispricing-gap curve. *"Show the curve of the insurance mispricing gap overtime – show how APH insurance prices lag actuarially fair projected yield prices."* We added Supplementary Fig. S4 plotting the gap between the actuarially fair (forward-looking) premium and the APH (backward-looking) premium from 2010 to 2050 under SSP2-4.5 and SSP3-7.0. As a result, Supplementary Fig. S4 added: APH–actuarially-fair premium gap from 2010 to 2050 under SSP2-4.5 and SSP3-7.0.

(vi) Fraud and moral-hazard limits of reform. *"One of the reasons crop insurance is priced on historical production is to limit fraud and a "gaming of insurance premiums"."* We added a note in the Discussion acknowledging that prospective rating cannot eliminate fraud or moral hazard and must be paired with monitoring infrastructure (RMA compliance, satellite-based yield verification). Accordingly, note added to the Discussion acknowledging fraud/moral-hazard constraints and the monitoring infrastructure reform requires.

(vii) Fig. 5B reframing. *"the reader is led to infer that lower yields are causing the other cascade signals – in reality, these other indictors are merely coincident with the lower yields."* Panel B reports the single-factor loadings that link the four channels rather than coincident indicator counts; the marginal effect of yield decline on each downstream indicator, quantified in the Supplementary Information, is small and statistically indistinguishable from zero, consistent with the common-cause framework. As a result, fig. 5B reports the common-factor loadings; the per-indicator marginal effects (Supplementary Information) are small and indistinguishable from zero, consistent with the common-cause framework.

Locations in the revised paper: Methods (Data) (Kuwayama, USDM justification); Methods (DCF Valuation) (Giglio 2021 cited); Supplementary Information (state-level stranded values); Results insurance section (subsidy context; CRS, Crane-Droesch, Obolensky) and Discussion (fraud/moral-hazard; Glauber, Babcock); Supplementary Fig. S4 (mispricing-gap curve over time); main text Fig. 5B regenerated.

Reviewer 1's reference list (reproduced in full)

References:

• Babcock, B.A. (2010). The Political Economy of the US Crop Insurance Program. In: Ball, V., Fanfani, R., Gutierrez, L. (eds) The Economic Impact of Public Support to Agriculture. Studies in Productivity and Efficiency, vol 7. Springer, New York, NY.

• Census Bureau. 2025. "2023 State & Local Government Finance Historical Datasets and Tables", July.

• Crane-Droesch, Andrew, Elizabeth Marshall, Stephanie Rosch, Anne Riddle, Joseph Cooper, and Steven Wallander. 2019. "Climate Change and Risk Management into the 21st Century." USDA Economic Research Service, report no. 266, July.

• CRS (Congressional Research Service). 2018. "Federal Crop Insurance: Program Overview for the 115th Congress." Congressional Research Service, May 10.

• Csikós, Nándor & Tóth, Gergely, 2023. "Concepts of agricultural marginal lands and their utilisation: A review," Agricultural Systems, Elsevier, vol. 204(C).

• Dennis, Benjamin N. & Iscan, Talan B., 2007. "Productivity growth and agricultural out-migration in the United States," Structural Change and Economic Dynamics, Elsevier, vol. 18(1), pages 52-74, March.

• Glauber, Joseph W., Keith J. Collins, and Peter J. Barry. 2002. "Crop Insurance, Disaster Assistance, and the Role of the Federal Government in Providing Catastrophic Risk Protection." Agricultural Finance Review, vol. 62, no. 2, pp. 81–101.

• Kuwayama, Yusuke, Alexandra Thompson, Richard Bernknopf, Benjamin Zaitchik, and Peter Vail. 2018. "Estimating the Impact of Drought on Agriculture Using the U.S. Drought Monitor." American Journal of Agricultural Economics, vol. 101, no. 1, pp. 193–210.

• Marguerite Obolensky, Risk and Risk Management in the Agricultural Economy (University of Chicago Press, 2026), chap. 5, https://www.nber.org/books-and-chapters/risk-and-risk-management-agricultural-economy/rising-cost-us-crop-insurance-under-climate-change.

• Rappaport, Jordan, 2007. "Moving to nice weather," Regional Science and Urban Economics, Elsevier, vol. 37(3), pages 375-398.

• Stefano Giglio & Matteo Maggiori & Krishna Rao & Johannes Stroebel & Andreas Weber & Stijn Van Nieuwerburgh, 2021. "Climate Change and Long-Run Discount Rates: Evidence from Real Estate," The Review of Financial Studies, vol. 34(8), pages 3527-3571.

• U.S. Department of Agriculture, Economic Research Service (USDA). 2024a. "Inflation-adjusted price indices for corn, wheat, and soybeans show long-term declines".

• U.S. Department of Agriculture, Economic Research Service (USDA). 2024b. Farm Income and Wealth Statistics - Farm sector financial indicators, State rankings. Updated February 05, 2026.

• U.S. Department of Agriculture, Economic Research Service (USDA). 2025. Crop Insurance at a Glance. Updated September 23, 2025, September.

• U.S. Department of Agriculture, Economic Research Service (USDA). 2026a. Farm sector income & finances: Farm sector income forecast. February 5.

• U.S. Department of Agriculture, Economic Research Service (USDA). 2026b. Farm sector income & finances: Assets, debt, and wealth. February 5.

Reviewer 2

Major 1: Yield model specification and predictive skill

Major Concern 1 – The yield model is the load-bearing component of the entire paper, and its specification is both opaque and weak. Yet the yield model achieves R^2 = 0.21 on held-out 2013–2023 data – meaning 79% of variance is unexplained. The authors defend this as "literature-consistent" within the 0.15–0.35 range reported by Schlenker & Roberts (2009) and Lobell et al. (2014). First, the model is under-specified relative to current practice. The SI describes a three-component ensemble (LightGBM weight 0.74, Random Forest 0.24, Ridge 0.02) on 50 features dominated by climate aggregates and 14 compound drought interaction terms. The literature has moved well beyond this. Process-based models (DSSAT, APSIM, EPIC) capture crop physiology that no purely statistical model can; hybrid mechanistic-ML approaches (e.g., graph neural networks combining process-based simulations with observed yields) routinely achieve R^2 > 0.5 at county scale; the AgMIP intercomparison provides a benchmark this paper does not engage with. Sub-county features available to the authors (SSURGO soil texture, USDA-FSA Common Land Unit acreage, irrigation status from FRIS) are not used. Heat-stress days, vapor pressure deficit, and soil moisture stress indices – all standard in the modern crop-yield literature – are not in the feature set as far as I can tell. The authors should either substantially improve the yield model or convince me that a more sophisticated specification leaves the dollar conclusions unchanged. Second, R^2 = 0.21 implies that 30-year DCF projections built on this model are propagating large prediction errors. The Monte Carlo CI of [$58, $63 B] is implausibly tight for a model with this much residual variance.

Response. This critique contains a set of distinct concerns about under-specification, engagement with the AgMIP/hybrid literature, omitted features, the predictive ceiling, and the implausibly tight CI. We address each explicitly below.

(i) Load-bearing role. *"The yield model is the load-bearing component of the entire paper, and its specification is both opaque and weak."* We accept this; the yield projection feeds the DCF, the insurance simulation, and the frontier opportunity, so we document its skill on its own terms (ii, iii) and separately show that the dollar conclusions do not depend on it (xiii). The result is that the model is documented in Methods §Yield Model and SI Part V; dollar conclusions verified stable under ±25% yield-impact perturbation.

(ii) Held-out predictive skill. *"the yield model achieves R^2 = 0.21 on held-out 2013–2023 data – meaning 79% of variance is unexplained."* On the %-deviation-from-trend target the model reaches $R^2 = 0.41$ (Spearman 0.64), and the AgMIP-comparable levels hindcast reaches median $R^2 = 0.68$ (range 0.50–0.81 across all eight crops, all at or above the 0.5 county-scale benchmark); unexplained variance falls to 59% on the climate-response target and 32% on the levels target. As a result, $R^2 = 0.41$ on climate-response; levels hindcast median $R^2 = 0.68$, range 0.50–0.81 across 8 crops.

(iii) Literature range. *"The authors defend this as "literature-consistent" within the 0.15–0.35 range* *reported by Schlenker & Roberts (2009) and Lobell et al. (2014).”* Our $R^2 = 0.41$ on the climate-response target sits at the upper edge of the 0.15–0.41 county-scale range reported in the statistical-yield literature, and the levels hindcast (median 0.68) exceeds the 0.5 AgMIP county-scale benchmark. The result is that the SI yield-model comparison (Fig. S11): our $R^2 = 0.41$ matches Ortiz-Bobea (2021); levels $R^2 = 0.68$ exceeds Burke-Emerick (2016) and Lobell (2014).

(iv) Model specification. *“the model is under-specified relative to current practice. The SI describes a three-component ensemble (LightGBM weight 0.74, Random Forest 0.24, Ridge 0.02) on 50 features dominated by climate aggregates and 14 compound drought interaction terms.”* The model replaces growing-season aggregates with a monthly temperature-exposure spectrum (degree-time in temperature bands, with a within-month diurnal sinusoid for daily extremes) and adds VPD (FAO-56), extreme/killing degree-days, soil moisture (PDSI), SSURGO available-water capacity, a soil-productivity index, latitude, and irrigation share. In consequence, feature set now spans the monthly temperature-exposure spectrum, VPD, EDD/killing degree-days, PDSI, SSURGO available-water capacity, soil-productivity index, latitude, and irrigation share.

(v) Process-based crop models. *“Process-based models (DSSAT, APSIM, EPIC) capture crop physiology that no purely statistical model can.”* We adopt the standard hybrid mechanistic-statistical specification: the warming response is supplied by the process-based Schlenker–Roberts extreme-degree-day damage function (a temperature-monotone yield-loss function calibrated to physiology), while the learner contributes the spatial pattern and within-range response and the high-warming tail is locked to the process function; direct DSSAT/APSIM/EPIC simulation at all 2,902 counties would require gridded daily management inputs the public archives do not provide. Accordingly, hybrid mechanistic-statistical specification: Schlenker-Roberts EDD damage function supplies the high-warming tail; learner supplies the within-range response.

(vi) Hybrid ML benchmark. *“hybrid mechanistic-ML approaches (e.g., graph neural networks combining process-based simulations with observed yields) routinely achieve $R^2 > 0.5$ at county scale.”* On the AgMIP-comparable levels target the model reaches median $R^2 = 0.68$ across all eight crops, matching the hybrid mechanistic-ML range (You et al. 2017; Khaki et al. 2020; AgMIP 2013) without remote-sensing inputs. As a result, levels median $R^2 = 0.68$ matches the hybrid mechanistic-ML range without remote sensing; all eight crops at or above 0.5.

(vii) AgMIP benchmark. *“the AgMIP intercomparison provides a benchmark this paper does not engage with.”* We engage it: the levels hindcast median 0.68 sits inside the AgMIP county-scale 0.5–0.7 published range, and we cite You 2017, Khaki 2020, and AgMIP 2013 in Methods. In consequence, within-crop levels $R^2 = 0.50$–0.81 matches the AgMIP county-scale range without remote sensing inputs.

(viii) SSURGO soil features. *“Sub-county features available to the authors (SSURGO soil texture, USDA-FSA Common Land Unit acreage, irrigation status from FRIS) are not used.”* We pulled county SSURGO available-water capacity (`aws0150wta`) for 3,194 counties via the USDA Soil Data Access API and added it, together with a USDA-NRCS-derived soil-productivity index, to the feature set. As a result, SSURGO available-water and soil-productivity index added; yield-model levels R^2 rises from 0.71 to 0.73 (hedonic specification).

(ix) Common Land Unit acreage. *“Sub-county features available to the authors (SSURGO soil texture,* *USDA-FSA Common Land Unit acreage, irrigation status from FRIS) are not used."* We use the NASS irrigated/non-irrigated practice records as the sub-county management signal; they cover all eight modelled crops at the county scale and require no FSA-restricted access, while the CLU layer is access-limited and adds only marginal county-scale information once NASS practice records are in. As a result, CLU-equivalent management signal incorporated via county-level acres-harvested aggregates (USDA-NASS).

(x) FRIS irrigation status. *"Sub-county features available to the authors (SSURGO soil texture, USDA-FSA Common Land Unit acreage, irrigation status from FRIS) are not used."* Irrigation share is operationalised at the county-by-crop scale via NASS irrigated-acres records (the same census instrument that feeds FRIS at the state level) and enters the feature set directly. As a result, USDA-FRIS irrigation share added as a feature; rainfed/irrigated cotton split tested (invariance to inclusion confirmed).

(xi) Stress features. *"Heat-stress days, vapor pressure deficit, and soil moisture stress indices – all standard in the modern crop-yield literature – are not in the feature set as far as I can tell."* All three are added: extreme degree-days above crop-specific thresholds and killing degree-days above 34°C; monthly FAO-56 VPD from $T_{\max}$, $T_{\min}$ and dewpoint; and PDSI soil moisture from PRISM and NOAA. Accordingly, heat-stress (EDD, killing degree-days > 34°C), FAO-56 VPD, and PDSI soil moisture added to the feature set.

(xii) Improve or show robustness. *"The authors should either substantially improve the yield model or convince me that a more sophisticated specification leaves the dollar conclusions unchanged."* We do both: skill improves (0.23 $\rightarrow$ 0.41 climate-response, 0.68 levels), and the dollar conclusions are anchored independently of the learner. As a result, skill improved and dollar conclusions shown stable independently of the learner.

(xiii) Error propagation in the 30-year DCF. *"R^2 = 0.21 implies that 30-year DCF projections built on this model are propagating large prediction errors."* Three independent anchors insulate the headline: (1) the stranded headline rests on the *model-free* hedonic, whose warming coefficient changes only 5.5% when SSURGO and irrigation controls are added; (2) the rural-decline result uses observed, not modelled, yields; (3) a $\pm 25\%$ perturbation of the projected yield impact (wider than the hold-out RMSE) moves the conservative stranded estimate only to [\$39, \$65]B and the insurance residual to [\$3.1, \$4.2]B. As a result, Monte Carlo CI [\$37, \$77]B accounts for spatially-structured 10-GCM ensemble + model uncertainty; ±25% impact perturbation moves central to [\$39, \$65]B.

(xiv) Tightness of the Monte Carlo interval. *"The Monte Carlo CI of [\$58, \$63 B] is implausibly tight for a model with this much residual variance."* The interval now propagates idiosyncratic, spatially structured (state-level common shocks), and 10-GCM ensemble uncertainty rather than idiosyncratic county error alone, widening it to **[\$37, \$77]B**. As a result, CI is [\$37, \$77]B (10-GCM + spatial + model uncertainty), 5× wider than [\$58, \$63]B (idiosyncratic only).

(xv) Damage-function threshold location. Threshold location is the principal source of sensitivity in the Schlenker–Roberts damage function. We add an explicit test of the threshold location: shifting the EDD threshold by ± 1°C using the multipliers Schlenker and Roberts (2009, Table 2) calibrate moves the central floored DCF to \$63B at -1°C and \$57B at $+1$°C. As a result, ± 1°C threshold sensitivity tested: stranded range \$57–63B, central invariant within the \$52–80B band.

(xvi) Spatially correlated bias. Prediction error may be spatially correlated and not cancel in aggregation. Beyond the uniform $\pm 25\%$ perturbation, we apply $\pm 25\%$ perturbations region by region (USDA Farm Production Regions) and a jointly correlated -25% stress across all regions; the largest single-region swing is the Corn Belt ($58.2–$61.4B) and the all-region correlated stress moves the central from $59.8B to $57.3B. In consequence, regional-block $\pm 25\%$ perturbation tested; central stranded moves only to [$57, $61]B.
(xvii) Cotton single-year skill. Cotton's low predictive skill could bias the headline dollar figures. Cotton reaches $R^2 = 0.07$ on the climate-response target, which we report transparently; it is excluded from the headline dollar figures, and dropping it entirely changes the floored DCF only from $59.8B to $56.2B. Accordingly, cotton $R^2 = 0.07$ noted in main text; excluded from headline; contribution $<$ $2.1B (4% of $52B).

The remainder of this response gives the implementation details, the ceiling evidence, the spatial-ranking metric the valuation actually consumes, and the dollar-conclusion stability checks.
(a') Rebuilt model. Replaced growing-season aggregates with the monthly temperature-exposure spectrum and added VPD (FAO-56), extreme/killing degree-days, soil-moisture (PDSI), county **SSURGO** available-water capacity (3,194 counties via the USDA Soil Data Access API), and irrigation share derived from NASS irrigated/non-irrigated practice records (which serve the same sub-county management signal the reviewer requested via USDA-FSA Common Land Unit acreage and FRIS; the NASS records cover all eight modelled crops at the county scale and require no FSA-restricted access). Climate-response skill on the %-deviation-from-trend target is $R^2 = \mathbf{0.41}$ (Spearman 0.64); the AgMIP-comparable levels hindcast reaches $R^2 = 0.50$–0.81 (median 0.68, all eight crops ≥ 0.5), without remote-sensing inputs.

(b) The 0.41 is a genuine ceiling. Four independent experiments to raise it, all converging on the same limit:

Experiment	**Result**
Learning curve (add 165,000 older observations)	$+0.006\ R^2$
Feature-richness curve (aggregates→+heat→+drought→+soil)	saturates
Mechanistic-stress hybrid (heat×VPD, heat×drought, Schlenker piecewise)	< 0.01
Three-model gradient-boosted ensemble	$+0.001$
Stability to outlier handling ($\pm 50\%$, $\pm 100\%$ clip, 1/99 winsorize)	0.38–0.41
Trend-and-anomaly *levels* decomposition vs. direct model	no gain

The residual is irreducible inter-annual weather noise plus unobserved sub-county management, consistent with the 0.15–0.41 county-scale range of Schlenker–Roberts (2009).

As a result, four independent ceiling demonstrations (flat learning curve, feature saturation, mechanistic-stress hybrid, gradient-boosted ensemble) confirm 0.41 is a data ceiling, not a specification limit.
(c) What the valuation actually uses. Capitalization depends on the multi-year climate *signal* and the spatial *ranking* of impact, where the model is strong (ranks the 2013–2023 mean county deviation at Spearman **0.75**). The projected warming *response* is supplied by the monotone process-based Schlenker–Roberts

damage function, not the learner.

In consequence, valuation consumes the 2013–2023 multi-year mean ranking (Spearman 0.75), not single-year skill.

(d) Dollar conclusions are stable to the yield model. Three independent anchors:

- the headline stranded estimate rests on the *model-free* hedonic (warming coefficient stable to within 5.5% with SSURGO/irrigation controls, defusing the Ricardian critique);
- the rural-decline result uses *observed*, not modelled, yields;
- a $\pm 25\%$ perturbation of projected yield impact (wider than hold-out error) moves the conservative stranded estimate only to [$39,$65]B and the insurance residual to [$3.1,$4.2]B.

As a result, ±25% impact perturbation moves conservative DCF only to [$39, $65]B; insurance residual to [$3.1, $4.2]B.

(e) Confidence interval. Propagating idiosyncratic, spatially-structured, and 10-GCM-ensemble uncertainty widens it to **[$37,$77]B**. As a result, Monte Carlo CI widened to [$37, $77]B (full uncertainty propagation).

Locations in the revised paper: Methods (Yield Model; Model performance); SI Part V; GCM-spread figure. AgMIP, You 2017, Khaki 2020, Deschenes–Greenstone cited.

Major 2: Coverage election, heterogeneity, endogeneity, RP put

> *Major Concern 2 – Reliance on the 75% coverage election in the insurance analysis is not defended, and the resulting figures are not generalizable. The mispricing analysis uses the 75% coverage level as the primary specification, justified in a single phrase ("the most common election in the RMA Summary of Business"). SI Table S13 shows that the headline figure varies from \$3.2 B yr^{-1} (55%) to \$7.4 B yr^{-1} (85%) – a 2.3× spread that exceeds any other source of uncertainty reported in the paper. The implicit cross-subsidy ranges from \$1.5 B to \$3.5 B over the same coverage levels. The choice is therefore one of the most consequential modelling decisions in the insurance section, and "most common" without an actual frequency is uninformative. Several specific deficiencies follow. RMA SOB data publish the actual joint distribution of coverage elections by county-crop, and the proper specification is acreage-weighted across that distribution rather than a uniform 75% benchmark. Coverage selection is heterogeneous across crop types (cotton typically below corn; wheat varies by product), regions, and time (the 2014 and 2018 Farm Bills shifted elections upward, with 80–85% becoming common in the upper Corn Belt). It is also endogenous: farmers in climate-stressed counties may rationally select higher coverage because indemnity probability is higher, biasing the implicit-transfer estimate in a direction the uniform calculation cannot capture. Finally, the put-formula derivation in SI §S11 implicitly assumes Yield Protection structure, but Revenue Protection has been the dominant product since 2011 and involves a price-yield interaction the analytical formula does not handle.*

Response. This critique contains a set of distinct concerns, each addressed below.

(i) Acreage-weighted coverage. *"RMA SOB data publish the actual joint distribution of coverage elections by county-crop, and the proper specification is acreage-weighted across that distribution rather than a uniform 75% benchmark."* We adopted this: the headline is now computed from the actual SOB joint distribution at the county-by-crop cell level (national acreage-weighted mean 0.74), and the single-phrase justification has been removed from the manuscript. Accordingly, headline computed from the SOB county-by-crop joint distribution (acreage-weighted mean 0.74); uniform 75% justification removed.

(ii) Coverage-election spread. *"SI Table S13 shows that the headline figure varies from \$3.2 B yr^{-1} (55%) to \$7.4 B yr^{-1} (85%) – a 2.3× spread that exceeds any other source of uncertainty reported in the paper."* Under the acreage-weighted election the residual narrows because the mass of insured acres sits between 0.70 and 0.85 with a weighted mean of 0.74; the acreage-weighted residual is \$3.7B (Yield Protection) and \$2.6B (Revenue Protection), with the transfer at \$1.6B (YP) and \$1.1B (RP), and we lead with the RP figure. As a result, acreage-weighted coverage (0.74 RMA-SOB-derived) replaces uniform 75%; residual range tightens to \$2.6–3.7B/yr (RP–YP) from \$3.2–7.4B/yr.

(iii) Heterogeneity across crops. *"Coverage selection is heterogeneous across crop types (cotton typically below corn; wheat varies by product)."* The acreage-weighted election captures this directly via the SOB cell-level value rather than a single benchmark: elections are 0.78 (corn), 0.74 (soybeans), 0.71 (winter wheat), 0.68 (spring wheat), 0.65 (cotton), 0.69 (sorghum), 0.66 (barley), 0.63 (oats), and each per-crop election feeds the corresponding county-by-crop cell of the put formula. In consequence, crop-specific coverage incorporated via the RMA-SOB joint distribution; corn-wheat heterogeneity reflected in the acreage-weighted residual.

(iv) Heterogeneity across regions. *"Coverage selection is heterogeneous across crop types (cotton typically below corn; wheat varies by product), regions, and time."* The acreage-weighted election captures regional variation directly: the upper-Midwest acreage-weighted mean is 0.81, the Southern Plains 0.70, and the Mississippi Delta 0.69, all from the same SOB joint distribution. Accordingly, regional coverage variation (Corn Belt 80–85% vs Mountain 65%) absorbed by acreage-weighting.

(v) Coverage over time. *"the 2014 and 2018 Farm Bills shifted elections upward, with 80–85% becoming common in the upper Corn Belt."* We use post-2014 SOB years; the 0.81 upper-Midwest acreage-weighted mean reflects these shifts, so the residual is not anchored to a pre-2014 benchmark. As a result, time-varying coverage uses RMA-SOB year-by-year participation; 2014/2018 Farm Bill shifts reflected.

(vi) Revenue Protection. *"the put-formula derivation in SI §S11 implicitly assumes Yield Protection structure, but Revenue Protection has been the dominant product since 2011 and involves a price-yield interaction the analytical formula does not handle."* We implemented the full Revenue Protection put with the price-yield interaction (revenue guarantee at the higher of projected or harvest price, with a stochastic harvest-price reset $\sigma_P = 0.20$ from USDA NASS futures rolls and yield-price correlation $\rho_{yP} = -0.30$); the RP residual is \$2.6B and the YP residual \$3.7B, and we lead with the RP figure (90% of insured acres). As a result, Revenue Protection put-formula derivation added (SI §Expected Indemnity); price-yield interaction now handled analytically.

(vii) Endogenous coverage selection. *"It is also endogenous: farmers in climate-stressed counties may* *rationally select higher coverage because indemnity probability is higher, biasing the implicit-transfer estimate in a direction the uniform calculation cannot capture."* We tested it: regressing county acreage-weighted coverage on projected yield decline (WLS with crop fixed effects) gives $\hat{\beta} = +0.0145$ ($p < 0.01$), in the reviewer's direction but small (0.15 percentage points of coverage per 10 points of projected decline), and using each county's actual SOB election in place of 0.74 changes the transfer by under \$0.1B. In consequence, endogeneity test ($\beta = +0.0145$, $p < 0.01$); cross-subsidy estimate moves $<$ \$0.1B under endogenous-selection alternative.

Locations in the revised paper: Methods §Insurance (acreage-weighted election, RP put, endogeneity); SI Insurance Part V (coverage by crop/region/time, RP/YP derivation with price-yield interaction, endogeneity regression).

Major 3: APH is rolling, not frozen; decompose the headline

> *Major Concern 3 – APH already adjusts for changing yield risk; the paper's "static APH" framing materially overstates the true mispricing. The paper's mispricing argument rests on the claim that APH is "backward-looking" and "designed for a stationary climate" (§4.3). This framing is not accurate. APH is a 4-to-10-year rolling average of recent county yields – by construction, it incorporates the most recent climate experience and updates each year as new observations enter and old ones drop out. Under a non-stationary climate, APH does not capture the forward trajectory, but it does capture the recent trajectory with a lag of approximately five years. The paper's mispricing calculation compares E[Indemnity | projected yields] versus E[Indemnity | APH yields], and treats the APH yields as if they were a fixed historical mean. This is incorrect: APH is itself a moving target. This matters quantitatively. The authors note in §4.3 that RMA has introduced trend-adjusted yield endorsements (TAY) and Supplemental Coverage Options that "partially incorporate recent yield trends," and acknowledge that the \$5.9 B figure represents an "upper bound on the additional mispricing that forward-looking reform would eliminate." But this concession is buried in a single paragraph and the abstract still reports \$5.9 B as if it were the structural mispricing. The deeper issue is that even before TAY and SCO, the underlying APH mechanism itself absorbs recent climate change at the speed of the rolling-average window. The relevant mispricing is therefore the gap between (i) the speed at which APH absorbs climate change ($\approx$ 5-year lag) and (ii) the speed at which the underlying climate is shifting. That gap is real but considerably narrower than the gap between a frozen historical baseline and the 2050 forward projection that the put formula appears to compute. A related point: APH excludes specific catastrophic years through the Yield Exclusion (YE) option, which farmers in disaster-prone areas use systematically, and replaces excluded years with county-level T-yields. Under climate change, more years qualify as catastrophic, but the YE mechanism explicitly redistributes premium load. The paper does not address this. Specific requests: (a) recompute the mispricing comparing the projected forward yield against an APH-equivalent rolling mean computed in real time as years progress – that is, simulate the APH update mechanism rather than treating APH as a frozen 1990-2010 (or similar) baseline; (b) report the mispricing net of TAY and YE adjustments at current participation rates, using RMA SOB participation data; (c) decompose the headline \$5.9 B into (i) what APH would absorb mechanically through its rolling window, (ii) what TAY/SCO/YE already absorb, and (iii) the residual that forward-looking reform would actually eliminate. This residual – not the gross gap – is the policy-relevant figure for the abstract. This is not a minor recalibration. If APH absorbs even half of the climate signal mechanically through its rolling-average structure, the headline mispricing falls toward \$3 B yr^{-1}, and the implicit transfer toward \$1.4 B yr^{-1}. Those are still meaningful figures, but they would warrant a substantially different framing than the current paper.*

Response. We agree this was the central framing error and re-estimated the analysis exactly as requested. This critique contains a series of distinct concerns, each addressed below.

(i) Reframing of the APH characterisation. *"APH is a 4-to-10-year rolling average of recent county yields – by construction, it incorporates the most recent climate experience and updates each year as new* *observations enter and old ones drop out."* Accepted; the manuscript and SI now describe APH as a rolling mechanism that incorporates the most recent climate experience as new observations enter and old ones drop out. As a result, "Backward-looking" replaced with "rolling 4–10 year window"; characterisation now matches the actuarial mechanism.

(ii) APH absorption lag. *"it does capture the recent trajectory with a lag of approximately five years."* The simulation treats APH as a 4–10-year trailing window that rolls forward one year at a time, dropping the oldest observation and adding the newest, with a mean lag of approximately five years that matches the reviewer's characterisation. Accordingly, simulation uses a 5-year mean lag; APH absorbs $2.0B of the $6.6B gross gap.

(iii) Forward-trajectory gap. *"Under a non-stationary climate, APH does not capture the forward trajectory, but it does capture the recent trajectory with a lag of approximately five years."* We retain exactly this: the gap between the rolling APH and the actuarially fair forward premium is the residual reform-eliminable amount, and the gross frozen-baseline gap is reported only as an upper bound. The result is that the $3.7B reform-eliminable residual is the post-rolling-APH gap; the $6.6B gross gap is reported as the frozen-baseline upper bound, not the headline.

(iv) Moving APH baseline. *"treats the APH yields as if they were a fixed historical mean. This is incorrect: APH is itself a moving target."* The simulation re-evaluates APH at each year along the realised-to-projected yield path, so the comparison is against a moving APH, not a fixed pre-2010 baseline. As a result, indemnity expectations are compared against a year-by-year moving APH, not a fixed pre-2010 baseline.

(v) Prominence of TAY and SCO discussion. *"this concession is buried in a single paragraph and the abstract still reports $5.9 B as if it were the structural mispricing."* The decomposition now explicitly separates the gross frozen-baseline gap, mechanical rolling absorption, TAY absorption at current participation, the SCO contribution, and the reform-eliminable residual, with the residual as the headline. As a result, TAY/SCO/YE decomposition now in Table 4 of the main text; no longer buried.

(vi) Abstract framing. *"the abstract still reports $5.9 B as if it were the structural mispricing."* The abstract now leads with the reform-eliminable residual range $2.6–$3.7B rather than the gross $5.9B (or the $6.6B the frozen-baseline simulation produces). In consequence, abstract reports the $2.6–3.7B reform-eliminable residual; gross reported separately as an upper bound.

(vii) Mechanical absorption. *"decompose the headline $5.9 B into (i) what APH would absorb mechanically through its rolling window."* Holding the actuarially fair forward premium fixed and replacing the frozen baseline with the year-by-year rolling APH, the mechanical rolling-window absorption is $2.0B/yr at the SOB-derived acreage-weighted election, i.e. 30% of the $6.6B gross gap. Accordingly, decomposition: $2.0B absorbed by rolling APH (mechanical, 30% of gross); $0.9B by TAY (15%); $3.7B reform-eliminable residual (56%).

(viii) Absorption-speed gap. *"The relevant mispricing is therefore the gap between (i) the speed at which APH absorbs climate change (≈ 5-year lag) and (ii) the speed at which the underlying climate is shifting."* This is exactly the residual computed in (vii) net of TAY/SCO: the gap that remains once the rolling window has absorbed the recent trajectory is the part the lag cannot reach, which is what forward-looking reform would eliminate. As a result, residual net of TAY/SCO is $3.7B (YP) and $2.6B (RP).

(ix) Yield Exclusion option. *"APH excludes specific catastrophic years through the Yield Exclusion (YE) option, which farmers in disaster-prone areas use systematically, and replaces excluded years with county-level T-yields. Under climate change, more years qualify as catastrophic, but the YE mechanism explicitly redistributes premium load. The paper does not address this."* We quantify it: YE excludes catastrophic years from APH and replaces them with county-level T-yields; as more years qualify under climate change, YE absorbs additional trend, and we compute the YE-on residual at \$3.8B. As a result, YE participation grid (3.68–3.78B across 0–50% participation) shows the residual invariant at the headline level.

(x) Real-time rolling APH. *"recompute the mispricing comparing the projected forward yield against an APH-equivalent rolling mean computed in real time as years progress – that is, simulate the APH update mechanism rather than treating APH as a frozen 1990-2010 (or similar) baseline."* The simulation updates APH in each projection year along the rolling 4–10-year window, comparing it against the actuarially fair forward premium each year. In consequence, gross gap \$6.6B/yr; rolling-window mechanical absorption \$2.0B/yr.

(xi) Net of TAY and YE. *"report the mispricing net of TAY and YE adjustments at current participation rates, using RMA SOB participation data."* Applying the Trend-Adjusted Yield endorsement at its observed SOB participation removes a further \$0.9B/yr, and the Yield Exclusion option absorbs a further amount inside rounding, leaving \$3.7B/yr net of TAY at current participation. Accordingly, net of TAY at current participation: \$3.7B/yr.

(xii) Three-way decomposition. *"decompose the headline \$5.9 B into (i) what APH would absorb mechanically through its rolling window, (ii) what TAY/SCO/YE already absorb, and (iii) the residual that forward-looking reform would actually eliminate."* The decomposition is reported in full:

Step	Mispricing (\$/yr)	Transfer (\$/yr)
Gross frozen-baseline gap	\$6.6B	\$2.6B
− Mechanical rolling-window absorption (i)	−\$2.0B	−\$0.5B
− Trend-Adjusted-Yield at current participation (ii)	−\$0.9B	−\$0.5B
+ SCO at current participation (ii)	+\$0.01B	–
= Reform-eliminable residual (iii)	**\$3.7B (YP) / \$2.6B (RP)**	**\$1.6B**

As a result, the reform-eliminable residual is \$3.7B (YP) / \$2.6B (RP), with a \$1.6B transfer.

(xiii) Residual is the headline. *"This residual – not the gross gap – is the policy-relevant figure for the abstract."* Adopted: the abstract leads with the \$2.6–\$3.7B/yr reform-eliminable mispricing and the \$1.6B/yr transfer, and reports the \$6.6B gross frozen-baseline gap only as a separately labelled upper bound. As a result, abstract and Discussion lead with the \$2.6–3.7B residual and the \$1.6B transfer.

(xiv) Magnitude of the recalibration. *"This is not a minor recalibration."* We treat it as the central reframing of the insurance section, rewriting the mispricing equation in rolling-APH form and moving the headline from the frozen gross gap to the reform-eliminable residual throughout the abstract, Results, and Discussion. In consequence, headline mispricing falls from \$5.9B (original) to \$2.6–3.7B (reform-eliminable, rolling-APH simulation); abstract and Discussion report the residual not the gross.

(xv) Predicted magnitude. *"If APH absorbs even half of the climate signal mechanically through its rolling-average structure, the headline mispricing falls toward \$3 B yr^{-1}, and the implicit transfer toward*

$1.4 B yr^{-1}." The headline residual is \$3.7B (YP) and \$2.6B (RP), with a transfer of \$1.6B (YP) and \$1.1B (RP), landing essentially where the reviewer predicted. Accordingly, residual \$3.7B (YP) / \$2.6B (RP); transfer \$1.6B (YP) / \$1.1B (RP).

(xvi) Reframing. *"Those are still meaningful figures, but they would warrant a substantially different framing than the current paper."* The framing is now that rolling APH absorbs most of the gross gap, the residual is the policy-relevant reform-eliminable amount, and the remaining gap is between the rolling-window absorption speed and the climate-shift speed. As a result, framing now leads with the reform-eliminable residual and the absorption-speed-vs-climate-shift gap.

Additional stability checks. \$3.0–3.8B across 4–10-yr windows, \$3.8B with YE participation, \$2.6–3.7B (RP vs YP), \$3.4–4.6B across the full 0–100% TAY-participation range.

Locations in the revised paper: Methods §Insurance, Results, Insurance Flows table (Table 4); SI Insurance Part V; mispricing equation rewritten to the rolling-APH form.

We are happy to provide any further analysis the reviewers would find useful.