

UC Santa Barbara

UC Santa Barbara Electronic Theses and Dissertations

Title

Environmental and Economic Consequences of Agriculture

Permalink

<https://escholarship.org/uc/item/00x2s0gh>

ISBN

9798171613686

Author

Blomqvist, Linus

Publication Date

2026-09-07

Peer reviewed|Thesis/dissertation

University of California

Santa Barbara

Environmental and Economic Consequences of Agriculture

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in Environmental Science and Management

by

Linus Blomqvist

Committee in charge:

Professor Andrew Plantinga, Chair

Professor Robert Heilmayr

Professor Kathy Baylis

September 2026

The Dissertation of Linus Blomqvist is approved.

Professor Robert Heilmayr

Professor Kathy Baylis

Professor Andrew Plantinga, Committee Chair

September 2026

Environmental and Economic Consequences of Agriculture

Copyright © 2026

by

Linus Blomqvist

Acknowledgements

First of all, I want to thank my committee — Andrew Plantinga, Robert Heilmayr, and Kathy Baylis — for their steadfast support through the ups and downs of this PhD. A thanks also goes out to all my friends at the Bren School and in the wider world, for making this journey more fun and less lonely. Thanks to my kids Anouk and Kiro for bringing so much joy to my life. Thanks also to my sister Petra, mother Giesela, and father Lasse for being there for me. I reserve my deepest gratitude for my wife Jessica, without whose unwavering love and encouragement I could never have completed this PhD.

During my PhD, I received funding from the Schmidt Family Foundation and from the University of Nebraska–Lincoln.

Chapters 1 (*Land-Use and Welfare Effects of Anti-Deforestation Policies*) and 2 (*Crop Yield Trends and Gaps*) are co-authored. For chapter 1, I contributed to the study design, carried out most of the analysis, and wrote the paper. Part of the methodology and code, in particular the simulation code, was developed jointly with Roberto Amaral-Santos and is also used in a related paper that he leads; I thank him for that collaboration. Robert Heilmayr and Andrew Plantinga contributed to the study design. All co-authors provided feedback along the way, as did Holly Gibbs and Lisa Rausch at the University of Wisconsin–Madison, whom I thank. For chapter 2, I developed the study design, carried out all the analysis, and wrote the paper. All co-authors commented at several stages of the work. Chapter 3 (*Drivers of Nitrogen Surplus in US Agriculture*) is sole-authored. It uses nitrogen input data provided by Stephen Ogle at Colorado State University, whom I thank.

I used Claude Code in all three chapters to assist with writing and debugging analysis code, as well as editing of prose. The research questions, analytical choices, interpretations, and conclusions are my own, as are any errors.

Vita of Linus Blomqvist

Education

2026	Ph.D. in Environmental Science and Management, University of California, Santa Barbara
2018	Master Environmental Science, Yale School of the Environment, Yale University
2011	B.A. in Management Studies, University of Cambridge
2008	Certificate in Social Sciences, The Open University

Professional Appointments

2022–present	Research Consultant, The Nature Conservancy
2018–2019	Research Associate, University of Tasmania
2014, 2023	Lone Mountain Fellow, Property and Environment Research Center
2011–2019	Director, Food and Agriculture Program, Breakthrough Institute

Publications

- Shah, S., D. Blaustein-Rejto, and L. Blomqvist (2025) “Future yield growth in Africa mitigates agricultural greenhouse gas emissions.” *Environmental Research Communications* 7: 015025.
- Blaustein-Rejto, D., N. Soltis, and L. Blomqvist (2023) “Carbon opportunity cost increases carbon footprint advantage of grain-finished beef.” *PLoS ONE* 18(12): e0295035.
- Lovering, J., M. Swain, L. Blomqvist, and R. R. Hernandez (2022) “Land-use intensity of energy production and tomorrow’s energy landscape.” *PLoS ONE* 17(7): e0270155.
- Blomqvist, L., L. Yates, and B. W. Brook (2020) “Drivers of increasing global crop production: A decomposition analysis.” *Environmental Research Letters* 15: 0940b6.
- Flies, E. J., B. W. Brook, L. Blomqvist, and J. C. Buettel (2018) “Forecasting future global food demand: A systematic review and meta-analysis of model complexity.” *Environment International* 120: 93–103.
- Swain, M., L. Blomqvist, J. McNamara, and W. J. Ripple (2017) “Reducing the environmental impact of global diets.” *Science of the Total Environment* 610: 1207.
- Brook, B. W., and L. Blomqvist (2016) “Innovations and limits in methods of forecasting global environmental change.” *Basic and Applied Ecology* 17(7): 565–575.
- Blomqvist, L., B. W. Brook, E. C. Ellis, P. M. Kareiva, T. Nordhaus, and M. Shellenberger (2013) “Does the shoe fit? Real versus imagined ecological footprints.” *PLoS Biology* 11(11): e1001700.

- Blomqvist, L., B. W. Brook, E. C. Ellis, P. M. Kareiva, T. Nordhaus, and M. Shellenberger (2013) “The ecological footprint remains a misleading metric of global sustainability.” *PLoS Biology* 11(11): e1001702.
- Brook, B. W., E. C. Ellis, M. P. Perring, A. W. Mackay, and L. Blomqvist (2013) “Does the terrestrial biosphere have planetary tipping points?” *Trends in Ecology and Evolution* 28(7): 396–401.

Book Chapters and Reports

- Blomqvist, L., B. Leonard, A. J. Plantinga, M. R. Sloggy, and J. J. Sanches (2024) *Local Effects of Stranded Public Lands*. US Forest Service Research Paper PSW-RP-271.
- Blomqvist, L., and R. D. Simpson (2017) “The value of ecosystem services: What is the evidence?” In P. Kareiva et al. (eds.), *Effective Conservation Science: Data Not Dogma*. Oxford: Oxford University Press.
- Blomqvist, L. (2016) “How humans spare nature.” In S. Regan (ed.), *Environmental Policy in the Anthropocene*. Bozeman, MT: Property and Environment Research Center.
- Blomqvist, L., T. Nordhaus, and M. Shellenberger (2015) *Nature Unbound: Decoupling for Conservation*. Oakland, CA: Breakthrough Institute.
- Nordhaus, T., M. Shellenberger, and L. Blomqvist (2012) *The Planetary Boundaries Hypothesis: A Review of the Evidence*. Oakland, CA: Breakthrough Institute.

Fields of Study

Major field: Environmental Science and Management, with an interdisciplinary emphasis in Economics and Environmental Science

Studies in agricultural and environmental economics with Professor Andrew Plantinga (committee chair), Professor Robert Heilmayr, and Professor Kathy Baylis, University of California, Santa Barbara.

Studies in environmental economics with Professor Robert Mendelsohn, Yale School of the Environment.

Abstract

Environmental and Economic Consequences of Agriculture

by

Linus Blomqvist

Agriculture produces nearly all human food and supports the livelihoods of a large share of the world's population. It is also among the leading causes of habitat loss, greenhouse gas emissions, and air and water pollution. What agriculture costs the environment, and what constraining it costs producers and consumers, are therefore central policy questions. This dissertation takes up both in three chapters.

Chapter 1 asks what a package of Brazilian anti-deforestation policies — the Amazon Soy Moratorium, cattle agreements, and 2012 Forest Code — costs the producers it restricts. A nested logit model of land use, estimated on a pixel-level panel of the Amazon and Cerrado, yields the compensating variation: those producers' welfare loss, or the money it would take to make them whole. That loss is 25 percent of their welfare, some 45 billion dollars in present value. The package retains 2.7 million hectares of Amazon forest, whose carbon justifies it at any social cost of carbon above 43 dollars per tonne.

Chapter 2 asks whether yields are rising fast enough to meet projected 2050 crop demand without further cropland expansion. Fitting ten piecewise-linear trend models to 34 crop-country yield series for maize, rice, wheat, and soybean over 1961–2024 shows that yield plateaus are rare — under 10 percent of global production for every crop. Where simulated potential yields are available, most systems, plateaued or not, remain well below their biophysical ceiling. Projected growth nonetheless falls far short of demand for maize and soybean, and closing that gap with land instead would take roughly 55 million hectares.

Chapter 3 asks what drives the US agricultural nitrogen surplus, and how much of the apparent efficiency gain is real. An exact index decomposition of a new nitrogen budget spanning 18 crops and 2,829 counties separates five drivers. A near-flat total since 1990 conceals two large offsetting forces: rising yields added 1.99 Tg N of surplus, while falling surplus per unit of output removed 1.40 Tg N. Nearly a third of the apparent gain in conventional nitrogen-use efficiency reflects shifts in crop mix and geography rather than improved field practice.

Table of Contents

Land-Use and Welfare Effects of Anti-Deforestation Policies: Evidence from the Brazilian Amazon	1
Introduction	1
Methods	7
Model Framework	7
Study Area and Land-Use Data	8
Rent	9
Specification and Estimation	10
Simulation	12
Compensating Variation	13
Consumer Surplus	14
Value of Avoided Carbon	15
Results	16
Model estimates	16
Land-use outcomes	18
Welfare cost	21
Discussion	23
References	32
Supplementary Information	37
Data details	37
Model details	44
Simulation details	45
Consumer surplus	46
 Crop Yield Trends and Gaps in the World's Major Cropping Systems, and Their Implications for Meeting Projected Global Demand	 49
Introduction	49
Methods	54
Data Sources and Processing	55
Model selection	58
Trend synthesis	62
Yield Gap Analysis	64
Global Yield Trends and Projection to 2050	66
Diagnostics	72
Results	73
Yield trends and classification	73
Yield gaps and closure rates	76
Global yield projection and 2050 demand	78
Discussion	81
References	97
Additional Methods	102

Dropped observations	102
Mexico irrigated maize	102
Potential yield data coverage	103
Model Diagnostics	104
Motivation	104
Residual diagnostics	104
Leave-one-out influential point check	107
Bootstrap knot stability	108
Diagnostic settings	109
Robustness checks	109
R1: minimum segment length 1, AICc, preferred model set	109
R2: minimum segment length 10, BIC, preferred model set	110
R3: minimum segment length 10, AICc, full model set	111
AICc vs. BIC: differences in classification	112
Effect of including the dropped smooth-curve models	112
Effect of minimum segment length constraint	113
Comparison with Grassini et al. (2013)	114
Additional Results	120
Production share by terminal-form class	120
Yield gaps and closure rates	120
Global average yield trends	121
Aggregate calorie-production index	123
Model coefficients	123
Drivers of Nitrogen Surplus in US Agriculture: A Decomposition Analysis	127
Introduction	127
Methods	132
Decomposition framework	132
Data sources and inputs	135
The field-balance panel	137
Results	139
Decomposition of the change	140
Harvested area and its distribution	141
Crop mix	142
Yields	144
Nitrogen balance intensity	144
Spatial synthesis	147
Within-field efficiency versus conventional NUE	147
Discussion	149
References	153
Supplementary information	157
Decomposition method	157
Data sources and processing	163
Attribution tables	195
Conventional NUE versus within-field efficiency	198

Land-Use and Welfare Effects of Anti-Deforestation Policies: Evidence from the Brazilian Amazon

Linus Blomqvist Roberto Amaral-Santos

Andrew J. Plantinga Robert Heilmayr

Bren School of Environmental Science and Management, University of California, Santa
Barbara

1 Introduction

The Brazilian Amazon holds one of the world's largest stores of forest carbon (Saatchi et al. 2011) and a large share of its terrestrial biodiversity (Ter Steege et al. 2013; Jenkins, Pimm, and Joppa 2013). Clearing the forest releases that carbon (Gatti et al. 2021) and destroys habitat of global significance. Slowing this loss is central to Brazil's commitments under the Paris Agreement and to international biodiversity targets (Brazilian Government 2024; Conference of the Parties to the Convention on Biological Diversity 2022).

Brazil has built a broad set of policies (Nepstad et al. 2014) to slow this clearing, which is driven primarily by cattle ranching and, indirectly, by soybean cropping (Barona et al. 2010; Song et al. 2021). Some are decades old, including protected areas, indigenous reserves, and the Forest Code — the country’s foundational land-use law, first enacted in 1965 (Soares-Filho et al. 2014). The Forest Code sets the basic rule for private land: owners of forested property in the Amazon must leave most of it standing. For decades that rule went largely unenforced. That changed in the 2000s, when the government overhauled its conservation policy in two waves (Assunção, Gandour, and Rocha 2015). Both put new weight on enforcing the law against illegal clearing. The first, in 2004, was an action plan for the Amazon (the PPCDAm) that built a satellite system to flag new clearing every two weeks and used those alerts to direct federal enforcement officers to it (Assunção, Gandour, and Rocha 2023). The second came in 2008, when the government concentrated enforcement on a published list of the worst-offending municipalities (Assunção and Rocha 2019) and made rural credit in the Amazon biome conditional on proof that the borrower complied with environmental law (Assunção, Gandour, and Rocha 2015). Two new supply-chain agreements followed. In 2006, major soy traders launched the Amazon Soy Moratorium (ASM), refusing to buy soy grown on Amazon land cleared after a cutoff later set at July 2008 (Heilmayr et al. 2020; Gibbs et al. 2015). In 2009, the largest slaughterhouses signed cattle agreements applying similar terms to beef (Alix-Garcia and Gibbs 2017). The 2012 revision of the Forest Code then converged on the same threshold, granting amnesty for illegal clearing that predated July 2008 while leaving later violations exposed to sanction (Soares-Filho et al. 2014). What unites the three is this common cutoff: each raises the cost of clearing Amazon forest after July 2008, whether by withdrawing market access (the ASM and cattle agreements) or, for illegal clearing, by legal sanction (the Forest Code). Because they draw the same line in the same biome, they bind producers as a single regime, which we call the policy package. The agreements cover much of the region’s soy and beef, and the Forest Code applies to all private land. The three are also linked in practice, not only in timing. The traders and slaughterhouses screen their suppliers using the government’s property

registry, its embargo lists and its satellite imagery, so the private agreements run on the state's monitoring system (Nepstad et al. 2014; Gibbs et al. 2015).

Their effect on clearing, though, is uneven across the three. It is clearest for the soy moratorium: a triple-difference evaluation attributes about 18,000 km² of avoided clearing to it over its first decade, roughly a third less than a no-moratorium counterfactual (Heilmayr et al. 2020). For the cattle agreements the evidence is mixed: an early evaluation found little net effect, gains on monitored suppliers lost to leakage (Alix-Garcia and Gibbs 2017), while more recent work finds they curbed deforestation where legally enforced (Da Mata, Dotta, and Severnini 2026). The Forest Code's registry yields only small reductions that soon fade (Azevedo et al. 2017).

These policies work by restricting what landowners may do with their own land. That restriction falls on a specific, bounded group: the farmers and ranchers of the frontier whose land it makes less valuable. The benefits are shared far more widely. A more stable climate and preserved biodiversity accrue to the world at large, and over a long horizon. This asymmetry is a familiar problem in the political economy of regulation. When costs fall on a small, organized group and benefits spread thinly across a very large one, the losers have both the incentive and the cohesion to resist, while the diffuse winners exert little countervailing pressure (Wilson 1980). A policy of this kind is politically fragile, even if a cost-benefit analysis would endorse it.

Exactly these dynamics are now playing out around the policy package. Brazil's competition authority has moved to suspend the ASM on antitrust grounds (Hanbury 2025). The state of Mato Grosso has passed laws penalizing firms that comply with it (WWF 2025). Producer associations have sued the signatory traders (Cassauwers 2025). The Forest Code's 2012 revision was itself the product of an agribusiness campaign to roll back the prior law (Soares-Filho et al. 2014), and the ruralista bloc continued to press against environmental enforcement under Bolsonaro (Ferrante and Fearnside 2019). How such conflicts resolve depends on that concentrated cost. When extending the soy moratorium into the Cerrado was negotiated, the central proposal was to pay soy farmers to forgo legally permitted clearing; producers rejected any moratorium

that did not compensate them, and downstream buyers pledged only limited funds (Bastos Lima and Persson 2020; Garrett et al. 2022). The size of that cost, and how it is measured, is the question this paper takes up.

That conservation restricts private land use at a cost to those it binds is a general finding, documented across many settings and measured in many ways. Studies have estimated the agricultural rents given up when forest is left standing (Börner and Wunder 2008). Others have measured the fall in land value when regulation removes the right to develop (Auffhammer et al. 2020; Bošković and Nøstbakken 2017), the revenue foregone by resource users placed under new limits (Booth et al. 2021; Ando et al. 1998), or the change in local income and poverty that follows a protected area (Ferraro et al. 2015; den Braber et al. 2024). One line of work goes further, valuing the cost of lost forest access to nearby households as the compensation they would require to accept it (Shyamsundar and Kramer 1996). These measures do not all point the same way; some conservation policies raise local incomes rather than lower them (Ferraro et al. 2015; den Braber et al. 2024). The cost of conservation is thus neither obvious nor uniform, and how it is measured shapes what one concludes.

That cost is best captured as a change in welfare. In economics, welfare is a measure of how well off a person is overall: their economic well-being, akin to utility or surplus. It is not the same as revenue, profit, or output: each of those measures a specific outcome, not how well off the producer is overall. Welfare instead values a producer's entire set of land-use options, not only the use finally chosen. A policy that lowers the value of those options makes the producer worse off even when observed land use does not change. Land value is the closest familiar analogue. A parcel's sale price reflects everything its owner could do with it, so a policy of this kind shows up in the price even when current use and current profit do not. The land-value studies noted above take this route, reading the cost of a regulation off what land sells for (Auffhammer et al. 2020; Bošković and Nøstbakken 2017).

For the Brazilian policies, the cost has been approached from several directions. Studies have

counted the area of clearable forest the ASM removes from use (Rausch and Gibbs 2021), estimated its effect on farmgate soy prices (Barrozo and Skidmore 2026), and traced the local economic activity that followed intensified enforcement (Merkus 2024). These point to a small or even favorable effect, but each captures a single dimension rather than the producer's overall position. Structural models capable of a full welfare measure have been built for Amazon land use, yet they have been aimed at policies that were never enacted: a hypothetical tax on deforestation or payments for ecosystem services (Souza-Rodrigues 2019), or the carbon-efficient allocation of land under an assumed carbon price (Araujo, Costa, and Sant'Anna 2026; Assunção et al. 2026). What has not been measured is the welfare cost of the enacted package itself, as the ASM, the cattle agreements, and the Forest Code actually bind producers. That is what this paper provides.

To do so, we estimate a structural model of land use. The model is a nested logit in which the owner of each pixel of land chooses among forest, pasture, and crops according to the rents each use offers (Train 2009; Lubowski, Plantinga, and Stavins 2006; McFadden 1981). Whether the policy package binds a pixel enters the model directly, through terms tied to the post-2008 restriction in the Amazon. We estimate the model on a pixel-level sample spanning the Amazon and Cerrado over 2000 to 2019. Because the three instruments coincide on the same cutoff and the same biome, we estimate their combined effect. We do not separate the contribution of any one instrument, and we make no claim about their individual effects. The welfare measure follows from the estimated model. We take producers as they stood in the base year, 2008, and compute the expected value of each one's full set of land-use options twice: once with the policy package in force, once with it removed. Both are evaluated at the same starting state, so their difference isolates the policy rather than any later divergence in land use or prices. That difference, expressed in money, is the compensating variation: the sum that would have to be paid to each producer to leave them as well off with the policy as without it. It is the standard money measure of welfare that economists use to value a policy change (Small and Rosen 1981; McFadden 1981; Dundas and Von Haefen 2020; Parsons, Plantinga,

and Boyle 2000). Because it is built from the value of the whole option set, it captures the loss that a revenue or profit comparison would miss.

To complete the cost-benefit comparison, we need three further quantities. The same estimated model supplies them when we simulate land use forward under the policy package and a no-policy counterfactual, with prices and yields free to adjust. None stands apart from the comparison: the first two bear on the benefit, the third on the cost. The first is how much clearing the package prevents. Prior evaluations have measured this one instrument at a time, as noted above; we do not claim a better estimate. Our interest is the combined figure the package produces, because the retained forest sets the value of the avoided carbon emissions, one side of the comparison this paper draws. The second question asks whether that figure overstates the gain. If forest saved in the Amazon is merely cleared elsewhere, the carbon benefit is smaller than it appears. Prior work finds that production does shift across the biome border but often lands on already-cleared ground, while price-mediated displacement within Brazil can offset part of the gain (Moffette and Gibbs 2021; Soterroni et al. 2019; Villoria et al. 2022). Our framework accounts for this displacement within the same model that produces the forest and welfare estimates. The third question bears on the cost. A restriction on clearing should, in principle, push producers to raise output on land they already farm (Thaler 2017), a response documented for cattle where enforcement has tightened (Koch et al. 2019) though not for displaced soy (Moffette and Gibbs 2021). This matters for our welfare measure in a specific way. The compensating variation compares producers' options in the base year, before prices adjust, so it does not embed the intensification that price adjustment would bring. Induced intensification therefore serves as a check on it. If the package induces substantial intensification, producers recover part of the forgone production, and our estimate is an upper bound on the true cost. If it induces little, the estimate is not biased in that direction.

Our results speak to both the cost and the benefit. The package lowers the monetized welfare of the producers it binds by about 25 percent, on the order of 45 billion US dollars in present

value, or roughly 3 billion dollars a year in perpetuity. To our knowledge this is the first welfare estimate of the enacted package. The loss is concentrated on frontier landowners, and they have little to offset it: we find almost no induced intensification, so producers recover little of the production the restriction prevents. On the benefit side, the package retains about 2.7 million hectares of Amazon forest relative to the counterfactual. We find almost no displacement of land use into the Cerrado, so this retained forest is close to a true reduction in clearing rather than clearing moved elsewhere. Keeping that forest standing avoids a one-time release of carbon worth on the order of 200 billion US dollars at a social cost of carbon of 190 dollars per tonne. Weighing the two against each other is a cost-benefit calculation, and by that test the package is justified for any social cost of carbon above roughly 43 dollars per tonne. But that test weighs only the totals. It is silent on who gains and who loses, and here the winners and losers are different people: the benefit is diffuse and global, while the cost is steep and falls on a narrow group of producers. It is this distribution, more than the overall balance, that will determine whether the package survives the challenges now before it. And because our cost measure is what would make a bound producer whole, it is also the size of the transfer that resolving that divide would take.

2 Methods

2.1 Model Framework

A risk-neutral, price-taking owner of pixel i chooses, in year t , the land use k that maximizes utility given the pixel's starting land use j , in a random-utility discrete-choice model of land use (Train 2009; Lubowski, Plantinga, and Stavins 2006; Lawler et al. 2014; Heilmayr, Echeverría, and Lambin 2020; Araujo, Costa, and Sant'Anna 2026). Utility splits into a component we observe, V_{ijkt} , and an i.i.d. type-I extreme-value error capturing pixel-specific unobservables,

which yields closed-form choice probabilities. What drives the choice is the rent each land use yields — the net agricultural return to the pixel — so the model requires a rent for every pixel, land use, and year.

A change in the return to one land use should not shift the odds among the others proportionally: clearing for crops and for pasture are closer substitutes than either is to staying in forest. We therefore do not impose independence of irrelevant alternatives (IIA) across all alternatives. Instead we use a nested logit with a forest nest and a cleared nest containing crops and pasture, imposing IIA only within nests.

The sections that follow construct the data the model requires — the land uses and the rents — and then give the specification and its estimation.

2.2 Study Area and Land-Use Data

Our unit of analysis is a 30×30 m pixel drawn on a 1 km fishnet grid with a random starting point, with annual land-use classifications for 2000–2019 taken from MapBiomas Collection 8 (MapBiomas Project 2024). To keep computation tractable, we use a 50% subsample of these pixels. The grid covers the portion of Brazil lying within 600 km of the Amazon biome boundary and falling inside either the Amazon or the Cerrado biome. Two considerations motivate this extent. First, this band follows the “Arc of Deforestation” along the southern and eastern edge of the Amazon biome (Fearnside 2005), the belt where agricultural expansion and land-cover change have long been concentrated. It takes in a large share of the country’s crop and pasture output. Second, restricting the sample to the Amazon and Cerrado biomes keeps the analysis centered on the soy- and pasture-dominated systems that are our focus. We collapse the MapBiomas classes into three land uses: forest, pasture, and crops. The constituent classes and the classes we exclude are detailed in Section A.1.

2.3 Rent

For each pixel i , land use k , and year t , we compute a rent π_{ikt} — the net agricultural return per hectare — as revenue minus cost. Revenue is a pixel-specific yield times a commodity price, $R_{ikt} = y_{ikt} p_{kt}$, where the prices p_{kt} are common to all pixels (in 2019 BRL per ton). Cost is the expense of transporting that output to the nearest export port, $C_{ikt} = tc_k d_i y_{ikt}$, where tc_k is the per-ton transport cost of commodity k and d_i the travel time from pixel i to port. We take transport cost as a proxy for total production cost, which introduces geographic variation in rents, since more remote pixels lose more value in transit and tend to face higher input costs. Rent is the difference,

$$\pi_{ikt} = R_{ikt} - C_{ikt} = y_{ikt} (p_{kt} - tc_k d_i).$$

Because it nets out transport but not the other costs of production (inputs, labor, machinery, and herd upkeep), this rent is a gross measure that overstates the true net return. Any welfare cost built on it is therefore an upper bound.

We test how much this matters by rescaling rents to an approximate net return, multiplying each land use’s rent by a profit margin: the share of revenue left after operating costs. The two margins come from separate sources. For crops we use a farm-budget study of typical soy and integrated crop-livestock farms in Mato Grosso (Dos Reis et al. 2020), and for cattle a production-cost panel for beef rearing and fattening in Rondônia (CNA SENAR 2024). We report the result as a robustness check rather than as our main estimate, because two farm budgets from particular states and years are a thin basis for rescaling rents across the whole study area and period. The margins, their sources, and the assumptions needed to apply them are detailed in Section A.1.3.

Yields vary geographically because pixel-level agronomic suitability from GAEZ (FAO 2022) scales a baseline observed yield. For crops that baseline is the municipal yield reported by IBGE (IBGE 2022). For pasture no such yield is reported, so we construct a national average: IBGE cattle carcass output (IBGE 2026) divided by MapBiomass pasture area. The crop rent is an area-weighted average across the ten crops that dominate production in the study area. Prices come from FAOSTAT and ESALQ (FAO 2026; ESALQ 2026). We convert these annual rents into perpetuities: a landowner values each land use by the present value of the rents it yields. That present value, computed from the average annual rent over the period, drives the choice rather than a single year’s return. The construction of yields, the crop basket, transport costs, and prices is detailed in Section A.1.2.

2.4 Specification and Estimation

We specify the observed component of utility as

$$V_{ijkt} = \beta_{1k} + \beta_2 \pi_{ikt} + \beta_{3k} \text{policy}_{it} + \beta_{4jk} \mathbf{1}\{\text{start} = j\} + \beta_{5jk} \text{policy}_{it} \mathbf{1}\{\text{start} = j\},$$

so choices respond to profit π_{ikt} through a single rent coefficient β_2 , to land-use intercepts β_{1k} , to the cost of converting out of the starting land use (β_{4jk}), and to the policy package both directly (β_{3k}) and interacted with the starting land use (β_{5jk}). Profit enters as a perpetuity, which assumes producers effectively expect recent rents to persist. The perpetuity treatment follows Lubowski, Plantinga, and Stavins (2006), who likewise base expected returns on an average of realized returns over the recent past, an assumption that smooths transitory shocks and avoids imposing an unobserved model of how producers forecast. Each of the three instruments targets the clearing of Amazon forest that was standing in 2008, once it is cleared after that year.

The soy moratorium and the cattle agreements deny market access to output from such land, while the revised Forest Code brings legal sanction against clearing that breaches the legal-reserve rules. The Forest Code’s July 2008 date is itself national, so the date alone does not separate the Amazon from the Cerrado. What separates them is the share of each property the law requires owners to keep forested, which is far higher in the Amazon and so puts much more post-2008 clearing there in breach. The policy dummy encodes this shared trigger, switching on after 2008 for Amazon pixels forested in 2008; because the instruments coincide on it they are collinear, so its coefficient is their combined effect rather than any one alone. The dummy also picks up the enforcement that stands behind the Forest Code. From 2008 the government concentrated enforcement on the worst-offending municipalities (Assunção and Rocha 2019) and, in the Amazon biome specifically, made rural credit conditional on proof of compliance with environmental law (Assunção, Gandour, and Rocha 2015). Both target illegal clearing on Amazon land after 2008, so they fall inside the treated group. We therefore read the coefficient as the effect of the post-2008 Amazon regime as a whole, not of the three named instruments alone. We use five-year periods, long enough to capture forest-to-crop transitions that seldom occur within a single year.

The policy enters through two coefficients because its effect on a parcel’s choice depends on the parcel’s starting land use. The direct term β_{3k} gives the policy’s effect on the appeal of converting to use k for the reference starting use, which is crops, while the interaction β_{5jk} shifts that effect for parcels that begin the period in use j . Because each is measured relative to the reference starting use, neither coefficient is interpretable on its own. Since the policy is designed to bind on parcels that are still forested, the quantity of interest is its effect on a forest-to- k conversion, the sum $\beta_{3k} + \beta_{5,\text{forest},k}$. We report and test that sum with a joint Wald test (Table 1) rather than interpreting the two coefficients separately.

These coefficients are identified from a difference-in-differences contrast built into the choice model. The dummy turns on only for parcels that were Amazon forest in 2008, and only after

2008: the *bound* parcels, those the policies restrict. So β_{3k} and β_{5jk} measure how their clearing behavior shifts once the policies take effect, beyond what their rents and the general transition structure predict. That shift is identified against unbound parcels (outside the Amazon, or not forested in 2008) and against the pre-2008 period. The approach differs from the usual use of an estimated choice model, which fixes the coefficients and simulates a *hypothetical* policy by perturbing prices; here we estimate the *enacted* policy’s effect directly. Reading the coefficient as a causal effect requires a parallel-trends assumption: absent the policies, the clearing propensities of bound and unbound parcels would have moved together after 2008, conditional on rents.

The model is estimated by maximum likelihood on one randomly chosen year per pixel, which avoids temporal autocorrelation among a pixel’s repeated observations. The transition-probability expression and the remaining estimation details are given in Section A.2.

2.5 Simulation

We simulate land use forward from 2008 in five-year steps through 2023, under a policy scenario and a no-policy counterfactual. At each step we turn the estimated transition probabilities into a realized land use for every pixel by drawing from a uniform distribution and assigning the use whose probability interval contains the draw.

Prices and yields are endogenous. Reallocating land changes the production of crops and cattle, which moves prices — through a price flexibility, the inverse of a demand elasticity — and, through that price change, yields, via a price-yield response on the intensive margin. Because the new yields themselves alter production, we iterate these adjustments within each step until production stabilizes, then carry the resulting prices, yields, and rents into the next step. The iteration lets the simulation capture the equilibrium response of both demand and supply to the policy-induced change in land use. We quantify uncertainty by drawing 100 coefficient vectors

from the model’s estimated distribution and re-running the simulation for each — a Monte Carlo approach — reporting 2.5th–97.5th-percentile intervals.

These simulations underlie our land-use results; the welfare measure (Section 2.6) is computed separately on the observed 2008 data and does not use simulation output. From the policy-versus-counterfactual contrast we read avoided deforestation (the difference in Amazon forest area between the two scenarios) and spillover (the same difference for forest outside the Amazon). We also isolate induced intensification, the extra output drawn from land already farmed as policy-driven scarcity raises prices and producers respond by lifting yields. We measure it as a double difference: the policy’s land-use effect with the yield response switched on (the full response), less the same effect with it switched off, so prices still move but yields do not (the price-only response). Their difference, in hectares, is the part of the policy’s land-use effect that this yield response accounts for. The adjustment equations, parameter values, convergence properties, and scenario design are given in Section A.3.

2.6 Compensating Variation

Our welfare measure is the money-metric change in landowners’ welfare from the land-use choice, a producer-side analogue of the compensating variation. Because a valid compensating variation compares two states that differ *only* in the policy, we evaluate it on the observed 2008 data, holding land uses and rents fixed and simply switching the policy dummy on and off; it therefore does not draw on the forward simulation. The relevant value is not the utility of the one use a producer ends up in, but the expected utility of the best of all uses open to them. The nested logit delivers this in closed form as the logsum, computed from the value of every option (forest, pasture, and crops) (Small and Rosen 1981). For pixel i , the compensating variation is the difference in this logsum, converted to money by dividing by the marginal utility of profit β_2 , between the policy (1) and no-policy (0) settings:

$$CV_i = \frac{1}{\beta_2} \left[\ln \left(\sum_m \exp(\lambda_m IV_{im}^1) \right) - \ln \left(\sum_m \exp(\lambda_m IV_{im}^0) \right) \right],$$

where λ_m is the dissimilarity parameter of nest m and the inclusive value of a nest is

$$IV_{im} = \ln \left(\sum_{k \in m} \exp(V_{ik}/\lambda_m) \right).$$

The forest nest holds a single alternative, so it is degenerate. Its dissimilarity parameter is then not separately identified — trivially one in a utility-consistent nested logit — so we do not estimate it (Hunt 2000). With forest as the reference alternative ($V_{\text{forest}} = 0$), its inclusive value is $IV_{\text{forest}} = 0$, and the compensating variation reduces to

$$CV_i = \frac{1}{\beta_2} \left[\ln(1 + \exp(\lambda_{\text{cleared}} IV_{i,\text{cleared}}^1)) - \ln(1 + \exp(\lambda_{\text{cleared}} IV_{i,\text{cleared}}^0)) \right].$$

We sum CV_i across pixels and scale to a landscape total in US dollars. We report the aggregate compensating variation with a Monte Carlo confidence interval, both as a present value (perpetuity) and as the equivalent annual flow, along with the proportional welfare loss for the bound group (Amazon pixels forested in 2008). Because the comparison holds rents fixed and the policy dummy is switched on only for bound pixels, unbound Amazon pixels and the Cerrado are unaffected by construction.

2.7 Consumer Surplus

The compensating variation captures landowners. Commodity consumers are a second group, reached not through the policy dummy but through prices: lower output raises prices, leaving consumers worse off. Because this channel is active only where the policy moves equilibrium

quantities and prices, we measure it on the forward simulation rather than the base-year data (the base-year compensating variation holds prices fixed and so has no consumer-surplus content by construction). The simulation embeds an inverse world demand curve (the price flexibilities of Section A.3, scaled to Brazil’s share of world output). The policy moves output along it from the no-policy equilibrium to the policy equilibrium at the final simulated period. The change in world consumer surplus is minus the area to the left of that demand curve between the two equilibrium prices, computed per commodity and summed. We report two estimators — a trapezoid on the two simulated equilibria and an exact constant-elasticity form, which we take as our primary measure — with Monte Carlo confidence intervals; the formulas are given in Section A.4.

2.8 Value of Avoided Carbon

We value the forest the policy retains as avoided carbon emissions. Clearing forest for pasture or crops releases the difference between the carbon stored in standing forest and in the land use that replaces it; keeping the forest avoids that release. We use land-use carbon stocks (above and below ground) of 159 tonnes per hectare for forest, 53 for pasture, and 52 for cropland, derived in a companion analysis (Amaral-Santos et al. 2026) from biomass carbon density (Spawn et al. 2020) and soil organic carbon (MapBiomas Project 2023). Avoided clearing in the Amazon is the area that stays forest under the policy but would have been cleared for pasture or cropland without it. We take it at the final period of the forward simulation (Section 2.5). We apply the per-hectare stock difference to this area, convert the retained carbon to carbon dioxide, and value it at a social cost of carbon of 190 US dollars per tonne, a round value between two recent central estimates at a 2 percent near-term discount rate: 193 dollars (Environmental Protection Agency 2023) and 185 dollars (Rennert et al. 2022). Dividing the producer compensating variation (Section 2.6) by these same avoided emissions gives the break-even social cost of carbon: the value at which the carbon benefit offsets the producer cost.

3 Results

3.1 Model estimates

Table 1: Nested logit coefficient estimates (preferred specification). Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

	Estimate
intercept crops	5.045*** (0.051)
intercept pasture	3.779*** (0.140)
rent	0.299*** (0.027)
policy:crops	-1.813*** (0.510)
policy:pasture	-1.822*** (0.548)
starting forest:crops	-8.671*** (0.093)
starting forest:pasture	-6.419*** (0.134)
starting pasture:crops	-3.546*** (0.133)
starting pasture:pasture	-0.768*** (0.137)
policy:starting forest:crops	1.224* (0.512)
policy:starting forest:pasture	1.548** (0.548)
policy:starting pasture:crops	0.244 (0.514)
policy:starting pasture:pasture	0.060 (0.549)
iv:cleared	0.448*** (0.041)

Table 1 reports the coefficient estimates for the preferred specification. The rent coefficient is positive and significant, meaning that an increase in the per-hectare rent of a land use raises

the probability that a parcel transitions into that use, as economic theory predicts. The large negative coefficients on the starting-land-use terms show that producers rarely clear standing forest: a forested parcel is far less likely to become crops or pasture than an already-cleared parcel is to stay in agriculture, consistent with the high cost of clearing forest. The inclusive-value parameter for the cleared nest is 0.45. A value between 0 and 1 confirms that crops and pasture are closer substitutes than either is to forest, since they share unobserved determinants, and keeps the model consistent with utility maximization (Train 2009). At 1 the nest would collapse to a simple logit, and outside that range the model would be inconsistent with utility-maximizing behavior.

Our central interest is in the policy coefficients, which capture the effect of the post-2008 anti-deforestation policy package on bound parcels (Amazon pixels forested in 2008). These coefficients are read against two reference categories: forest as the ending use and crops as the starting use. The direct term (policy:crops, β_3) says that, for a parcel starting in crops, the policy makes staying in crops less likely than reverting to forest; the interaction (policy:starting forest:crops, β_5) adjusts that effect for a parcel starting in forest instead. Neither term on its own is the policy's effect on clearing. For a forest parcel, that effect is the sum of the two, $\beta_3 + \beta_5$ — the policy's effect on clearing forest to crops, and, from the matching pasture terms, forest to pasture. We therefore test the combined policy effect on each managed land use with a joint Wald test. The combined effect on conversion to crops is negative and statistically distinguishable from zero ($\chi^2(1) = 219.54, p < 0.001$), as is the combined effect on conversion to pasture ($\chi^2(1) = 294.15, p < 0.001$). In other words, the policies reduced the probability that a forested parcel was cleared for either crops or pasture, which is their intended effect.

3.2 Land-use outcomes

We simulate land use under the policy and under a no-policy counterfactual and report the difference (policy minus counterfactual) at the final simulated period, with Monte Carlo confidence intervals.

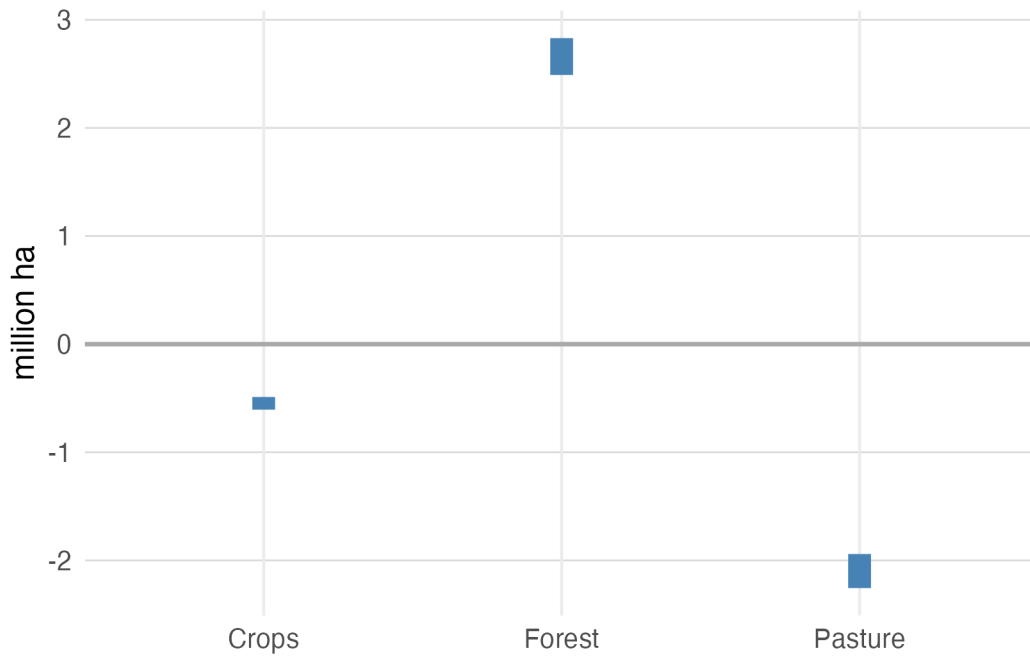


Figure 1: Effect of the anti-deforestation policies on land-use area in the Amazon (million ha). Policy minus no-policy counterfactual; thick bars are 95% Monte Carlo confidence intervals.

In the Amazon, the policies increased forest area by about 2.7 million hectares (Figure 1), a gain that is mirrored by reductions of roughly 2.1 million hectares of pasture and 0.6 million hectares of cropland; all three differences are significant. In proportional terms (Figure 2), forest is about 8% larger under the policies, while cropland is about 9% smaller and pasture about 8% smaller. The forest gain is thus drawn predominantly from avoided pasture expansion in absolute terms, even though cropland is the most affected use proportionally. This pattern is consistent with the policies reducing forest-to-pasture and forest-to-crop conversions on bound parcels.

The Cerrado tells the opposite story (Figure 3): the estimated area effects for all three uses are

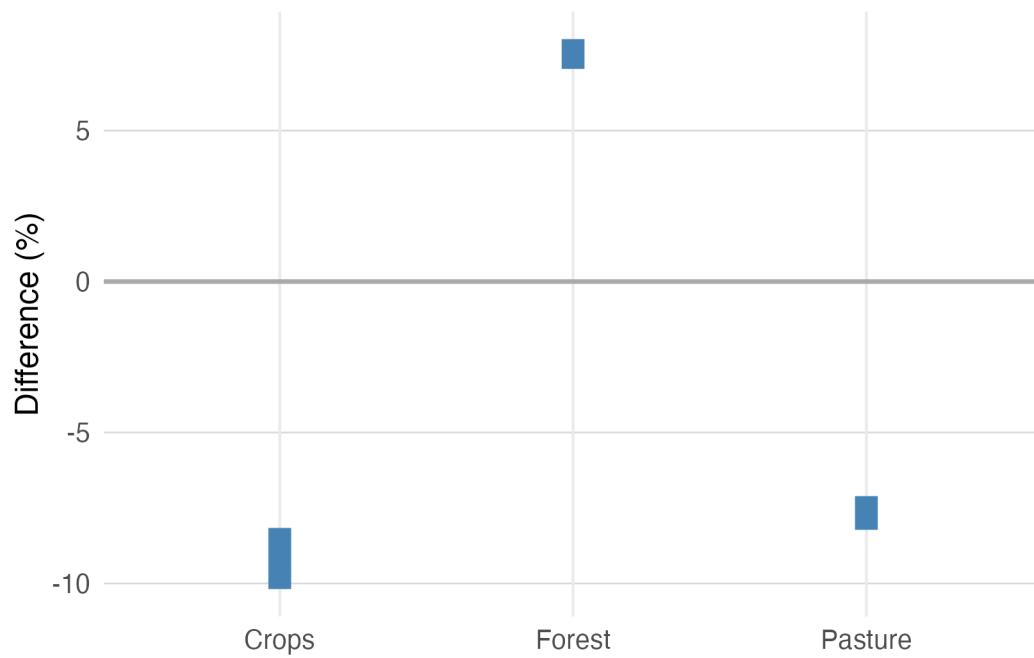


Figure 2: Proportional effect on land-use area in the Amazon (%). Policy minus counterfactual, relative to the counterfactual area; 95% Monte Carlo confidence intervals.

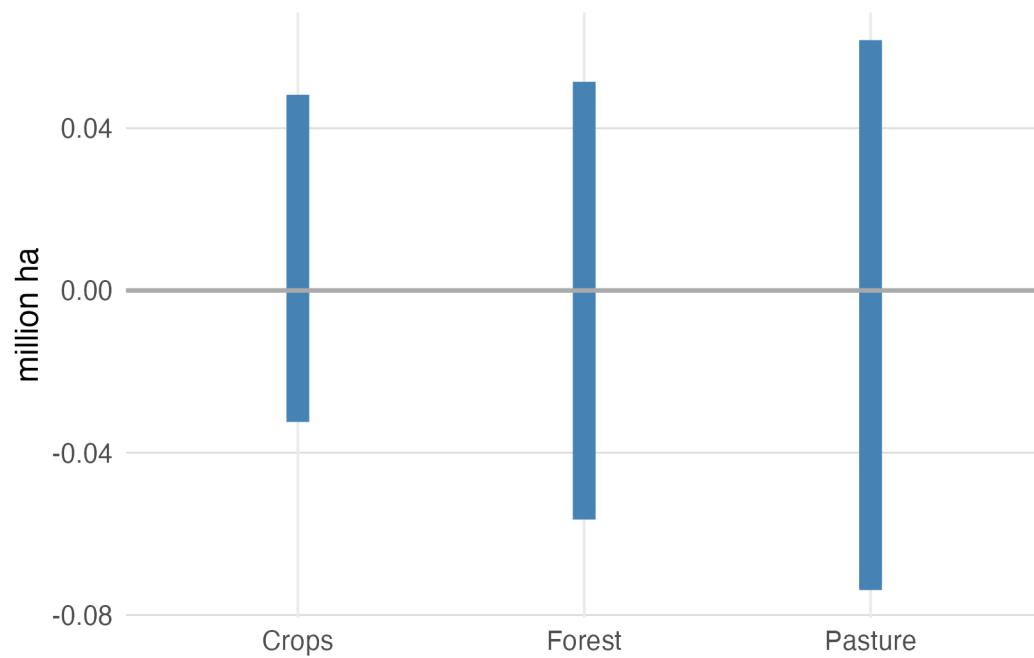


Figure 3: Effect of the anti-deforestation policies on land-use area in the Cerrado (million ha). Policy minus counterfactual; 95% Monte Carlo confidence intervals.

small and their confidence intervals span zero, so we cannot distinguish them from no effect. The policies have no direct effect in the Cerrado by construction, since they apply only inside the Amazon biome. The absence of any indirect effect indicates negligible cross-biome leakage. For crops, the Amazon production cut is too small to move prices appreciably, so nothing pulls Cerrado land into production. For cattle, the beef price does rise, but Cerrado pasture area is almost unresponsive to price, so again no additional land is drawn in.

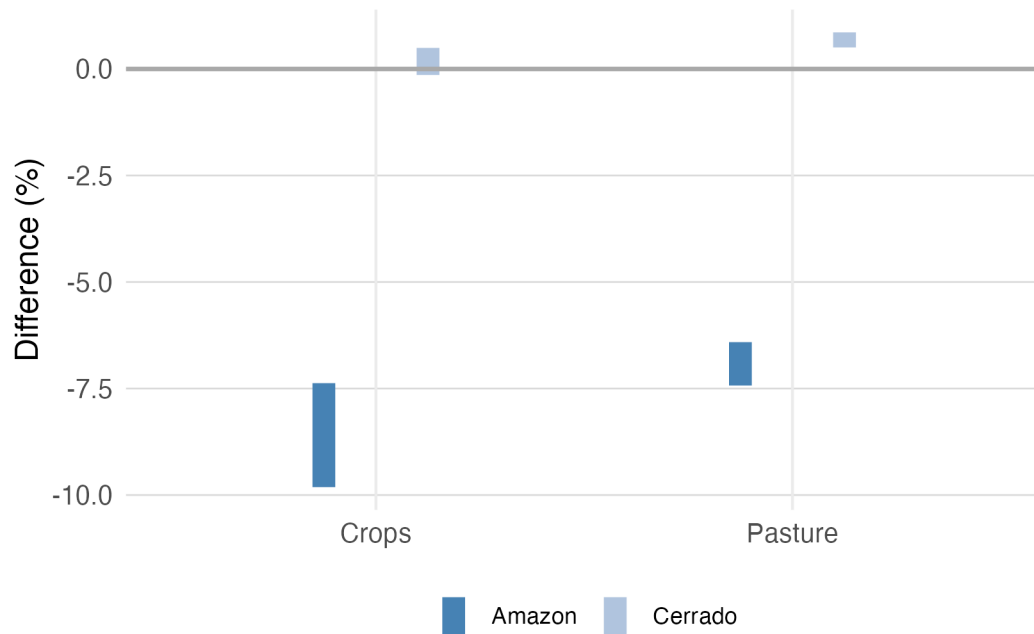


Figure 4: Proportional change in commodity production (%) by biome under the policy (demand destruction). Policy minus counterfactual; 95% Monte Carlo confidence intervals.

The reduction in agricultural land lowers commodity production (Figure 4). In the Amazon, crop production is about 9% lower and cattle production about 7% lower under the policies than under the counterfactual, while the corresponding changes in the Cerrado are close to zero. Because output falls, prices rise, and this is the channel through which the policies reach commodity consumers; we quantify the resulting loss in world consumer surplus below.

We also report induced intensification — the hectares of the policy’s land-use effect attributable to producers raising yields on land already farmed (Section 2.5) — in Table 2. Only the forest

Table 2: Induced intensification in the Amazon: double difference in land-use area (ha), isolating the yield-elasticity contribution to the policy effect. 95% Monte Carlo confidence intervals.

Land use	Double difference (ha)	Lower CI	Upper CI	% of counterfactual
Forest	2,200	495	3,000	0
Crops	1,400	-2,400	3,105	0
Pasture	-3,600	-5,305	905	0

effect is marginally significant, and all three are negligible relative to the headline area effects above, so induced intensification is a minor channel here.

3.3 Welfare cost

We measure the welfare cost to landowners as the base-year (2008) compensating variation: the difference in the monetized expected utility (logsum) of the land-use choice with the policy on versus off. Both are evaluated on the observed data, so the two scenarios are compared at the same rents and starting land uses. The total compensating variation is -45.1 bn USD (95% CI -55.8, -37.5), a present-value (perpetuity) welfare loss equivalent to an annual flow of 2.9 bn USD. Because rent here nets out transport but not other production costs, this is an upper-bound estimate; we quantify how much that matters below and discuss it further in the Discussion.

The loss falls entirely on the bound group — Amazon pixels forested in 2008, where the policy restricts clearing. Because the base-year compensating variation holds prices fixed, the policy's only welfare channel runs through the policy dummy, which is switched on only for these pixels. Unbound Amazon pixels and the Cerrado are unaffected by construction: any effect on them would operate through prices, captured separately by the consumer-surplus result below. In proportional terms, the policy reduces the bound group's welfare by 25% (95% CI 24, 28%), relative to its no-policy opportunity value.

Rescaling rents to an approximate net return lowers the dollar cost but leaves its proportional size intact. Applying crop and cattle margins of 43% and 38% (Section A.1.3), the compensating variation becomes -19.5 bn USD (95% CI -23.5, -16.1), about 0.43 times the gross figure. The proportional loss to the bound group is 25.3%, against 25.4% on gross rents — a difference too small to read as a change in behavior. The adjustment thus changes what the loss is worth, not how much of their opportunity value the bound group gives up. This check covers the producer compensating variation only: the consumer surplus and carbon figures that follow rest on the forward simulation, which we do not re-run on net returns.

The compensating variation captures landowners only. Commodity consumers are reached through prices, which move in the forward simulation (Section 2.5): the policy's lower output raises prices along a world demand curve — the quantity and price changes consumer surplus requires. This reduces world consumer surplus by 6.7 bn USD, integrating the constant-elasticity demand curve the simulation uses; a functional-form-agnostic trapezoid gives a smaller 3.0 bn USD (Section 2.7).

Set against these costs is the value of avoided deforestation. The carbon retained by the policy in the Amazon is worth 196.9 bn USD at a social cost of carbon of 190 USD/t CO₂ (Section 2.8). Equivalently, the producer compensating variation is offset by the carbon value at a break-even social cost of carbon of 43 USD/t CO₂.

Taken together, these components imply that the policies impose a substantial private cost while delivering a much larger global benefit. The producer compensating variation (45.1 bn USD) and the consumer surplus loss (6.7 bn USD, or 3.0 bn USD by the trapezoid estimator) are both real welfare losses concentrated on producers in the bound group and on commodity consumers worldwide. Set against them, the value of the avoided-deforestation carbon (197 bn USD at 190 USD/t CO₂) exceeds the combined producer and consumer losses by roughly fourfold. On these components the policies are welfare-improving in aggregate, and remain so for any social cost of carbon above the break-even value of 43 USD/t CO₂. This aggregate gain, however,

masks a sharp distributional divide: the costs fall on a specific group of landowners while the carbon benefits accrue globally. The accounting here is also partial: it omits the value of avoided biodiversity loss and other ecosystem services from forest retention, as well as any administrative and enforcement costs of the policies. Each component the accounting includes — the producer compensating variation, the consumer surplus loss, and the carbon value — is an appropriate welfare measure; the accounting is incomplete in what it covers, not in how those components are measured. The cost-benefit comparison should therefore be read as one across the components we can monetize, not a complete social welfare calculation. The omitted administrative and enforcement costs are minor next to these welfare flows: federal spending on preventing and controlling deforestation and fires was about 72 million US dollars in 2023 (Climate Policy Initiative 2024).

4 Discussion

This paper measures what a package of Brazil’s anti-deforestation policies — the Amazon Soy Moratorium, the cattle agreements, and the 2012 Forest Code — costs the producers it binds. It lowers their monetized welfare by about 25 percent, on the order of 45 billion US dollars in present value, or roughly 3 billion dollars a year in perpetuity. Two strands of prior work bear on the cost of conservation, but neither measures what we do. One tracks the policies’ realized effects — clearable area, farmgate prices, local incomes — but is not a welfare measure. The other is a welfare measure, but from structural models of idealized instruments such as an optimal deforestation tax, not the enacted package. To our knowledge ours is the first to combine the two: a welfare measure of the package as it actually bound producers, measured after the fact.

Three careful studies of the policies’ realized effects each track a different margin, and none finds a large loss, which would seem to put a low cost on the package. Rausch and Gibbs

(2021) measure the forest that the soy moratorium keeps out of soy production, though it is soy-suitable and legally clearable under the Forest Code. They find it small: about 49,000 hectares on existing soy farms, roughly one percent of Amazon soy area, and concentrated on the few that hold such land. Barrozo and Skidmore (2026) find no measurable decline in average farmgate soy prices in the Amazon relative to the Cerrado, though that municipal average can mask a decline for the excluded farmers offset by gains for compliant ones. Merkus (2024) finds that tightened enforcement of the Forest Code raised local incomes rather than lowering them. But even a muted or favorable realized effect is consistent with the large welfare cost we find. The policy lowers the value of clearing a bound parcel for crops and for pasture, so the land is worth less to its owner even when the policy does not change the use it is put to. A producer who would clear for pasture with or without the policy is a case in point: that conversion is worth less under the policy, yet a comparison of realized land use would show no change. These studies also each isolate a single instrument, the soy moratorium in two cases and enforcement of the Forest Code in the third, whereas our cost is that of the whole package, so part of the difference is a wider scope, not only a different measure. And because our estimate is built on the imperfect compliance actually observed rather than on a full ban, it is if anything a conservative measure of the burden a fully enforced package would impose.

The structural welfare models each estimate the cost of an efficient instrument — such as an optimal tax — that conserves far more forest than the package does. The costs they report to producers span a wide range, and not all are welfare losses. Souza-Rodrigues (2019) finds an optimal land-use tax that raises forest cover from 44 to 80 percent of private land at about 480 million dollars a year in lost farmer surplus, below our cost. Araujo, Costa, and Sant’Anna (2026) find that a first-best carbon tax foregoes some 356 billion dollars in private profits, far above ours, to add roughly 120 million hectares of forest. Assunção et al. (2026) put the loss at about 345 billion dollars, though as agricultural output rather than surplus, for a large-scale reforestation that returns nearly all the biome’s farmland, about 15 percent of its area, to forest. Set against our 2.7 million hectares of avoided deforestation, these are the costs of conserving

vastly different amounts of forest, and do not line up with our estimate.

The cost we measure is the cost of a blunt instrument: the soy moratorium and cattle agreements withdraw market access from Amazon land cleared after 2008, and the Forest Code sanctions illegal clearing, each binding regardless of a parcel's rent. Because it ignores rent, the instrument blocks even the most valuable conversions — the ones a price instrument, which lets producers pay to clear, would allow. The package is command-and-control in economic form, even though its moratoria work through private markets rather than the state; in Souza-Rodrigues's model, reaching a forest target that way costs about eight times more than the equivalent tax. An instrument that spared the highest-value parcels would likely secure a similar forest gain at far lower cost to producers.

Our compensating variation (Section 2.6) is static, holding prices and yields at their 2008 values, so it omits the general-equilibrium adjustments the policy sets off. Two such offsets could in principle shrink the welfare loss: the higher market price the policy's supply cut brings to producers who keep selling, and the price-mediated intensification of land already farmed. Our forward simulation (Section 2.5) lets prices and yields adjust. The price-mediated intensification of existing land is small for both crops and cattle. The price gain is small for crops: Amazon crop production falls about 9 percent (Figure 4), but the Amazon supplies only about a seventh of the market's crops, so this is barely one percent of total supply, and the crop price rises less than one percent. Cattle is more concentrated in the Amazon — roughly two-fifths of supply — and its demand is far less elastic, so the 7 percent Amazon cut, about 2.6 percent of total supply, raises the beef price roughly 5 percent. That gain to cattle producers who keep selling is not negligible. The static measure is therefore close to the general-equilibrium loss for crops, and somewhat overstates it for cattle. A second intensification channel lies outside the model, one that needs no price change: a producer blocked from clearing may work existing land harder as its shadow value rises. Stocking-rate studies document this response where enforcement has tightened (Koch et al. 2019; Garrett et al. 2018), and the cattle agreements

produced it too, raising productivity in place of clearing (Da Mata, Dotta, and Severnini 2026). Rational producers intensify only where the gain exceeds the added cost, so this response recovers part of the welfare loss. Because our estimate leaves that recovery out, it overstates the true loss and is to that extent an upper bound.

On the benefit side, the package retains about 2.7 million hectares of Amazon forest relative to a no-policy counterfactual. Valued as avoided carbon at a social cost of carbon of 190 dollars per tonne, that forest is worth on the order of 200 billion US dollars. The retained area is of the same order of magnitude as the largest single-instrument estimate in the literature: a triple-difference evaluation credits the soy moratorium alone with roughly 1.8 million hectares of avoided clearing over 2006 to 2016 (Heilmayr et al. 2020). The two designs rest on different identifying assumptions and were built for different purposes, so the agreement is reassuring.

The carbon value of the avoided deforestation is real only if the clearing is genuinely avoided, not relocated elsewhere. Whether it is depends on leakage, and we find little. Production is not displaced across the biome border into the Cerrado. In our model, two forces hold production inside the Amazon. Crop prices barely move, so no crop land is pulled across into the Cerrado. The beef price does rise, but Cerrado pasture area is almost unresponsive to price, so no pasture follows either. Beyond our study region the crop argument still holds — a price that scarcely moves cannot drive international crop displacement. The cattle argument does not: the beef price rises about 5 percent, which could induce production elsewhere, and our single-region sample cannot bound that response. International cattle leakage is thus a channel we cannot rule out, one a multi-region model would be needed to quantify. Most evaluations of the soy moratorium's leakage track direct, parcel-level displacement: soy relocating onto already-cleared pasture, largely without new clearing (Moffette and Gibbs 2021). Ours instead runs through prices, so the apt comparison is with equilibrium models of that channel. Using a computable general equilibrium model, Villoria et al. (2022) report that about half of the deforestation the soy moratorium avoids reappears elsewhere in Brazil. That rate follows from how

the policy is represented there: a complete halt on conversion, but only in the municipalities where committed traders dominate, which leaves uncovered Amazon forest as a low-cost, unrestrained margin to absorb the displaced production. We instead apply an empirically estimated policy effect across all Amazon forest standing in 2008, removing that low-cost local margin, so displaced production finds no unrestrained forest to absorb it.

The welfare measure we provide lets us compare the producer cost and the carbon benefit on a single scale. In carbon terms, the producer cost is about 43 dollars for every tonne of carbon dioxide the package keeps from the atmosphere. The carbon benefit therefore exceeds the producer cost for any social cost of carbon above that figure, far below the roughly 190 dollars at which recent authoritative estimates place it (Section 2.8), so by the standard efficiency test the package is justified. Araujo, Costa, and Sant’Anna (2026) and Assunção et al. (2026) reach the same verdict for their idealized instruments, the former with net social gains above 1.6 trillion dollars, the latter with carbon payments that more than repay the agricultural income foregone. But the test weighs only the totals. It is silent on who gains and who loses, and here they are different groups. The benefit is diffuse and largely global; the cost is steep and falls on a specific group of frontier landowners.

Our cost measure is the size of the transfer that would make bound producers whole. Not all of it, though, is a transfer anyone would be willing to make. Some of the producers the package binds cleared forest in breach of the Forest Code: in Mato Grosso, at least 627 soy properties cleared illegally while the ASM was in force (Gibbs et al. 2015). The ASM withdraws market access according to when land was cleared, not according to whether the clearing was lawful (Gibbs et al. 2015), so it reaches them too. Paying them would mean compensating the loss of an option they never lawfully held. The schemes negotiated for the Cerrado in the late 2010s drew exactly this line. Soy traders and their buyers proposed a conversion-free sourcing agreement modelled on the ASM, to be signed together with a scheme paying farmers to conserve native vegetation. Payments would go only to farmers whose vegetation exceeds the legal re-

serve requirement and who hold a permit to clear it (Bastos Lima and Persson 2020; Garrett et al. 2022). Compensation, on that design, is for a legal option given up, not an illegal one foreclosed. That narrows what a transfer would have to cover, though we cannot say by how much. Rausch and Gibbs (2021) come closest to measuring the legal option, and their evidence points two ways. Of roughly 14,600 soy farms in the Amazon, only 185 held forest that was both suitable for soy and clearable under the Forest Code — the 49,000 hectares noted above. For an incumbent soy producer, then, the foreclosed legal option is tiny. But the same study counts 3.6 million hectares of legally clearable, soy-suitable forest on properties that grow no soy, some seventy times as much. Our cost is measured over land rather than over farms, so that area falls inside it. Our policy variable does not distinguish parcels by their legal-reserve status, and the evidence here covers soy rather than cattle.

The Cerrado negotiations also show what a workable transfer would have to get right, and where the attempt has so far failed. That the sourcing agreement was to be conditional on the payment scheme is itself evidence for our argument: the negotiators treated compensation as the price of restricting legal clearing (Bastos Lima and Persson 2020). Compensation was not enough to win producers over, though. Brazil's main soy farmer association left the talks rather than consider any restriction on legal clearing, paid for or not (Bastos Lima and Persson 2020; Garrett et al. 2022). But tying eligibility to a clearing permit gives farmers with surplus vegetation a reason to seek permits they had not previously wanted, so as to be paid for not using them — a key reason Greenpeace gave for leaving the negotiations (Bastos Lima and Persson 2020). Five-year contracts leave the vegetation exposed as soon as they lapse, where the Forest Code's little-used environmental servitude would fix its status for fifteen years or more (Bastos Lima and Persson 2020). And paying soy farmers alone excludes the communities and smallholders who have conserved most, which is both inequitable and corrosive of the scheme's standing (Bastos Lima and Persson 2020; Garrett et al. 2022). Garrett et al. (2022) conclude that payments work best not on their own but alongside a Cerrado-wide market exclusion. They would go to the poorest soy households, and would pay for restoring cleared land as well as for forgone

clearing. Better still, they argue, is to run the payments at the level of states or municipalities, so that every land user is eligible and not soy farmers alone. Restoration is also the instrument that fits the producer who cleared illegally, because it finances a return to compliance rather than paying for the breach.

One caution qualifies all of this. That the beneficiaries could cover the cost is a claim about the carbon value, not a claim about anyone's willingness to pay it. Garrett et al. (2022) price a payment scheme for Cerrado soy at more than 2.6 billion dollars a year to companies and the federal government. The food companies that pledged support for the Cerrado Manifesto set a fundraising goal of 50 million a year, which at best would hold about 2.5 percent of the most suitable legally clearable area for five years. The distance between the cost of compensation and the funding goal is the practical obstacle to the arrangement our results point toward. The gap has a familiar cause. Climate stability is a global public good, so no single company or government captures enough of the gain to justify covering the cost on its own, and each does better waiting for others to pay — which is why a benefit worth several times the cost can still go unfunded.

We note several limitations. Because the three instruments share the post-2008 Amazon cut-off and are collinear in the data, we estimate their combined effect and cannot decompose it. Our results should not be read as evidence on any single instrument, and separating them would require variation the enacted policies do not provide. The same holds for the enforcement that stands behind them. From 2008 the government concentrated enforcement on the worst-offending municipalities and made rural credit in the Amazon biome conditional on environmental compliance (Assunção and Rocha 2019; Assunção, Gandour, and Rocha 2015). What we measure is therefore the cost of the whole post-2008 regime in the Amazon, not of the three named instruments on their own. This bears on how our figure should be read against the challenges now facing the ASM. Were it struck down, the Forest Code, the property registry and federal enforcement would all remain, and producers would recover only part of what we

measure. Our estimate is the cost of the regime, and an upper bound on what removing any single piece of it would give back.

The leakage and intensification results depend on price flexibilities we adopt from the literature rather than estimate. The intensification result depends further on the yield response to prices, which for cattle we set equal to the crop response, though limited evidence suggests cattle may respond more strongly (Nascimento et al. 2019; Skidmore et al. 2022). Because the beef price moves materially, a stronger cattle yield response would raise the intensification we find in rough proportion; what bounds it is the assumed yield response, not the price. Direct estimates of the price flexibilities and yield responses would put these results on firmer empirical footing. Our leakage result also treats the policy as a uniform disincentive across every parcel of Amazon forest standing in 2008, whereas real coverage and enforcement vary in space. Where they are weaker, lightly bound forest could absorb displaced clearing much as the uncovered municipalities do in Villoria et al. (2022), so within-Amazon relocation could be larger and our net forest gain smaller. Leakage can also occur outside our study region (the Amazon–Cerrado band within 600 km of the Amazon biome boundary) if displaced production resurfaces in agricultural areas elsewhere in Brazil or abroad. If this is the case, our net forest gain would be an overstatement. Because that leakage runs through world commodity prices, capturing it would require a multi-region equilibrium model of agricultural markets, in which prices and land use adjust jointly across regions rather than within a single one.

Our compensating variation is also static. As we argue above from the forward simulation, the general-equilibrium offsets it leaves out are small for crops but non-trivial for cattle, where the beef-price gain is real, so the measure is close to the net cost for crops and somewhat overstates it for cattle. But because it holds 2008 values fixed, it sits on a different timeline from the carbon benefit, which accrues over the simulated policy periods. A dynamic, forward-looking model, in which producers anticipate the price and yield paths the policy sets off, would fold in the offsets and place cost and benefit on a common timeline. The rent behind the welfare cost

is also incomplete on the cost side: it subtracts transport but not the other costs of production, so it measures a gross rather than a net return. This overstates how much the policy lowers the value of bound land and adds to the reasons the cost is an upper bound. Rescaling rents by crop and livestock profit margins cuts the dollar cost to roughly two fifths of the figure we report, but leaves the proportional loss to the bound group almost exactly where it was. What the gross measure gets wrong, then, is what the foregone options are worth, not how large a share of their value the policy takes away. We keep the gross estimate as our headline because the margins rest on a small number of farm budgets from particular states and years, and because the simulation behind the carbon and consumer-surplus results runs on gross rents, so only the producer cost can be restated consistently. Finally, we value the environmental benefit through avoided carbon alone, at a single social cost of carbon, leaving unpriced the biodiversity and other non-carbon benefits that motivate the policies. The benefit we report is thus a lower bound, and pricing those benefits would only widen its margin over the producer cost.

The package of anti-deforestation policies we study retains a large area of forest at a real and concentrated cost to the producers it binds. The benefits already outweigh the costs in total; what will decide the policy's future is not that balance but how the costs and benefits are shared. Concentrated costs of this kind make a policy politically fragile, and that fragility is already driving the effort to suspend the soy moratorium. But the compensation that would make the bound producers whole is smaller than the benefit their forest secures for everyone else, and the part of it anyone would actually owe — what the law itself allowed producers to do — is smaller again. Those who gain could cover it and still be better off than with no policy at all. So far they have not offered to. Setting the cost and the benefit side by side takes a welfare measure, which the realized-effect studies do not provide and the structural models provide only for idealized instruments, not the policy as enacted. Measuring the cost is the first step toward sharing it.

References

- Adjemian, Michael K., and Aaron Smith. 2012. "Using USDA Forecasts to Estimate the Price Flexibility of Demand for Agricultural Commodities." *American Journal of Agricultural Economics* 94 (4): 978–95. <https://doi.org/10.1093/ajae/aas032>.
- Alix-Garcia, Jennifer, and Holly K. Gibbs. 2017. "Forest Conservation Effects of Brazil's Zero Deforestation Cattle Agreements Undermined by Leakage." *Global Environmental Change* 47 (November): 201–17. <https://doi.org/10.1016/j.gloenvcha.2017.08.009>.
- Amaral-Santos, Roberto, Linus Blomqvist, Laurel Abowd, Jason Jon Benedict, Catharina Latka, Kathy Baylis, Robert Heilmayr, and Andrew J. Plantinga. 2026. "Better Together? Impacts and Coordination of Terrestrial Carbon Sequestration Initiatives."
- Ando, Amy, Jeffrey Camm, Stephen Polasky, and Andrew Solow. 1998. "Species Distributions, Land Values, and Efficient Conservation." *Science* 279 (5359): 2126–28. <https://doi.org/10.1126/science.279.5359.2126>.
- Araujo, Rafael, Francisco Costa, and Marcelo Sant'Anna. 2026. "Efficient Conservation of the Brazilian Amazon: Estimates from a Dynamic Model." *The Review of Economic Studies* 93 (1): 72–105. <https://doi.org/10.1093/restud/rdaf031>.
- Assunção, Juliano, Clarissa Gandour, and Romero Rocha. 2023. "DETER-ing Deforestation in the Amazon: Environmental Monitoring and Law Enforcement." *American Economic Journal: Applied Economics* 15 (2): 125–56. <https://doi.org/10.1257/app.20200196>.
- Assunção, Juliano, Clarissa Gandour, and Rudi Rocha. 2015. "Deforestation Slowdown in the Brazilian Amazon: Prices or Policies?" *Environment and Development Economics* 20 (6): 697–722. <https://doi.org/10.1017/S1355770X15000078>.
- Assunção, Juliano, Lars Peter Hansen, Todd Munson, and José A Scheinkman. 2026. "Carbon Prices, Forest Conservation and Reforestation in the Brazilian Amazon."
- Assunção, Juliano, and Romero Rocha. 2019. "Getting Greener by Going Black: The Effect of Blacklisting Municipalities on Amazon Deforestation." *Environment and Development Economics* 24 (2): 115–37. <https://doi.org/10.1017/S1355770X18000499>.
- Auffhammer, Maximilian, Maya Duru, Edward Rubin, and David L. Sunding. 2020. "The Economic Impact of Critical-Habitat Designation: Evidence from Vacant-Land Transactions." *Land Economics* 96 (2): 188–206. <https://doi.org/10.3368/le.96.2.188>.
- Azevedo, Andrea A., Raoni Rajão, Marcelo A. Costa, Marcelo C. C. Stabile, Marcia N. Macedo, Tiago N. P. Dos Reis, Ane Alencar, Britaldo S. Soares-Filho, and Rayane Pacheco. 2017. "Limits of Brazil's Forest Code as a Means to End Illegal Deforestation." *Proceedings of the National Academy of Sciences* 114 (29): 7653–58. <https://doi.org/10.1073/pnas.1604768114>.
- Barona, Elizabeth, Navin Ramankutty, Glenn Hyman, and Oliver T Coomes. 2010. "The Role of Pasture and Soybean in Deforestation of the Brazilian Amazon." *Environmental Research Letters* 5 (2): 024002. <https://doi.org/10.1088/1748-9326/5/2/024002>.
- Barrozo, Marcos, and Marin Skidmore. 2026. "Are Sustainability Pledges Anticompetitive?"
- Bastos Lima, Mairon G., and U. Martin Persson. 2020. "Commodity-Centric Landscape Governance as a Double-Edged Sword: The Case of Soy and the Cerrado Working Group in Brazil." *Frontiers in Forests and Global Change* 3 (March): 27. <https://doi.org/10.3389/ffgc.2020.00027>.

- Berry, Steven, and Wolfram Schlenker. 2011. “Technical Report for the ICCT: Empirical Evidence on Crop Yield Elasticities.”
- Booth, Hollie, Dale Squires, Irfan Yulianto, Benaya Simeon, Muhsin, Luky Adrianto, and Eleanor Jane Milner-Gulland. 2021. “Estimating Economic Losses to Small-Scale Fishers from Shark Conservation: A Hedonic Price Analysis.” *Conservation Science and Practice* 3 (9): e494. <https://doi.org/10.1111/csp2.494>.
- Börner, J, and S Wunder. 2008. “Paying for Avoided Deforestation in the Brazilian Amazon: From Cost Assessment to Scheme Design.” *International Forestry Review* 10 (3): 496–511. <https://doi.org/10.1505/for.10.3.496>.
- Bošković, Branko, and Linda Nøstbakken. 2017. “The Cost of Endangered Species Protection: Evidence from Auctions for Natural Resources.” *Journal of Environmental Economics and Management* 81 (January): 174–92. <https://doi.org/10.1016/j.jeem.2016.08.003>.
- Brazilian Government. 2024. “Brazil’s NDC: National Determination to Contribute and Transform.”
- Cassauwers, Tom. 2025. “Brazil’s Soya Moratorium Slowed Amazon Deforestation. Now It’s Challenged.” *AlJazeera*. <https://www.aljazeera.com/economy/2025/11/6/brazils-soya-moratorium-slowed-amazon-deforestation-now-its-challenged>.
- Climate Policy Initiative. 2024. “Landscape of Climate Finance for Land Use in Brazil 2021–2023.” Climate Policy Initiative.
- CNA SENAR. 2024. “Custo de Produção Da Bovinocultura de Corte (Recría e Engorda), Em Colorado Do Oeste/RO.” CNA SENAR.
- Conference of the Parties to the Convention on Biological Diversity. 2022. “Decision Adopted by the Conference of the Parties to the Convention on Biological Diversity.”
- Da Mata, Daniel, Mario Dotta, and Edison R. Severnini. 2026. “Industrial Activity, State Capacity, and Deforestation: Evidence from Brazil.” NBER.
- de Santana, Antonio Cordeiro. 1999. “Recent Changes in the Relations of Brazilian Meat Demand.” *Brazilian Review of Agricultural Economics and Rural Sociology* 37 (2).
- den Braber, Bowry, Johan A. Oldekop, Katie Devenish, Javier Godar, Christoph Nolte, Marina Schmoeller, and Karl L. Evans. 2024. “Socio-Economic and Environmental Trade-Offs in Amazonian Protected Areas and Indigenous Territories Revealed by Assessing Competing Land Uses.” *Nature Ecology & Evolution* 8 (8): 1482–92. <https://doi.org/10.1038/s41559-024-02458-w>.
- Dos Reis, Júlio César, Mariana Y. T. Kamoi, Daniel Latorraca, Rafael F. F. Chen, Miqueias Michetti, Flávio Jesus Wruck, Rachael D. Garrett, Judson Ferreira Valentim, Renato De Aragão Ribeiro Rodrigues, and Saulo Rodrigues-Filho. 2020. “Assessing the Economic Viability of Integrated Crop-Livestock Systems in Mato Grosso, Brazil.” *Renewable Agriculture and Food Systems* 35 (6): 631–42. <https://doi.org/10.1017/S1742170519000280>.
- Dundas, Steven J., and Roger H. Von Haefen. 2020. “The Effects of Weather on Recreational Fishing Demand and Adaptation: Implications for a Changing Climate.” *Journal of the Association of Environmental and Resource Economists* 7 (2): 209–42. <https://doi.org/10.1086/706343>.
- Environmental Protection Agency. 2023. “Report on the Social Cost of Greenhouse Gases: Estimates Incorporating Recent Scientific Advances.”
- ESALQ. 2026. “Agricultural Prices.” <https://www.cepea.org.br/en>.
- FAO. 2022. “GAEZ Data Portal.” <https://gaez.fao.org/pages/data-viewer>.

- . 2026. “FAOSTAT.” <https://www.fao.org/faostat>.
- Fearnside, Philip M. 2005. “Deforestation in Brazilian Amazonia: History, Rates, and Consequences.” *Conservation Biology* 19 (3): 680–88. <https://doi.org/10.1111/j.1523-1739.2005.00697.x>.
- Ferrante, Lucas, and Philip M Fearnside. 2019. “Brazil’s New President and ‘Ruralists’ Threaten Amazonia’s Environment, Traditional Peoples and the Global Climate.” *Environmental Conservation* 46 (4): 261–63. <https://doi.org/10.1017/S0376892919000213>.
- Ferraro, Paul J., Merlin M. Hanauer, Daniela A. Miteva, Joanna L. Nelson, Subhrendu K. Pattanayak, Christoph Nolte, and Katharine R. E. Sims. 2015. “Estimating the Impacts of Conservation on Ecosystem Services and Poverty by Integrating Modeling and Evaluation.” *Proceedings of the National Academy of Sciences* 112 (24): 7420–25. <https://doi.org/10.1073/pnas.1406487112>.
- Garrett, R. D., J. Grabs, F. Cammelli, F. Gollnow, and S. A. Levy. 2022. “Should Payments for Environmental Services Be Used to Implement Zero-Deforestation Supply Chain Policies? The Case of Soy in the Brazilian Cerrado.” *World Development* 152 (April): 105814. <https://doi.org/10.1016/j.worlddev.2022.105814>.
- Garrett, R. D., I. Koh, E. F. Lambin, Yann le Polain de Waroux, J. H. Kastens, and J. C. Brown. 2018. “Intensification in Agriculture-Forest Frontiers: Land Use Responses to Development and Conservation Policies in Brazil.” *Global Environmental Change* 53 (November): 233–43. <https://doi.org/10.1016/j.gloenvcha.2018.09.011>.
- Gatti, Luciana V., Luana S. Basso, John B. Miller, Manuel Gloor, Lucas Gatti Domingues, Henrique L. G. Cassol, Graciela Tejada, et al. 2021. “Amazonia as a Carbon Source Linked to Deforestation and Climate Change.” *Nature* 595 (7867): 388–93. <https://doi.org/10.1038/s41586-021-03629-6>.
- Gibbs, Holly K., L. Rausch, J. Munger, I. Schelly, D. C. Morton, P. Noojipady, B. Soares-Filho, P. Barreto, L. Micol, and N. F. Walker. 2015. “Brazil’s Soy Moratorium.” *Science* 347 (6220): 377–78.
- Hanbury, Shanna. 2025. “Brazil Soy Deal That Curbs Amazon Deforestation to Be Suspended in 2026.” *Mongabay*. <https://news.mongabay.com/short-article/2025/10/brazil-soy-deal-that-curbs-amazon-deforestation-to-be-suspended-in-2026/>.
- Heilmayr, Robert, Cristian Echeverría, and Eric F. Lambin. 2020. “Impacts of Chilean Forest Subsidies on Forest Cover, Carbon and Biodiversity.” *Nature Sustainability* 3 (9): 701–9. <https://doi.org/10.1038/s41893-020-0547-0>.
- Heilmayr, Robert, Lisa Rausch, Jacob Munger, and Holly Gibbs. 2020. “Brazil’s Amazon Soy Moratorium Reduced Deforestation.” *Nature Food* 1: 801–10.
- Hunt, Gary L. 2000. “Alternative Nested Logit Model Structures and the Special Case of Partial Degeneracy.” *Journal of Regional Science* 40 (1): 89–113. <https://doi.org/10.1111/0022-4146.00166>.
- IBGE. 2022. “Produção Agrícola Municipal.” <https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9117-producao-agricola-municipal-culturas-temporarias-e-permanentes.html>.
- . 2026. “Pesquisa Trimestral Do Abate de Animais.” <https://sidra.ibge.gov.br/tabela/1092>.
- Jenkins, Clinton N., Stuart L. Pimm, and Lucas N. Joppa. 2013. “Global Patterns of Terrestrial Vertebrate Diversity and Conservation.” *Proceedings of the National Academy of Sciences*

- 110 (28). <https://doi.org/10.1073/pnas.1302251110>.
- Koch, Nicolas, Erasmus K H J zu Ermgassen, Johanna Wehkamp, Francisco J B Oliveira Filho, and Gregor Schwerhoff. 2019. “Agricultural Productivity and Forest Conservation: Evidence from the Brazilian Amazon.” *American Journal of Agricultural Economics* 101 (3): 919–40. <https://doi.org/10.1093/ajae/aay110>.
- Lawler, Joshua J., David J. Lewis, Erik Nelson, Andrew J. Plantinga, Stephen Polasky, John C. Withey, David P. Helmers, Sebastián Martinuzzi, Derric Pennington, and Volker C. Radeloff. 2014. “Projected Land-Use Change Impacts on Ecosystem Services in the United States.” *Proceedings of the National Academy of Sciences* 111 (20): 7492–97. <https://doi.org/10.1073/pnas.1405557111>.
- Lubowski, Ruben N., Andrew J. Plantinga, and Robert N. Stavins. 2006. “Land-Use Change and Carbon Sinks: Econometric Estimation of the Carbon Sequestration Supply Function.” *Journal of Environmental Economics and Management* 51 (2): 135–52. <https://doi.org/10.1016/j.jeem.2005.08.001>.
- MapBiomass Project. 2023. “Annual Mapping of Soil Organic Carbon Stock in Brazil 1985–2021 (Beta Collection). Map Collection.” <https://doi.org/10.58053/MapBiomass/DHAYLZ>.
- . 2024. “MapBiomass Collection 8.”
- McFadden, Daniel. 1981. “Econometric Models of Probabilistic Choice.” In *Structural Analysis of Discrete Data with Econometric Applications*, 198–272. Cambridge, MA: MIT Press.
- Merkus, Erik. 2024. “The Economic Consequences of Environmental Enforcement: Evidence from an Anti-Deforestation Policy in Brazil.” *World Development* 181 (September): 106646. <https://doi.org/10.1016/j.worlddev.2024.106646>.
- Moffette, Fanny, and Holly K. Gibbs. 2021. “Agricultural Displacement and Deforestation Leakage in the Brazilian Legal Amazon.” *Land Economics* 97 (1): 155–79. <https://doi.org/10.3368/le.97.1.040219-0045R>.
- Nascimento, Nathália, Thales A. P. West, Jan Börner, and Jean Ometto. 2019. “What Drives Intensification of Land Use at Agricultural Frontiers in the Brazilian Amazon? Evidence from a Decision Game.” *Forests* 10 (6): 464. <https://doi.org/10.3390/f10060464>.
- Nepstad, Daniel, David McGrath, Claudia Stickler, Ane Alencar, Andrea Azevedo, Briana Swette, Tathiana Bezerra, et al. 2014. “Slowing Amazon Deforestation Through Public Policy and Interventions in Beef and Soy Supply Chains.” *Science* 344 (6188): 1118–23. <https://doi.org/10.1126/science.1248525>.
- Parsons, George R., Andrew J. Plantinga, and Kevin J. Boyle. 2000. “Narrow Choice Sets in a Random Utility Model of Recreation Demand.” *Land Economics* 76 (1): 86–99. <https://doi.org/10.2307/3147259>.
- Rausch, Lisa L., and Holly K. Gibbs. 2021. “The Low Opportunity Costs of the Amazon Soy Moratorium.” *Frontiers in Forests and Global Change* 4 (March). <https://doi.org/10.3389/ffgc.2021.621685>.
- Rennert, Kevin, Frank Errickson, Brian C. Prest, Lisa Rennels, Richard G. Newell, William Pizer, Cora Kingdon, et al. 2022. “Comprehensive Evidence Implies a Higher Social Cost of CO₂.” *Nature* 610 (7933): 687–92. <https://doi.org/10.1038/s41586-022-05224-9>.
- Saatchi, Sassan S., Nancy L. Harris, Sandra Brown, Michael Lefsky, Edward T. A. Mitchard, William Salas, Brian R. Zutta, et al. 2011. “Benchmark Map of Forest Carbon Stocks in Tropical Regions Across Three Continents.” *Proceedings of the National Academy of*

- Sciences* 108 (24): 9899–9904. <https://doi.org/10.1073/pnas.1019576108>.
- Shyamsundar, Priya, and Randall A. Kramer. 1996. “Tropical Forest Protection: An Empirical Analysis of the Costs Borne by Local People.” *Journal of Environmental Economics and Management* 31 (2): 129–44. <https://doi.org/10.1006/jeem.1996.0036>.
- Skidmore, Marin Elisabeth, Kaitlyn M Sims, Lisa L Rausch, and Holly K Gibbs. 2022. “Sustainable Intensification in the Brazilian Cattle Industry: The Role for Reduced Slaughter Age.” *Environmental Research Letters* 17 (6): 064026. <https://doi.org/10.1088/1748-9326/ac6f70>.
- Small, Kenneth A., and Harvey S. Rosen. 1981. “Applied Welfare Economics with Discrete Choice Models.” *Econometrica* 49 (1): 105–30.
- Soares-Filho, Britaldo, Raoni Rajão, Marcia Macedo, Arnaldo Carneiro, William Costa, Michael Coe, Hermann Rodrigues, and Ane Alencar. 2014. “Cracking Brazil’s Forest Code.” *Science* 344 (6182): 363–64. <https://doi.org/10.1126/science.1246663>.
- Song, Xiao-Peng, Matthew C. Hansen, Peter Potapov, Bernard Adusei, Jeffrey Pickering, Marcos Adami, Andre Lima, et al. 2021. “Massive Soybean Expansion in South America Since 2000 and Implications for Conservation.” *Nature Sustainability* 4 (9): 784–92. <https://doi.org/10.1038/s41893-021-00729-z>.
- Soterroni, Aline C., Fernando M. Ramos, Aline Mosnier, Joseph Fargione, Pedro R. Andrade, Leandro Baumgarten, Johannes Pirker, et al. 2019. “Expanding the Soy Moratorium to Brazil’s Cerrado.” *Science Advances* 5 (7): eaav7336. <https://doi.org/10.1126/sciadv.aav7336>.
- Souza-Rodrigues, Eduardo. 2019. “Deforestation in the Amazon: A Unified Framework for Estimation and Policy Analysis.” *The Review of Economic Studies* 86 (6): 2713–44. <https://doi.org/10.1093/restud/rdy070>.
- Spawn, Seth A., Clare C. Sullivan, Tyler J. Lark, and Holly K. Gibbs. 2020. “Harmonized Global Maps of Above and Belowground Biomass Carbon Density in the Year 2010.” *Scientific Data* 7 (1): 112. <https://doi.org/10.1038/s41597-020-0444-4>.
- Ter Steege, Hans, Nigel C. A. Pitman, Daniel Sabatier, Christopher Baraloto, Rafael P. Salomão, Juan Ernesto Guevara, Oliver L. Phillips, et al. 2013. “Hyperdominance in the Amazonian Tree Flora.” *Science* 342 (6156): 1243092. <https://doi.org/10.1126/science.1243092>.
- Thaler, Gregory M. 2017. “The Land Sparing Complex: Environmental Governance, Agricultural Intensification, and State Building in the Brazilian Amazon.” *Annals of the American Association of Geographers* 107 (6): 1424–43. <https://doi.org/10.1080/24694452.2017.1309966>.
- Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. 2nd ed. New York, NY: Cambridge University Press.
- Victoria, Daniel De Castro, Ramon Felipe Bicudo Da Silva, James D. A. Millington, Valeri Katerinchuk, and Mateus Batistella. 2021. “Transport Cost to Port Though Brazilian Federal Roads Network: Dataset for Years 2000, 2005, 2010 and 2017.” *Data in Brief* 36 (June): 107070. <https://doi.org/10.1016/j.dib.2021.107070>.
- Villoria, Nelson, Rachael Garrett, Florian Gollnow, and Kimberly Carlson. 2022. “Leakage Does Not Fully Offset Soy Supply-Chain Efforts to Reduce Deforestation in Brazil.” *Nature Communications* 13 (1): 5476. <https://doi.org/10.1038/s41467-022-33213-z>.
- Wilson, James Q. 1980. *The Politics of Regulation*. New York: Basic Books.
- WWF. 2025. “2025 Retrospective: Soy Moratorium Under Attack.” <https://www.wwf.org.br>

A Supplementary Information

A.1 Data details

A.1.1 Land-use categories

We discard the classes that fall outside our three land uses — herbaceous and shrubby vegetation, non-vegetated land, water, and non-observed pixels — and aggregate the remainder into forest, pasture, and crops. The forest category combines forest formation, savanna formation, mangrove, floodable forest, wooded sandbank vegetation, and forest plantation. Pasture comprises the pasture class together with mosaic-of-uses, which we treat as predominantly pasture. The crop category aggregates soybean with sugar cane, rice, cotton, coffee, citrus, palm oil, and other temporary and perennial crops.

A.1.2 Rent construction

For a given pixel i , land use k , and year t , per-hectare revenue is yield times price, $R_{ikt} = y_{ikt} p_{kt}$, with y_{ikt} in tons per hectare and p_{kt} in 2019 BRL per ton and common across pixels. Crop prices come from FAOSTAT (FAO 2026) and ESALQ (ESALQ 2026), and cattle prices from ESALQ. The associated cost is the expense of moving that output to its destination,

$$C_{ikt} = tc_k d_i y_{ikt},$$

where tc_k is the per-ton transport cost for commodity k and d_i the travel time from pixel i to the nearest export port. Rent is the difference, which we convert to a perpetuity for use in the model:

$$\pi_{ikt} = R_{ikt} - C_{ikt} = y_{ikt} (p_{kt} - tc_k d_i).$$

Transport cost stands in for total production cost; it generates geographic variation in rents, since more remote pixels lose more value in transit and tend to face higher input costs. Because it leaves out the other costs of production, this rent is a gross measure and overstates the true net return, so any welfare cost built on it is an upper bound. Section A.1.3 reports what happens when rents are rescaled to an approximate net return.

Crop yields. Pixel-level crop yields combine agronomic potential with observed municipal yields. From FAO’s Global Agro-Ecological Zones (GAEZ) (FAO 2022) we obtain attainable yield at the pixel, AY_{ik} , and the municipal-average attainable yield, $\overline{AY}_{m(i)k}$, where $m(i)$ denotes the municipality containing pixel i . Taking their ratio defines a yield fraction,

$$\theta_{ik} = \frac{AY_{ik}}{\overline{AY}_{m(i)k}}.$$

Pixels with missing attainable yield are treated as unsuitable for the crop — inspection against observed production supported this — and assigned a value of zero. Multiplying the yield fraction by the observed municipal-average yield $\overline{Y}_{m(i)kt}$ (IBGE 2022) gives the pixel- and year-specific yield, $y_{ikt} = \theta_{ik} \overline{Y}_{m(i)kt}$.

Crop basket. The crop land use aggregates the ten crops that account for most production in the study area: soybean, rice, bean, sugarcane, maize, cassava, potato, wheat, tomato, and cotton. For each crop we compute its municipal area share from harvested-area data (IBGE

2022), $\omega_{m(i)kt} = A_{m(i)kt} / \hat{A}_{m(i)t}$, where $A_{m(i)kt}$ is the area of crop k and $\hat{A}_{m(i)t}$ the total area of the ten crops in the municipality. Because the shares sum to one, area-weighted crop revenue is the sum of the individual crops' area-weighted revenues, $r_{ikt} = y_{ikt} p_{kt} \omega_{m(i)kt}$, and crop profit can be written

$$\pi_{i,\text{crops},t} = \sum_k y_{ikt} p_{kt} \omega_{m(i)kt} - tc_{\text{crops},t} d_{it} \sum_k y_{ikt} \omega_{m(i)kt}.$$

Pasture yields. For pasture we begin from the national-average carcass yield (tons per hectare), computed from IBGE production (IBGE 2026) and MapBiomass pasture area (MapBiomass Project 2024), $\bar{Y}_{\text{pasture},t}$, and scale it by a pixel-level suitability fraction $\rho_{it} = S_{it} / \bar{S}$, where S_{it} is GAEZ pasture suitability at the pixel and $\bar{S}$ its national average. This yields $y_{i,\text{pasture},t} = \rho_{it} \bar{Y}_{\text{pasture},t}$.

Transport costs. Travel time d_i is the shortest time from a pixel to an export port, by road or by a road–river combination. Using port and riverway datasets from the Brazilian Ministry of Transportation, we computed riverway travel times from each non-export port to the export ports, which serve as the network's exit points. An export port is one that shares a municipality with one of the 13 export ports identified by Victoria et al. (2021) or that lies on the coast. Paraná-basin ports with no riverway connection to a Brazilian export port were instead routed to the Paraguay–Argentina–Brazil border. Road travel times between the centroids of 50×50 km grid cells and the ports — including the export ports themselves, not only inland ports — were taken from OpenStreetMap (2020). Each pixel inherits the times of its nearest grid cell, and we assign it the smallest; because roads run directly to export ports, this minimum is effectively the faster of an all-road route and a road-then-river route. To monetize travel time, we regressed observed soy transport costs for 32 municipality pairs (USDA) on travel time and used the fitted relationship to predict the per-ton soy cost tc_{soy} , which we apply to all crops. For

cattle, the single available cost observation (Arima et al. 2005, adjusted for inflation) implied a cost roughly 9% above soy, so $tc_{\text{cattle}} = 1.09 tc_{\text{soy}}$; this likely understates the true ratio.

A.1.3 Profit margins

Our rent nets out transport but no other production cost, so it is a gross measure. To gauge how much that inflates the welfare cost, we rescale each land use's rent by a profit margin — the share of revenue left after operating costs — and re-estimate the model and the compensating variation on the rescaled rents. The two margins come from unrelated sources, and we do not pool them.

Crops: 42.7%. From a farm-budget study of typical farms in Mato Grosso (Dos Reis et al. 2020), we combine the reported soy productivity and operational cost for the continuous crop typical farm with the soy prices in the same paper, for crop years 2005/06 through 2011/12, and average the resulting annual margins. The operational cost covers inputs, labor, and machinery; it excludes land, capital depreciation, interest, taxes, and overhead. Excluding land is what we want, since our rent is itself the return to land, but excluding capital and overhead makes this a gross margin and so an upper bound on the true profit share.

One choice in that calculation drives most of the uncertainty in the figure. The study reports costs by crop year, which straddles two calendar years, but prices by calendar year. Pairing a harvest's cost with a price therefore means choosing which of the two years it sold in. Soy in Mato Grosso is planted late in the first year and harvested and largely sold in the second, so we use the second year's price. The choice matters because soy prices roughly doubled over the study period, so moving the price window by a year changes revenue substantially while leaving costs untouched. Using the first year's price gives 32.4% instead of 42.7%, and averaging the two gives 38.6%.

Cattle: 37.6%. Our cattle margin comes from a production-cost panel for beef rearing and fattening in Colorado do Oeste, Rondônia, for the 2023 accounting year (CNA SENAR 2024). The panel reports a gross margin of 18.2% of revenue. We do not use that figure, and because the correction is large, the reasoning is worth setting out in full.

The farm in the panel does not raise cattle from birth. It buys weaned calves, grows them on pasture and through a short finishing phase, and sells them for slaughter. Over the accounting year it spent R\$162,500, freight included, on calves. That purchased weight comes to 678 arrobas (15 kg of carcass weight each, the unit Brazilian cattle prices are quoted in), against 1,124 arrobas grown on its own pasture. Roughly two fifths of what it sells, then, was grown on someone else's land.

Three prices govern its accounts. It pays R\$239.68 for each arroba it buys in. It can grow an arroba itself for R\$131.00 — its own operating costs divided by the weight it produced. And it sells at R\$210.00 per arroba. That it pays more per arroba than it sells for is not an anomaly: weight on a lean young animal is dearer than weight on a finished one, because the buyer is paying for the animal's remaining growing potential as well as its present mass.

The published margin treats the entire R\$162,500 as a cost. But R\$108.68 of every purchased arroba — the gap between the R\$239.68 paid and the R\$131.00 it would cost to grow — is not a cost of producing beef. It is a payment to the rancher who raised the calf, and most of it is that rancher's own profit. Counting it as a cost yields the margin of a business that buys in most of its raw material, rather than the margin of producing beef.

That distinction matters here because our model contains no finishing businesses. It contains hectares of pasture, and our pasture yield is carcass weight produced per hectare, at whatever stage of the animal's life that weight is added. The margin we need is therefore the margin on producing cattle weight on pasture across the whole chain, from calf to slaughter, not the margin earned by the last operator in that chain.

We therefore restate the panel's accounts on that basis. We value all 1,802 arrobas sold at R\$210. We charge the farm's own operating costs of R\$147,293, which exclude the animal purchase. And we charge the 678 bought-in arrobas at what it costs to grow that weight rather than what was paid for it, taking this farm's own R\$131.00 per arroba as the closest available proxy for the upstream operation. The resulting margin is 37.6%, which reduces to the difference between the selling price and the cost of producing an arroba, as a share of the selling price: $(210 - 131)/210$.

The imputation of the upstream cost is the weak point, and it brackets the answer rather than pinning it. Charging the bought-in weight at the R\$239.68 actually paid returns the panel's published 18.2%, but that figure implicitly assumes the calf producer earned nothing and that the whole purchase price was cost. Our 37.6% assumes instead that the calf producer's costs match this farm's. The true whole-chain margin lies between, and we treat the lower figure as a floor rather than as an alternative estimate.

One element does transfer cleanly. The R\$210 selling price is for finished cattle, the same basis as the ESALQ series our model uses for cattle revenue, so the margin and the rents it rescales rest on the same price concept.

Applying them. We multiply each land use's rent by its margin. Because the prices behind our revenue are farm-gate — the price received by producers, from which the cost of reaching port has already been deducted — the margin applies to revenue net of transport rather than to gross revenue. Applying it to gross revenue and subtracting transport separately would charge the freight twice. Both source margins are measured on that same basis: the crop study's prices are those received by producers in the Mid-North of Mato Grosso, at the farm or local delivery point (J. C. dos Reis, personal communication, 2026), and the cattle panel's revenue is what the farm received at the gate.

What the check does not settle. Three caveats, in rough order of how much they matter.

First, the crop margin is a soy margin, while our crop rent is a composite of ten crops. Soy is the largest of them: across the municipalities of our study area it is about half of the ten crops' combined value of production (IBGE 2022). The other half is spread across the remaining nine, the largest shares being maize (15%), then cotton and sugarcane at about 10% each. The same farm-budget study puts maize's margin at 31.5%, well below soy's, so 42.7% is likely at the upper end of what a composite margin would be. For the other eight crops we have no estimate at all.

Second, both margins are transferred from single studies of particular places and years to a study area spanning two biomes and two decades. The cattle margin carries a further mismatch, and it runs in a known direction. The margin is measured on a working commercial operation, but the pasture revenue it rescales is built from national carcass output spread over all mapped pasture, including land that is degraded or barely stocked. Our median pixel produces roughly an eighth of that farm's output per hectare. That gap is not a sign that our yield measure is wrong. The two simply measure different things: one a managed herd, the other an average hectare. Because part of the cost of holding a hectare does not fall as its output falls, a margin measured on a commercial operation is too generous for an average hectare, so 37.6% should be read as an upper bound. Sizing the overstatement would require knowing how much of the cost is fixed per hectare rather than incurred per unit of output, which the panel does not report.

Third, the rescaled model fits marginally worse than the gross one — the log-likelihood falls by about 11 at the same number of parameters — so the data, if anything, mildly prefer gross rents.

For these reasons we report the rescaled compensating variation as a robustness check and keep the gross estimate as our headline. The check covers the producer compensating variation only. Consumer surplus and the value of avoided carbon both derive from the forward simulation, which runs on gross rents; restating them consistently would require re-running that simulation on rescaled rents, which we do not do here. Rescaling the producer cost while leaving the

carbon benefit on its gross basis would compare quantities built on different measures of rent, so we do not restate the break-even social cost of carbon.

A.2 Model details

The model is static; we retain the time subscript t only because the underlying data span 2000–2019. A risk-neutral, price-taking owner of pixel i chooses the land use k that maximizes utility,

$$U_{ijkt} = V_{ijkt} + \epsilon_{ijkt},$$

with V_{ijkt} the observed component given in the main text and ϵ_{ijkt} an i.i.d. type-I extreme-value error representing pixel-specific unobservables. The specification of V has 13 coefficients, plus a nest dissimilarity parameter.

Profit as a perpetuity. The profit term π_{ikt} is the present value of a perpetual rent stream, formed from the average rent over the preceding five years discounted at about 6.5%. Because expected rents are built from recently realized rents, agents effectively expect them to persist.

Transition probabilities. Under the type-I extreme-value errors and the forest/cleared nesting (the cleared nest containing crops and pasture), the probability that pixel i moves from starting land use j to ending land use k in year t is

$$P_{ijkt} = \frac{e^{V_{ijkt}/\lambda_{n(k)}} \left(\sum_{m \in B_{n(k)}} e^{V_{ijmt}/\lambda_{n(k)}} \right)^{\lambda_{n(k)}-1}}{\sum_l \left(\sum_{m \in B_l} e^{V_{ijmt}/\lambda_l} \right)^{\lambda_l}},$$

where $n(k)$ is the nest containing alternative k , B_l the set of alternatives in nest l , and $\lambda_l \in (0, 1]$ the dissimilarity parameter measuring the independence of unobserved utility within nest l ($\lambda_l = 1$ recovers a standard logit). The model is estimated by maximum likelihood; to avoid temporal autocorrelation, each pixel contributes a single, randomly chosen year to the likelihood.

A.3 Simulation details

Production and the initial shock. Total production of land use k is $\text{Prod}_{kt} = A \sum_{i: \text{use}=k} y_{ikt}$, where A is the (common) area each pixel represents. Reallocating land at the start of a step changes this total; we write the resulting proportional change as Δprod , the shock that initiates the within-step iteration.

Price and yield feedback. Let μ be the price flexibility and η the yield-price elasticity for a land use. At iteration m we form a price multiplier and a yield multiplier,

$$M_m^P = 1 + \mu \Delta\text{prod}_m, \quad M_m^Y = 1 + \mu \eta \Delta\text{prod}_m,$$

and update prices by M^P , yields by M^Y , revenue by $M^P M^Y$, the per-hectare transport cost by M^Y (it scales with yield), and rent as revenue minus transport cost. Because area is fixed within a step, production changes only through yield, so the next shock is $\Delta\text{prod}_{m+1} = M_m^Y - 1 = \mu \eta \Delta\text{prod}_m$. We iterate until $|\Delta\text{prod}| < 0.001$. The recursion is geometric with ratio $\mu\eta$, so it converges only when $|\mu\eta| < 1$.

Parameter values. The price flexibility for cattle is the inverse of the -0.271 Brazilian meat demand elasticity (de Santana 1999), i.e. $\mu = -3.69$; we do not adjust it, as the Brazilian meat market is largely domestic. For crops we start from the -1.03 U.S. soy price flexibility of

Adjemian and Smith (2012), convert it to a global flexibility by dividing by the U.S. 40% share of the world soy market, and rescale by Brazil’s 27% share, giving $\mu = -0.7$. The rescaling is needed because a given proportional change in Brazilian output moves global supply less than the U.S. equivalent. The yield-price elasticity is $\eta = 0.2$ for both crops and pasture (Berry and Schlenker 2011), applied on the intensive margin only. With these values the pasture iteration has $|\mu\eta| \approx 0.74$: it converges, but a pasture yield elasticity above about 0.27 would cause divergence, which limits how large a cattle yield response we can explore.

Scenarios and reported quantities. We run four scenarios, all with the price response active, crossing the policy (on/off) with the yield response (on/off). Scenario 1 is policy on with the yield response on; scenario 2 is policy off with the yield response on; scenario 3 is policy on with the yield response off; and scenario 4 is policy off with the yield response off. Avoided deforestation and spillover are the scenario-1-minus-scenario-2 differences in forest area inside and outside the Amazon, respectively — the policy effect with both responses active in each arm. Induced intensification is the double difference

$$[\text{area}(1) - \text{area}(2)] - [\text{area}(3) - \text{area}(4)],$$

i.e. the policy effect on area with the yield response active minus the policy effect with it switched off, leaving the change in area attributable to the intensification the policies induce.

A.4 Consumer surplus

The policy reaches commodity consumers through prices: it withdraws agricultural land, lowering output and raising prices. We measure the resulting change in world consumer surplus on the forward simulation, which moves output along an inverse demand curve to two horizon equilibria — $(P_{\text{on}}, Q_{\text{on}})$ under the policy and $(P_{\text{off}}, Q_{\text{off}})$ without. The market price responds

to production summed across biomes, so for each commodity we aggregate quantity Q and revenue R over the Amazon and Cerrado at the final simulated period and form the price as $P = R/Q$. The change in consumer surplus is minus the area to the left of the demand curve between the two prices,

$$\Delta CS = - \int_{P_{\text{off}}}^{P_{\text{on}}} Q(P) dP,$$

which we compute per commodity — pasture and crops, the two endogenously priced land uses — and sum. We report two estimators of this integral. The first is a trapezoid on the two simulated equilibria, which uses only the realized prices and quantities and is agnostic about functional form,

$$\Delta CS \approx -\frac{1}{2}(Q_{\text{on}} + Q_{\text{off}})(P_{\text{on}} - P_{\text{off}}).$$

The second assumes constant-elasticity demand and integrates it exactly. Writing f for the price flexibility (the inverse-demand slope $d \ln P / d \ln Q$ of Section A.3) and $\varepsilon = 1/f$ for the corresponding demand elasticity,

$$\Delta CS = -\frac{R_{\text{on}} - R_{\text{off}}}{\varepsilon + 1} = -(R_{\text{on}} - R_{\text{off}}) \frac{f}{f + 1}.$$

We take the constant-elasticity estimator as our primary measure, since it integrates the same constant-flexibility demand curve the simulation moves along; the trapezoid, agnostic about functional form, serves as a comparison. The gap between them measures sensitivity to functional form. It is largest for inelastic demand (pasture), where the steep demand curve makes the trapezoid's straight line understate the area. There the *level* of consumer surplus is infinite, but the *change* between two finite quantities is well defined, and both estimators difference

between the two equilibria rather than integrating down to zero. Simulated production already carries the pixel-representative area, so the raw surplus change requires only the BRL-to-USD conversion, not the per-hectare-to-landscape scaling applied to the compensating variation. We attach a Monte Carlo confidence interval to each estimator from the coefficient draws of the simulation.

Crop Yield Trends and Gaps in the World's Major Cropping Systems, and Their Implications for Meeting Projected Global Demand

Linus Blomqvist¹

Patricio Grassini²

Kathy Baylis¹

Kenneth G. Cassman²

¹University of California, Santa Barbara

²University of Nebraska–Lincoln

1 Introduction

Cropland expansion typically comes at the expense of forests and other natural ecosystems, releasing large carbon stores and destroying biodiversity, but raising crop yields can offset the need for it by producing more food on existing land. Quantifying this land-sparing effect, Villoria (2019) estimates that yield gains between 1990 and 2010 alone spared approximately 173 million ha of natural habitat from conversion. Isolating the contribution of improved crop varieties specifically, Baldos et al. (2025) estimate that their adoption in developing countries avoided roughly a thousand threatened species extinctions and about 5 billion tonnes of CO₂-equivalent emissions over 1961–2015. Beyond land, rising agricultural productivity held down real food prices through most of the twentieth century (Alston, Beddow, and Pardey 2009). Yield growth therefore affects habitat loss, greenhouse gas emissions, and the cost of food.

Global crop demand is projected to roughly double between 2005 and 2050, driven by population growth and dietary shifts toward more land-intensive foods, especially meat (Tilman et al. 2011). Meeting this demand through cropland expansion would cause large-scale habitat loss and greenhouse gas emissions. Meeting demand without large-scale habitat conversion therefore requires continued yield growth (Cassman 1999; Foley et al. 2011). Yields are not the only margin of adjustment: a shortfall would also raise food prices — dampening demand, most sharply for low-income consumers — and encourage substitution toward other sources such as seafood (Costello et al. 2020). But cropland expansion is the adjustment with the greatest environmental stakes, and it is the one we focus on here.

Whether yields can keep pace is an empirical question, and three complementary lines of research bear on it. Yield trend analysis characterizes how yields have moved over time, and where that leaves them now: still rising, or plateaued. Yield gap analysis compares observed yields to their potential (the yield attainable under ideal management, whether rainfed or irrigated) to gauge how much room remains to raise output on existing land. And global yield projection extrapolates recent trends forward and weighs them against future demand. The first two read the same question from different angles: one asks how yields have moved, the other how much room is left. Each bears on the third, since the trend models supply the trajectories that are extrapolated and the size of the remaining gap indicates whether the extrapolated growth is biophysically possible.

Yield trend analysis has produced worrying signs that growth is slowing or stopping in some of the world's most important cropping systems. Grassini, Eskridge, and Cassman (2013) documented yield plateaus (periods in which yields have entirely stopped growing) in productive systems including rice in China and wheat across northwest Europe. Across all the systems they examined, plateaus covered 33% of global rice production and 27% of wheat, but only 5% of maize. Ray et al. (2012) analyzed roughly 2.5 million census observations globally and found yields no longer increasing in regions producing 43% of global rice, 44% of wheat, 21%

of maize, and 18% of soybean. Each study fits a set of candidate models (competing mathematical descriptions of how a yield series has evolved, such as a straight line, a curve, or several joined segments) and keeps the best-fitting one. Each, however, has notable limitations, falling into two groups. The first concerns the model set and selection rule: Grassini, Eskridge, and Cassman (2013) capped their models at two segments, ruling out more than one change in trend even though such breaks are plausible over the 60-plus-year series now available. They also chose among candidates with an error-based decision rule rather than a formal information criterion. Ray et al. (2012) applied a formal criterion but only to a set of polynomial models, including a cubic that Grassini, Eskridge, and Cassman (2013) argue “has no biophysical justification.” The second concerns coverage: Grassini, Eskridge, and Cassman (2013) covered only three crops, omitting soybean, and both studies’ data end well over a decade ago: Grassini, Eskridge, and Cassman (2013) in 2010, Ray et al. (2012) in 2008. The years since have seen continued expansion of irrigation, rising cropping intensity, and ongoing agronomic and genetic innovation, all of which have reshaped these systems. Extending the series to the present is needed to show whether the trends these studies identified have persisted.

Yield gap analysis asks a complementary question: how much room remains between observed yields and the potential a location could support. Potential yield (Y_p) — the yield achievable when water and nutrients are non-limiting and pests are controlled — and its rainfed analogue, water-limited yield (Y_w), define this upper bound for existing crop varieties and technology (Van Ittersum et al. 2013). Countries achieving only a low percentage of their potential yield retain substantial biophysical headroom. Those approaching 80% of Y_p or Y_w face a harder constraint: that threshold marks the practical “attainable yield” ceiling under realistic economic and agronomic conditions (Cassman et al. 2003; Van Ittersum et al. 2013). The distinction matters because it points to different responses: yields far below potential signal a management shortfall, where the gap can be closed through improved on-farm practice, whereas raising the biophysical ceiling itself requires genetic improvement or agronomic innovation (Cassman et al. 2003).

A large body of work has mapped these yield gaps and found substantial unrealized potential across many major cropping systems (Licker et al. 2010; Neumann et al. 2010; Mueller et al. 2012; Zhang et al. 2025), although the gaps are far from uniform: some intensively managed systems already operate close to their potential (YPYGA 2026). Estimates also depend on how the ceiling is defined, from top-down empirical proxies, such as the highest yields observed within zones of similar climate or a fitted statistical frontier, to bottom-up crop simulation grounded in local weather and soils (Van Ittersum et al. 2013). These assessments are largely static, however, quantifying the gap at a point in time without asking whether yields are still climbing toward the ceiling or have already stalled.

A smaller strand of yield gap analysis does connect gaps to trajectories, and it is the closest precedent for the present study. Grassini, Eskridge, and Cassman (2013) hypothesize that the plateaus they observe in East Asian rice and northwest European wheat reflect yields approaching the biophysical ceiling, but do not estimate potential yields themselves, inferring proximity from external literature for a handful of cases. Gerber et al. (2024) extend this globally, finding that yield gaps are already closing across 84% of rice-growing and 56% of wheat-growing area — a pattern they term “ceiling pressure” — and that such closure predicts subsequent stagnation. But their attainable yield, defined as the 95th percentile of observed yields in similar conditions, approximates the biophysical ceiling only where farmers already operate near their potential. Where all farmers are constrained by limited inputs, markets, or knowledge, it systematically underestimates the ceiling, so that the measured limit reflects management rather than biology (Van Ittersum et al. 2013; Couedel et al. 2025), and can even fall below actual farm yields in some locations (Rattalino Edreira et al. 2021). The convergence Gerber et al. document nonetheless carries a real signal where leading farmers already operate near the biophysical limit. Carracelas et al. (2023) pair a formal model-selection criterion for yield trends with crop-simulation potential yields to show that Uruguayan rice has plateaued well below its ceiling, demonstrating that the two can be combined to separate management-limited from ceiling-limited plateaus. But they do so for only a single crop-country case.

Global yield projection extends recent trends to 2050 and compares the result with future demand. At the global scale, Ray et al. (2013) find that observed growth rates for maize, rice, wheat, and soybean (averaging 0.9–1.6% per year) fall well short of the roughly 2.4% per year (non-compounding) needed to double production by 2050 without cropland expansion. But their projection fits a single linear trend over 1989–2008 to each of about 13,500 sub-national units and aggregates the resulting rates, and a single trend over one window cannot detect break-points within it. Where yields decelerated or plateaued after early growth, the fitted slope reflects the full-window average and overstates future production; where they accelerated, the reverse holds.

We apply consistent methods to all three (yield trend analysis, yield gap analysis, and global yield projection) for the major crops and their leading producers, addressing the limitations of prior studies. We do this by fitting 10 candidate models — piecewise linear models with one to three segments, any of which may be flat (a yield plateau) — to FAOSTAT yield data (1961–2024) for all countries each accounting for more than 2% of global production across four crops (maize, rice, wheat, and soybean). We select the best model by the small-sample-adjusted Akaike Information Criterion (AICc), which rewards goodness of fit while penalizing added parameters. Together, the included countries account for 75% of global maize production, 82% of rice, 76% of wheat, and 92% of soybean. We constrain each model segment to fit a minimum of ten years so that the estimates are not driven by short-run movements in weather or prices. We classify each series by the terminal form of its trend and systematically identify yield growth acceleration, whether after a plateau or after a linear segment. We then compare observed yields to potential yield benchmarks from the Yield Potential and Yield Gap Atlas (YPYGA 2026) to compute actual yields as a percentage of potential and yield gap closure rates. Set against the fitted trajectories, these let us ask, where yields have plateaued, whether the binding constraint is the biophysical ceiling or unrealized management potential. Finally, we project global average yields to 2050 and assess the gap relative to crop-specific demand projections. As an aggregate complement, we collapse the four crops into a single calorie-weighted index

of production, whose projected growth we compare against Tilman et al. (2011)’s projection of total crop-calorie demand.

We find that plateaus are rare — accounting for less than 10% of global production for every crop — and that yield growth resumption after a plateau is at least as common as ongoing stagnation. Most crop-country series, including those with yield plateaus, retain substantial biophysical headroom. Yields are below 50% of potential for about half the sample and yield gap closure rates are low, suggesting that in most systems today the binding constraint is not the biophysical ceiling. This leaves substantial room for management improvements to contribute to yield advances alongside genetic advances. Two systems are exceptions, and they are large ones: maize in the United States and rice in China are close enough to their attainable ceilings that, at current rates of gap closure, the ceiling could become the binding constraint within a few decades. Together they supply 35% of world maize and 28% of world rice. Even with this widespread headroom, projected yield growth at current trajectories falls dramatically short of 2050 demand for maize and soybean; roughly matches it for wheat; and exceeds it only for rice. In aggregate, a calorie-weighted production index for the four crops is projected to rise only a little more than half as much as the roughly 100% growth in crop-calorie demand that Tilman et al. (2011) project for 2005–2050. Meeting that demand at these projected yield trajectories would require substantial cropland expansion.

2 Methods

The analysis proceeds in three parts. First, we fit statistical models to national-level yield time series for major producing countries of four crops. We characterize each series by its terminal form (linear increase, linear decrease, or plateau), by any acceleration in yield growth, and by the recent rate of growth (*yield trend analysis*). Second, we compare observed yields to potential yield benchmarks to assess remaining headroom and the pace of gap closure (*yield gap analysis*). Third, we fit models to global average yields and extrapolate to 2050, comparing the

result against crop-specific demand projections. As an aggregate complement, we also compare it against Tilman et al. (2011)’s projection of total crop-calorie demand, via a single calorie-weighted index of production (*global yield projection*).

2.1 Data Sources and Processing

Data for both the yield trend analysis and the global yield projection — including yield, production, and harvested area for maize, rice, soy, and wheat — were obtained from FAOSTAT (FAO 2026). The yield trend analysis uses country-level data restricted to major producers (described below); the global yield projection uses production-weighted averages across all countries in the dataset. The FAO data cover 192 territorial units for the years 1961–2024, although not all years are present for all crop-country series. Yields are measured on a per-harvest basis, so a one-hectare field that is cropped twice contributes two harvested hectares; the reported yield is total annual production divided by this harvested area, equal to the average yield per individual harvest. The FAO yield data are national averages and as such mix irrigated and rainfed yields. FAOSTAT figures are compiled from national statistical agencies and therefore carry measurement error that varies across countries and years and that we cannot independently quantify. This error tends to be larger in lower-income countries, where yields may rest on incomplete farm reports or eye estimates, and it is not always random. Reporting incentives can bias the figures upward (Desiere, Staelens, and D’Haese 2016; See et al. 2015). For the major maize, rice, and wheat producers that dominate our sample, however, FAO yields have been cross-checked against independent micro-level estimates and found to agree closely (Gollin, Lagakos, and Waugh 2014). To the extent that residual error is idiosyncratic year-to-year noise rather than systematic bias, it should have little effect on our main results, since model selection responds to sustained multi-year patterns rather than to individual observations.

To focus the yield trend analysis on major producers, we included only countries accounting for more than 2% of global production over the period 2000–2024 for each crop. This threshold was chosen to capture a large majority of global production while keeping the number of countries

analytically tractable. Using a multi-year average instead of a single year avoids distortion from production anomalies in any one year. The 2000–2024 window reflects the modern production landscape without giving weight to historical producers (such as the former Soviet Union) that are no longer among the major players. Table 1 lists the selected countries and their production shares, along with the total share of world production covered by the entire sample.

Table 1: Countries included in the analysis, by crop. Values are each country’s average share of global production over 2000–2024. The bottom row shows the total share of global production covered.

Country	Maize	Rice	Soy	Wheat
Argentina	3%		15%	2%
Australia				3%
Bangladesh		7%		
Brazil	8%		30%	
Canada				4%
China	22%	28%	6%	17%
France				5%
Germany				3%
India	2%	22%	4%	13%
Indonesia		8%		
Mexico	3%			
Myanmar		4%		
Pakistan				3%
Paraguay			3%	
Philippines		2%		
Russia				9%
Thailand		5%		
Türkiye				3%
Ukraine	2%			3%
United Kingdom				2%
United States	35%		35%	8%
Vietnam		6%		
Total	75%	82%	92%	76%

Based on anomalies in a few self-reported series — implausibly flat runs or isolated outliers — we manually removed a small number of observations after visual inspection; these cases and the reasoning behind each are documented in Section A.1.

In China, FAO rice yields blend the main crop with the ratoon crop, a second harvest regrown from the first crop’s stubble. Using disaggregated main-crop and ratoon yields for 2004–2023

(provided by Shaobing Peng and Xing Yu, Huazhong Agricultural University), we re-ran the model-selection procedure on the main-crop series. Comparing that fit with the blended-series fit from the main analysis gauges how ratoon expansion affects the apparent trend.

For the yield gap analysis, potential yield benchmarks were obtained from YPYGA (2026), accessed via the REST API in January 2026. YPYGA provides site-based estimates of potential yield (Y_p , the yield achievable under irrigated conditions) and water-limited potential yield (Y_w , the yield achievable under rainfed conditions). These are aggregated to the country level by harvest year, using weights based on each station's representation of the country's cropped area, so the country-level benchmark inherits more uncertainty where station coverage is sparse. To weight these benchmarks by production system where needed, we used gridded harvested area data from MIRCA (Portmann, Siebert, and Döll 2010; Kebede et al. 2025), which provides crop-specific 5-arcminute rasters of irrigated and rainfed harvested area.

YPYGA estimates these benchmarks with process-based crop growth simulation models, not from statistical envelopes or maximum observed yields, following the protocol of Van Ittersum et al. (2013). For each crop, a daily-step model simulates crop development and biomass accumulation as a function of solar radiation, temperature, atmospheric CO₂, and cultivar traits governing the length of the growing period. The simulation is driven by observed daily weather, local soil properties, and the dominant cropping-system management (sowing date, cultivar maturity, and plant density). Potential yield Y_p is simulated with water and nutrients assumed non-limiting and biotic stresses controlled, so that growth is bounded only by climate and the crop's genetic potential. Water-limited yield Y_w additionally constrains growth by rainfed water supply, and so also depends on soil water-holding capacity and terrain. Simulations are run at reference weather stations in the main production areas of relatively homogeneous agro-climatic zones and then upscaled to the country level, weighting each zone by its share of national harvested area. Because the simulations hold cultivar, management, and technology at recent levels, the benchmark reflects the biophysical ceiling of the current cropping system

rather than yields attainable through future genetic or agronomic advances.

For the yield gap analysis, we replace FAO’s blended maize yields for Mexico with irrigation-specific yields from Mexico’s agricultural statistical service (SIAP), enabling a like-for-like comparison with YPYGA’s irrigated benchmark; the data and their processing are described in Section A.2.

2.2 Model selection

To mathematically characterize yield trends for the included crop-country series, we assembled a candidate model set spanning the plausible trajectories of crop yield over time (Table 2; Figure 1). Previous work has shown that crop yields typically grow approximately linearly, sometimes with structural breakpoints — the years at which adjacent linear segments join, hereafter *knots* — where the growth rate shifts (Grassini, Eskridge, and Cassman 2013). We therefore restricted the preferred model set to forms composed of one or more linear segments, where any segment may be flat (a plateau). We extended the set used by Grassini, Eskridge, and Cassman (2013) to include models with up to three linear segments (two breakpoints): most of our time series span 64 years, making multiple breakpoints plausible, and visual inspection of several series confirmed this. We did not go beyond three segments, for three reasons. We saw little evidence of more than two breaks across the different series. Each additional segment enlarges the candidate set and adds parameters that AICc penalizes, and the 10-observation minimum-segment constraint (below) already limits how finely even a series this long can be divided.

This restriction excludes the three smooth-curve forms sometimes used in yield-trend analyses — the quadratic, exponential, and quadratic-plateau — none of which were visually supported by the data we examined. With certain parameter values these forms can also closely resemble piecewise linear trajectories, complicating classification (see below). The effect of re-including all three is explored in Table 12, Table 13, and Table 16.

To avoid spurious plateaus and other breakpoints, we restricted each segment to a minimum

of 10 observations. This ensures that each segment captures a meaningful medium-term trend instead of reacting to a short run of exceptional years driven primarily by weather. The choice of a decade-long minimum segment is a compromise between a shorter minimum that would still admit segments driven by short-run weather fluctuations and a longer one that would preclude genuine decadal-scale regime shifts. In this way, breakpoints were more likely to reflect genuine structural changes in yield trends. Table 8, Table 9, and Table 17 report how the classifications change when the constraint is removed.

The null and single-segment linear models were fit by ordinary least squares; all other models were fit by nonlinear least squares using the Levenberg-Marquardt algorithm. To select among the fitted models, we used the small-sample-adjusted Akaike Information Criterion (AICc), which balances goodness-of-fit against model complexity, penalizing models with more parameters to guard against overfitting. We used the small-sample-adjusted version since this correction is non-trivial for some of the shorter series in our dataset (Hurvich and Tsai 1989). All models assume Gaussian i.i.d. errors, with the log-likelihood computed from the residual sum of squares as $-n/2 \cdot \log(\text{RSS}/n)$ (up to an additive constant); this common likelihood form makes AICc values directly comparable across the OLS- and NLS-fitted models. The best model is the one with the lowest AICc score. Because AICc is evaluated across the full candidate set, the number of linear segments (one, two, or three) is itself selected by the data rather than fixed in advance. Models with AICc values within 2 units of the best-fitting model ($\Delta\text{AICc} \leq 2$) were treated as statistically indistinguishable from it, following the conventional threshold of Burnham and Anderson (2002). The effect of instead using the Bayesian Information Criterion (BIC) — which penalizes complex models more — is reported in Table 10, Table 11, Table 14, and Table 15.

For all models containing knots, knot locations were estimated by nonlinear least squares initialized from a grid of 32 evenly spaced candidate positions (capped at $\lfloor n/2 \rfloor$ for shorter series) spanning the interior of each series. The value 32 balances solution coverage against computa-

tional cost, giving approximately one starting value per two observations for a typical 64-year series. For single-knot models, each grid position was used in turn as a starting value; for two-knot models, all series of grid positions were tried. The algorithm ran to convergence from each starting point; solutions where the fitted knot position had drifted such that any segment contained fewer than 10 observations were discarded, and the solution with the lowest AICc among the remainder was retained.

These procedures assume independent, homoscedastic errors. Crop yield series commonly exhibit positive autocorrelation, which reduces the effective sample size below the nominal n . Because AICc penalties are based on n , autocorrelation may cause them to be too lenient, potentially favouring more complex models than would be selected under a correct effective-sample-size adjustment. This primarily affects model selection; the direction of bias with respect to plateau classification is ambiguous, since both plateau-ending and non-plateau-ending multi-segment models are more complex than their simpler alternatives. Homoscedasticity may also be violated for series with substantial yield growth, since absolute yield variability tends to increase as average yields rise. We assess both assumptions on a series-by-series basis in Section B; see also Section 2.6.

Table 2: Mathematical form of candidate models. a : intercept; b, b_1, b_2, b_3 : slope(s); k, k_1, k_2 : knot(s).

Model	Abbrev.	Form
Null	N	$y = a$
Linear	L	$y = a + bx$
Linear-linear	LL	$y = \begin{cases} a + b_1x & \text{if } x \leq k \\ a + b_1k + b_2(x - k) & \text{if } x > k \end{cases}$
Linear-plateau	LP	$y = \begin{cases} a + bx & \text{if } x \leq k \\ a + bk & \text{if } x > k \end{cases}$
Plateau-linear	PL	$y = \begin{cases} a & \text{if } x \leq k \\ a + b(x - k) & \text{if } x > k \end{cases}$
Linear-linear-linear	LLL	$y = \begin{cases} a + b_1x & \text{if } x \leq k_1 \\ a + b_1k_1 + b_2(x - k_1) & \text{if } k_1 < x \leq k_2 \\ a + b_1k_1 + b_2(k_2 - k_1) + b_3(x - k_2) & \text{if } x > k_2 \end{cases}$
Plateau-linear-linear	PLL	$y = \begin{cases} a & \text{if } x \leq k_1 \\ a + b_1(x - k_1) & \text{if } k_1 < x \leq k_2 \\ a + b_1(k_2 - k_1) + b_2(x - k_2) & \text{if } x > k_2 \end{cases}$
Linear-linear-plateau	LLP	$y = \begin{cases} a + b_1x & \text{if } x \leq k_1 \\ a + b_1k_1 + b_2(x - k_1) & \text{if } k_1 < x \leq k_2 \\ a + b_1k_1 + b_2(k_2 - k_1) & \text{if } x > k_2 \end{cases}$
Plateau-linear-plateau	PLP	$y = \begin{cases} a & \text{if } x \leq k_1 \\ a + b(x - k_1) & \text{if } k_1 < x \leq k_2 \\ a + b(k_2 - k_1) & \text{if } x > k_2 \end{cases}$
Linear-plateau-linear	LPL	$y = \begin{cases} a + b_1x & \text{if } x \leq k_1 \\ a + b_1k_1 & \text{if } k_1 < x \leq k_2 \\ a + b_1k_1 + b_2(x - k_2) & \text{if } x > k_2 \end{cases}$

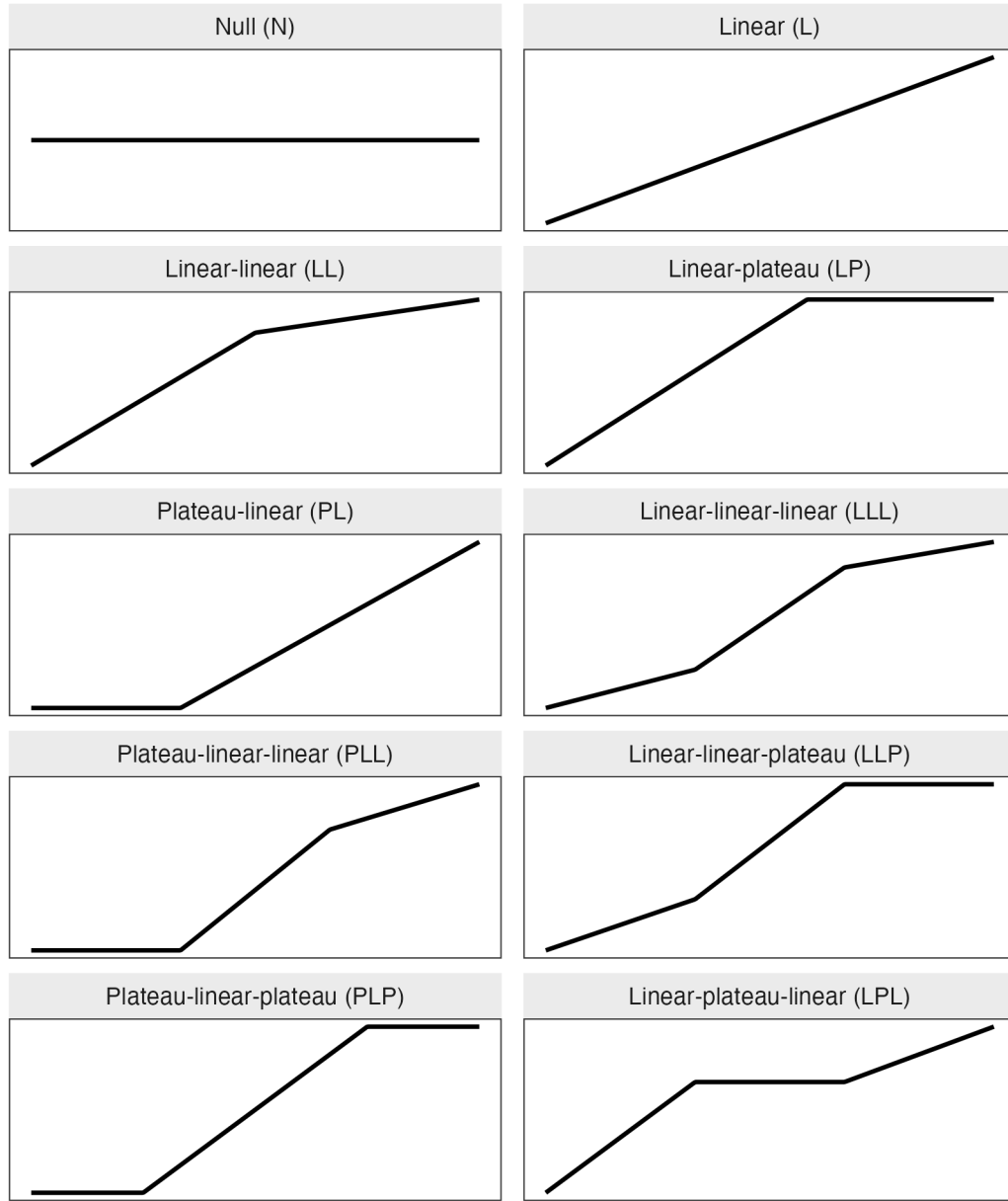


Figure 1: Illustrative shapes of the candidate models. Each panel shows a representative trajectory for one model form; any linear segment may instead have a negative slope, representing a declining yield trend.

2.3 Trend synthesis

For the yield trend analysis, we synthesized model fits in three ways to aid interpretation.

First, each series was assigned a *terminal-form class* based on the form its trend takes at the

endpoint of the time series: linear (increasing or decreasing), plateau, or unclassified. The intercept-only null model is grouped with the plateau class, as it implies a constant yield with no trend. Where models with different terminal forms fell within $\Delta\text{AICc} \leq 2$, the series was assigned to the “unclassified” class.

Second, we identified *acceleration* in yield growth — an increase in the fitted slope from one segment of the model to the next — distinguishing two contexts in which it occurs. An acceleration is *after a plateau* when a flat segment (slope zero) is followed by a segment with positive slope, and *after a linear segment* when a positively-sloped segment is followed by a steeper positively-sloped one ($0 < b_i < b_{i+1}$). Acceleration is assessed at the level of individual knots, so a single series can exhibit both kinds: for example, a plateau-linear-linear (PLL) series that resumes growth after a plateau and then accelerates further between its two linear segments. We report the share of world production exhibiting each kind (Table 4). A series counts toward a given kind only if every candidate model within $\Delta\text{AICc} \leq 2$ contains at least one acceleration event of that kind, so that the pattern is robust to model-selection uncertainty. Because the two kinds are assessed independently, a series can count toward both and the columns are not mutually exclusive.

Third, we derived the recent rate of yield growth. We first took the slope of the fitted trend at the terminal point in the time series (typically 2024). For a linear model, this was simply the slope coefficient. For multi-segment models, it was the slope of the last segment. For any model ending in a plateau, the slope was 0 by definition. We then divided this slope by the model’s fitted yield in that same year, giving a non-compounding annual growth rate in percent per year. Expressing the trend this way makes the pace of growth comparable across series with very different yield levels.

2.4 Yield Gap Analysis

To contextualize yield trends relative to biophysical ceilings, we compared observed yields to YPYGA potential yield benchmarks. Observed yields are the FAO series throughout, except for Mexico maize, for which we used SIAP irrigated yields (see data sources above) to enable a like-for-like comparison with the irrigated benchmark. For each crop-country series in the yield trend analysis, we assigned a water regime — irrigated, rainfed, or mixed — based on the dominant production system in that country. We retained series for which there are at least five observations of potential yield from 2010 onwards for that regime (a minimum we judged sufficient for a reliable multi-year average; Table 7). For series designated as irrigated or rainfed, the corresponding single benchmark (Y_p or Y_w) was used directly; this treats the country as 100% the dominant regime, and small minority shares of the other regime are not represented in the benchmark. For mixed-regime series, we computed a harvested-area-weighted average of both benchmarks using MIRCA. For each crop and country, we summed irrigated (A_{ir}) and rainfed (A_{rf}) harvested area over all grid cells within the country's boundary and computed the irrigated fraction $f_{ir} = A_{ir}/(A_{ir} + A_{rf})$. The weighted potential yield was then $f_{ir} \cdot Y_p + (1 - f_{ir}) \cdot Y_w$. For some countries in our mixed sample, MIRCA's underlying irrigation data is itself imputed from national AQUASTAT totals downscaled via area equipped for irrigation rather than directly observed at the crop level, so f_{ir} for these series carries additional uncertainty. We used the 2015 vintage of the MIRCA rasters because it falls roughly in the middle of the post-2010 period over which potential yields are averaged, making it a reasonable representation of the irrigated share across that window. Shifts in irrigation extent since then are not captured. A sensitivity analysis shifting the irrigated fraction by ± 15 percentage points altered closure rates by at most 0.3 percentage points per year, suggesting results are not sensitive to uncertainty in the 2015 irrigation fractions.

The potential yield benchmark (Y^*) for each series was taken as the average of the regime-appropriate potential yield estimate (Y_p , Y_w , or the MIRCA-weighted average) over the avail-

able post-2010 YPYGA years, reducing sensitivity to any single anomalous year. This window does not span the full length of the FAO yield series, an unavoidable temporal mismatch given YPYGA's available coverage.

To match the temporal reference period of the YPYGA average, we computed the actual yield of each crop-country series ($\hat{Y}$) as the yield predicted by the best-fit trend model averaged over the same set of years. Using model-predicted rather than raw observed yields removes interannual weather noise and ensures that the comparison reflects the underlying trend rather than a single year's conditions.

The absolute yield gap is $Y^* - \hat{Y}$ (potential yield benchmark minus actual yield). We report actual yield as a percentage of the potential yield benchmark, $100 \times \hat{Y} / Y^*$: values approaching 100% indicate that observed yields are near the potential yield.

The rate of yield gap closure was computed as the terminal slope of the fitted yield trend divided by the yield gap, expressed as a percentage per year. The terminal slope is the same quantity reported in Table 5, except for Mexico maize (see Section A.2). This metric captures how quickly the remaining gap between observed and potential yield is being closed. A high closure rate implies that a country is approaching its potential yield and faces an elevated risk of a yield plateau in the near term. A low or zero rate indicates that yields have already stagnated or that other constraints are limiting progress. For series whose best-fit model ends in a plateau, the closure rate is zero by construction, as the terminal slope is structurally fixed at zero; it is not an independent measure of stagnation. We also report an attainable closure rate, which uses 80% of Y^* as the ceiling instead of Y^* itself, following the conventional observation that yields rarely exceed this level under realistic economic and agronomic conditions (Cassman et al. 2003; Van Ittersum et al. 2013). Where multiple models fall within $\Delta AICc \leq 2$, all unique closure rate values are shown separated by “/”. For crop-country series whose terminal-form class is uncertain (unclassified), closure rates are omitted, consistent with the treatment of unclassified series in Table 5.

YPYGA’s potential yield simulations vary year to year, but only with observed weather: they hold the cropping system, technology, and genetics fixed at recent levels and reconstruct earlier years on climate alone. The resulting series therefore captures interannual weather variation, not any gradual shift in the biophysical ceiling itself, the quantity a true potential-yield trend would measure. That ceiling need not be constant: genetic gains and management improvements tend to raise it, while a changing climate may lower it, and the net direction is not known a priori. Because we treat the potential yield benchmark Y^* as a static ceiling, using a single multi-year average rather than a trend, our closure rates compare a moving actual-yield trend against a fixed benchmark. If potential yields have risen over the period, the true gap is closing more slowly than these rates imply; if they have fallen, the reverse holds. We return to this limitation in the discussion.

2.5 Global Yield Trends and Projection to 2050

To assess whether yields are on track to meet future demand, we compared the projected growth in each crop’s global average yield against independent projections of 2050 demand. Published projections of global food demand vary widely (Falcon, Naylor, and Shankar 2022), so we drew on three sources instead of one. The commodity-by-region projections of Falcon, Naylor, and Shankar (2022) are our primary benchmark for the crops it covers or for which its figures can be derived; IMPACT (Robinson et al. 2015; Rosegrant et al. 2024) is the benchmark for wheat, which Falcon et al. do not report. Falcon et al. project demand for individual commodities by applying regional income elasticities to SSP2 population and income growth, following the empirical regularities of Engel’s and Bennett’s laws. They report total utilization (food, feed, and industrial uses) as a percentage change from 2019 to 2050 (their Table 8). We treat these as our primary benchmark for two reasons. First, they measure demand *pressure* — the quantity that would be consumed at constant real prices — rather than a supply-reconciled market equilibrium. For a test of whether yields can keep pace, unconstrained demand is the appropriate comparison, because an equilibrium projection has already been reconciled with the very supply

constraints under examination. Second, they resolve feed and food uses crop by crop, capturing differentiated dynamics — such as declining per-capita rice consumption in Asia versus rising feed and industrial demand for maize — that aggregate or food-only projections obscure. Their figures are reported on slightly different commodity bases for some crops (milled rather than paddy rice, and excluding biofuel), as noted in Table 6.

For comparison we also report two equilibrium-based projections. The first is IFPRI’s IMPACT model (International Model for Policy Analysis of Agricultural Commodities and Trade), a partial-equilibrium model of global agricultural markets developed at the International Food Policy Research Institute (Robinson et al. 2015; Rosegrant et al. 2024). IMPACT decomposes demand into food, feed, biofuel, and other (industrial, seed, waste) uses; we use total demand across all uses, expressed as a percentage increase over a 2020 baseline. The model is calibrated to a three-year FAOSTAT average centered on 2019, with projections running from 2020 to 2050 under the SSP2 (“middle of the road”) socioeconomic scenario and the RCP 7.0 climate-forcing pathway. That pathway is driven by the IPSL-CM6A-LR general circulation model and bias-corrected via the ISIMIP3b protocol, and the reference scenario assumes no CO₂ fertilization effect on yields. The second is the FAO projection of Alexandratos and Bruinsma (2012), which reports world production for 2050 (their Table 4.12); at the global level production equals total demand across all uses. These are absolute levels (Mt) on a 2005/07 base, which we convert to a growth rate over 2020 (below). Because IMPACT and FAO both report equilibrium quantities already reconciled with feasible production, their demand is lower than — and not directly comparable to — the unconstrained demand pressure that Falcon, Naylor, and Shankar (2022) measure.

Falcon, Naylor, and Shankar (2022) do not report whole-soybean demand, so we derive it. Soybeans are consumed almost entirely after crushing, which yields oil and meal in nearly fixed proportions (about 20% oil and 80% meal by weight), so whole-bean demand tracks the demand for crush. Crush is set by whichever co-product is binding: because palm oil

increasingly governs vegetable-oil prices, soybean oil is effectively a by-product, and crush is pulled instead by demand for the protein meal fed to livestock and aquaculture (Falcon, Naylor, and Shankar 2022). We therefore tie soybean demand to feed demand rather than to soybean oil, whose reported growth of 46% (Falcon, Naylor, and Shankar 2022, Table 8) understates whole-bean demand for this reason. Falcon, Naylor, and Shankar (2022) project total meat and fish consumption to rise by about 67% over 2019–2050 (their Table 7), concentrated in the most feed- and meal-intensive products (poultry, +131%; fish, +66%). They project a corresponding increase in feed-grain demand of roughly 64%. Their feed-sector analysis applies a marginal feed-to-output ratio of about four to the projected rise in meat production (an increase of about 880 Mt against the roughly 1,380 Mt of feed used in 2019), a figure they note is comparable to an independent FAO estimate of 64%. Taking soybean meal as a complementary protein feed whose demand grows broadly with the feed sector, we adopt a whole-bean demand increase of approximately 64% over 2019–2050, with a plausible range of roughly 60–75%. Several factors bias this central estimate upward, toward the upper end of that range. The marginal feed-to-output ratio of four assumes the meat mix keeps shifting toward poultry, so a greater role for grain-finished (feedlot) beef would raise it. The underlying projections cover meat and fish but exclude dairy and eggs (Falcon, Naylor, and Shankar 2022), omitting the considerable soybean meal used in dairy rations; aquaculture, an increasingly meal-intensive use, is treated only loosely; and soy-based biodiesel is excluded. Each of these would push soybean demand above our central figure. Because Table 6 places all demand projections on a common 2020 base rather than Falcon et al.’s 2019 base, this 64% increase appears there as approximately 53%. The difference is larger for soybeans than for the other crops: global soybean production rose by roughly 7% between 2019 and 2020, because 2019 was an unusually low year. That raised the 2020 base against which the 2050 target is measured.

Global average yield for each crop was calculated as total world production divided by total world harvested area in a given year, including all countries in the FAO dataset. We applied the model-fitting procedure described above to each crop’s series and extrapolated the terminal

trend of the best-fit model to 2050. A linear best fit gives a linear extrapolation, a linear-linear best fit extrapolates using the second-segment slope, and a plateau best fit gives a flat line. To place yields and demand on a common footing, we expressed every quantity as a percentage change over 2020. The projected yield increase is the fitted model's 2050 value relative to its 2020 value. Each demand projection is a percentage increase in production over recorded 2020 production. Falcon et al.'s 2019-based percentages and FAO's absolute 2050 production levels are rebased to 2020 using recorded global production for those years. Assuming no cropland expansion, the required percentage increase in production equals the required percentage increase in yield. Holding harvested area constant is a normative choice: it defines the yield trajectory needed to meet demand without converting further land, with its attendant climate and biodiversity costs. IMPACT and FAO both project substantial area expansion alongside yield growth to 2050, so the required-yield trajectory we compute is stricter than their internal equilibria. Because the comparison is made crop by crop, it holds each crop's own area fixed, not just their combined total. That is the stronger of the two assumptions, since relative prices reallocate land among crops. Each crop's required-yield figure is therefore the trajectory that would meet its demand from its current area, not a projection of how the market would clear. Our yield extrapolations are illustrative trend projections rather than forecasts. They carry parameter and model-selection uncertainty, which we quantify below, and, more fundamentally, they assume that recent trends continue — an assumption that biophysical ceilings, the pace of technology diffusion, agricultural policy shifts, and climate change could each violate. They are best read as a benchmark: the yield level that would be reached in 2050 if the recent past persisted, against which the demand projections can be judged. The question they answer is whether *current* trajectories are on track, not where yields will actually land. The demand projections carry their own uncertainties, including the choice of socioeconomic and (for IMPACT) climate scenario, and, for our soybean figure, the derivation above.

We attach a 95% interval to each yield projection, capturing the two statistically tractable sources of uncertainty. The first is *model-selection* uncertainty. A plateau best fit projects

a flat line while a linear or linear-linear fit projects a continuing slope, so the extrapolation depends heavily on which model is selected. The leading models are often separated by only a small margin in AICc. We represent this with Akaike weights over the full candidate set, $w_i = \exp(-\Delta_i/2) / \sum_j \exp(-\Delta_j/2)$, where Δ_i is model i 's AICc relative to the best-fitting model. The second is *parameter* uncertainty: within a given model the coefficients — including the slopes that a thirty-year extrapolation magnifies — are estimated with error. We combine the two in a Monte Carlo. In each of 2,000 draws we sample a model in proportion to its Akaike weight, draw that model's coefficients from their estimated sampling distribution $\mathcal{N}(\hat{\beta}, \hat{\Sigma})$, and compute the projected 2050 increase. The 2.5th and 97.5th percentiles of the draws form the interval. The band is a confidence band on the projected mean trend, not a prediction interval for a single year's realized yield. It can be wide or asymmetric — sometimes reaching down to near-zero growth — when a plateau and a still-rising model both carry appreciable weight. We propagate the same draws through the aggregate calorie index and the implied-cropland calculation below, so those intervals are consistent with the per-crop ones. The interval captures sampling and model-selection uncertainty but not the structural risk that recent trends break, which no purely statistical interval can express and which we treat as the separate caveat above. Coefficient sampling for the nonlinear segmented models is a large-sample approximation, and the fits assume independent errors whereas yield residuals are serially correlated, so the intervals are best read as indicative rather than exact.

The comparison above matches each crop to its own demand projection. As a complementary, aggregate check, we collapsed the four crops into a single measure of calorie production and compared its projected growth against a projection of total crop-calorie demand. For the demand benchmark we used Tilman et al. (2011), who estimate that global demand for crop calories will rise by approximately 100% ($\pm 11\%$) from 2005 to 2050. Tilman et al. measure demand as the total caloric content of all crops used as human food or as livestock and fish feed, inclusive of production and post-harvest losses (that is, total utilization expressed in calories). That measure is the appropriate counterpart to an index that counts every crop's calories

regardless of end use. Because it captures feed demand in full, this estimate is generally higher than projections that count only direct human food consumption. A recent meta-analysis places those in the range of a 35–56% increase from 2010 to 2050, about 51% under a middle-of-the-road SSP2 scenario (Van Dijk et al. 2021). It agrees closely, however, with the one comparable projection that likewise counts feed use: Fukase and Martin (2020) independently obtain an increase of about 100%, albeit measured on a feed-weighted cereal-equivalent rather than a strictly caloric basis. We use 2005 as the base year on both sides of this comparison, which avoids rescaling Tilman et al.’s published figure.

To construct the calorie index we converted each crop’s global production to calories using energy-content factors from the USDA FoodData Central database (per 100 g: maize 365, soybeans 446, wheat 327, milled rice 365). Rice requires a processing conversion, because production is recorded for the raw crop while the calorie value refers to a processed form: paddy rice was converted to milled rice using a milling yield of 0.67. We then took each crop’s recorded 2005 production, scaled it by the projected growth in its fitted global yield from 2005 to 2050, converted the resulting 2005 and 2050 production figures to calories, and summed across crops. The projected growth in total calorie production is the 2050 sum relative to the 2005 sum. As in the by-crop comparison, harvested area is held fixed, so production grows at the yield rate. The calculation is equivalent to a base-weighted (Laspeyres) index whose aggregate growth rate is a weighted average of the individual crops’ yield growth, with weights given by their shares of base-year calorie production. The energy-content factors therefore affect only the weighting, not any single crop’s growth rate.

To make this aggregate shortfall concrete, we expressed it as the cropland expansion it implies. The fixed-area identity can be inverted: if projected calorie production grows by a factor $1 + g$ while demand grows by $1 + d$, then meeting demand at trend yields instead requires harvested area to scale by the multiplier $(1 + d)/(1 + g)$. Applying Tilman et al.’s demand growth as d , we multiplied the four crops’ combined 2005 harvested area by this factor to obtain the area

that, worked at trend yields, would meet 2050 demand. Subtracting the area actually harvested in 2020 gives the additional cropland required beyond what is in use today. We benchmark against 2020 instead of the 2005 demand baseline because much of the 2005–2020 supply response already took the form of real area expansion, so the increment relative to 2020 is the land response still outstanding. The calculation is an illustrative accounting identity, not a forecast: it holds demand exogenous, whereas low yields would in practice raise prices and curb demand. It is also a lower bound, since new cropland is typically more marginal and lower-yielding than land already in production, so closing a given production gap would take more area than uniform scaling implies.

The four crops account for roughly 64% of the total crop calories in Tilman et al.’s global accounting (their Table S2: maize 24%, rice 17%, wheat 16%, and soybeans 7%) (Tilman et al. 2011), so the index is representative of, but not identical to, all-crop demand. The benchmark is also a 2011-vintage projection whose income and population assumptions predate more recent data; we use it unaltered, on its native 2005 base.

2.6 Diagnostics

We performed three types of diagnostics, presented in more depth in Section B. First, we inspected residual plots for each crop-country series, examining residuals against fitted values, residuals against time, and Q-Q plots, to assess whether model assumptions of homoscedasticity and approximate normality were met, and whether any systematic misfit remained. Second, we conducted a leave-one-out (LOO) analysis for all series whose best-fit model contained a knot, re-running full model selection after removing each observation in turn. This assessed whether knot locations were sensitive to individual data points and whether the terminal-form class selected on the full dataset was robust to the exclusion of any single observation. Third, we applied a pairs bootstrap (200 replicates) to the same knot-containing series, re-running full model selection on each resample and recording the distribution of estimated knot locations. When the 95% confidence interval on knot position is wide, it indicates uncertainty about the

timing of a structural break. (The LOO analysis addresses the separate question of whether a structural break exists at all.)

3 Results

3.1 Yield trends and classification

Relatively few crop-country series — 6 of the 34 — unambiguously end in a plateau (Table 3); these include maize in Argentina, rice in Myanmar, soy in India, and wheat in Argentina, Canada, and Germany. Two features of the fitted models stand out otherwise. First, multiple breakpoints are common: 13 series are best described by three-segment models. Second, in many (20) cases more than one model falls within $\Delta\text{AICc} \leq 2$, so the exact model is uncertain, but a clear story can still emerge from the plausible set. Rice in Myanmar, for example, has LLP/LP, both describing linear growth (with or without a breakpoint) followed by a plateau.

Table 3: Best-fit model(s) for each crop-country series. Where multiple models fall within $\Delta\text{AICc} \leq 2$, all are listed best-first separated by /. Full model names along with their abbreviations are given in Table 2; full model coefficients are given in Table 23.

country	Maize	Rice	Soy	Wheat
Argentina	LLP		LLP/LL/LP/LLL	PLP
Australia				LL/LPL/PLL
Bangladesh		PLL		
Brazil	LLL/LL		LL/L	
Canada				LLP
China	LLL/LL	LL	LPL	L/LL
France				LP/LL
Germany				LP/LLP
India	PLL/LLL	LLL	LLP/PLP	LLL
Indonesia		LL		
Mexico	LL/LLL			
Myanmar		LLP/LP		
Pakistan				L/LL
Paraguay			LLP/PLL	
Philippines		PLL/LL		
Russia				LL/PL
Thailand		LLP/LLL		
Türkiye				LPL
Ukraine	L			PL/LL
United Kingdom				PLL/LP/LL
United States	L		LL	L

(continued)

country	Maize	Rice	Soy	Wheat
Vietnam	LLL/PLL			

Weighted by production, plateaus remain a small share of the total: series ending in a plateau account for less than 10% of world production for every crop (Figure 2). Most global production instead falls in the linear-increase class, and only a small share is unclassified. The underlying percentages are given in Table 20.

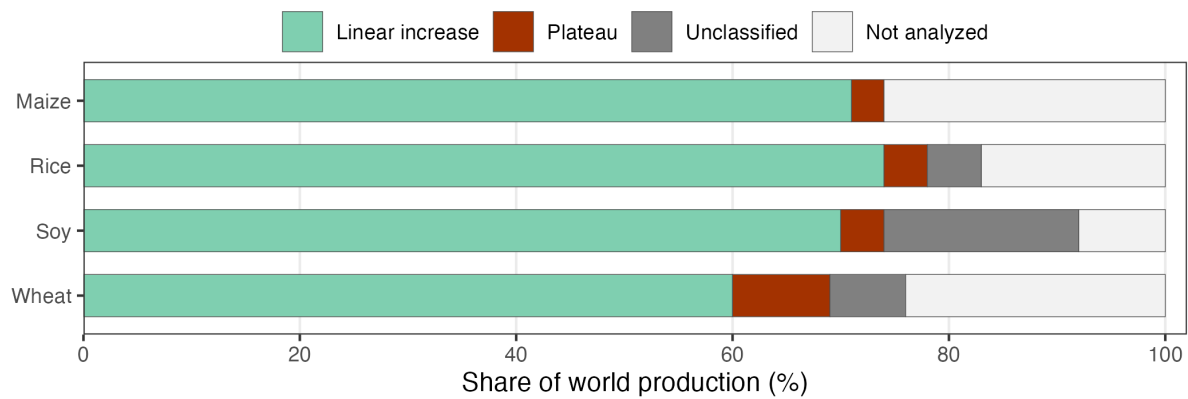


Figure 2: Share of world production accounted for by countries in each terminal-form class, by crop. The terminal-form class is defined by the trend in the final segment of the fitted model; a series is ‘unclassified’ when models within $\Delta AICc \leq 2$ span more than one terminal-form class. ‘Not analyzed’ is the share of world production in countries below the 2% inclusion threshold, so each bar sums to 100%. Percentages are tabulated in Table 20.

Beyond their terminal form, many series show *acceleration*, an increase in yield growth from one segment of the fitted model to the next. Table 4 reports the share of world production exhibiting acceleration in each of two contexts: after a plateau, where growth resumes following a flat period, and after a positively-sloped linear segment. Because acceleration is assessed knot by knot, a single series can exhibit both, so the two shares can overlap and are not additive (see Section 2). Acceleration is more common after a linear segment, but the amount occurring after a plateau is non-trivial: the extent of yield-growth resumption following a plateau is comparable in magnitude to the share of production with current plateaus (Figure 2).

Table 4: Share of global production (%) exhibiting yield growth acceleration, by context and crop. *After plateau*: a flat segment is followed by a segment with positive slope. *After linear segment*: a positively-sloped segment is followed by a steeper positively-sloped one. A series' production share is counted toward a context only if every model within $\Delta\text{AICc} \leq 2$ contains at least one such event. Because acceleration is assessed at the knot level, a series can contribute to both contexts, so the two shares may overlap and are not additive.

Crop	After plateau	After linear segment
Maize	0	16
Rice	7	34
Soy	6	35
Wheat	5	17

Expressing the recent trend as a non-compounding annual growth rate (the terminal slope relative to the current fitted yield) makes the pace of growth comparable across crops and countries of very different yield levels (Table 5). In most series it is modest: recent growth is below 2% per year in all but 5 of the 34 series: maize in India and Ukraine, and wheat in Australia, Russia, and Ukraine.

Growth is typically slow where it matters most: among the largest producers. In maize, the two dominant producers — China (22% of world production) and the United States (35%) — grow at 1.0% and 1.0% per year. In rice, China (28%) and India (22%) together supply roughly half of world production; China's rate of 0.6% is especially low, and India's is 1.5%. In soybeans, the United States (35%) grows at 1.0% and Brazil (30%) somewhat faster, at 1.3/1.2%. Wheat is more dispersed and its growth rates the most variable, several exceeding 2%; the two largest producers, China (17%) and India (13%), grow at 1.4% and 1.8% respectively.

Table 5: Recent yield trend at the end of each series, expressed as a non-compounding annual growth rate (% yr⁻¹): the terminal slope of the fitted model relative to its fitted yield at the end of the series. Zero indicates a plateau. Where multiple models fall within $\Delta\text{AICc} \leq 2$, all unique values are shown best-first separated by “/”; see Table 3 for the corresponding model set.

country	Maize	Rice	Soy	Wheat
Argentina	0.0		0.0/0.8/-0.8	0.0
Australia				3.7/2.5/3.6
Bangladesh		1.6		

(continued)

country	Maize	Rice	Soy	Wheat
Brazil	1.1/2.0		1.3/1.2	
Canada				0.0
China	1.0	0.6	1.0	1.4
France				0.0/-0.1
Germany				0.0
India	2.3	1.5	0.0	1.8
Indonesia		0.6		
Mexico	1.5/1.6			
Myanmar		0.0		
Pakistan				1.2/1.0
Paraguay			0.0/0.2	
Philippines		1.1/0.6		
Russia				2.7/2.2
Thailand		0.0/-0.2		
Türkiye				1.3
Ukraine	2.3			2.2/2.3
United Kingdom				0.4/0.0/0.2
United States	1.0		1.0	0.8
Vietnam		0.8		

3.2 Yield gaps and closure rates

Two things matter for whether a yield gap represents usable headroom: how far a series currently sits below its potential, and how fast it is moving toward it. Figure 3 shows both, with the underlying figures tabulated in Table 21.

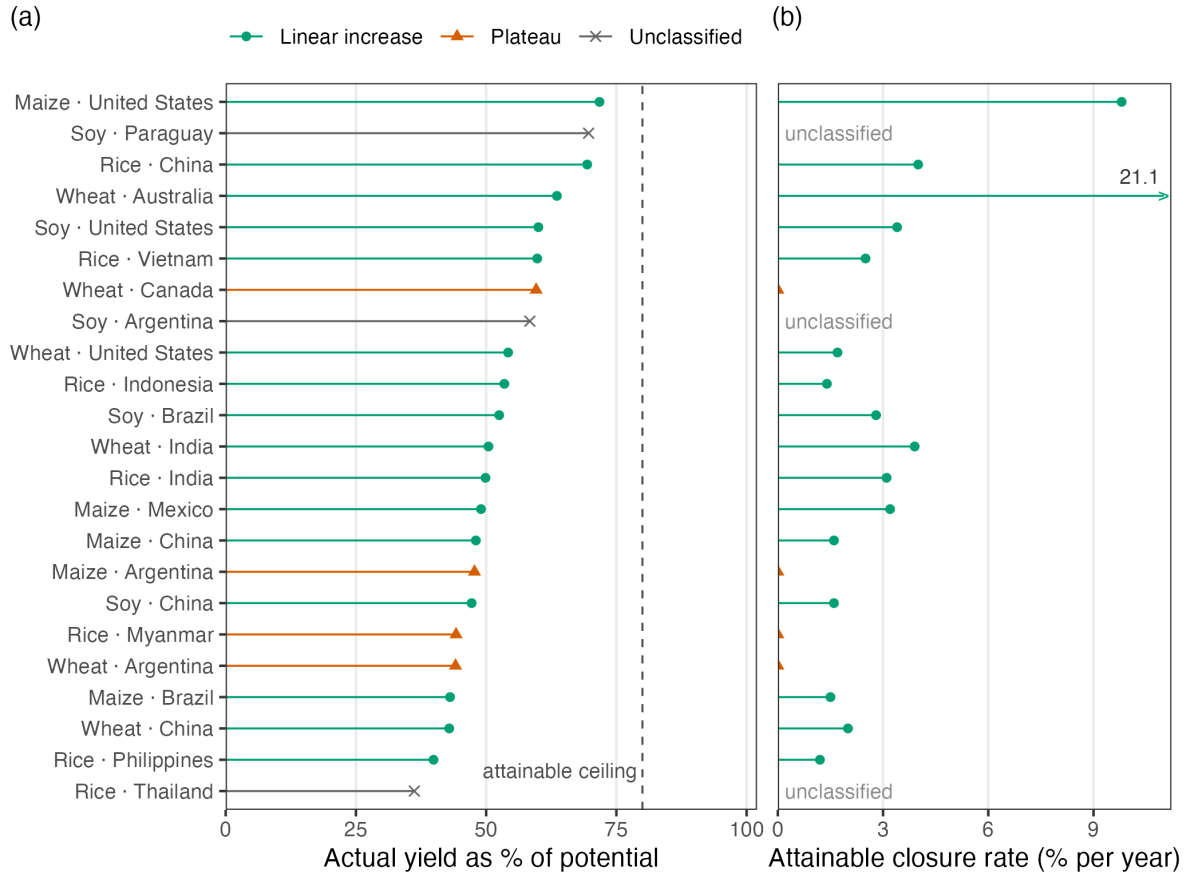


Figure 3: Yield gaps and the pace at which they are closing, for crop-country series with available YPYGA data. Series are ordered by their exploitation of potential yield and coloured by terminal-form class. **(a)** Actual yield as a percentage of potential yield; actual yield is the model-predicted average over the post-2010 years with available YPYGA data. The dashed line marks 80% of potential, the practical attainable ceiling. **(b)** Percentage of the remaining attainable gap closed per year, taking the value from the best-fitting model; plateaued series close no gap and sit at zero. Series whose terminal-form class is uncertain (models within $\Delta AICc \leq 2$ spanning more than one class) have no closure rate and are marked ‘unclassified’. Wheat in Australia closes its attainable gap faster than the axis extends; its value is printed at the arrow.

Actual yields are below 50% of potential for 11 of the 23 series (Figure 3, panel a), indicating that observed yields would need to increase by more than 60% before approaching the attainable yield ceiling. Three series stand out at approximately 70% of potential: maize in the United States, rice in China, and soy in Paraguay. At this level, yields would only need to grow by roughly 10–15% to reach the attainable ceiling.

Yields are closing their gaps slowly in the vast majority of cases (Figure 3, panel b). Measured

against the full potential yield, the rate of gap closure (the share of the remaining gap closed each year) is below 2% for every classified series except maize in the United States, wheat in Australia, and wheat in India (Table 21). Against the attainable ceiling of 80% of potential, the attainable closure rate is 4% or less in every case but two. Those are maize in the United States, at about 10% (the highest among series without model uncertainty), and wheat in Australia, at 14–21% depending on the model. At these rates, yields in most systems could continue growing for two decades or more before reaching the attainable level.

The series that have plateaued — and thus have zero closure rates — still sit well below their potential. Maize and wheat in Argentina and rice in Myanmar are below 50% of potential, and wheat in Canada is at around 60%, all well below the attainable level.

Three series — rice in Thailand and soy in Argentina and Paraguay — are unclassified, with competing models within $\Delta AICc \leq 2$ pointing to different terminal patterns (linear growth versus plateau). No closure rates are reported for these series.

3.3 Global yield projection and 2050 demand

At current trajectories, projected yield growth falls short of 2050 demand for maize and soy, roughly matches it for wheat, and exceeds it only for rice (Figure 4 and Table 6; per-crop plots in Figure 6). Measured against our primary benchmark — the demand projections of Falcon, Naylor, and Shankar (2022) — the maize and soy shortfalls are 28 and 24 percentage points respectively. They remain at least 25 and 19 points even at the optimistic end of the projection interval, so they do not hinge on which model is selected. Rice exceeds its much lower demand growth by 25 points (95% interval 24–26); wheat, which Falcon et al. do not cover, roughly matches the IMPACT projection, with a gap of 2 percentage points whose interval (-2 to 7) straddles zero.

Table 6: Projected increase in global average yield between 2020 and 2050 (extrapolating the best-fit model, with its 95% projection interval; Section 2) vs. the yield increase required to meet crop-specific 2050 demand projections from three sources (IMPACT, FAO, and Falcon et al.), assuming no cropland expansion. All figures are percentage changes over a 2020 base. Dashes mark crops a source does not report.

Crop	Projected increase (%)	95% interval	IMPACT	A&B	Falcon
Maize	34	(0, 37)	48	2	62
Rice	26	(25, 27)	16	7	1
Soy	29	(18, 34)	30	9	53*
Wheat	40	(36, 45)	38	11	—

Falcon et al.: rice is reported on a milled basis (our production is paddy); maize excludes ethanol. A&B (Alexandratos and Bruinsma): absolute 2050 production levels, rebased here from their 2005/07 base. * Soy Falcon figure is an authors' estimate following Falcon et al.'s joint-product approach; Falcon et al. report soybean oil only.

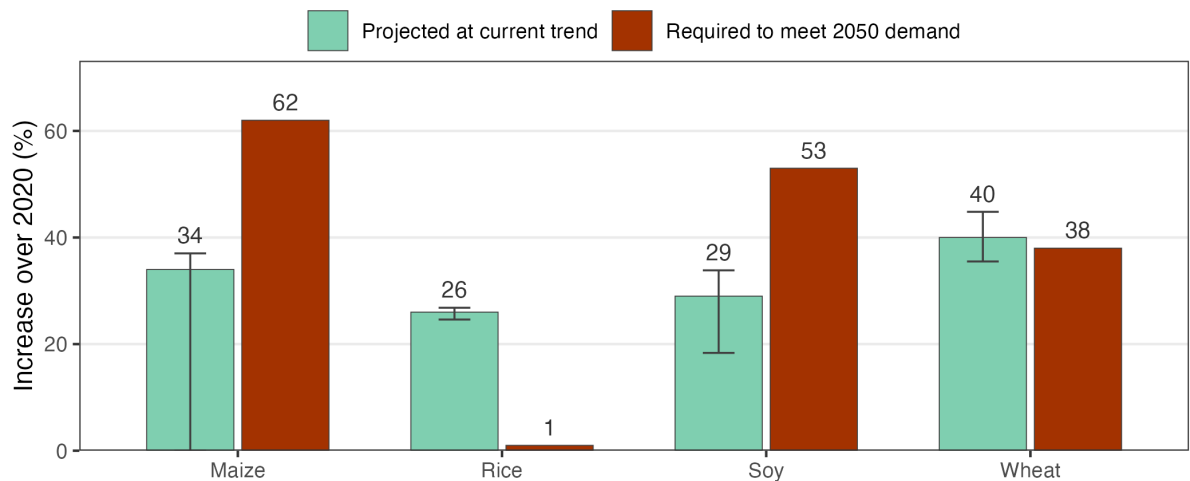


Figure 4: Projected increase in global average yield between 2020 and 2050 against the increase required to meet 2050 demand, by crop. Both are percentage changes over a 2020 base, and the demand comparison assumes no cropland expansion. The projected bar extrapolates the best-fitting model; its whisker is the 95% projection interval (2.5–97.5 percentiles of a Monte Carlo combining AICc model-averaging weights with parametric coefficient uncertainty; Section 2). The required bar uses the primary demand benchmark only: Falcon, Naylor, and Shankar (2022) for maize, rice, and soy, and IFPRI IMPACT for wheat, which Falcon et al. do not cover. Table 6 gives all three demand sources; Figure 6 shows the underlying yield trajectories.

Aggregating the four crops into a single calorie-weighted index of production gives a similar verdict at the aggregate level: under the no-cropland-expansion assumption, projected calorie production rises by about 60% between 2005 and 2050 (95% interval 43–63%). That is only a little over half of the roughly 100% ($\pm 11\%$) growth in total crop-calorie demand projected by Tilman et al. (2011) over the same period. The shortfall is about 40 percentage points (95% interval 37–57), and it holds across all four crops and across the full projection interval (Table 22). This $\sim 100\%$ figure is measured from a 2005 base: because calorie-weighted production of the four crops already grew by about 43% between 2005 and 2020, rebasing Tilman et al. (2011)’s projection onto the 2020 base used in Table 6 implies about 40% further growth to 2050. That is comparable to the crop-specific demand projections there rather than far above them. The larger headline number reflects the earlier base year, not a more demanding assumption.

Expressed as land, this shortfall is substantial. Meeting Tilman et al.’s demand at trend yields would require the four crops’ combined harvested area to scale by a factor of 1.25 (the ratio of demanded to projected calorie production). That means going from about 615 million ha in 2005 to about 768 million ha in 2050, an expansion of roughly 153 million ha. About 98 million ha of this has already occurred: harvested area for these crops grew from 615 to 713 million ha between 2005 and 2020. That expansion accounted for roughly 41% of the growth in these crops’ calorie-weighted production over the same period. The residual (the further expansion needed between 2020 and 2050) is roughly 55 million ha, or 8% above the 2020 area. Its 95% interval is about 40–149 million ha, which widens sharply at the upper end because the land requirement rises steeply if yields grow more slowly. This is an illustrative accounting identity rather than a forecast, since low yields would in practice curb demand through higher prices. It is also a lower bound, since newly converted land tends to be more marginal and lower-yielding than land already in production (Andrade et al. 2022).

Results are broadly robust to the three model-selection choices we varied one at a time: minimum segment length, selection criterion (AICc vs. BIC), and inclusion of the three smooth-

curve forms (quadratic, exponential, and quadratic-plateau): the share of world production with a terminal plateau stays below 15% for every crop under all specifications, and the few series whose classification shifts move mostly into “unclassified” rather than reversing direction (Section C).

A further robustness check addressed the case of rice in China. FAO yield data for China reflect a blended average of the primary (main) crop and the ratoon crop (a second harvest grown from harvested stubble) whose area has expanded substantially since 2004 (Peng and Yu 2026). In the preferred model for the full FAO series (1961–2024), blended yields grow at $41 \text{ kg ha}^{-1} \text{ yr}^{-1}$ from approximately 1993 onward; fitting the same procedure to the main-crop series from Peng and Yu (2026) gives a steeper linear trend of $54 \text{ kg ha}^{-1} \text{ yr}^{-1}$ over 2004–2023. The blended series thus understates the main crop’s yield growth, as expected. FAO counts the second (ratoon) harvest as additional area, so the blended yield is an area-weighted average of the main crop and the lower-yielding ratoon crop, and the growing ratoon share pulls the measured rate down. The two slopes also span different windows. This understatement does not, however, change the trajectory that matters for our classification: both series show Chinese rice yields clearly still rising, with no sign of a plateau. Furthermore, ratoon production was negligible before 2004 — ratoon harvested area was only 0.3 Mha in 2004 compared to 28.1 Mha for the main crop (1.1%) — so the blending has minimal effect on the pre-2004 portion of the FAO series. By 2023, the ratoon share had grown to 1.4 Mha out of 27.5 Mha (4.8%), confirming that ratoon production only became non-trivial during the period covered by the robustness check.

4 Discussion

How fast crop yields rise is a central determinant of whether rising food demand can be met without clearing more land. To see where the major cropping systems stand, we characterized the trajectory of national yields for the leading producers of maize, rice, wheat, and soy. We

then gauged how far those yields sit below their biophysical ceilings and how fast the remaining gap is closing, and projected current trends against mid-century demand.

The overall picture that emerges is more optimistic than that painted by earlier assessments of yield stagnation. Across the major crop-country series we analyzed, yield plateaus are the exception rather than the rule: only 6 of 34 series unambiguously end the time series in a plateau. Terminal plateaus account for just 3–4% of global production for maize, rice, and soy, and 9% for wheat. This strict count captures only unambiguous plateaus; a few additional series, notably wheat in Western Europe, show effectively flat recent yields without being formally classified as plateaus (discussed below). This stands in contrast to earlier characterizations of widespread yield stagnation in major cropping systems. Grassini, Eskridge, and Cassman (2013) reported plateaus representing 33% of global rice production, 27% of global wheat production, and 5% of global maize production. Ray et al. (2012) found that yields were no longer increasing in regions accounting for 43% of global rice, 44% of global wheat, 21% of global maize, and 18% of global soybean. Our methodological improvements (particularly the longer time series, the inclusion of three-segment models that can capture multiple regime shifts within a single series, and formal information-criterion model selection) together yield a picture in which terminal plateaus are uncommon.

Closer examination of the Grassini, Eskridge, and Cassman (2013) results shows that the comparison plays out differently across the three crops they cover. For maize, the two analyses largely agree at the aggregate level: 3% of global production in our analysis versus 5% in Grassini, Eskridge, and Cassman (2013). The countries where plateaus are detected differ between the two studies, but there is agreement that the phenomenon is not widespread. For rice, China alone accounts for 28% of global rice production (averaged over 2000–2024) and thus for nearly all of Grassini, Eskridge, and Cassman (2013)’s 33% plateau share, making the disagreement effectively a single case. For wheat, the picture is more dispersed: India (13% of global production) is the largest plateau-classified producer but accounts for only about half of

Grassini, Eskridge, and Cassman (2013)’s 27% plateau share, with most of the remainder concentrated in Western Europe (discussed below). The cases of China rice and India wheat share a common mechanism. Applying our selection procedure to data truncated at 2010 and excluding three-segment models (more closely approximating Grassini, Eskridge, and Cassman (2013)’s setup) leaves a linear-plateau and a two-segment linear (LL) trajectory effectively indistinguishable (see Section C). Grassini, Eskridge, and Cassman (2013)’s strict $p < 0.01$ significance requirement for breakpoint parameters, combined with the absence of two-knot models in their candidate set, resolves this ambiguity toward the plateau. Under our preferred specification (series extended to 2024 and three-segment models allowed) post-2010 growth makes a plateau implausible and AICc accordingly favors a multi-segment linear model (LL for rice in China, LLL for wheat in India). Of the differences between the two analyses, the length of the series is both the most consequential and the most straightforward to interpret: the additional years are not a matter of methodological preference but of evidence, and they show that yields resumed rising. The model-set and selection-rule differences matter only in resolving a genuinely ambiguous pre-2010 series; once the post-2010 data are in hand, that ambiguity disappears.

The remaining plateau share for wheat in Grassini, Eskridge, and Cassman (2013) is concentrated in Western Europe, where they found plateaus across all five countries they analyzed — a pattern consistent with earlier wheat-specific analyses (Brisson et al. 2010; Lin and Huybers 2012). Our sample includes the three Western European producers above the 2% threshold (France, Germany, and the United Kingdom) and all three show recent yield growth at or near zero (Table 5), consistent with that stagnation. Their formal terminal-form classification differs: Germany an unambiguous plateau, France and the United Kingdom unclassified because competing models cannot distinguish a plateau from a shallow linear trend. But this reflects statistical uncertainty about the exact form rather than disagreement about the substance, since growth in all three has effectively stalled. We cannot, however, test whether this stagnation reflects proximity to the biophysical ceiling, as Grassini, Eskridge, and Cassman (2013) hypoth-

esize, because YPYGA has too few post-2010 potential-yield observations for these countries (fewer than five each) to include them in our yield-gap analysis.

The other unclassified series tell a similar story. Rice in Thailand and soybean in Paraguay also show recent growth at or near zero, so treating them as stagnant rather than merely unclassified would nudge the effectively-flat share upward. But only modestly: they account for just 4.5% of global rice and 2.6% of global soybean. Soybean in Argentina (15.5% of world production) is a genuine outlier: an unusually noisy series whose competing models disagree even on the sign of recent growth. Its best-fitting model ends in a plateau, so counting it as one would add roughly 15 percentage points to soybean’s plateau share; because the evidence is genuinely ambiguous, we leave it unclassified.

The discrepancy with Ray et al. (2012) has a different character: several methodological differences contribute, and we cannot cleanly apportion the gap among them. First, their headline “not witnessing yield increases” combines plateaus with declining and never-improving yields, a broader category than our terminal plateaus. (Note, however, that we observe neither null cases nor declining yields anywhere in our sample, so our equivalent to “not witnessing yield increases” is simply the sum of plateau and unclassified in Table 20.) Second, the analysis is computed at sub-national resolution (~13,500 political units, equally weighted) rather than at the national level. Their own intercomparison (Supplementary Table S3) illustrates the effect. For India wheat, sub-national analysis classifies 70% of area as stagnating, while applying the same method at the national level classifies 100% as increasing, with similar patterns for the United States, China, and Argentina wheat. Third, with no multi-segment alternative in their model set, a decelerating-but-still-rising series — which we would fit as a two- or three-segment linear model with positive but decreasing segment slopes — can be best fit by a downward-opening quadratic whose vertex lies inside the data window. Such a series is then classified as stagnating even though the underlying trajectory is still rising at the endpoint. Fourth, their data end in 2008, so post-2008 resumption, like that documented above for China rice and India

wheat, would not have been captured. We cannot weight these factors precisely, but they do not contribute equally. The sub-national resolution and the restrictive model set are the clearest sources of the gap: each demonstrably leads Ray et al. (2012)’s procedure to classify more series as non-increasing. The resolution difference in particular means the two studies are partly answering different questions, since sub-national stagnation — real and policy-relevant as it is — is distinct from whether national yields have reached a terminal plateau. The remaining two differences are weaker: the broader category is definitional rather than substantive (our comparable measure is the plateau-plus-unclassified share noted above). The shorter window would matter only if post-2008 resumption proved widespread, which our two documented cases cannot establish.

The two cases where we can document a resumption following an apparent plateau (China rice and India wheat) both turn upward in the years just after Grassini, Eskridge, and Cassman (2013)’s data end, and this timing invites a tentative explanation. Those years close a roughly three-decade decline in real agricultural commodity prices and open with the 2007–08 food price spike and its 2010–11 echo, which revived investment attention to agriculture after decades of relative neglect (Headey and Fan 2008). Sustained higher prices plausibly spur the investments that raise yields (producers intensify on existing land, not only expand area) and several studies estimate a positive though modest own-price elasticity of *yield* (roughly 0.04–0.17 globally, about 0.23 for US maize) (Haile, Kalkuhl, and Von Braun 2016; Miao, Khanna, and Huang 2016). Its magnitude is contested — other work finds the aggregate supply response comes mainly through area, with little net yield effect (Roberts and Schlenker 2013; Berry and Schlenker 2011) — and it is in any case partly offset by the heightened volatility of the spike years (Haile, Kalkuhl, and Von Braun 2016). It is therefore a contributing factor rather than a clean cause. If these two documented resumptions are part of a broader wave of post-2008 yield-growth resumption, that wave would likewise fall outside Ray et al. (2012)’s window, which ends in 2008 — though the sub-national resolution of their analysis prevents us from testing this directly.

The persistence of yield growth across most crop-country series in our sample likely reflects a combination of factors whose relative weight our analysis is not designed to disentangle. The leading candidates are the following. One is continued investment in agricultural research and development, both public and private, which has been a primary driver of past productivity growth (Alston, Beddow, and Pardey 2009). This includes advances in plant genetics and agronomy, along with the extension services and input industries that translate research into farm-level practice. A second is the policy environment: price support mechanisms, crop insurance, and access to credit lower barriers to adopting yield-enhancing technologies and practices, both by reducing downside risk and by making capital available for investment (mechanisms long documented in the technology-adoption literature (Feder, Just, and Zilberman 1985; Foster and Rosenzweig 2010)). Political stability and functioning institutions provide the broader preconditions. Third, structural factors play a role: where the share of harvested area under irrigation has expanded, aggregate national yields rise mechanically, since irrigated yields typically exceed rainfed yields (Mueller et al. 2012; Van Ittersum et al. 2013). A shift in the mix raises the average even if within-regime productivity is unchanged.

Continued yield growth is more striking for maize and wheat, because the net effect of climate change on their yields has been negative. Lobell, Schlenker, and Costa-Roberts (2011) estimate that 1980–2008 temperature and precipitation trends reduced global maize and wheat production by roughly 4% and 6% respectively, relative to a counterfactual without those trends. Rising CO₂ over the same period raised C₃-crop yields (wheat, rice, soy) by about 3% through the fertilization effect but did essentially nothing for C₄ maize. Net climate-plus-CO₂ effects were therefore roughly –4% for maize and –2% for wheat, while rice and soy came out net positive, at about +3% and +1% respectively. That maize and wheat yields nonetheless rose across most systems underscores that the other drivers (investment, policy, and structural change) more than offset the net climatic drag on those two crops. These global averages also mask regional heterogeneity: where climate trends have run favorable, they have instead lifted yields. In the irrigated maize systems of Nebraska, for instance, Rizzo et al. (2022) attribute nearly half of the

2005–2018 yield gain to a favorable decadal trend toward cooler, longer grain-filling seasons, together with rising CO₂.

Where plateaus do occur, our yield gap analysis shows that they generally arise well below the attainable yield ceiling. Among the four series in the yield gap sample with zero closure rates, maize and wheat in Argentina and rice in Myanmar are below 50% of potential and wheat in Canada is at around 60%, all indicating substantial remaining biological headroom. A constraint other than the biophysical ceiling (socioeconomic, policy, management, or some combination) must therefore be responsible. This pattern qualifies Grassini, Eskridge, and Cassman (2013)’s framing of plateaus as ceiling-driven phenomena: ceiling-driven plateaus may indeed occur in some intensive systems, but they are not the dominant story across the major producers in our sample.

That plateaus are rarely ceiling-driven rests on the series we can both classify and measure against a potential-yield benchmark. Two of the cases most likely to be ceiling-driven are harder to assess. First, wheat in Western Europe has too few recent YPYGA potential-yield observations for us to say how close those yields sit to their ceiling. Second, soybean in Paraguay does sit near its ceiling (about 70% of potential) with near-flat yields, but its terminal form is unclassified. These do not change the broad picture, however. Terminal plateaus are a small share of production in every crop (from 3% for maize to 9% for wheat), so the ceiling-versus-management question concerns a minor slice to begin with. The other unclassified series we can measure (rice in Thailand and soybean in Argentina) sit well below their ceilings, at about 36% and 58% of potential. The unresolved cases are themselves small, with France and the United Kingdom together accounting for 7.3% of world wheat and Paraguay for 2.6% of world soybean.

A further finding reinforces this picture: yield growth resumption after plateaus occurs in a number of cases. A handful of crop-country series (across rice, soy, and wheat) exhibit acceleration following a historical plateau (Table 4), and for several of these crops the share

of global production with post-plateau resumption exceeds the share with current terminal plateaus. This directly contradicts Grassini, Eskridge, and Cassman (2013)’s observation that “there was no case of a return to the previous rising yield trend after a statistically significant upper yield plateau occurred.” Recognizing post-plateau resumption matters for how current plateaus should be interpreted: in some cases what looks like terminal stagnation may turn out to be a transitional pause. More broadly, the prevalence of multi-segment best-fit models — with regime changes including both decelerations and accelerations within the same series — indicates that a country’s current yield trend is not a permanent state. Growth can slow, stall, or resume, so neither an ongoing plateau nor a phase of rapid growth should be read as a fixed endpoint.

Beyond the plateau cases, the broader yield gap picture is one of substantial remaining headroom. Actual yields are below 50% of potential for about half of the crop-country series with available YPYGA data, meaning that observed yields would need to increase by more than 60% before approaching the attainable yield ceiling. Closure rates — the percentage of the remaining yield gap closed per year at current trends — are correspondingly modest. Of the 16 classified series with non-zero attainable closure rates (computed against 80% of potential yield as the ceiling), 14 are below 5% and 9 below 3%. That implies timelines of more than two decades, and for most series more than three, before yields reach the attainable level. These figures rest on a fixed potential yield benchmark (the average of post-2010 YPYGA estimates) rather than on a dynamic measure of how potential yields are themselves changing. If varietal improvement and agronomic innovation continue to raise the ceiling, the real gap is closing more slowly than our metric indicates and the implied time to reach the attainable level is longer.

How fast that ceiling is actually moving, however, is poorly constrained, because trends in yield potential are notoriously hard to measure. The standard approach — growing a historical series of cultivars side by side and regressing yield on year of release — conflates true gains in yield potential with *maintenance breeding*. Maintenance breeding is the continual varietal replace-

ment needed merely to offset the “yield erosion” that older cultivars suffer as pests, diseases, soils, and climate drift away from the conditions under which they were selected (Cassman et al. 2003; Peng et al. 2010). Because older cultivars are handicapped in a contemporary environment, this comparison overstates the underlying rise in potential. Once the two components are separated the picture differs sharply by crop. For tropical rice and temperate maize, controlled comparisons find little or no detectable increase in yield potential over several decades (Cassman 1999; Cassman et al. 2003; Peng et al. 2010), and the evidence is thinner still for soybean, where erosion-controlled comparisons are largely absent (Andrade et al. 2026). Wheat is the exception: Andrade et al. (2026) attribute roughly 40 of a measured $73 \text{ kg ha}^{-1} \text{ yr}^{-1}$ genetic improvement to higher yield potential (the remainder to maintenance breeding) in the systems they studied, a pace comparable to the yield growth many major wheat systems have achieved. Such gains are hard-won, since raising yield potential ultimately requires increasing canopy photosynthesis, a highly constrained target (Cassman 1999), but for wheat they are evidently real.

The contrast between wheat and the other crops bears directly on our closure rates. For rice and maize, a near-stationary ceiling leaves the static YPYGA benchmark largely unbiased — including for the two series closest to their ceilings, maize in the United States and rice in China, where an accurate benchmark matters most. For wheat, a ceiling that is itself rising means our closure rates overstate how fast the gap is narrowing and understate the time to reach the attainable level. That bias points toward more remaining headroom rather than less, and the one rapidly closing wheat series we single out, Australia, we already attribute to a probable one-off step change rather than convergence on a fixed ceiling.

One line of work tries to track the moving ceiling empirically rather than inferring it from cultivar trials. Gerber et al. (2024) attempt a dynamic estimation of attainable yields using a quantile regression approach and report that yield gaps are already closing across 84% of rice-growing area and 56% of wheat-growing area, a pattern they term “ceiling pressure.” Their

framework can in principle decompose gap dynamics into movements in actual and attainable yields, an advantage over our static benchmark. However, their definition of attainable yield as the 95th percentile of observed yields under similar biophysical conditions systematically underestimates the biophysical ceiling where farmers are widely constrained by limited inputs, markets, or knowledge (Van Ittersum et al. 2013; Couedel et al. 2025). Where the empirical 95th percentile reflects management rather than biology, convergence toward it does not necessarily indicate approach to the biophysical ceiling.

How much of this headroom is realistically available depends on economic conditions, not agronomy alone. A gap between observed and attainable yield need not indicate inefficiency: at prevailing output prices, input costs, and production risk, operating below the attainable ceiling can be privately optimal for farmers. A wide gap can therefore persist as an equilibrium rather than a shortfall waiting to be captured (Van Dijk et al. 2017; Gatti et al. 2023). The biophysical gap we measure is therefore best read not as production that farmers could profitably obtain today, but as a diagnostic of which constraint binds. Where it is wide, the limiting factor is economic and managerial rather than biological. It is thus in principle movable by higher output prices (which carry their own consumer-welfare costs), lower input costs, reduced production risk, stronger credit and institutions, or technology that raises yield per unit cost. A binding biophysical ceiling would yield to none of these, only to advances that raise the ceiling itself. Neither route is costless: closing an economic gap takes real investment and changed incentives, just as raising the ceiling takes frontier R&D. The distinction we draw is therefore about the kind of intervention required — prices, credit, risk instruments, and extension versus genetic and agronomic advance. The two routes are not, however, equally tractable. Closing a management gap draws on interventions whose agronomic content is already known and whose effects appear within a few seasons: nutrient and water management, pest control, sowing date and plant density, and the extension and input-supply systems that carry them to farmers. Raising the ceiling is a frontier research problem that operates on breeding cycles of a decade or more, and one that, as the erosion-controlled comparisons above indicate, has produced little measurable

movement in temperate maize or tropical rice over several decades. Where a wide gap remains, the more reliable near-term route to higher yields is therefore the management one. That is not to say it is cheap or politically simple: the structural conditions that hold gaps open are themselves slow to change.

Two systems depart from this pattern of wide gaps. In both, the biophysical ceiling, though not yet the binding constraint, is close enough to check further growth before long. Maize in the United States and rice in China are at 72% and 69% of potential respectively, meaning yields would only need to grow by roughly 10–15% to reach the attainable ceiling. US maize moreover has an attainable closure rate of about 10%, the highest in our sample outside of wheat in Australia (discussed below). These two are best read as the leading edge of the process described above rather than as counterexamples to it. A system that closes its gap quickly eventually converts a management-limited constraint into a biophysical one, and US maize — the smallest gap and the fastest closure rate in our sample — is furthest along that path. Neither has plateaued. Both are still on rising terminal trends, and approaching the attainable level implies growth that decelerates as the marginal returns to further intensification fall, not an abrupt halt.

The weight of these two cases depends on how they are counted. Counted by series, headroom is clearly the rule: 11 of the 23 series in our yield gap sample are below 50% of potential. Weighted by production the balance is closer, because the near-ceiling systems are among the largest. Those 11 series account for 41% of the production covered by the sample, while maize in the United States and rice in China alone account for 30% — 20% of world production of the four crops. Because of that weight, if yield growth slows in these two systems the effect on global production will be disproportionately large. This is the most direct empirical support in our analysis for the concern raised by Grassini, Eskridge, and Cassman (2013) about intensive systems approaching their biological ceiling, although it is more narrowly scoped — confined to two specific series rather than entire crop categories.

One series deserves a separate caveat: wheat in Australia has the highest attainable closure rate in our sample, but this likely overstates the underlying trajectory. The rapid yield growth observed since 2016 appears to reflect a one-off package of innovations, chiefly the widespread adoption of earlier planting, using long-coleoptile varieties sown into dry soil ahead of the rains. The size of the gain is consistent with the yield benefits projected for early-sowing systems (Hunt et al. 2019; Zhao et al. 2022). The recent yield levels likely represent a genuine step change in productivity, but the steep *rate* of increase is unlikely to persist once adoption of this technology package is complete. The closure rate itself is hard to interpret: the YPYGA benchmark for Australian wheat covers 2003–2017 and largely predates the step change. Our closure rate uses the post-2016 slope, a temporal mismatch that mixes a pre-step-change ceiling with a post-step-change rate. It is also unclear whether the recent innovations represent gap closure against a fixed ceiling or have themselves raised attainable yields, since agronomic practice can affect both.

The projected shortfall to mid-century is not, for the most part, a ceiling story: it arises from the rate at which yields are rising across all four crops, not from a limit on how high they can rise. In the two systems whose gaps are already narrow, faster growth would increasingly have to come from raising the ceiling rather than from closing a gap, and those are among the largest suppliers of maize and rice. Extrapolating the fitted trends to 2050 and comparing them with independent demand projections suggests that current trajectories will, for most crops, not keep pace. Measured primarily against the commodity-by-region projections of Falcon, Naylor, and Shankar (2022), projected yield growth falls short for maize and soy and roughly matches projected demand for wheat. Those projections represent demand pressure rather than a supply-reconciled market equilibrium, and are therefore the appropriate benchmark for asking whether yields can keep up. Only rice clearly exceeds its requirement, and it does so because Asian per-capita rice consumption is declining (Bennett’s law) rather than because its yields are rising unusually fast. Rice’s relatively favorable position is reinforced on the supply side: ratoon cropping, a second harvest taken from land already in rice, has been expanding in China

(Peng and Yu 2026) and offers additional scope to raise output without converting new land. Its margin should not, however, be read as slack available to the crops that fall short. It arises on the demand side rather than from unusually fast yield growth, and the land it frees lies mostly in monsoon Asia, where paddy systems do not readily convert to the rainfed crops whose demand is growing fastest. The maize and soy shortfalls are wide enough to persist across the full 95% projection interval (Section 2), so they do not depend on which trend model best fits each series. The crops that drive demand growth, the feed and industrial crops maize and soy, are precisely those whose projected supply falls furthest behind. The aggregate picture is the same. Combining the four crops into a single calorie-weighted index, projected calorie production rises about 60% between 2005 and 2050 (95% interval 43–63%). That is only a little more than half of the roughly 100% growth in total crop-calorie demand projected by Tilman et al. (2011) over the same period, a shortfall that remains across the entire interval. Because these comparisons hold harvested area fixed, they describe the yield trajectory that would be required to meet demand without converting additional land, a normative benchmark rather than a prediction.

Filling the gap that yield growth leaves unmet would require the four crops' cropland to reach roughly 1.25 times its 2005 extent by 2050. Much of that increase is already under way, with about 98 million ha added between 2005 and 2020 and a further 55 million ha implied on current trends. The land conversion that continued yield growth is meant to forestall is thus already partly occurring, and a persistent shortfall would prolong it. Because newly converted land tends to be less productive than land already farmed, even these figures are conservative. The residual alone amounts to an area of new cropland approaching the size of Ukraine.

Several caveats apply to these conclusions. First, the trend analysis assumes Gaussian i.i.d. errors; positive autocorrelation in yield series may make the AICc penalty too lenient, potentially favoring more complex models than would be selected under a correct effective-sample-size adjustment. Second, FAO yields are self-reported national statistics with measurement error

that we cannot independently quantify, and the 2% production threshold restricts the sample to major producers, leaving the long tail of smaller ones, whose trajectories may differ, outside the analysis. Third, classifications could in principle depend on choices in the model-selection procedure (the candidate set, the selection criterion, and the minimum-segment constraint) but robustness checks varying each of these (Section C) leave the main results unchanged.

The yield gap analysis carries further caveats. YPYGA potential yields aggregate a finite set of reference weather stations weighted by cropped-area coverage, so countries with sparser station networks inherit more uncertainty in the benchmark itself. YPYGA estimates are also unavailable for some crop-country series in our sample, restricting the yield gap analysis to a subset, and the post-2010 YPYGA window does not match the full length of the FAO series. For series assigned a single water regime, we treat the country as 100% that regime, ignoring small minority shares. For mixed-regime series we weight the irrigated and rainfed benchmarks using MIRCA harvested-area fractions. Four of the eleven mixed-regime series (rice in Indonesia, Thailand, Myanmar, and the Philippines) rest on MIRCA inputs that are themselves imputed from AQUASTAT national totals downscaled by area-equipped-for-irrigation rather than directly observed. The 2015 vintage of MIRCA does not capture later shifts in irrigation extent. To bound this uncertainty, we perturbed the MIRCA fractions by ± 15 percentage points for these series and recomputed closure rates; the largest shift was 0.3 percentage points per year (wheat in China), and the SE Asia rice series moved by less than 0.02 percentage points per year.

Finally, the comparison to 2050 demand rests on assumptions on both sides. Our yield projections are illustrative extrapolations of the terminal trend of each best-fit model, not forecasts. We quantify their parameter and model-selection uncertainty with a 95% interval (Section 2), computed by Monte Carlo over Akaike-weighted models and their sampling distributions. More fundamentally, they assume recent trends continue, which no such interval captures and which biophysical ceilings, the pace of technology diffusion, policy shifts, and climate

change could each disrupt. In some systems, moreover, part of the recent gain reflects favorable climate trends rather than durable progress, so extrapolating it risks overstating future yields (Rizzo et al. 2022). The demand projections are likewise model-based and depend on assumptions about population, income growth, dietary change, and trade. They also depend on the base year: older projections measured from a mid-2000s base, such as Alexandratos and Bruinsma (2012), have been partly overtaken by events, so that rebased to 2020 their remaining implied growth to 2050 is small. That low residual reflects a stale base year rather than evidence that little additional growth is required. We benchmark the aggregate calorie index against Tilman et al. (2011) because its total-utilization, all-crop-calorie basis matches our index, which counts every crop's calories regardless of end use; estimates restricted to direct food consumption measure a different quantity and are not directly comparable. The fixed-area comparison is itself a normative device rather than a prediction. The per-crop comparisons rest on a further restriction: each holds that crop's own harvested area fixed, whereas relative prices move land between crops. That mechanism is not hypothetical. Between 2000 and 2024, harvested area grew by 93% for soybean and 53% for maize (the two crops whose projected demand outruns their yield growth) against 12% for rice and 2% for wheat. A crop-level shortfall is therefore better read as the yield growth that crop would need from its current area than as a claim that its demand goes unmet, since part of the adjustment would come as area. The aggregate calorie index is largely insulated from this, because reallocation among the four crops is internal to it: moving a hectare between them changes total calorie production only through the difference in calories per hectare. In practice a yield shortfall would not translate one-for-one into cropland expansion: higher prices would dampen consumption (most steeply for the land-intensive animal products that drive much of the projected demand growth) and shift some demand toward substitutes such as seafood (Costello et al. 2020). Realized conversion would therefore likely fall below the implied-area figure. These adjustments are not costless, however: the price increases that clear the market fall hardest on low-income consumers, which is itself a reason to value the yield growth that would avoid them. Finally, the four crops we aggregate represent

roughly two-thirds of global crop calories rather than the entire food supply.

The two findings pull in different directions. The widespread biological headroom and the rarity of ceiling-constrained plateaus suggest that there is more scope for management-side improvements (adoption of best practices, irrigation, fertility management, and pest control) to contribute to global production growth than a picture of widespread ceiling-driven stagnation would imply. At the same time, the projection comparison above shows that current rates of progress — whatever their drivers — will not suffice to meet projected demand. The two findings are reconcilable: because the shortfall reflects how fast yields are rising rather than a biophysical limit on how high they can rise, it is in most systems closable in principle. That the binding constraint is economic and managerial today does not mean it will remain so: wherever a gap closes quickly, the constraint eventually converts into a biophysical one, and the intervention that would relieve it changes accordingly. Our findings describe which constraint binds now, not a claim that ceilings never bind. Where gaps remain wide, though, closing them calls on interventions that are already well understood. Meeting 2050 demand without cropland expansion, with its substantial environmental costs (Villoria 2019; Baldos et al. 2025), therefore depends on accelerating the closure of yield gaps that in most systems remain wide open.

For policymakers and the research community, our results suggest that the binding constraint on yield growth is in most cases not currently biophysical but a combination of socioeconomic, institutional, and management factors that determine how rapidly farmers close the gap to the attainable ceiling. Understanding why this gap closes faster in some systems than others, and where the remaining headroom is most accessible, should be a priority for future research. So should the complementary problem: where gaps are already narrow, further growth requires moving the ceiling itself. It is reasonable to expect some self-correction here, since proximity to the ceiling, together with the price pressure implied by the demand shortfall above, is precisely the condition under which the returns to ceiling-raising research are highest. It may induce the investment needed. That response cannot be assumed, though: it operates on breeding-

cycle lags of a decade or more. The erosion-controlled evidence discussed above offers little assurance for the two crops where the need is most immediate, since measured yield potential in temperate maize and tropical rice has changed little over several decades. Sustained investment in raising yield potential is therefore needed alongside the gap-closing agenda.

References

- Alexandratos, Nikos, and Jelle Bruinsma. 2012. “World Agriculture Towards 2030/2050: The 2012 Revision.” June 2012. Rome, Italy: FAO.
- Alston, Julian M., Jason M. Beddow, and Philip G. Pardey. 2009. “Agricultural Research, Productivity, and Food Prices in the Long Run.” *Science* 325 (5945): 1209–10. <https://doi.org/10.1126/science.1170451>.
- Andrade, José F., Kenneth G. Cassman, Juan I. Rattalino Edreira, Fahmuddin Agus, Abdullahi Bala, Nanyan Deng, and Patricio Grassini. 2022. “Impact of Urbanization Trends on Production of Key Staple Crops.” *Ambio* 51 (5): 1158–67. <https://doi.org/10.1007/s13280-021-01674-z>.
- Andrade, José F., Jianguo Man, Juan P. Monzon, Juan I. Rattalino Edreira, Shen Yuan, Romulo P. Lollato, Abelardo J. De La Vega, et al. 2026. “Maintenance Breeding and Breeding for Yield Potential Both Contribute to Genetic Improvement in Wheat Yield.” *Nature Communications* 17 (1): 2078. <https://doi.org/10.1038/s41467-026-69936-6>.
- Baldos, Uris Lantz C., Alfredo Cisneros-Pineda, Keith O. Fuglie, and Thomas W. Hertel. 2025. “Adoption of Improved Crop Varieties Limited Biodiversity Losses, Terrestrial Carbon Emissions, and Cropland Expansion in the Tropics.” *Proceedings of the National Academy of Sciences* 122 (6): e2404839122. <https://doi.org/10.1073/pnas.2404839122>.
- Berry, Steven, and Wolfram Schlenker. 2011. “Technical Report for the ICCT: Empirical Evidence on Crop Yield Elasticities.”
- Brisson, Nadine, Philippe Gate, David Gouache, Gilles Charmet, François-Xavier Oury, and Frédéric Huard. 2010. “Why Are Wheat Yields Stagnating in Europe? A Comprehensive Data Analysis for France.” *Field Crops Research* 119 (1): 201–12. <https://doi.org/10.1016/j.fcr.2010.07.012>.
- Burnham, Kenneth P., and David R. Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd ed. NY: Springer New York.
- Carracelas, Gonzalo, Nicolas Guilpart, Kenneth G. Cassman, and Patricio Grassini. 2023. “Distinguishing Between Yield Plateaus and Yield Ceilings: A Case Study of Rice in Uruguay.” *Field Crops Research* 292 (March): 108808. <https://doi.org/10.1016/j.fcr.2023.108808>.
- Cassman, Kenneth G. 1999. “Ecological Intensification of Cereal Production Systems: Yield Potential, Soil Quality, and Precision Agriculture.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (11): 5952–59.
- Cassman, Kenneth G., Achim Dobermann, Daniel T. Walters, and Haishun Yang. 2003. “Meet-

- ing Cereal Demand While Protecting Natural Resources and Improving Environmental Quality.” *Annual Review of Environment and Resources* 28: 315–58. <https://doi.org/10.1146/annurev.energy.28.040202.122858>.
- Costello, Christopher, Ling Cao, Stefan Gelcich, Miguel Á. Cisneros-Mata, Christopher M. Free, Halley E. Froehlich, Christopher D. Golden, et al. 2020. “The Future of Food from the Sea.” *Nature* 588 (7836): 95–100. <https://doi.org/10.1038/s41586-020-2616-y>.
- Couedel, Antoine, Romulo P. Lollato, Sotirios V. Archontoulis, Fatima A. Tenorio, Fernando Aramburu-Merlos, Juan I. Rattalino Edreira, and Patricio Grassini. 2025. “Statistical Approaches Are Inadequate for Accurate Estimation of Yield Potential and Gaps at Regional Level.” *Nature Food* 6: 715–24.
- Desiere, Sam, Lotte Staelens, and Marijke D’Haese. 2016. “When the Data Source Writes the Conclusion: Evaluating Agricultural Policies.” *The Journal of Development Studies* 52 (9): 1372–87. <https://doi.org/10.1080/00220388.2016.1146703>.
- Falcon, Walter P., Rosamond L. Naylor, and Nikhil D. Shankar. 2022. “Rethinking Global Food Demand for 2050.” *Population and Development Review* 48 (4): 921–57. <https://doi.org/10.1111/padr.12508>.
- FAO. 2026. “FAOSTAT.” Food and Agriculture Organization of the United Nations. <https://www.fao.org/faostat>.
- Feder, Gershon, Richard E. Just, and David Zilberman. 1985. “Adoption of Agricultural Innovations in Developing Countries: A Survey.” *Economic Development and Cultural Change* 33 (2): 255–98. <https://doi.org/10.1086/451461>.
- Foley, Jonathan A., Navin Ramankutty, Kate A. Brauman, Emily S. Cassidy, James S. Gerber, Matt Johnston, Nathaniel D. Mueller, et al. 2011. “Solutions for a Cultivated Planet.” *Nature* 478 (7369): 337–42. <https://doi.org/10.1038/nature10452>.
- Foster, Andrew D., and Mark R. Rosenzweig. 2010. “Microeconomics of Technology Adoption.” *Annual Review of Economics* 2 (1): 395–424. <https://doi.org/10.1146/annurev.economics.102308.124433>.
- Fukase, Emiko, and Will Martin. 2020. “Economic Growth, Convergence, and World Food Demand and Supply.” *World Development* 132 (August): 104954. <https://doi.org/10.1016/j.worlddev.2020.104954>.
- Gatti, Nicolas, Michael Cecil, Kathy Baylis, Lyndon Estes, Jordan Blekking, Thomas Heckeley, Noemi Vergopolan, and Tom Evans. 2023. “Is Closing the Agricultural Yield Gap a ‘Risky’ Endeavor?” *Agricultural Systems* 208 (May): 103657. <https://doi.org/10.1016/j.agsy.2023.103657>.
- Gerber, James S., Deepak K. Ray, David Makowski, Ethan E. Butler, Nathaniel D. Mueller, Paul C. West, Justin A. Johnson, et al. 2024. “Global Spatially Explicit Yield Gap Time Trends Reveal Regions at Risk of Future Crop Yield Stagnation.” *Nature Food* 5 (2): 125–35. <https://doi.org/10.1038/s43016-023-00913-8>.
- Gollin, Douglas, David Lagakos, and Michael E. Waugh. 2014. “Agricultural Productivity Differences Across Countries.” *American Economic Review* 104 (5): 165–70. <https://doi.org/10.1257/aer.104.5.165>.
- Grassini, Patricio, Kent M. Eskridge, and Kenneth G. Cassman. 2013. “Distinguishing Between Yield Advances and Yield Plateaus in Historical Crop Production Trends.” *Nature Communications* 4 (December): 1–11. <https://doi.org/10.1038/ncomms3918>.
- Haile, Mekbib G., Matthias Kalkuhl, and Joachim Von Braun. 2016. “Worldwide Acreage and

- Yield Response to International Price Change and Volatility: A Dynamic Panel Data Analysis for Wheat, Rice, Corn, and Soybeans.” *American Journal of Agricultural Economics* 98 (1): 172–90. <https://doi.org/10.1093/ajae/aav013>.
- Headey, Derek, and Shenggen Fan. 2008. “Anatomy of a Crisis: The Causes and Consequences of Surging Food Prices.” *Agricultural Economics* 39 (s1): 375–91. <https://doi.org/10.1111/j.1574-0862.2008.00345.x>.
- Hunt, James R., Julianne M. Lilley, Ben Trevaskis, Bonnie M. Flohr, Allan Peake, Andrew Fletcher, Alexander B. Zwart, David Gobbett, and John A. Kirkegaard. 2019. “Early Sowing Systems Can Boost Australian Wheat Yields Despite Recent Climate Change.” *Nature Climate Change* 9 (3): 244–47. <https://doi.org/10.1038/s41558-019-0417-9>.
- Hurvich, Clifford M, and Chih-Ling Tsai. 1989. “Regression and Time Series Model Selection in Small Samples.” *Biometrika* 76 (2): 297–307.
- Kebede, Endalkachew Abebe, Kevin Ong’are Oluoch, Stefan Siebert, Piyush Mehta, Sarah Hartman, Jonas Jägermeyr, Deepak Ray, et al. 2025. “A Global Open-Source Dataset of Monthly Irrigated and Rainfed Cropped Areas (MIRCA-OS) for the 21st Century.” *Scientific Data* 12 (1): 208. <https://doi.org/10.1038/s41597-024-04313-w>.
- Licker, Rachel, Matt Johnston, Jonathan A. Foley, Carol Barford, Christopher J. Kucharik, Chad Monfreda, and Navin Ramankutty. 2010. “Mind the Gap: How Do Climate and Agricultural Management Explain the ‘Yield Gap’ of Croplands Around the World?” *Global Ecology and Biogeography* 19 (6): 769–82. <https://doi.org/10.1111/j.1466-8238.2010.00563.x>.
- Lin, M, and P Huybers. 2012. “Reckoning Wheat Yield Trends.” *Environmental Research Letters* 7 (2): 024016. <https://doi.org/10.1088/1748-9326/7/2/024016>.
- Lobell, David B., Wolfram Schlenker, and Justin Costa-Roberts. 2011. “Climate Trends and Global Crop Production Since 1980.” *Science* 333 (6042): 616–20. <https://doi.org/10.1126/science.1204531>.
- Miao, Ruiqing, Madhu Khanna, and Haixiao Huang. 2016. “Responsiveness of Crop Yield and Acreage to Prices and Climate.” *American Journal of Agricultural Economics* 98 (1): 191–211. <https://doi.org/10.1093/ajae/aav025>.
- Mueller, Nathaniel D., James S. Gerber, Matt Johnston, Deepak K. Ray, Navin Ramankutty, and Jonathan A. Foley. 2012. “Closing Yield Gaps Through Nutrient and Water Management.” *Nature* 490 (7419): 1–28. <https://doi.org/10.1038/nature11420>.
- Neumann, Kathleen, Peter H. Verburg, Elke Stehfest, and Christoph Müller. 2010. “The Yield Gap of Global Grain Production: A Spatial Analysis.” *Agricultural Systems* 103 (5): 316–26. <https://doi.org/10.1016/j.agsy.2010.02.004>.
- Peng, Shaobing, Jianliang Huang, Kenneth G. Cassman, Rebecca C. Laza, Romeo M. Visperas, and Gurdev S. Khush. 2010. “The Importance of Maintenance Breeding: A Case Study of the First Miracle Rice Variety-IR8.” *Field Crops Research* 119 (2-3): 342–47. <https://doi.org/10.1016/j.fcr.2010.08.003>.
- Peng, Shaobing, and Xing Yu. 2026. “Disaggregated Main-Crop and Ratoon-Crop Rice Yield Data, China, 2004–2023.” Personal Communication.
- Portmann, Felix T., Stefan Siebert, and Petra Döll. 2010. “MIRCA2000—Global Monthly Irrigated and Rainfed Crop Areas Around the Year 2000: A New High-Resolution Data Set for Agricultural and Hydrological Modeling.” *Global Biogeochemical Cycles* 24 (1): GB1011. <https://doi.org/10.1029/2008GB003435>.

- Rattalino Edreira, Juan I., José F. Andrade, Kenneth G. Cassman, Martin K. Van Ittersum, Marloes P. Van Loon, and Patricio Grassini. 2021. “Spatial Frameworks for Robust Estimation of Yield Gaps.” *Nature Food* 2 (10): 773–79. <https://doi.org/10.1038/s43016-021-00365-y>.
- Ray, Deepak K., Nathaniel D. Mueller, Paul C. West, and Jonathan A. Foley. 2013. “Yield Trends Are Insufficient to Double Global Crop Production by 2050.” Edited by John P. Hart. *PLoS ONE* 8 (6): e66428. <https://doi.org/10.1371/journal.pone.0066428>.
- Ray, Deepak K., Navin Ramankutty, Nathaniel D. Mueller, Paul C. West, and Jonathan A. Foley. 2012. “Recent Patterns of Crop Yield Growth and Stagnation.” *Nature Communications* 3 (1293): 1293. <https://doi.org/10.1038/ncomms2296>.
- Rizzo, Gonzalo, Juan Pablo Monzon, Fatima A. Tenorio, Réka Howard, Kenneth G. Cassman, and Patricio Grassini. 2022. “Climate and Agronomy, Not Genetics, Underpin Recent Maize Yield Gains in Favorable Environments.” *Proceedings of the National Academy of Sciences* 119 (4): e2113629119. <https://doi.org/10.1073/pnas.2113629119>.
- Roberts, Michael J., and Wolfram Schlenker. 2013. “Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate.” *American Economic Review* 103 (6): 2265–95. <https://doi.org/10.1257/aer.103.6.2265>.
- Robinson, S., D. Mason-D’Croz, S. Islam, T. B. Sulser, R. D. Robertson, T. Zhu, A. Gueneau, G. Pitois, and M. W. Rosegrant. 2015. “The International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT): Model Description for Version 3.” 1483. Washington, D.C.: International Food Policy Research Institute.
- Rosegrant, Mark W., Timothy B. Sulser, Shahnaila Dunston, Abhijeet Mishra, Nicola Cenacchi, Yohannes Gebretsadik, Richard Robertson, Timothy Thomas, and Keith Wiebe. 2024. “Food and Nutrition Security Under Changing Climate and Socioeconomic Conditions.” *Global Food Security* 41 (June): 100755. <https://doi.org/10.1016/j.gfs.2024.100755>.
- See, Linda, Steffen Fritz, Liangzhi You, Navin Ramankutty, Mario Herrero, Chris Justice, Inbal Becker-Reshef, et al. 2015. “Improved Global Cropland Data as an Essential Ingredient for Food Security.” *Global Food Security* 4 (March): 37–45. <https://doi.org/10.1016/j.gfs.2014.10.004>.
- Tilman, David, Christian Balzer, Jason Hill, and Belinda L. Befort. 2011. “Global Food Demand and the Sustainable Intensification of Agriculture.” *Proceedings of the National Academy of Sciences of the United States of America of the United States of America* 108 (50): 1–5. <https://doi.org/10.1073/pnas.1116437108>.
- Van Dijk, Michiel, Tom Morley, Roel Jongeneel, Martin Van Ittersum, Pytrik Reidsma, and Ruerd Ruben. 2017. “Disentangling Agronomic and Economic Yield Gaps: An Integrated Framework and Application.” *Agricultural Systems* 154 (June): 90–99. <https://doi.org/10.1016/j.agsy.2017.03.004>.
- Van Dijk, Michiel, Tom Morley, Marie Luise Rau, and Yashar Saghai. 2021. “A Meta-Analysis of Projected Global Food Demand and Population at Risk of Hunger for the Period 2010–2050.” *Nature Food* 2 (7): 494–501. <https://doi.org/10.1038/s43016-021-00322-9>.
- Van Ittersum, Martin K., Kenneth G. Cassman, Patricio Grassini, Joost Wolf, Pablo Tittonell, and Zvi Hochman. 2013. “Yield Gap Analysis with Local to Global Relevance-A Review.” *Field Crops Research* 143: 4–17. <https://doi.org/10.1016/j.fcr.2012.09.009>.
- Villoria, Nelson B. 2019. “Technology Spillovers and Land Use Change: Empirical Evidence from Global Agriculture.” *American Journal of Agricultural Economics* 101 (3): 870–93.

- <https://doi.org/10.1093/ajae/aay088>.
- YPYGA. 2026. “Yield Potential and Yield Gap Atlas (YPYGA).” <https://yieldgap.com/>.
- Zhang, Wenmeng, Yangfeng Zou, Wulue Zhou, Cheng Li, Lijun Zuo, Lijuan Miao, and Xuefeng Cui. 2025. “Current Status and Obstacles of Narrowing Yield Gaps of Four Major Crops.” *Journal of the Science of Food and Agriculture* 105 (1): 42–53. <https://doi.org/10.1002/jsfa.13801>.
- Zhao, Zhigan, Enli Wang, John A. Kirkegaard, and Greg J. Rebetzke. 2022. “Novel Wheat Varieties Facilitate Deep Sowing to Beat the Heat of Changing Climates.” *Nature Climate Change* 12 (3): 291–96. <https://doi.org/10.1038/s41558-022-01305-9>.

A Additional Methods

A.1 Dropped observations

After visual inspection of the data for the included crop-country combinations, a small number of observations were removed. These decisions were made case by case through visual inspection informed by agronomic reasoning, without applying formal numerical thresholds. For soy in India, the reported yields for 1961–1971 were dropped because they were identical or near-identical across the entire span, with a near-zero variance inconsistent with the rest of the sample. For soy in Paraguay, the reported yields for 1964, 1968, and 1969 were dropped because they were roughly twice those of the surrounding period, which we judged agronomically implausible.

A.2 Mexico irrigated maize

For the yield gap analysis, we replace FAO data on average maize yields in Mexico with irrigation-specific yield data from Mexico’s agricultural statistical service (Servicio de Información Agroalimentaria y Pesquera, SIAP), covering 1980–2024. This enables a like-for-like comparison with YPYGA’s irrigated-specific benchmark. For 2003–2024, SIAP reports data at the municipal level; we aggregated to the national level by summing irrigated harvested area and production volume across all municipalities and computing yield as their ratio. For 1980–2002, SIAP reports data at the national level directly. Throughout, records were filtered to irrigated grain maize (modalidad Riego, cultivo Maíz grano). No level discontinuity is apparent at the 2002/2003 join.

This substitution also affects the closure-rate calculation. The terminal slope used to compute Mexico maize’s rate of gap closure is taken from the SIAP irrigated-only fit rather than the FAO-based fit reported in Table 5, to match the irrigated YPYGA benchmark. The terminal slopes for all other crop-country observations are those reported in Table 5.

A.3 Potential yield data coverage

Table 7 lists, for each crop-country series included in the potential yield analysis, the number of post-2010 years with available YPYGA potential yield data and the benchmarking method used. “Rainfed” indicates the water-limited potential yield (Y_w) was used directly; “Irrigated” indicates the irrigated potential yield (Y_p) was used directly; “Weighted (ir + rf)” indicates a MIRCA-derived area-weighted average of Y_p and Y_w was computed, applied to countries where both irrigation systems contribute substantially to production.

Table 7: YPYGA potential yield data coverage for each crop-country series. *YPYGA years*: number of post-2010 years with available potential yield data used to compute the average. *Benchmark*: potential yield benchmarking method — rainfed (Y_w), irrigated (Y_p), or MIRCA area-weighted average of both.

Crop	Country	GYGA years	Benchmark
Maize	Argentina	12	Rainfed
Maize	Brazil	8	Rainfed
Maize	China	5	Weighted (ir + rf)
Maize	Mexico	12	Irrigated
Maize	United States	13	Weighted (ir + rf)
Rice	China	5	Irrigated
Rice	India	6	Weighted (ir + rf)
Rice	Indonesia	6	Weighted (ir + rf)
Rice	Myanmar	10	Weighted (ir + rf)
Rice	Philippines	7	Weighted (ir + rf)
Rice	Thailand	8	Weighted (ir + rf)
Rice	Vietnam	9	Irrigated
Soy	Argentina	12	Rainfed
Soy	Brazil	9	Rainfed
Soy	China	10	Weighted (ir + rf)
Soy	Paraguay	13	Rainfed
Soy	United States	9	Weighted (ir + rf)
Wheat	Argentina	12	Rainfed
Wheat	Australia	8	Rainfed
Wheat	Canada	9	Rainfed
Wheat	China	7	Weighted (ir + rf)
Wheat	India	5	Weighted (ir + rf)
Wheat	United States	9	Rainfed

B Model Diagnostics

B.1 Motivation

Model selection via AICc identifies the best-fitting model *among the candidates considered*, but does not verify that the selected model is adequate in an absolute sense. In particular, AICc-based comparison assumes Gaussian i.i.d. errors and will still reward a model that happens to absorb an anomalous observation by placing a breakpoint on it. Three complementary diagnostics are therefore applied to every crop-country series after model selection. All diagnostic plots are available upon request from the author.

B.2 Residual diagnostics

For each crop-country series, three plots are produced from the residuals of the best-fit model:

1. **Residuals vs fitted values.** A loess smoother is overlaid. Points scattered randomly around zero across the full range of fitted values indicate a good fit. A U-shape or arch suggests the functional form is misspecified (e.g., a linear model on data that actually curve). A funnel shape, with spread increasing as the fitted value rises, indicates heteroskedasticity: yields in high-production years are more variable than in low ones. Heteroskedasticity is common in yield data: absolute yield variability tends to increase as average yields rise, so a series spanning several decades of yield growth will often show this pattern. It violates the equal-variance assumption underlying the Gaussian log-likelihood used by AICc, which means AICc values are biased and comparisons between models are less reliable. If heteroskedasticity is pronounced, a log transformation of yield or a weighted least-squares approach would be more appropriate.

Inspection of the residuals-vs-fitted plots across all crop-country series showed generally random scatter around zero. The clearest exception was rice in Bangladesh (best model: PL), where the loess smoother traced a wave shape across the fitted value range. The

wave is not a major concern, for two reasons. First, the PL model assigns all observations in the plateau phase the same fitted value, creating a dense cluster of points at a single x-position. The loess smoother then connects from this cluster across the sparse linear-phase observations, giving it high leverage to produce an apparent wave that reflects noise rather than a systematic misfit. Second, inspection of the raw points — ignoring the smoother — shows residuals that are approximately symmetric around zero throughout. The series does exhibit large absolute residuals ($\pm 300 \text{ kg ha}^{-1}$ against fitted values up to $\sim 5500 \text{ kg ha}^{-1}$), reflecting high year-to-year yield variability in Bangladesh, but this is a property of the data rather than evidence of model misspecification.

Maize in Ukraine (best model: PLL) showed the clearest evidence of heteroskedasticity, with residuals in the plateau phase (fitted values $\sim 2700\text{--}3000 \text{ kg ha}^{-1}$) noticeably tighter than in the high-yield linear phase (fitted values $\sim 6500\text{--}7000 \text{ kg ha}^{-1}$, residuals reaching $\pm 1200 \text{ kg ha}^{-1}$). Some of this apparent spread at the upper end reflects a small number of observations and likely genuine agro-economic volatility during the post-Soviet transition and subsequent intensification period. For this analysis the practical consequences are limited. Since all 10 candidate models are fit on the same yield scale, heteroskedasticity affects all of them equally and does not distort the relative AICc rankings in a way that would favour one functional form over another. The main cost is a slight underestimation of the AICc penalty for complex models. The natural remedies would be a log transformation of yield — well-motivated since yields tend to grow at roughly constant percentage rates — or weighted least squares. Both would require re-running the full analysis and are not pursued here.

2. **Residuals vs time.** This is the most consequential plot for yield data. Points should bounce around zero throughout the series with no systematic pattern. A run of positive residuals followed by a run of negative ones (or vice versa) suggests autocorrelation. A residual trend indicates the model has not captured the full shape of the yield trajectory,

and a sudden jump suggests a structural break that the selected model missed.

The clearest pattern across all crop-country series was an M-shape in the residuals-vs-time plot for rice in Myanmar (best model: PLL), with roughly 10-year stretches of same-sign residuals cycling twice across the ~60-year series. That cycling is a signature of positive autocorrelation, and likely reflects a combination of inherent year-to-year autocorrelation in yield series (weather regimes, multi-year policy effects) and remaining model misspecification. The actual trajectory may have more inflection points than a three-segment piecewise model can capture. Myanmar's particular political and economic history (including major shocks in 1988 and post-2021) may amplify this pattern through genuine structural breaks and reporting artefacts that no smooth piecewise model handles well. The practical implication is that the effective sample size is smaller than the nominal observation count, since neighbouring years carry redundant information. Because AICc penalties are calculated using the nominal n , they are too small, and model selection may favour more complex models more readily than the actual information content warrants. However, since all 10 candidate models are subject to the same autocorrelation equally, the relative AICc rankings are similarly affected across the board and the model that minimises RSS will still be selected. The main residual concern is that knot locations may be partially shaped by autocorrelation structure rather than genuine breaks in the yield trend.

3. **Normal Q–Q plot.** Standardised residuals are plotted against the theoretical Gaussian quantiles; points close to the diagonal indicate approximate normality. An S-shape means the residuals are lighter-tailed than Gaussian; points curving away from the line at both ends indicate heavy tails; curvature at one end indicates skew. The main practical concern is heavy tails or strong skew, since these deflate the log-likelihood and make AICc comparisons less reliable. Mild departures are common and not very consequential — AICc is fairly robust to modest non-normality when all competing models share the same

error assumption.

A small number of crop-country series showed a pattern where lower-tail points fell well below the diagonal while the upper tail tracked the line or fell slightly below it, indicating a left-skewed distribution with a heavy lower tail. The clearest instance was maize in the United States, where two or three points in the lower left sat around a sample quantile of -4 against a theoretical quantile of -2.5 . The pattern reflects years of severe yield shortfall far below the fitted trend. For US maize, those were likely drought years such as 2012. Agricultural data commonly show this asymmetry: catastrophic yield failures (drought, frost, pest outbreak) produce crashes well below trend, while good years tend to be only moderately above it. The departure is not a major concern. It is driven by a handful of observations rather than a systematic distributional shift, and since the same extreme observations are present for all 10 candidate models, the relative AICc rankings are not materially affected.

B.3 Leave-one-out influential point check

For every crop-country series whose best model contains one or more breakpoints (“knot models”), each observation is removed in turn and full model selection (all 10 candidates, AICc, quadratic/exponential/quadratic-plateau excluded, minimum segment length 10) is run on the reduced dataset. Two quantities are recorded for each LOO replicate: which model was selected, and — if the same model was selected as on the full dataset — the estimated knot year(s).

The LOO check separates two distinct failure modes. *Knot instability* occurs when the same model wins on the reduced dataset but the estimated knot year shifts substantially; this indicates that the breakpoint location is driven by a single observation rather than by a genuine change in trend. *Model instability* occurs when a different model is preferred on the reduced dataset; in this case comparing knot positions is not meaningful. By recording which model won on each

LOO replicate, the diagnostics can flag both types of instability independently.

For each series, two panels are produced. The first is a scatter plot of knot year(s) against the removed observation index, coloured by whether the same model was selected (grey) or a different model (red), with the full-data estimate shown as a dashed red line. The second is a bar chart of which models were selected across all LOO replicates, with the full-data best model highlighted in blue. A series is flagged if any same-model replicate shifts the estimated knot year by more than five years.

B.4 Bootstrap knot stability

To characterise the sampling uncertainty around each fitted knot year, a pairs bootstrap is applied: year–yield pairs are resampled with replacement 200 times. Full model selection (all 10 candidates, AICc, quadratic/exponential/quadratic-plateau excluded, minimum segment length 10) is run on each replicate. Only replicates where the same model is selected as on the full dataset contribute to the knot distribution, for the same reason as in the LOO check: if a different model wins, comparing knot positions is not meaningful. The plot title reports the number of same-model replicates as a fraction of total replicates.

For each knot, a histogram of same-model estimates is produced with the original full-data estimate (red dashed line) and the 2.5th and 97.5th percentiles (navy dotted lines) marked. A 95% bootstrap confidence interval wider than ten years is flagged in the plot title.

A wide interval signals that the data do not precisely constrain the timing of the trend change, even if model selection strongly prefers a breakpoint model over a linear one. In those cases the knot year should not be interpreted as a precise historical date.

Note on temporal autocorrelation. The pairs bootstrap treats observations as exchangeable and ignores autocorrelation in the yield series. The simplification is appropriate here because the goal is to assess *knot location sensitivity* under plausible perturbations of the data, not to make formal inference on the mean function. A block bootstrap would be more appropriate for

the latter.

B.5 Diagnostic settings

Each series is bootstrapped 200 times. Leave-one-out and bootstrap refits use a coarser knot grid than the main model fitting, 16 points instead of 32, to keep run time tractable (about 40 minutes in total). That is acceptable because the goal is stability assessment, not precise global optimisation. Bootstrap confidence intervals wider than ten years, and leave-one-out knot shifts larger than five years, are flagged in the plot titles.

C Robustness checks

We consider three alternative specifications (R1–R3) varying one parameter at a time from the preferred specification, and two comparisons (G1–G2) with parameter choices set to approximate the Grassini, Eskridge, and Cassman (2013) setup.

C.1 R1: minimum segment length 1, AICc, preferred model set

Table 8: R1 (minimum segment length 1, AICc, preferred model set): best-fit model(s) per crop-country series.

country	Maize	Rice	Soy	Wheat
Argentina	LLP		LLL/PLL/LLP/LL	PLP/PLL/LLP/LLL
Australia				LLP
Bangladesh		PLL		
Brazil	LLP		LLP/LL/PL/L/LPL	
Canada				LLP
China	LPL/LLL	LLL	LLL/LPL	LLL
France				LPL/LP
Germany				LP/LLP
India	PLL/LLL	PLL/LLL	LLP/LLL/PLP	LPL
Indonesia		LLL		
Mexico	LL/LLP/LPL/PLL/LLL			
Myanmar		LLL/PLL		
Pakistan				LPL/PLL/LLL
Paraguay			PLP	
Philippines		LLL		
Russia				PL/LL/PLL/LLL/LLP
Thailand		LLP/LLL		

(continued)

country	Maize	Rice	Soy	Wheat
Türkiye				LPL
Ukraine	PLP/PL			LPL/PL/PLL/LL
United Kingdom				PLP/PLL
United States	L		LLP	L/LPL
Vietnam		LLL/PLL		

Table 9: R1 (minimum segment length 1, AICc, preferred model set): share of global production (%) by terminal-form class.

crop	linear increase	plateau	unclassified	not analyzed
Maize	59	11	5	25
Rice	77	0	5	18
Soy	6	37	49	8
Wheat	47	11	18	24

C.2 R2: minimum segment length 10, BIC, preferred model set

Table 10: R2 (minimum segment length 10, BIC, preferred model set): best-fit model(s) per crop-country series.

country	Maize	Rice	Soy	Wheat
Argentina	LLP		LP/LL	PLP/L
Australia				LL
Bangladesh		PLL		
Brazil	LL		L	
Canada				LLP
China	LL	LL	LPL	L
France				LP
Germany				LP
India	PLL	LLL/PL	LLP/PLP/L	LLL
Indonesia		LL		
Mexico	LL			
Myanmar		LP		
Pakistan				L
Paraguay			LLP/PLL	
Philippines		LP/LL/PLL		
Russia				LL/PL
Thailand		LLP		
Türkiye				LPL
Ukraine	L			PL
United Kingdom				LP
United States	L		LL/PL	L
Vietnam		PLL/LLL		

Table 11: R2 (minimum segment length 10, BIC, preferred model set): share of global production (%) by terminal-form class.

crop	linear increase	plateau	unclassified	not analyzed
Maize	71	3	0	26
Rice	71	8	2	19
Soy	70	0	22	8
Wheat	60	15	2	23

C.3 R3: minimum segment length 10, AICc, full model set

Table 12: R3 (minimum segment length 10, AICc, full model set): best-fit model(s) per crop-country series.

country	Maize	Rice	Soy	Wheat
Argentina	LLP		Q/LLP/LL	PLP
Australia				LL/LPL/PLL
Bangladesh		PLL		
Brazil	LLL/LL		Q/LL/L	
Canada				LLP
China	LLL/LL	LL	LPL	L/Q/LL
France				LP/LL
Germany				LP/LLP
India	PLL/LLL	LLL/Q	LLP/PLP	LLL
Indonesia		LL		
Mexico	Q/LL			
Myanmar		LLP/LP		
Pakistan				L/Q/LL
Paraguay			LLP/PLL	
Philippines		PLL/Q/LL		
Russia				Q/E
Thailand		LLP/LLL		
Türkiye				LPL
Ukraine	L			Q/PL
United Kingdom				PLL/LP/LL
United States	L/Q		E/Q	L/Q
Vietnam		LLL/PLL		

Table 13: R3 (minimum segment length 10, AICc, full model set): share of global production (%) by terminal-form class.

crop	linear increase	plateau	unclassified	not analyzed
Maize	34	3	37	26
Rice	49	4	29	18
Soy	6	4	83	7
Wheat	19	9	48	24

C.4 AICc vs. BIC: differences in classification

6 crop-country series receive a different terminal-form class depending on whether AICc or BIC is used as the selection criterion (with all other settings equal to the preferred specification: minimum segment length 10).

Table 14: Crop-country series assigned to different terminal-form classes under AICc vs. BIC (minimum segment length 10, preferred model set).

Crop	Country	AICc class	BIC class
Rice	Philippines	linear	unclassified
Rice	Thailand	unclassified	plateau
Soy	India	plateau	unclassified
Wheat	Argentina	plateau	unclassified
Wheat	France	unclassified	plateau
Wheat	United Kingdom	unclassified	plateau

Table 15: Number of crop-country series classified as unclassified under each criterion.

Criterion	Unclassified
AICc	5
BIC	5

C.5 Effect of including the dropped smooth-curve models

11 crop-country series receive a different terminal-form class when the three smooth-curve forms (quadratic, exponential, and quadratic-plateau) are included as candidates (all other settings equal to the preferred specification: minimum segment length 10, AICc).

Table 16: Crop–country series assigned to different terminal-form classes between the preferred set and the full set including quadratic, exponential, and quadratic-plateau (minimum segment length 10, AICc).

Crop	Country	Preferred set	Full set
Maize	Mexico	linear	unclassified
Maize	United States	linear	unclassified
Rice	India	linear	unclassified
Rice	Philippines	linear	unclassified
Soy	Brazil	linear	unclassified
Soy	United States	linear	unclassified
Wheat	China	linear	unclassified
Wheat	Pakistan	linear	unclassified
Wheat	Russia	linear	unclassified
Wheat	Ukraine	linear	unclassified
Wheat	United States	linear	unclassified

C.6 Effect of minimum segment length constraint

24 crop-country series had an unconstrained best-fit model with at least one segment shorter than 10 observations. Of these, 6 switched from a model ending in a plateau to one ending in a linear segment, and 3 switched in the opposite direction, when the constraint was applied.

Table 17: Crop–country series where the unconstrained best-fit model (R1: minimum segment length 1, AICc, preferred model set) had at least one segment shorter than 10 observations. Segment lengths refer to the unconstrained best-fit model; the constrained best model is from the preferred specification (minimum segment length 10, AICc). Model forms use the abbreviations given in Table 2.

Crop	Country	Best model (unconstrained)	Segment lengths	Best model (constrained)
Maize	Brazil	LLP	31 / 27 / 6	LLL
Maize	China	LPL	35 / 6 / 23	LLL
Maize	Ukraine	PLP	8 / 19 / 6	L
Rice	China	LLL	17 / 8 / 39	LL
Rice	India	PLL	8 / 40 / 16	LLL
Rice	Indonesia	LLL	18 / 5 / 41	LL
Rice	Myanmar	LLL	16 / 4 / 44	LLP
Rice	Philippines	LLL	14 / 7 / 43	PLL
Soy	Argentina	LLL	18 / 41 / 5	LLP
Soy	Brazil	LLP	29 / 29 / 6	LL
Soy	China	LLL	39 / 7 / 18	LPL
Soy	India	LLP	26 / 9 / 29	LLP
Soy	Paraguay	PLP	25 / 9 / 30	LLP
Soy	United States	LLP	52 / 3 / 9	LL
Wheat	Argentina	PLP	10 / 45 / 9	PLP
Wheat	Australia	LLP	59 / 2 / 3	LL

(continued)

Crop	Country	Best model (unconstrained)	Segment lengths	Best model (constrained)
Wheat	Canada	LLP	42 / 13 / 9	LLP
Wheat	China	LLL	17 / 6 / 41	L
Wheat	France	LPL	35 / 27 / 2	LP
Wheat	India	LPL	40 / 5 / 19	LLL
Wheat	Pakistan	LPL	54 / 5 / 5	L
Wheat	Russia	PL	8 / 25	LL
Wheat	Ukraine	LPL	5 / 10 / 18	PL
Wheat	United Kingdom	PLP	9 / 26 / 29	PLL

C.7 Comparison with Grassini et al. (2013)

Grassini, Eskridge, and Cassman (2013) found widespread yield plateaus across major crop-producing countries, whereas the present analysis finds very few. Several mechanisms could explain this discrepancy.

First, the two analyses cover different country sets. Grassini, Eskridge, and Cassman (2013) do not apply a formal production-share threshold; their 36 cases were selected to cover major producing areas globally and include three multi-country African regions (East, Central, and West Africa) for maize. No individual African country reaches our 2% global production threshold for maize, so African maize trends appear in Grassini et al.'s analysis but not in ours.

Second, the analyses cover different time periods: Grassini, Eskridge, and Cassman (2013) used data up to 2010, whereas the preferred specification here extends to 2024. The G1 and G2 runs (Table 18 and Table 19) truncate the series at 2010 while holding the model set and selection criterion at their preferred values, and plateau classifications remain rare. What this establishes is narrower than it may appear: shortening the data on its own does not reproduce Grassini et al.'s plateaus. That test does not show that the ending year is unimportant. Two series drive the disagreement: China rice and India wheat. For both, the crop-by-crop analysis below finds that a plateau reading requires the 2010 cutoff *together with* the narrower model set and a selection rule that resolves near-ties toward the plateau. The three differences act jointly rather than any one of them dominating.

Third, the two analyses differ in model set and selection criterion. Grassini, Eskridge, and Cassman (2013) evaluate six candidate models: linear, piecewise-linear, linear-upper plateau, linear-lower plateau, quadratic-plateau, and exponential. They select among them on root mean squared error (RMSE), with a 5% RMSE difference threshold to declare one model superior and a requirement that all parameters be statistically significant ($p < 0.01$). The present analysis evaluates 10 models and uses AICc for selection. Three consequences follow. First, our larger model set includes multiple piecewise-linear shapes with two or three segments (LL, LLL, PL, LPL, PLL), which can capture decelerating yield growth without invoking a hard plateau. Grassini et al.'s single piecewise-linear model offers fewer alternatives to the plateau classification for decelerating series, making plateau models relatively more competitive. Second, Grassini et al.'s requirement that breakpoint parameters be significant at $p < 0.01$ means borderline piecewise models are discarded, potentially leaving a plateau model as the best surviving alternative; AICc-based selection has no such significance filter. Third, when two models fall within 5% RMSE of each other, Grassini et al. assign a single classification, taking the lower-RMSE model as the tiebreaker. The present analysis instead leaves such cases unclassified when the competing models span different terminal-form classes, so some series that Grassini et al. classify as plateaus would be unclassified here.

Finally, Grassini, Eskridge, and Cassman (2013) use sub-national data for the United States (separate regions for rice, wheat, and maize). That disaggregation allows regional plateaus, such as rice in California, to be detected where they would be diluted or hidden in the national aggregates used in the present analysis.

A crop-by-crop comparison helps illustrate which of these mechanisms matter in practice. For rice, both analyses include Bangladesh, China, India, Indonesia, Philippines, Thailand, and Vietnam. The United States does not appear in either sample (Grassini et al. use California and South-Central regions, which also fall below our 2% production threshold at the national level). Grassini et al. additionally include Japan and Republic of Korea, both below our thresh-

old and together accounting for a small share of global rice production. Among Grassini et al.'s rice plateau findings, those in California and Republic of Korea reflect small-production systems and contribute little to the aggregate discrepancy. The key divergence is China. Grassini et al. classify China rice as a linear-upper plateau (LUP, equivalent to our LP), whereas the present analysis assigns a two-segment linear (LL) model. Both LL and LP are available in each framework, since the LL model corresponds directly to Grassini et al.'s piecewise-linear (PW) model, so model availability alone does not explain the disagreement. As the further test below shows, the time period, the selection criterion, and the presence of two-knot models in the candidate set all contribute. The left panel of Figure 5 shows why: China rice yield growth slowed around 1995–2010, producing a pattern that resembles a plateau when the time series ends there, and the linear-plateau model fitted to the truncated series flattens out accordingly. Yield growth resumed strongly after 2010, however, and our data extending to 2024 capture this continuation, resolving the apparent plateau as a cyclical fluctuation around an LL trend. Grassini et al. explicitly note that China rice exhibits severe positive serial autocorrelation, which can affect estimated variances and plateau detection. The only rice series classified as a plateau in the present analysis is Myanmar, which does not appear in Grassini et al.'s sample.

A test illustrates how time period and model set interact for China rice. We repeat the analysis with data truncated at 2010 and two-knot models excluded from the candidate set, approximating Grassini et al.'s framework. LP then has the lowest AICc, but the two-segment linear (LL) model is within $\Delta\text{AICc} = 0.19$. Under the present analysis's classification rules this case would be unclassified rather than assigned to either class. In other words, with data ending in 2010 and no two-knot alternatives, LP and LL are empirically indistinguishable, and Grassini et al.'s selection procedure resolves the ambiguity toward the plateau. LL only pulls clearly ahead of LP once post-2010 resumed growth provides unambiguous evidence for a continuing positive slope.

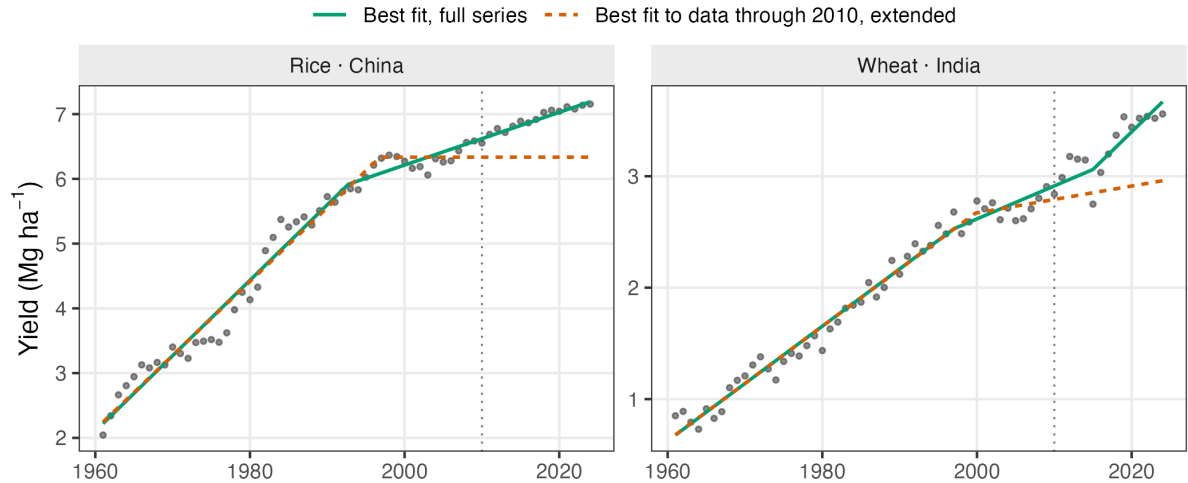


Figure 5: What a pre-2010 reading implied, against what subsequently happened, for the two series that drive the disagreement with Grassini, Eskridge, and Cassman (2013). Points are FAO yields. The solid line is the best-fitting model under the preferred specification, fitted to the full series. The dashed line is the best-fitting model under the approximation of Grassini, Eskridge, and Cassman (2013)’s framework described above — data truncated at 2010, two-knot models excluded — extended past the cutoff (dotted vertical line) to show the trajectory it implied. That model is a linear-plateau (LP) for China rice and a two-segment linear (LL) with a near-flat second segment for India wheat. In each series LP and LL are within $\Delta AICc$ 0.2 of each other, so the choice between them is close to arbitrary. In both series the post-2010 observations rise well above the implied trajectory.

The wheat comparison is less straightforward. The two analyses share seven countries: China, India, Canada, Australia, France, Germany, and Argentina. The present analysis additionally includes Russia, Pakistan, Ukraine, and Turkey — large producers absent from Grassini et al.’s sample — bringing coverage to 76.3% of global wheat production in 2020; Grassini et al.’s sample captures 56% (base year unspecified). Grassini et al. also include Denmark and the Netherlands (both below our 2% production threshold), and analyze the United States as three sub-national Great Plains regions rather than the national aggregate used here. Despite covering a smaller share of production, Grassini et al. find plateaus in 27% of their sample, compared to 9% in the present analysis.

Several factors drive this gap. The main case is India: Grassini et al. classify India wheat as a linear-upper plateau (LUP), whereas the present analysis assigns a three-segment linear (LLL)

model. As with China rice, the right panel of Figure 5 shows wheat yields in India dipping between approximately 2000 and 2010, a pattern that resembles a plateau when the time series ends there. Growth resumed after 2010, and the present analysis, which extends to 2024, captures the dip as the middle segment in an LLL model rather than as a plateau. A smaller contribution comes from France, which the present analysis leaves unclassified (LP and LL models within $\Delta\text{AICc} \leq 2$) but Grassini et al. classify as a linear-upper plateau. Finally, Grassini et al. find plateaus in Denmark and the Netherlands — two high-intensity wheat-producing systems — but both fall below our 2% production threshold and do not appear in the present analysis.

An analogous exercise for India wheat is less clear-cut. With data through 2010 and two-knot models excluded, the best-fitting model is linear-linear (LL), but the post-knot slope is small (12 kg/ha/year) and statistically indistinguishable from zero ($p = 0.204$). The AICc difference between LL and linear-plateau (LP) is only 0.02, meaning the two models are effectively indistinguishable by this criterion. Under the present analysis's own classification rules, which flag cases as unclassified when competing models within $\Delta\text{AICc} \leq 2$ span different terminal-form classes, India wheat would be unclassified rather than linear. Grassini et al.'s requirement that all estimated parameters meet $p < 0.01$ eliminates LL — the post-knot slope falls well short of this threshold — again resolving the ambiguity toward the plateau. As with China rice, the pre-2010 yield series is genuinely ambiguous between LP and LL, and Grassini et al.'s selection procedure resolves that ambiguity toward the plateau.

The China rice and India wheat cases thus reveal a consistent pattern: under Grassini et al.'s conditions — data through 2010, no two-knot models, and a strict significance filter — both series produce LP/LL near-ties that their selection procedure resolves toward the plateau. Under the present analysis, post-2010 data make the continued growth trend unambiguous, and a larger model set provides additional alternatives to the plateau. The discrepancy between the two analyses therefore reflects not a fundamental disagreement about the underlying trends, but a

combination of a shorter time series, a more limited model set, and a selection criterion that consistently breaks ambiguous cases in favor of the plateau.

For maize, the two analyses reach similar aggregate results — 3% and 5% of global production classified as a plateau — though through different specific cases. The present analysis finds a plateau for Argentina under a linear-linear-plateau (LLP) model, where the first two segments (acceleration then slower growth) correspond closely to the piecewise model with increasing rate (PW+) that Grassini et al. identify. The plateau is the final segment and falls largely after 2000, so it would not have been detectable with data ending in 2010. Grassini et al. find plateaus for France, Italy, and East Africa (a multi-country aggregate), none of which appear in the present analysis: France and Italy fall below our 2% production threshold, and no individual East African country reaches it. The near-identical aggregate plateau shares thus reflect different underlying cases rather than genuine agreement.

C.7.1 G1: data up to 2010, minimum segment length 10

Table 18: G1 (data up to 2010, minimum segment length 10, AICc, preferred model set): share of global production (%) by terminal-form class.

crop	linear increase	linear decrease	plateau	unclassified	not analyzed
Maize	75	0	0	0	25
Rice	82	0	0	0	18
Soy	80	3	6	4	7
Wheat	61	0	8	7	24

C.7.2 G2: data up to 2010, minimum segment length 1

Table 19: G2 (data up to 2010, minimum segment length 1, AICc, preferred model set): share of global production (%) by terminal-form class.

crop	linear increase	linear decrease	plateau	unclassified	not analyzed
Maize	71	0	0	3	26
Rice	77	0	5	0	18
Soy	84	3	0	6	7
Wheat	46	0	8	22	24

D Additional Results

D.1 Production share by terminal-form class

Table 20 gives the percentages plotted in Figure 2.

Table 20: Share of global production (%) accounted for by countries in each terminal-form class, by crop. The terminal-form class is defined by the trend in the final segment of the fitted model; a series is ‘unclassified’ when models within $\Delta\text{AICc} \leq 2$ span more than one terminal-form class. The ‘not analyzed’ column is the share of world production in countries below the 2% inclusion threshold, so rows sum to 100.

crop	linear increase	plateau	unclassified	not analyzed
Maize	71	3	0	26
Rice	74	4	5	17
Soy	70	4	18	8
Wheat	60	9	7	24

D.2 Yield gaps and closure rates

Table 21 gives the series-level figures plotted in Figure 3, together with the potential and actual yield levels and the closure rate measured against full potential yield and the attainable ceiling.

Table 21: Potential yield, actual yield, actual yield as a percentage of potential, rate of yield gap closure, and rate of attainable yield gap closure for crop-country series with available YPYGA data. Potential and actual yields are in Mg ha^{-1} . Actual yield is the model-predicted average over the post-2010 years with available YPYGA data. Closure rate is the percentage of the yield gap closed per year; attainable closure rate uses 80% of potential yield as the ceiling. Series whose terminal-form class is uncertain are labelled “unclassified”, and no closure rate is reported for them. Where a classified series has more than one model within $\Delta\text{AICc} \leq 2$, each distinct rate implied by those models is shown, separated by “/”.

Crop	Country	Pot. yield	Actual yield	% of pot.	Closure rate	Att. closure rate
Maize	Argentina	14.2	6.8	48	0	0
Maize	Brazil	11.1	4.8	43	1/1.9	1.5/2.9
Maize	China	12.0	5.7	48	1	1.6/1.7
Maize	Mexico	16.4	8.0	49	1.9	3.2
Maize	United States	14.5	10.4	72	2.8	9.8
Rice	China	9.7	6.7	69	1.4	4
Rice	India	7.2	3.6	50	1.9	3.1
Rice	Indonesia	9.3	5.0	53	0.8	1.4
Rice	Myanmar	8.8	3.9	44	0	0

(continued)

Crop	Country	Pot. yield	Actual yield	% of pot.	Closure rate	Att. closure rate
Rice	Philippines	9.5	3.8	40	0.8/0.4	1.2/0.7
Rice	Thailand	8.3	3.0	36	unclassified	unclassified
Rice	Vietnam	9.4	5.6	60	1.3	2.5
Soy	Argentina	4.7	2.7	58	unclassified	unclassified
Soy	Brazil	5.7	3.0	52	1.6/1.4	2.8/2.5
Soy	China	3.9	1.8	47	1	1.6
Soy	Paraguay	3.7	2.6	70	unclassified	unclassified
Soy	United States	5.2	3.1	60	1.7	3.4
Wheat	Argentina	6.6	2.9	44	0	0
Wheat	Australia	3.0	1.9	64	9.5/6.1/9.4	21.1/13.5/20.8
Wheat	Canada	5.3	3.1	60	0	0
Wheat	China	11.8	5.1	43	1.3	2/1.9
Wheat	India	5.9	3.0	50	2.3	3.9
Wheat	United States	5.7	3.1	54	1	1.7

D.3 Global average yield trends

Figure 6 shows production-weighted global average yields for each crop from 1961 to the most recent available year, fitted with the same model selection procedure used in the main analysis. The dashed line projects the best-fitting model to 2050 and the shaded band is the 95% projection interval described in Section 2. The red diamond marks the 2050 demand target, expressed as the required increase over the crop's 2020 yield, from Falcon, Naylor, and Shankar (2022) for maize, rice, and soy and from IFPRI IMPACT for wheat (which Falcon et al. do not cover).

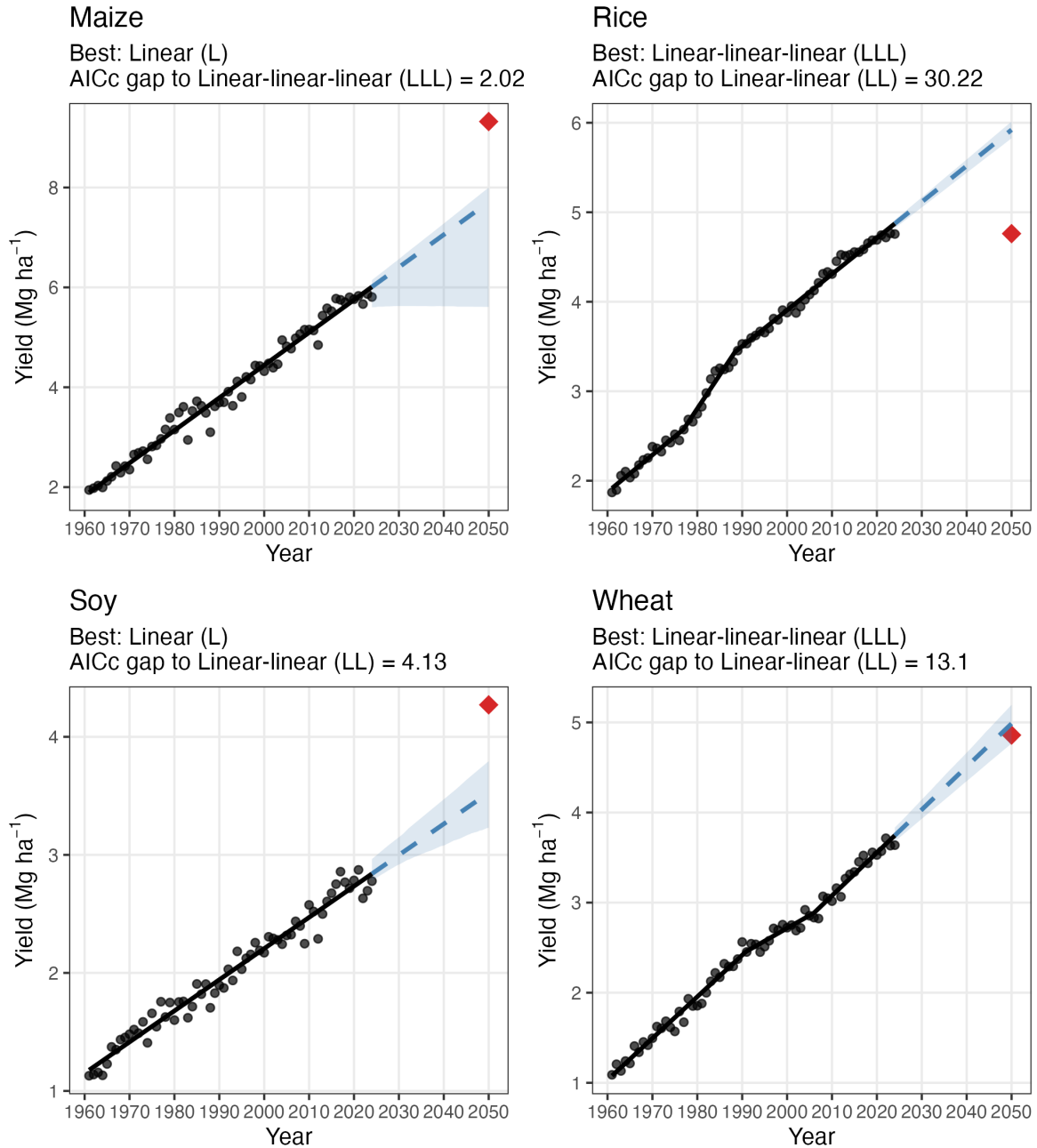


Figure 6: Global average yield trends and projections to 2050. Points are production-weighted observed yields; solid line is the best-fitting model; dashed line projects the trend to 2050. The shaded band is the 95% projection interval (2.5–97.5 percentiles of a Monte Carlo that combines AICc model-averaging weights with parametric coefficient uncertainty); it widens where the leading models diverge in their extrapolated trend. The red diamond is the 2050 demand target, expressed as the required increase over the crop’s 2020 yield, from Falcon, Naylor, and Shankar (2022) for maize, rice, and soy and from IFPRI IMPACT for wheat (which Falcon et al. do not cover).

D.4 Aggregate calorie-production index

Table 22 reports the projected increase in each crop's global average yield between 2005 and 2050 (extrapolating the best-fit model, with the 95% projection interval from Section 2) alongside its share of the four crops' combined calorie production in 2005. The calorie-weighted aggregate in the final row is the projected growth in total calorie production from the four crops, holding harvested area fixed. It is the benchmark compared against Tilman et al. (2011)'s projected ~100% growth in total crop-calorie demand over 2005–2050 in Section 3. Production is converted to calories using energy-content factors from the USDA FoodData Central database, with paddy rice converted to milled rice (milling yield 0.67). The four crops account for roughly 64% of the total crop calories in Tilman et al.'s global accounting.

Table 22: Projected increase in global average yield between 2005 and 2050 and share of 2005 calorie production, by crop, with the calorie-weighted aggregate. Yields are extrapolated from the best-fit model; parenthetical ranges are the 95% projection interval (Section 2). The aggregate growth is the calorie-share-weighted average of the per-crop yield growth, equivalent to a base-weighted (Laspeyres) calorie-production index. Conversion factors are from USDA FoodData Central; paddy rice is converted to milled rice.

Crop	Yield increase (%)	95% interval	Calorie share (%)
Maize	62	(18, 68)	36
Rice	44	(42, 46)	22
Wheat	74	(64, 82)	29
Soy	51	(38, 59)	13
All 4 crops (calorie-weighted)	60	(43, 63)	100

D.5 Model coefficients

Table 23 lists the estimated coefficients from the best-fitting model for each crop-country series under the preferred specification. Knot values represent years elapsed since the first observation year for each series.

Table 23: Estimated coefficients from the best-fitting model for each crop-country series (preferred specification).

Country	Crop	Model	Coefficient	Value
Argentina	Maize	LLP	intercept	1614.98
Argentina	Maize	LLP	slope_1	75.66
Argentina	Maize	LLP	slope_2	241.42
Argentina	Maize	LLP	knot1	32.84
Argentina	Maize	LLP	knot2	44.00
Brazil	Maize	LLL	intercept	1197.06
Brazil	Maize	LLL	slope_1	24.11
Brazil	Maize	LLL	slope_2	130.43
Brazil	Maize	LLL	slope_3	62.12
Brazil	Maize	LLL	knot1	30.43
Brazil	Maize	LLL	knot2	54.00
China	Maize	LLL	intercept	1193.64
China	Maize	LLL	slope_1	91.04
China	Maize	LLL	slope_2	125.87
China	Maize	LLL	slope_3	62.12
China	Maize	LLL	knot1	15.36
China	Maize	LLL	knot2	30.00
India	Maize	PLL	intercept	1030.14
India	Maize	PLL	slope_2	34.29
India	Maize	PLL	slope_3	80.95
India	Maize	PLL	knot1	44.65
India	Maize	PLL	knot2	16.54
Mexico	Maize	LL	intercept	925.85
Mexico	Maize	LL	slope_pre	34.22
Mexico	Maize	LL	slope_post	61.80
Mexico	Maize	LL	knot	27.00
Ukraine	Maize	L	intercept	2017.49
Ukraine	Maize	L	slope	168.74
United States	Maize	L	intercept	3991.05
United States	Maize	L	slope	116.69
Bangladesh	Rice	PLL	intercept	1679.17
Bangladesh	Rice	PLL	slope_2	50.58
Bangladesh	Rice	PLL	slope_3	87.26
Bangladesh	Rice	PLL	knot1	12.66
Bangladesh	Rice	PLL	knot2	33.00
China	Rice	LL	intercept	2219.99
China	Rice	LL	slope_pre	116.52
China	Rice	LL	slope_post	40.72
China	Rice	LL	knot	31.73
India	Rice	LLL	intercept	1423.09
India	Rice	LLL	slope_1	21.68
India	Rice	LLL	slope_2	46.68
India	Rice	LLL	slope_3	67.37
India	Rice	LLL	knot1	12.94
India	Rice	LLL	knot2	48.17
Indonesia	Rice	LL	intercept	1430.29
Indonesia	Rice	LL	slope_pre	99.38
Indonesia	Rice	LL	slope_post	34.85
Indonesia	Rice	LL	knot	26.96

(continued)

Country	Crop	Model	Coefficient	Value
Myanmar	Rice	LLP	intercept	1282.68
Myanmar	Rice	LLP	slope_1	62.99
Myanmar	Rice	LLP	slope_2	44.83
Myanmar	Rice	LLP	knot1	22.96
Myanmar	Rice	LLP	knot2	49.00
Philippines	Rice	PLL	intercept	1377.45
Philippines	Rice	PLL	slope_2	91.96
Philippines	Rice	PLL	slope_3	44.99
Philippines	Rice	PLL	knot1	9.95
Philippines	Rice	PLL	knot2	21.00
Thailand	Rice	LLP	intercept	1741.72
Thailand	Rice	LLP	slope_1	10.50
Thailand	Rice	LLP	slope_2	61.91
Thailand	Rice	LLP	knot1	44.64
Thailand	Rice	LLP	knot2	29.25
Vietnam	Rice	LLL	intercept	1935.52
Vietnam	Rice	LLL	slope_1	8.52
Vietnam	Rice	LLL	slope_2	107.03
Vietnam	Rice	LLL	slope_3	48.28
Vietnam	Rice	LLL	knot1	18.60
Vietnam	Rice	LLL	knot2	50.62
Argentina	Soy	LLP	intercept	858.24
Argentina	Soy	LLP	slope_1	57.03
Argentina	Soy	LLP	slope_2	24.58
Argentina	Soy	LLP	knot1	18.00
Argentina	Soy	LLP	knot2	54.00
Brazil	Soy	LL	intercept	989.17
Brazil	Soy	LL	slope_pre	31.39
Brazil	Soy	LL	slope_post	43.71
Brazil	Soy	LL	knot	27.00
China	Soy	LPL	intercept	682.63
China	Soy	LPL	slope_1	27.37
China	Soy	LPL	slope_3	20.37
China	Soy	LPL	knot1	37.00
China	Soy	LPL	knot2	47.25
India	Soy	LLP	intercept	1056.15
India	Soy	LLP	slope_1	-14.70
India	Soy	LLP	slope_2	36.45
India	Soy	LLP	knot1	25.00
India	Soy	LLP	knot2	35.16
Paraguay	Soy	LLP	intercept	1621.95
Paraguay	Soy	LLP	slope_1	-7.86
Paraguay	Soy	LLP	slope_2	111.20
Paraguay	Soy	LLP	knot1	23.37
Paraguay	Soy	LLP	knot2	33.41
United States	Soy	LL	intercept	1616.42
United States	Soy	LL	slope_pre	17.91
United States	Soy	LL	slope_post	35.32
United States	Soy	LL	knot	22.07
Argentina	Wheat	PLP	intercept	1367.15
Argentina	Wheat	PLP	slope_2	35.66

(continued)

Country	Crop	Model	Coefficient	Value
Argentina	Wheat	PLP	knot1	9.85
Argentina	Wheat	PLP	knot2	54.00
Australia	Wheat	LL	intercept	1122.29
Australia	Wheat	LL	slope_pre	13.96
Australia	Wheat	LL	slope_post	103.10
Australia	Wheat	LL	knot	53.89
Canada	Wheat	LLP	intercept	1436.27
Canada	Wheat	LLP	slope_1	21.23
Canada	Wheat	LLP	slope_2	72.98
Canada	Wheat	LLP	knot1	41.00
Canada	Wheat	LLP	knot2	54.00
China	Wheat	L	intercept	526.80
China	Wheat	L	slope	87.61
France	Wheat	LP	intercept	2638.46
France	Wheat	LP	slope	124.83
France	Wheat	LP	knot	34.73
Germany	Wheat	LP	intercept	2935.45
Germany	Wheat	LP	slope	114.38
Germany	Wheat	LP	knot	39.84
India	Wheat	LLL	intercept	677.39
India	Wheat	LLL	slope_1	51.42
India	Wheat	LLL	slope_2	29.58
India	Wheat	LLL	slope_3	67.49
India	Wheat	LLL	knot1	36.00
India	Wheat	LLL	knot2	54.00
Pakistan	Wheat	L	intercept	757.96
Pakistan	Wheat	L	slope	38.06
Russia	Wheat	LL	intercept	1505.39
Russia	Wheat	LL	slope_pre	36.05
Russia	Wheat	LL	slope_post	88.58
Russia	Wheat	LL	knot	20.00
Türkiye	Wheat	LPL	intercept	963.05
Türkiye	Wheat	LPL	slope_1	42.43
Türkiye	Wheat	LPL	slope_3	40.91
Türkiye	Wheat	LPL	knot1	37.82
Türkiye	Wheat	LPL	knot2	25.27
Ukraine	Wheat	PL	intercept	2840.29
Ukraine	Wheat	PL	slope	103.01
Ukraine	Wheat	PL	knot	14.99
United Kingdom	Wheat	PLL	intercept	4044.67
United Kingdom	Wheat	PLL	slope_2	199.05
United Kingdom	Wheat	PLL	slope_3	28.79
United Kingdom	Wheat	PLL	knot1	26.75
United Kingdom	Wheat	PLL	knot2	11.12
United States	Wheat	L	intercept	1740.34
United States	Wheat	L	slope	25.47

Drivers of Nitrogen Surplus in US

Agriculture: A Decomposition Analysis

Linus Blomqvist

1 Introduction

Synthetic nitrogen fertilizer underpins the modern food supply. The Haber–Bosch synthesis enables production of roughly 40% of the protein in the human diet, and an estimated two billion people are alive today because of it (Galloway et al. 2008). That achievement has come by profoundly reshaping the global nitrogen cycle. Human activities now create reactive nitrogen — the chemically and biologically active compounds, such as ammonia, nitrate, and nitrous oxide, as opposed to the inert N_2 that makes up most of the atmosphere — far faster than natural terrestrial processes do. Anthropogenic creation rose from about 15 Tg N yr^{-1} in 1860 to roughly 187 Tg N yr^{-1} in 2005; since 1970, this creation of reactive nitrogen has outpaced human population growth (Galloway et al. 2008).

Once created, reactive nitrogen cascades — moving in sequence from one environmental compartment to the next — through the atmosphere, soils, and waters, contributing to a series of environmental problems before it is eventually denitrified back to N_2 (Galloway et al. 2008). In fresh and coastal waters, nitrogen enrichment drives eutrophication and oxygen depletion.

The northern Gulf of Mexico hosts one of the world's largest coastal “dead zones”, reaching some 20,000 km² each summer. The zone is fed by a roughly threefold rise in the Mississippi River's nitrate flux since the mid-twentieth century (Rabalais, Turner, and Wiseman 2002). In the atmosphere, reactive nitrogen is a major contributor to fine particulate matter, and nitrous oxide is both the third most important long-lived greenhouse gas and the largest remaining anthropogenic threat to stratospheric ozone. On land, nitrogen deposition ranks among the leading drivers of biodiversity loss (De Vries 2021). Reactive nitrogen also bears directly on human health: nitrate leaching from farmland into groundwater contaminates drinking water across many agricultural regions (De Vries 2021), with adverse health outcomes documented at concentrations below current regulatory limits (Ward et al. 2018; Stayner et al. 2017; Semprini 2025).

Agriculture is the dominant source of this pollution. In the United States it accounts for more than half of nitrate, ammonia, and nitrous-oxide emissions (Ribaudo et al. 2011), and most of the nitrogen reaching the Gulf originates on cropland (Rabalais, Turner, and Wiseman 2002). The damages are large enough that, across the European Union, the social cost of agricultural nitrogen pollution (€35–230 billion yr⁻¹) has been estimated to exceed the value nitrogen adds in primary production (€20–80 billion yr⁻¹) (Van Grinsven et al. 2013). These harms trace to a single quantity: the nitrogen balance, or surplus. It is the sum of all nitrogen inputs to a field — synthetic fertilizer, manure, biological fixation, and atmospheric deposition — minus the nitrogen removed in harvest. What is left in the field is at risk of loss to air and water, potentially driving the cascade above (X. Zhang et al. 2015; Lassaletta et al. 2014). The balance measures that risk rather than the loss itself. Not all of the surplus is lost, and not all of what is lost does harm. Some accumulates in soil organic matter rather than escaping. Some is denitrified in the field to inert N₂, which ends the cascade instead of feeding it (Galloway et al. 2008). The balance is therefore best read as a measure of pressure on the environment, not of damage done (Oenema, Kros, and De Vries 2003).

There is potential to narrow the balance without sacrificing production by using nitrogen more efficiently, raising the share of applied nitrogen that ends up in the crop rather than lost. How well this is done is conventionally judged by nitrogen-use efficiency (NUE): the ratio of nitrogen output to nitrogen input, alongside the related recovery and partial-balance ratios of the agronomic literature (Fixen et al. 2015). NUE has the appeal of any ratio: it is dimensionless and comparable across crops and countries. But it has two limitations that bear directly on our task.

The first limitation is a flaw in the metric: it confounds efficiency with composition. Aggregated to a region or a nation, NUE is a ratio of summed outputs to summed inputs, or equivalently an input-weighted average of the efficiencies of individual crops in individual places. Such an average can mask large underlying variation: in the United States, only about half of the major crops show rising efficiency while the rest decline (J. Zhang, Cao, and Lu 2021), and outcomes vary widely across regions and crops (Lu et al. 2019). It also moves whenever its weights move. A region can raise its aggregate NUE by shifting its crop mix toward inherently less nitrogen-intensive crops, or by shifting the geographic distribution of output toward regions where the same crops score higher, with no change whatever in field management. What we want to see — how efficiently each crop is grown in each place, the level at which farm practice actually operates — is what the aggregate averages away. This confounding is well documented. The share of low-efficiency fruit and vegetable production alone explains roughly half the NUE gap between China and the United States (X. Zhang et al. 2015), crop mix is a significant statistical determinant of national NUE (Ren et al. 2022), and the same confound pervades the recovery and partial-balance ratios (Bouwman, Van Dreht, and Van der Hoek 2005; Fixen et al. 2015).

Existing studies recognise the confound without removing it. They quantify composition cross-sectionally but never decompose it from the US time trend they read as improved management (X. Zhang et al. 2015), are unable to separate crop mix at all for want of crop-level data (Yan,

Xia, and Ti 2022), or attribute drivers through statistical models that do not partition the aggregate change into exact, additive components (Liu et al. 2025).

The second limitation is inherent to any efficiency ratio: NUE can say efficiency improved, but not whether the improvement was big enough to matter. The surplus is a mass: production multiplied by the surplus each unit of production leaves behind. It falls only if that surplus per unit declines faster than production grows. Weighing them requires both on the same scale, and the scale that matters is mass (kilograms of nitrogen). NUE cannot be put in kilograms. A ten percent rise in NUE does not convert into surplus avoided, so it cannot be set against the surplus that a larger harvest adds. Whether efficiency gained more than production cost cannot be answered in NUE's terms.

What is missing is a method that overcomes both limitations at once: separating the efficiency signal from the compositional channels and putting it on the same scale as production. We decompose the change in the aggregate US agricultural nitrogen balance between 1990 and 2022 into five components, defined over every crop and county. These are total harvested area (H), its distribution across counties (S_i), the mix of crops within counties (S_{ij}), crop yield (Y_{ij}), and the nitrogen balance intensity (NBI_{ij}), the field nitrogen balance per unit of output. Multiplied together, the first four give production: total area, split to the county and then to the crop, times yield. The fifth is the surplus each unit of production leaves behind. We claim no new measure of a field's efficiency. For one crop in one county, NBI_{ij} is an exact transform of that cell's NUE and carries the same information; any result here maps back to the NUE scale. What is new is the form, and it gives the NBI_{ij} term two advantages over the aggregate NUE it replaces. The first, and more consequential, is clean aggregation: because crop mix and geographic distribution are carried by their own terms, the NBI_{ij} contribution is free of the composition that moves that aggregate. The second is a common scale: because NBI_{ij} is a surplus per unit of output, its contribution comes out in kilograms of nitrogen, the units of the balance itself. It can therefore be weighed directly against the production terms.

This separation is what makes the accounting useful, in three ways. The first is separating a real improvement in nitrogen management from an apparent one. Because an aggregate efficiency ratio can rise through a change in crop mix or geography alone, a reported gain may reflect nothing that changed on any field. Isolating the field channel shows whether farmers are in fact leaving less nitrogen behind per unit of output. The second is that two regions can see their surplus rise or fall by the same amount through different channels. In one, the change came from nitrogen management. In the other, from the crop mix. A regional efficiency ratio moves in both cases and cannot say which channel was at work. The third is seeing how large the underlying changes are. A change in the national surplus is the net of the five factor effects, which can offset or reinforce one another. The size of the national change says nothing about the size of the changes that produced it.

We focus on the United States for two reasons. It is a major agricultural producer, and its nitrogen pollution carries the wide-ranging environmental and health costs set out above. It is also one of the few countries with an inventory that resolves nitrogen inputs — synthetic fertilizer and manure — to individual crops within each county (Hoskovec et al. 2026). Onto this base we layer the remaining inputs, biological fixation and atmospheric deposition, and subtract the nitrogen removed in harvest, giving a comprehensive nitrogen budget at the crop $\times$ county $\times$ year level. Input data of this kind have seldom been used to study aggregate nitrogen trends. Bringing them to the national balance, within an exact decomposition, is part of what is new here.

Applied to the United States from 1990 to 2022, the decomposition finds that field-level efficiency did improve, though by less than conventional measures imply, and by not quite enough to keep the national surplus from rising. The overstatement is compositional. Geographic distribution barely moved the surplus, but crop mix pushed it up by about 0.5 Tg N. Through an aggregate NUE, this crop-mix shift would read as improved efficiency, because it favours crops that rate high on the ratio. The decomposition assigns that shift to the crop-mix term

instead. What is left in the nitrogen balance intensity term is a field-level gain, smaller than an aggregate NUE would show. Even so, in kilograms it was substantial. The near-flat national surplus is the near-cancellation of two large forces: rising yields pushed it up by about 2.0 Tg N, while improving intensity pulled it down by about 1.4 Tg N. Efficiency took back most of what yields added, but not all, and the surplus still rose by about 0.8 Tg N. Holding the surplus even level would have taken more improvement than three decades delivered. The gain was also geographically narrow, a Corn Belt phenomenon. On the expanding acreage of the Plains, intensity moved the other way.

2 Methods

2.1 Decomposition framework

We attribute changes in aggregate US agricultural nitrogen balance to their underlying drivers using additive logarithmic-mean Divisia index (LMDI-I) decomposition (Ang and Liu 2001; Ang 2005). Total field nitrogen balance B (kg N yr⁻¹) is a net measure: the nitrogen added to a field from all sources, less the nitrogen removed in the harvested crop. A positive value is a surplus, a negative value a deficit. It is summed in each year over sub-categories indexed by county i and crop j . Every quantity is year-specific; the year subscript t is suppressed throughout to reduce clutter and reinstated only where the exposition refers to specific years. Each cell-level balance is written as a product of five factors:

$$B = \sum_{ij} B_{ij}, \quad B_{ij} = H \cdot S_i \cdot S_{ij} \cdot Y_{ij} \cdot \text{NBI}_{ij},$$

where H is total US harvested area, $S_i = H_i/H$ is county i 's share of that area, $S_{ij} = H_{ij}/H_i$ is crop j 's share of county i 's harvested area, $Y_{ij} = P_{ij}/H_{ij}$ is yield, and $\text{NBI}_{ij} = B_{ij}/P_{ij}$

is the nitrogen balance intensity (H_{ij} and P_{ij} being the cell's harvested area and production). For the headline decomposition we compare two *period-averaged states*: cell-level means of area, production, and B over a base window (1990–1994) and a terminal window (2018–2022), connected by a single step. The decomposition expresses the change between them, ΔB , as the exact sum of five additive factor effects with no residual. Three of these are area effects: a scale effect (H , the extensive margin of harvested area), a geographic-distribution effect (S_i , shifts in which counties hold the nation's area), and a crop-mix effect (S_{ij} , changes in the crop portfolio within counties). The remaining two are a yield effect (Y_{ij} , output per unit area) and a nitrogen-balance-intensity effect (NBI_{ij}). LMDI-I is consistent in aggregation, so county- and state-level contributions sum exactly to the national effect, which we exploit for sub-national attribution. We prefer this single-step comparison over a chained, year-on-year decomposition, for a reason specific to a signed quantity. The nitrogen balance is a surplus in most cells but a deficit in some, and a substantial share sit near the break-even point, changing sign — surplus to deficit or back — from one year to the next. When a cell's balance changes sign, the standard weighting cannot be formed, and the decomposition falls back on a rule that loads the cell's entire change onto the NBI effect. In a chained decomposition these fallback attributions accumulate year on year without cancelling, producing a cumulative NBI effect that swings by roughly 2.0 Tg N with the choice of terminal year. That swing is annual noise in near-zero cells, not a genuine trend. Averaging over a five-year window at each end removes the problem: near-zero cells average to a stable sign, so sign changes between the two averaged states are rare. For a long-run structural comparison the quantity of interest is in any case the net change between periods. The two-window design follows established practice for long-run structural LMDI comparisons (Kastner et al. 2012). A chained year-on-year decomposition is also computed over 1990–2022 and reported as a trajectory sensitivity that shows within-period dynamics the period comparison cannot resolve. The complete specification — the logarithmic-mean weighting, the analytical-limit treatment of crop cells that are born or die between years, the handling of sign-changing balance, and the computational algorithm — is

given in the Supplementary Information.

The nitrogen balance intensity is defined at the crop $\times$ county level as balance per unit of output, $\text{NBI}_{ij} = B_{ij}/P_{ij}$. It is a monotone transform of the conventional nitrogen-use-efficiency ratio $\text{NUE}_{ij} = N_{\text{out},ij}/N_{\text{in},ij}$: because $N_{\text{out},ij} = n_j P_{ij}$ for a crop-specific nitrogen content n_j , $\text{NBI}_{ij} = n_j(1/\text{NUE}_{ij} - 1)$, so any result maps back to the NUE scale. Knowing n_j — the harvest-removal coefficient (Table S4) — we compute the change in aggregate NUE from the same data and split it into genuine within-field efficiency and composition (SI). For one cell the two are equivalent measures. They diverge on aggregation, where n_j makes NBI comparable across crops and places the efficiency signal in the units of the balance (SI). Because the scale, geographic, crop-mix, and yield channels are carried by the separate H , S_i , S_{ij} , and Y_{ij} factors, the NBI effect isolates within-cell change in the surplus left per unit of output, net of area, geography, mix, and yield.

The cell-level balance is the nitrogen input minus the nitrogen removed in harvest,

$$B_{ij} = N_{\text{in},ij} - N_{\text{out},ij} = (N_{\text{fert},ij} + N_{\text{man},ij} + \text{BNF}_{ij} + N_{\text{dep},ij}) - N_{\text{out},ij},$$

the net nitrogen entering the field from synthetic fertilizer, manure, biological fixation, and atmospheric deposition that is not exported in grain or fiber. The balance is constructed identically to the field nitrogen balance (N_{sur}) of X. Zhang et al. (2015). A positive B_{ij} is nitrogen lost to the environment plus any net accumulation in the soil organic pool, and we reserve the phrase “losses to the environment” for that positive interpretation. Net change in the soil organic-nitrogen pool is not modelled separately but is absorbed into B , as in the standard soil-surface nitrogen balance (Oenema, Kros, and De Vries 2003; Van Beek, Brouwer, and Oenema 2003) and in X. Zhang et al. (2015). Reading the balance as environmental loss therefore assumes this term is small relative to the surplus, an acknowledged limitation.

2.2 Data sources and inputs

Nitrogen application. Synthetic fertilizer and manure nitrogen are taken from the nitrogen application data developed for the US National Greenhouse Gas Inventory (Hoskovec et al. 2026), a county $\times$ crop $\times$ year reconstruction covering 1990–2023 and built on the USDA National Resources Inventory survey. We combine fertilizer and manure and average the six multiply-imputed realizations to a headline applied input $N_{\text{appl},ij} = N_{\text{fert},ij} + N_{\text{man},ij}$, retaining the individual realizations for an imputation-sensitivity check. Application rates are modelled crop- and location-specifically with a gradient-boosting model for corn, cotton, soybeans, and wheat, with sorghum borrowing corn’s rates and barley and oats wheat’s. The remaining crops receive coarser average rates from survey data for lumped categories such as vegetables and horticulture. A top-down fertilizer-sales total is imposed as an over-application constraint across all crops. This coarse treatment of the non-major crops contributes to the elevated negative-balance prevalence in the specialty-crop tail discussed below. Because the inventory’s survey points are fixed in space, whether a given county is sampled at all is a property of the county rather than of the county-year, a feature the panel construction relies on.

Harvested area and production. Area and production come from the USDA National Agricultural Statistics Service (NASS) QuickStats database, queried at county and state level over 1990–2023 through the `rnassqs` interface (Potter 2019) for both the annual survey and the quinquennial Census of Agriculture. The two sources are coalesced cell by cell — survey where published, census filling its gaps — and all reported quantities are converted to a common mass basis using standard test weights. NASS withholds county observations when fewer than three operations report, so a substantial part of the pipeline is devoted to recovering the withheld values. A state-residual rule distinguishes genuine zeros from suppressed values, and log-linear interpolation between census anchors fills the resulting gaps. A yield-bound coupling step reconciles county sums to published state totals, bounding implausible imputed yields, and the poorly-covered potato and sweet-potato series are imputed separately. The Supplementary

Information describes these steps in full. The crops analysed, and the four NRI categories excluded for lack of usable county-level NASS coverage, are listed in Table S1. The analysis period is 1990–2022, with 2023 dropped as an incomplete NASS year. The poorly-covered potato and sugar-beet series have no 2023 values once their annual survey coverage ends and the 2022 census provides their last anchor, so retaining 2023 would register spurious crop disappearances.

Nitrogen removal in harvest. Nitrogen exported in the harvested product is production times a crop-specific removal coefficient, $N_{\text{out},ij} = P_{ij} n_j$, on a US agronomic harvested-product-removal basis: International Plant Nutrition Institute values for the major field crops (Roberts et al., n.d.) and Bouwman, Van Dreht, and Van der Hoek (2005) for the specialty crops, with potato and tobacco from crop-specific sources (Table S4). Winter and spring wheat carry distinct coefficients to reflect their differing grain protein. Cotton’s coefficient is the nitrogen in lint plus the cottonseed removed at ginning, which is the correct quantity for a field nitrogen balance: a lint-only figure would understate removal roughly thirtyfold.

Biological fixation. Legumes acquire much of their nitrogen by symbiotic fixation, which we add to their inputs ($\text{BNF}_{ij} = 0$ for non-legumes) to align the input definition with X. Zhang et al. (2015) and to prevent the spurious deficits that would otherwise arise from omitting a legume’s dominant nitrogen source. Whether fixed nitrogen belongs on the input side depends on the question. Ours is how much reactive nitrogen a field receives and does not export, and fixed nitrogen is reactive nitrogen subject to the same environmental fate as fertilizer, so it counts. A farm-gate accounting restricted to purchased nitrogen would ask instead what fertilizer policy can directly affect, and would place legume fields in deficit. The applied-only variant reported in the Supplementary Information shows what that choice does to the efficiency record. Because the pipeline observes only grain nitrogen, fixation is reconstructed by scaling N_{out} up to whole-plant nitrogen through a nitrogen harvest index and a belowground multiplier, then applying the fraction of plant nitrogen derived from the atmosphere. Parameters follow Herridge, Peoples,

and Boddey (2008) and Salvagiotti et al. (2008) (Table S5). The headline estimate includes belowground nitrogen, under which US soybean fixes on the order of 200 kg N ha^{-1} and is approximately nitrogen-neutral, consistent with that literature; an aboveground-only variant is retained for sensitivity analysis.

Atmospheric deposition. Deposition of reactive nitrogen (wet + dry, oxidized + reduced) is added to every cell as $N_{\text{dep},ij} = d_i H_{ij}$, where d_i is the deposition rate in county i (uniform across crops) and H_{ij} the cell’s harvested area. Because no single fine-resolution product spans the study period, we use the NADP Total Deposition product for 2000–2022 (Schwede and Lear 2014; National Atmospheric Deposition Program 2025) and the NACID reconstruction for 1990–1999 (Hember 2018), the latter harmonized to the TDep level over their overlap. Each annual raster is reduced to an area-weighted county mean. National cropland-county deposition is about $8\text{--}10 \text{ kg N ha}^{-1} \text{ yr}^{-1}$ and declines over the period as oxidized-nitrogen emissions fall, and a constant-back-cast sensitivity confirms the pre-2000 splice is immaterial to the results.

2.3 The field-balance panel

These components assemble into a sparse crop $\times$ county $\times$ year panel: 18 crops — NASS commodity $\times$ class, so winter and spring wheat enter separately — across 2,829 counties over 1990–2022, about 331,000 crop-county-year cells. The panel contains only cells with positive harvested area. Cells absent from an entire window (born after the base or died before the terminal) are handled by the analytical-limit rule in the period comparison, while persistently absent cells contribute nothing and are omitted. Because computing B_{ij} requires a nitrogen value from the inventory, the usable universe is bounded by inventory coverage. We therefore gate the panel on county-level membership of the fixed inventory sample frame. That gate keeps cells in surveyed counties even where a crop is unrecorded in a given year (a genuine input zero), and drops the small share of area in counties the inventory never samples.

Covering the entire harvested area is intrinsic to the question: the decomposition attributes the change in the *national* balance, and a crop-mix effect can only capture shifts among crops that are all in the panel. The specialty-crop tail is nonetheless a small share of area and nitrogen, so including it makes the decomposition complete without materially affecting the headline effects. The yield and balance-intensity effects rest on the well-measured major crops, as the robustness suite confirms.

Under the headline specification B is positive across the great majority of harvested area, as expected of a nitrogen balance. Negative balance touches about 11% of harvested area, concentrated in the low-input small grains and a few specialty crops together with cells in which the inventory records no input at all. Wheat alone accounts for about 64% of the negative-balance area, with barley and oats adding a further 9%. The Supplementary Information shows the negative balances to be an input-side phenomenon — low recorded application relative to harvest removal, stable across all six nitrogen imputations — rather than an artefact of the removal coefficients. They are in part a real drawdown of rotational or prior-year nitrogen that an annual, crop-specific balance does not carry. Because the change in B is explained between two period-averaged states and the four activity ratios (H , S_i , S_{ij} , Y) come entirely from NASS area and production, these cells do not materially affect the headline yield and balance-intensity conclusions.

The nitrogen input is not wholly confined to the NBI effect, however: all five factor effects share a common logarithmic-mean weight equal to the cell balance, so the input enters each effect through that weight. The weighting matters chiefly for the crop-mix effect, whose sign for a given crop depends on whether that crop is a net surplus or a net deficit. The small grains, whose balances are least certain, are precisely the crops carrying negative crop-mix contributions. That part of the crop-mix attribution therefore inherits their balance uncertainty, an issue we take up in the Discussion. A bounding sensitivity that removes the suspect small-grain negatives leaves the yield and balance-intensity effects essentially unchanged and shifts only the crop-

mix effect. A robustness suite is reported in the Supplementary Information: it substitutes each of the six nitrogen imputations, varies the fixation and deposition specifications, and removes the elevated-negative crops.

3 Results

We decompose the change in the aggregate US agricultural nitrogen balance between 1990 and 2022 into five multiplicative drivers using an additive LMDI-I index decomposition. The drivers are total harvested area (H), the distribution of that area across counties (S_i), the mix of crops grown within each county (S_{ij}), crop yields (Y_{ij}), and the nitrogen balance intensity (NBI_{ij}). NBI is the field N surplus or deficit generated per unit of crop production. The first three terms describe the *structure* of US agriculture (how much land is farmed, where, and in what crops); the last two describe its *intensity* (how much is produced per hectare and how much surplus or deficit that production leaves behind). All results use the canonical period-averaged single-step estimator, comparing the 1990–1994 base window with the 2018–2022 terminal window.

The national balance was a large and persistent N surplus throughout the study period (Figure 1). It stood at 5.6 Tg N in 1990 and 7.0 Tg N in 2022, fluctuating year to year with weather and crop markets. It fell to a low of 5.3 Tg N in 2009 and spiked to 7.5 Tg N in the 2012 drought year, but showed no pronounced secular trend in either direction. Averaged over the base window the balance was 6.1 Tg N, rising modestly to 6.9 Tg N over the terminal window, a net increase of +0.83 Tg N. This near-stationarity at the national level masks substantial and partly offsetting changes underneath, both across drivers and across space. County-level changes were almost evenly divided between increases (1394 counties) and decreases (1389), and were spatially concentrated: the most active 5% of counties account for 31.3% of the total absolute change (Figure 4, panel A).

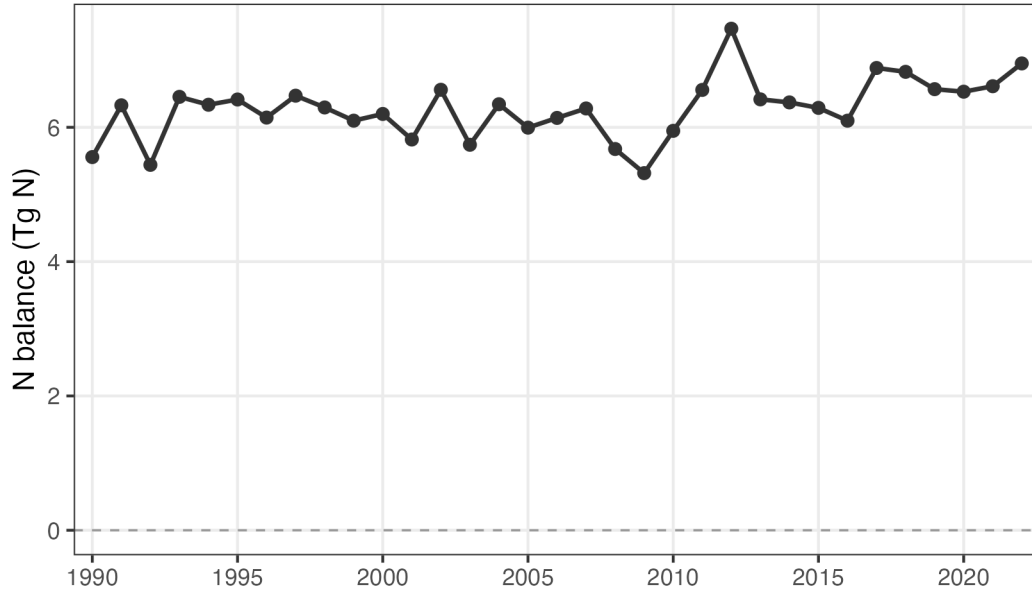


Figure 1: Aggregate US agricultural nitrogen balance, 1990–2022. Positive values indicate a national N surplus (inputs exceed outputs); negative values indicate a deficit.

3.1 Decomposition of the change

The decomposition attributes the +0.83 Tg N net change to the five drivers (Figure 2). Two effects dominate and nearly cancel. Rising yields (Y_{ij}) added +1.99 Tg N — by far the largest single effect — as greater crop production per hectare, with the surplus left per unit of output (NBI) held fixed, scaled up the absolute surplus. Pulling in the opposite direction, falling nitrogen balance intensity reduced the balance by -1.40 Tg N: the surplus left behind per unit of output declined, signalling improved technical nitrogen efficiency. The crop-mix effect (S_{ij}) added a further +0.50 Tg N, while the two area terms were small and negative: total harvested area (H) -0.11 Tg N and its geographic redistribution (S_i) -0.15 Tg N. The modest net increase is therefore what remains after rising output and improving efficiency nearly cancel. We take the four structural and yield effects in turn before turning to nitrogen balance intensity, the metric this decomposition is built to isolate.

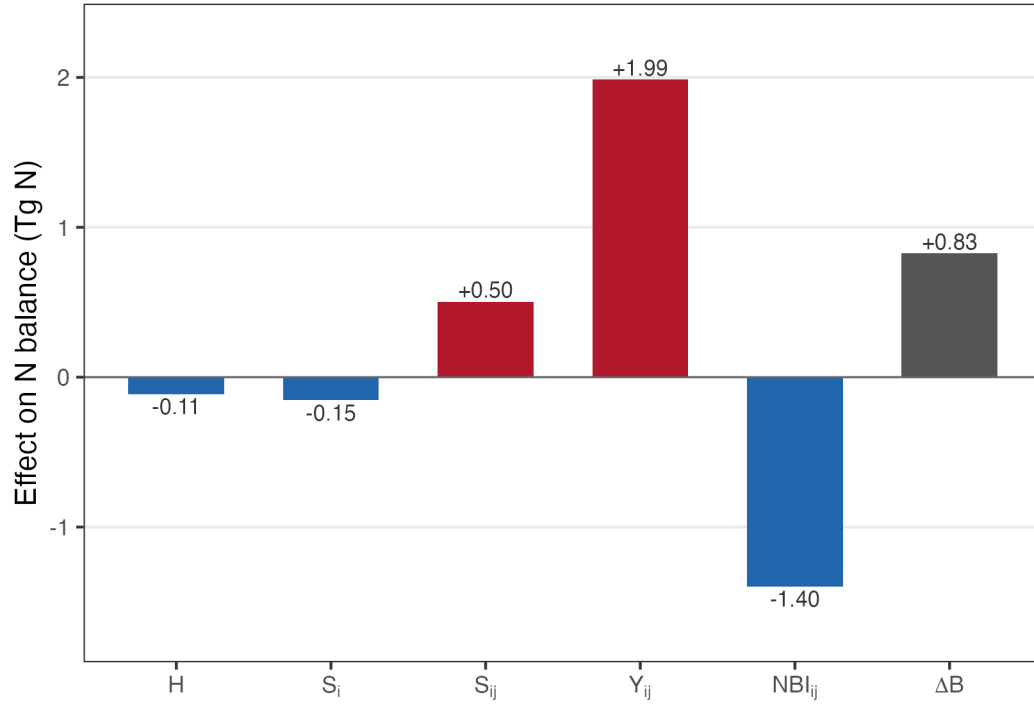


Figure 2: Additive LMDI decomposition of the change in aggregate US agricultural nitrogen balance (period-averaged single-step, base 1990–1994 vs. terminal 2018–2022). Each bar is one factor effect; red bars increase the N surplus, blue bars reduce it. The rightmost bar (ΔB) is the algebraic sum of the five factor effects and equals the observed difference between the terminal and base period-average balances. The factor effects are robust to nitrogen-input uncertainty: across the six multiply-imputed realizations of the applied-nitrogen inventory each varies by roughly 1% or less around its central value (Table S9). Values in Tg N.

3.2 Harvested area and its distribution

Total harvested area fell 3.3% between the base and terminal windows, from 94.2 to 91.1 million ha, contributing a small downward -0.11 Tg N to the balance. The redistribution of that area across counties (S_i) contributed a further -0.15 Tg N (Figure 4, panel B). This term is positive where a county’s share of national area grew and negative where it shrank, with each county’s magnitude scaled by the size of the balance it carries. A given proportional shift in area share therefore moves the national total most in counties already handling a large absolute surplus (or deficit). The losses form a coherent band along the southern Plains and West: Texas (-0.16 Tg

N), California (-0.12 Tg N), and Oklahoma (-0.09 Tg N), extending into the Mississippi Delta and the Pacific Northwest. The gains are concentrated across the Corn Belt and northern Plains, where harvested area expanded. On net the losses in these higher-surplus southern and western regions slightly outweighed the gains, but neither area term is a major driver of the national change.

3.3 Crop mix

Changes in the mix of crops grown within counties raised the balance by +0.50 Tg N (Figure 3). This effect is dominated by the expansion of soybeans (+0.54 Tg N) and, to a lesser extent, corn (+0.28 Tg N), partly offset by contractions in sorghum, winter wheat, and oats. The soybean signal warrants emphasis because it is easily misread. Although soybeans fix much of their own nitrogen, the biological fixation we credit as an input makes a hectare of soybeans a net N surplus on our accounting, a convention that follows from the question the balance asks. The surplus is a consequence of that convention, not a finding that legumes carry a high surplus relative to other crops. Because the decomposition attributes each crop's mix effect to its own balance and its own change in area share, soybeans' expansion registers as a positive contribution wherever they gain share. That contribution is independent of the crop displaced, which carries its own, separate term. The net effect is positive here because the crops that ceded share are lower-surplus small grains (the partial offsets noted above), so the reallocation toward soybeans raises the aggregate balance rather than lowering it. Spatially, the crop-mix effect is largest in the northern Plains — North Dakota alone contributes +0.13 Tg N (Figure 4, panel C) — where soybean and corn cultivation expanded onto land formerly in wheat and small grains. This northern-Plains signal carries one caveat: it arises partly from wheat and small grains *leaving* the local crop mix, and the field balances of those crops are the least well-determined in the panel. Their frequent small deficits are partly an artefact of the fertilizer inventory's fixed-point coverage rather than a measured drawdown (Methods; SI negative-balance audit).

Most of the small grains’ negative crop-mix contribution in fact comes from *positive*-balance acreage losing share — correctly registering the contraction of a surplus crop — rather than from the disputed negative cells, which are few and nudge the effect slightly the other way. Correcting those negatives toward zero or positive values can therefore only trim the crop-mix effect, never inflate it: a bounding sensitivity that removes them leaves the sign of the effect and its soybean-driven core intact. The soybean contribution is unchanged, deriving from soybeans’ own surplus rather than the crops they displace. The sensitivity does trim the national crop-mix effect, by about 20%, and the Dakotas’ contributions by a similar margin (about 17%; Discussion). The *direction* of the crop-mix effect on the northern Plains is therefore robust, and the headline its upper bound; only the precise magnitude is uncertain.

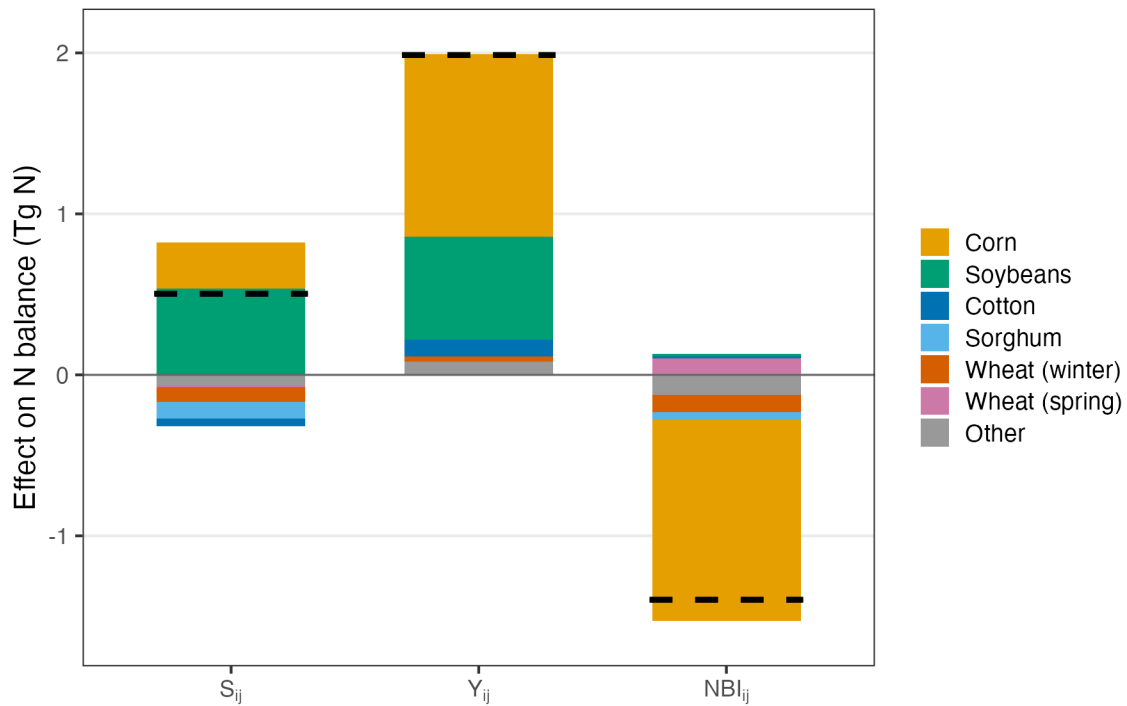


Figure 3: Crop-level attribution of the three crop-indexed factor effects (period-averaged single-step decomposition). Each bar shows the contribution of individual crops (and an aggregate ‘Other’ category) to the crop-mix (S_{ij}), yield (Y_{ij}), and nitrogen balance intensity (NBI_{ij}) effects. Positive segments increase the N surplus; negative segments reduce it. The dashed horizontal line marks the net factor total. Units: Tg N.

3.4 Yields

Rising yields are the single largest driver, adding +1.99 Tg N (Figure 3). The effect is overwhelmingly a corn-and-soybean phenomenon: corn (+1.14 Tg N) and soybeans (+0.64 Tg N) together account for 89.3% of it. Because yield enters the balance multiplicatively, more output per hectare at a given nitrogen balance intensity scales up the absolute surplus; decades of sustained yield growth in the major commodity crops therefore translate into a large upward effect even with efficiency improving. Geographically the yield effect is the most broadly distributed of the five: it is positive in 79.8% of cropping counties, with almost no counties showing a yield-driven decline anywhere east of the Rockies (Figure 4, panel D). It is a near-universal intensification rather than a regional one, largest in the Corn Belt core: Iowa (+0.28 Tg N), Illinois (+0.23 Tg N), and Minnesota.

3.5 Nitrogen balance intensity

The nitrogen balance intensity effect — the change in field N surplus per unit of crop production, holding area, crop mix, and yield fixed — is the central result of the decomposition and the novel metric it isolates. Falling NBI reduced the national balance by -1.40 Tg N, the only large effect pushing the surplus down and the principal counterweight to rising yields (Figure 3). Corn alone accounts for 89.4% of the total NBI effect (-1.25 Tg N), reflecting decades of improving nitrogen efficiency in the crop that receives the most fertilizer and has the best data. The two wheat classes nearly cancel (winter -0.11, spring +0.10 Tg N), confirming that the aggregate NBI improvement does not rest on the noisier wheat cells (SI). Spatially the effect forms a sharp core–periphery contrast (Figure 4, panel E). A dense, near-contiguous zone of improvement covers the Corn Belt core, where 90.4% of counties in Iowa, Illinois, Indiana, Minnesota, and Nebraska improve, strongest in Illinois (-0.31 Tg N). That zone is ringed by a patchier periphery on the northern and western Plains where NBI instead *worsens* on expanding, lower-efficiency

acreage. The same Corn Belt counties that are deepest red in panel D (largest yield gains) are deepest blue in panel E (largest efficiency gains), which makes the yield-versus-efficiency tension visible at the county scale.

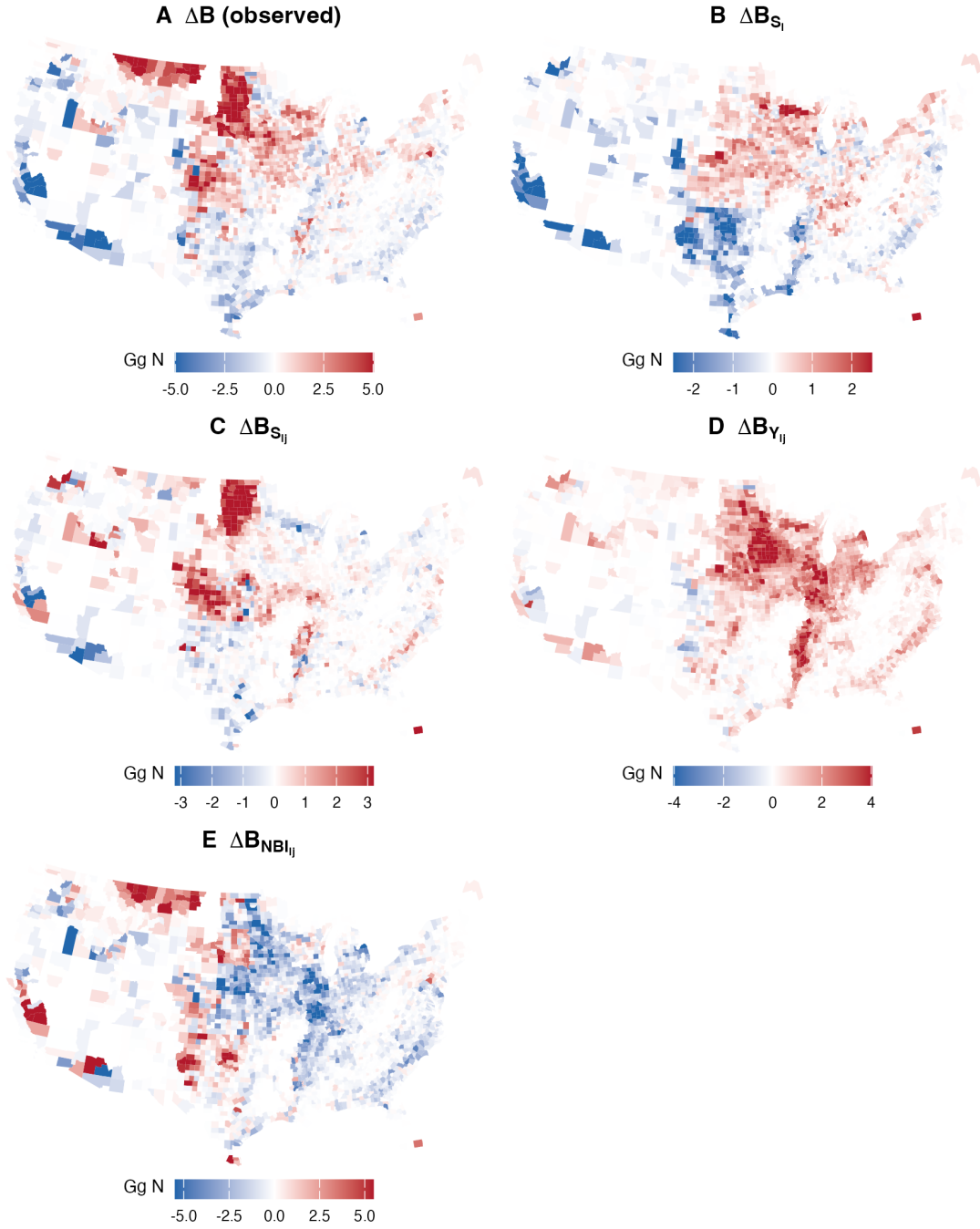


Figure 4: County-level contributions to the change in aggregate N balance (period-averaged single-step decomposition, base 1990–1994 vs. terminal 2018–2022). **(A)** Observed ΔB . **(B–E)** Factor contributions: county share of area (S_i), crop mix (S_{ij}), yield (Y_{ij}), and nitrogen balance intensity (NBI_{ij}). Blue = reduces the N surplus; red = increases the N surplus. Units: Gg N per county. Scale bars are symmetric about zero and capped at the 98th percentile of absolute values.

3.6 Spatial synthesis

The two dominant effects — rising yields, which push the surplus up, and improving nitrogen balance intensity, which pulls it down — are strongest in the same counties, so they offset locally instead of each shaping a separate region (Figure 4). In the core Corn Belt, the largest yield-driven increases (+1.23 Tg N across Iowa, Illinois, Indiana, Minnesota, Nebraska, Missouri, Ohio, and Wisconsin) are substantially offset *within the same counties* by improving nitrogen balance intensity (-1.12 Tg N). The net change there (+0.49 Tg N) is far smaller than either gross effect, the national yield-versus-efficiency tension reproduced at regional scale. The largest net increases occur instead on the northern Plains, in South Dakota (+0.24 Tg N) and North Dakota (+0.17 Tg N), driven by crop-mix expansion onto former wheat land (the magnitude of this crop-mix-driven increase carries the small-grain-balance uncertainty noted above). The two states differ in what happened to efficiency. North Dakota's increase is almost purely a crop-mix effect (+0.13 Tg N) with roughly neutral NBI, whereas in South Dakota the expansion is compounded by worsening NBI (+0.10 Tg N) on the newly cultivated acreage. Worsening efficiency is specific to South Dakota here and does not extend across the northern Plains. Net decreases concentrate in the South and West — most visibly Texas and California — where the dominant signal is a loss of harvested-area share. The aggregate near-stationarity of the national balance is thus the sum of three distinct regional regimes: an intensifying but efficiency-moderated Corn Belt, a northern Plains shifting from wheat to corn and soybeans, and a contracting South and West.

3.7 Within-field efficiency versus conventional NUE

The NBI effect differs from a conventional nitrogen-use-efficiency (NUE) reading in two ways. It is in the units of the balance, so it can be weighed against the production terms in kilograms

of nitrogen, which is what the preceding sections have been doing. And it is free of the compositional shifts that move an aggregate NUE, which we can now quantify directly. Aggregate NUE, or harvested N removed per unit of N applied, is exactly an input-weighted mean of cell-level efficiencies. The same additive LMDI therefore splits its change into a within-cell *intensity* effect (genuine field-level efficiency change) and a *structural* effect carrying only the shifting mix of crops and counties (SI). Between the base and terminal windows, aggregate US NUE rose 13.1% (from 56.7% to 64.2%). Only about two-thirds of that improvement is genuine within-field change (intensity +0.050); the remaining 31% (structural +0.023) is pure composition, the system shifting weight toward inherently more efficient crops and regions with no change in management. This structural gain and the crop-mix effect (Figure 3) are the same soybean-led shift, and the two are consistent. Soybeans carry a small positive balance per unit of output. Because that balance is positive, their expansion pushes the surplus up. Because it is small relative to their large nitrogen input, their efficiency ratio sits well above the average, so the same expansion lifts aggregate NUE. A crop can add surplus and still register as more efficient than average. A NUE-based assessment would thus overstate the on-farm efficiency gain by nearly half (45%). The bias is not an artefact of our broad input definition. Under the conventional farm-gate definition (synthetic plus manure N only), composition does not inflate the trend but creates it. Apparent NUE rises 35.9% even as genuine within-field efficiency *falls* (-0.116), the aggregate gain driven entirely by expansion of low-input legumes (SI). NBI, defined on each crop-county's own balance per unit of output, is unaffected by these compositional shifts. Its improvement of -1.40 Tg N is what within-field efficiency actually contributed to the national balance, free of composition and in kilograms of nitrogen.

4 Discussion

Our decomposition resolves the near-flat national trend in the US nitrogen surplus into two kinds of drivers. Structural factors (the scale of production, its geographic distribution, and the crop mix) together moved the surplus little. The near-stationarity instead reflects two intensive forces that nearly cancel. Rising yields pushed the surplus up, and improving nitrogen balance intensity — the surplus left per unit of crop produced, a measure of technical efficiency — pulled it down by nearly as much. Our contribution is the accounting. To our knowledge this is the first exact, additive attribution of the change to its structural and intensive parts. It is also the first to isolate within-field efficiency from the crop-mix and geographic shifts that a national efficiency figure conflates. The efficiency literature has long recognised this confound and measured it in cross-section. What it has not done is divide a change in national efficiency over time into the part from field practice and the part from composition (X. Zhang et al. 2015).

The geographic distribution of production, in particular, contributed just -0.15 Tg N. This is a genuine null result, and not an obvious one. The nitrogen surplus is highly uneven across the country, so a large reshuffling of production toward or away from high-surplus counties could in principle have moved the national total substantially. In fact the shifts in where crops are grown roughly cancelled, and the surplus-weighted redistribution nets to almost nothing.

The improvement in nitrogen balance intensity comes mostly from corn, densest in the Corn Belt core (Figure 4). That the signal is corn-led matters for confidence in it: corn is the best-measured crop in the panel, so the effect rests on the firmest part of the record.

This efficiency gain should be read with its counterpart in view. Better management removed about 70% of the surplus that rising yields added, most of it but not all. Prior studies of the US trajectory credit the restraint to yields alone: rising yields recover more nitrogen from a given input, and so hold the surplus down (X. Zhang et al. 2015; J. Zhang, Cao, and Lu 2021). Our decomposition places that restraint with NBI, while the yield term becomes a separate scale

effect pushing the surplus up. The disagreement is one of accounting, and the two accounts describe the same underlying change. The net trends do differ, though. Those studies report the US surplus flat-to-declining since the 1990s, where we find a slight rise. The difference likely reflects how the inputs are built: top-down fertilizer-consumption totals in their case, a crop- and county-level survey inventory in ours.

Separating management from composition makes a large correction to the estimated efficiency gain. A substantial share of the apparent US efficiency gain is composition rather than management, and under a farm-gate definition composition accounts for essentially all of it. The mechanism is one this literature already names: the spread of soybeans, whose expansion lifts aggregate efficiency with no change in field practice (X. Zhang et al. 2015). Nitrogen balance intensity is built from the surplus this literature already works with (Lassaletta et al. 2014; Bouwman, Van Dreht, and Van der Hoek 2005). We resolve that surplus to the crop and county and divide it by output rather than by area, which is what holds composition fixed. That resolution lets us partition what the closest precedents could only describe: county and state budgets show the modest national rise in crop efficiency concealing opposing movements across crops, corn improving as several small grains decline (J. Zhang, Cao, and Lu 2021; Lu et al. 2019). The decomposition turns that description into an exact accounting: corn-led, Corn-Belt-centred, set against a Plains periphery where intensity worsens on expanding acreage. NBI does not replace NUE. For a single crop in a single county, the two are the same measurement, and NUE is the more familiar way to report it. They differ only when aggregated. A rise in a regional or national NUE can come from better management in the field or from a shift toward crops and places that score higher on the ratio, and the ratio cannot say which.

Rising yields keep raising the amount of abatement that efficiency must deliver just to hold the surplus level. Yields added +1.99 Tg N to the surplus over the period. Improving nitrogen balance intensity took -1.40 Tg N back, about 70% of the yield-driven increase, which is why the surplus rose slightly instead of holding level. Because both terms are in kilograms of nitrogen,

the gap can be read off directly: within-field efficiency would have had to deliver roughly 59% more abatement than it did merely to hold the national surplus flat, and more still to reduce it. Every further gain in yield raises that bar. This is a steep target for the voluntary, incentive-based US approach, which has curbed nutrient loading only modestly: the Gulf of Mexico dead zone reached its largest recorded extent in 2017, decades into the conservation-programme era (Del Rossi et al. 2023; Ribaud et al. 2011). Efficiency improved substantially, but an improvement on this scale delivers only near-stability in the national surplus. Whether that bar can be cleared is not something our accounting settles. Evidence from national borders suggests the yield cost of lower pollution is small. Wuepper et al. (2020) find that countries causing 35% less nitrogen pollution than their neighbours show yield gaps only about 1% larger, and that national factors move nitrogen pollution far more than they move yields. That comparison runs between countries instead of along one country's trend, so it does not measure the US margin. It does suggest the binding constraint is institutional rather than agronomic.

The spatial results show where the surplus is growing and why. Of the three regional regimes, the surplus rose fastest on the northern Plains, while the Corn Belt's increases were largely offset by improving efficiency. But the northern Plains is not a single case. North Dakota and South Dakota both saw large increases — +174 and +243 Gg N — and in both the expansion came onto former wheat land. North Dakota's rise is almost entirely crop mix (+132 Gg N), its fields essentially unchanged (+7 Gg N). In South Dakota, worsening within-field intensity is the single largest term (+96 Gg N), ahead of crop mix itself (+73 Gg N). The same increase therefore has different causes in the two states. North Dakota's surplus follows from what is grown there. South Dakota's follows from how its fields are managed.

An efficiency ratio cannot tell these two cases apart. A state-level NUE would move in both places, but it would move for both reasons at once: a crop switch and a change in field practice are indistinguishable in a ratio of sums. And because the ratio is unitless, it could not rank the two states by the nitrogen at stake even if it could separate the causes. Conservation spending

is cost-effective only when it reaches the acres most in need, which voluntary enrolment does not guarantee (Ribaud et al. 2011; Del Rossi et al. 2023). Finding those acres means knowing both what is wrong and how many kilograms are involved.

Several limitations bound the interpretation. The most consequential concerns the crop-mix effect. Because every factor effect is weighted by the cell balance, each depends on the modelled nitrogen inputs — robustly so for the well-measured major crops that carry the yield and intensity effects, but not for the crop-mix effect, whose negative contributions come from the small grains, the crops whose balances are least certain in the panel. The bound here is one-sided, however. Most of the small grains’ negative contribution comes from positive-balance acreage losing share rather than from the disputed negative cells, so correcting those negatives can only trim the crop-mix effect, never inflate it. A sensitivity that sets every negative wheat, barley and oats balance to zero leaves the yield and intensity effects essentially unchanged and shrinks the national crop-mix effect by about 20%. The northern-Plains crop-mix effect is therefore robust in direction, and the headline figure is an upper bound on its size.

The balance is an accounting construct, not a measured loss. A positive balance is nitrogen at risk of loss to air and water plus any net accumulation in the soil organic pool, which we do not model separately; reading the balance as environmental loss therefore assumes that accumulation is small relative to the surplus (Oenema, Kros, and De Vries 2003; Van Beek, Brouwer, and Oenema 2003). A negative balance means nitrogen is drawn from the soil faster than inputs replace it. These deficits are real (Bouwman, Van Dreht, and Van der Hoek 2005), but in our panel they cluster in the specialty-crop tail, where coarse application rates for the non-major crops exaggerate them. The budget also omits some minor inputs, including net soil-organic-matter mineralization, residue carryover, irrigation-water nitrogen, and seed nitrogen. The accounting is also annual and crop-specific, so it does not trace nitrogen across a rotation or between years. Where a legume or a fertilized preceding crop leaves nitrogen for the next — winter wheat after soybean, say — the balance books the fixed or residual nitrogen to the

first crop and registers the drawdown as a deficit in the second. This sharpens the crop-level attribution already flagged for the small grains, but leaves the national total unaffected (it sums over both crops) and the trend robust: the headline compares five-year averages and explains the change, so a stable carryover offset cancels.

Finally, the estimator compares two period averages in a single step. This is deliberate: a chained year-on-year decomposition is unstable for a signed balance that is near zero in many cells. The trade-off is that a single-step comparison is blind to the path between the endpoints, which is why we also report the chained decomposition as a trajectory sensitivity.

Two extensions follow. The first is evaluative. Because nitrogen balance intensity isolates within-field efficiency from composition, it gives a clean signal for asking whether an observed improvement, whatever its source, actually reached the field. An aggregate surplus or an aggregate NUE cannot answer this, since both move with crop mix and geography.

The second is comparative. The same decomposition could be applied to other national systems, but it depends on a nitrogen budget resolved to crop and region. Area and production at that resolution are published for many countries; crop- and location-specific nitrogen inputs are not, and we found no international equivalent to the budget assembled here. Applying it elsewhere therefore depends on nitrogen budgets that most countries do not yet publish.

References

- Ang, B. W. 2005. "The LMDI Approach to Decomposition Analysis: A Practical Guide." *Energy Policy* 33 (7): 867–71. <https://doi.org/10.1016/j.enpol.2003.10.010>.
- Ang, B. W., and F. L. Liu. 2001. "A New Energy Decomposition Method: Perfect in Decomposition and Consistent in Aggregation." *Energy* 26 (6): 537–48. [https://doi.org/10.1016/S0360-5442\(01\)00022-6](https://doi.org/10.1016/S0360-5442(01)00022-6).
- Ang, B. W., and Na Liu. 2007a. "Handling Zero Values in the Logarithmic Mean Divisia Index Decomposition Approach." *Energy Policy* 35 (1): 238–46. <https://doi.org/10.1016/j.enpol.2005.11.001>.

- . 2007b. “Negative-Value Problems of the Logarithmic Mean Divisia Index Decomposition Approach.” *Energy Policy* 35 (1): 739–42. <https://doi.org/10.1016/j.enpol.2005.12.004>.
- Baston, Daniel. 2025. “Exactextractr: Fast Extraction from Raster Datasets Using Polygons.” <https://doi.org/10.32614/CRAN.package.exactextractr>.
- Blomqvist, Linus, Luke Yates, and Barry W Brook. 2020. “Drivers of Increasing Global Crop Production: A Decomposition Analysis.” *Environ. Res. Lett* 15: 940–46. <https://doi.org/10.1088/1748-9326/ab9e9c>.
- Bouwman, AF, G. Van Drecht, and KW Van der Hoek. 2005. “Global and Regional Surface Nitrogen Balances in Intensive Agricultural Production Systems for the Period 1970-2030.” *Pedosphere* 15 (2): 137–55.
- De Vries, Wim. 2021. “Impacts of Nitrogen Emissions on Ecosystems and Human Health: A Mini Review.” *Current Opinion in Environmental Science & Health* 21 (June): 100249. <https://doi.org/10.1016/j.coesh.2021.100249>.
- Del Rossi, Gemma, Mohammad Mainul Hoque, Yongjie Ji, and Catherine L. Kling. 2023. “The Economics of Nutrient Pollution from Agriculture.” *Annual Review of Resource Economics* 15 (1): 105–30. <https://doi.org/10.1146/annurev-resource-111820-021317>.
- Fixen, Paul, Frank Brentrup, Tom W. Bruulsema, Fernando Garcia, Rob Norton, and Shamie Zingore. 2015. “Nutrient/Fertilizer Use Efficiency: Measurement, Current Situation and Trends.” In *Managing Water and Fertilizer for Sustainable Agricultural Intensification*, 8–38. International Fertilizer Industry Association (IFA), International Water Management Institute (IWMI), International Plant Nutrition Institute (IPNI), and International Potash Institute (IPI).
- Galloway, James N, Alan R Townsend, Jan Willem Erisman, Mateete Bekunda, Zucong Cai, John R Freney, Luiz a Martinelli, Sybil P Seitzinger, and Mark a Sutton. 2008. “Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions.” *Science (New York, N.Y.)* 320 (5878): 889–92. <https://doi.org/10.1126/science.1136674>.
- Hember, Robbie A. 2018. “Spatially and Temporally Continuous Estimates of Annual Total Nitrogen Deposition over North America, 1860–2013.” *Data in Brief* 17 (April): 134–40. <https://doi.org/10.1016/j.dib.2017.12.052>.
- Herridge, David F., Mark B. Peoples, and Robert M. Boddey. 2008. “Global Inputs of Biological Nitrogen Fixation in Agricultural Systems.” *Plant and Soil* 311 (1-2): 1–18. <https://doi.org/10.1007/s11104-008-9668-3>.
- Hoskovec, L., T. Liu, S. M. Ogle, Y. A. Oshan, S. Spencer, A. Keyser, S. Williams, and F. J. Breidt. 2026. “Agricultural Land Management Practices in the Conterminous United States from 1980-2023.” Manuscript under Review.
- Kastner, Thomas, Maria Jose Ibarrola Rivas, Wolfgang Koch, and Sanderine Nonhebel. 2012. “Global Changes in Diets and the Consequences for Land Requirements for Food.” *Proceedings of the National Academy of Sciences* 109 (18): 6868–72. <https://doi.org/10.1073/pnas.1117054109>.
- Lassaletta, Luis, Gilles Billen, Bruna Grizzetti, Juliette Anglade, and Josette Garnier. 2014. “50 Year Trends in Nitrogen Use Efficiency of World Cropping Systems: The Relationship Between Yield and Nitrogen Input To Cropland.” *Environmental Research Letters* 9 (10): 105011. <https://doi.org/10.1088/1748-9326/9/10/105011>.
- Liu, Ji, Hai Wang, Josep Penuelas, Juan Mou, Manuel Delgado-Baquerizo, Jordi Sardans, Fer-

- nando Coello, et al. 2025. “Global-Scale Prevalence of Low Nutrient Use Efficiency Across Major Crops.” *Nature Communications* 16 (1): 11036. <https://doi.org/10.1038/s41467-025-66019-w>.
- Lu, Chaoqun, Jien Zhang, Peiyu Cao, and Jerry L. Hatfield. 2019. “Are We Getting Better in Using Nitrogen?: Variations in Nitrogen Use Efficiency of Two Cereal Crops Across the United States.” *Earth’s Future* 7 (8): 939–52. <https://doi.org/10.1029/2019EF001155>.
- Murrell, T. Scott. 2008. “Measuring Nutrient Removal, Calculating Nutrient Budgets.” In *ASA, CSSA, and SSSA Books*, edited by Sally Logsdon, Dave Clay, Demie Moore, and Teferi Tsegaye, 1st ed., 159–82. Wiley. <https://doi.org/10.2136/2008.soilsciencestepbystep.c13>.
- National Atmospheric Deposition Program. 2025. “Total Deposition Maps, Version 2025.01.” Wisconsin State Laboratory of Hygiene.
- Oenema, Oene, Hans Kros, and Wim De Vries. 2003. “Approaches and Uncertainties in Nutrient Budgets: Implications for Nutrient Management and Environmental Policies.” *European Journal of Agronomy* 20 (1-2): 3–16. [https://doi.org/10.1016/S1161-0301\(03\)00067-4](https://doi.org/10.1016/S1161-0301(03)00067-4).
- Potter, Nicholas. 2019. “Rnassqs: An R Package to Access Agricultural Data via the USDA National Agricultural Statistics Service (USDA-NASS) ‘Quick Stats’ API.” *Journal of Open Source Software* 4 (43): 1880. <https://doi.org/10.21105/joss.01880>.
- Rabalais, Nancy N, R Eugene Turner, and William J Wiseman. 2002. “Gulf of Mexico Hypoxia, a.k.a. ”The Dead Zone”.” *Annu. Rev. Ecol. Syst.* 33: 235–63.
- Ren, Chenchen, Xiuming Zhang, Stefan Reis, and Baojing Gu. 2022. “Socioeconomic Barriers of Nitrogen Management for Agricultural and Environmental Sustainability.” *Agriculture, Ecosystems & Environment* 333 (August): 107950. <https://doi.org/10.1016/j.agee.2022.107950>.
- Ribaudo, Marc, LeRoy Hansen, Mike J. Livingston, Roberto Mosheim, James Williamson, and Jorge Delgado. 2011. “Nitrogen in Agricultural Systems: Implications for Conservation Policy.” ERR-127. USDA. <https://doi.org/10.2139/ssrn.2115532>.
- Roberts, Trent, Wes Kirkpatrick, Nathan Slaton, and Rick Norman. n.d. “Estimating Nutrient Removal for Row Crops Grown in Arkansas.” Little Rock, AR: University of Arkansas Division of Agriculture.
- Salvagiotti, F., K. G. Cassman, J. E. Specht, D. T. Walters, A. Weiss, and A. Dobermann. 2008. “Nitrogen Uptake, Fixation and Response to Fertilizer N in Soybeans: A Review.” *Field Crops Research* 108 (1): 1–13. <https://doi.org/10.1016/j.fcr.2008.03.001>.
- Schwede, Donna B., and Gary G. Lear. 2014. “A Novel Hybrid Approach for Estimating Total Deposition in the United States.” *Atmospheric Environment* 92 (August): 207–20. <https://doi.org/10.1016/j.atmosenv.2014.04.008>.
- Semprini, Jason. 2025. “Early Prenatal Nitrate Exposure and Birth Outcomes: A Study of Iowa’s Public Drinking Water (1970-1988).” *PLOS Water* 4 (6): e0000329. <https://doi.org/10.1371/journal.pwat.0000329>.
- Shober, Amy L. 2024. “Nitrogen Removal by Delaware Crops.” Factsheet. Newark, DE: University of Delaware Cooperative Extension. <https://www.udel.edu/academics/colleges/canr/cooperative-extension/fact-sheets/nitrogen-removal-delaware-crops/>.
- Stayner, Leslie Thomas, Kirsten Almberg, Rachael Jones, Judith Graber, Marie Pedersen, and Mary Turyk. 2017. “Atrazine and Nitrate in Drinking Water and the Risk of Preterm Delivery and Low Birth Weight in Four Midwestern States.” *Environmental Research* 152

- (January): 294–303. <https://doi.org/10.1016/j.envres.2016.10.022>.
- Van Beek, C. L., L. Brouwer, and O. Oenema. 2003. “The Use of Farmgate Balances and Soil Surface Balances as Estimator for Nitrogen Leaching to Surface Water.” *Nutrient Cycling in Agroecosystems* 67 (3): 233–44. <https://doi.org/10.1023/B:FRES.00000003619.50198.55>.
- Van Grinsven, Hans J. M., Mike Holland, Brian H. Jacobsen, Zbigniew Klimont, Mark A. Sutton, and W. Jaap Willems. 2013. “Costs and Benefits of Nitrogen for Europe and Implications for Mitigation.” *Environmental Science & Technology* 47 (8): 3571–79. <https://doi.org/10.1021/es303804g>.
- Ward, Mary, Rena Jones, Jean Brender, Theo De Kok, Peter Weyer, Bernard Nolan, Cristina Villanueva, and Simone Van Breda. 2018. “Drinking Water Nitrate and Human Health: An Updated Review.” *International Journal of Environmental Research and Public Health* 15 (7): 1557. <https://doi.org/10.3390/ijerph15071557>.
- Wuepper, David, Solen Le Clech, David Zilberman, Nathaniel Mueller, and Robert Finger. 2020. “Countries Influence the Trade-Off Between Crop Yields and Nitrogen Pollution.” *Nature Food* 1: 713–19.
- Yan, Xiaoyuan, Longlong Xia, and Chaopu Ti. 2022. “Temporal and Spatial Variations in Nitrogen Use Efficiency of Crop Production in China.” *Environmental Pollution* 293 (January): 118496. <https://doi.org/10.1016/j.envpol.2021.118496>.
- Zhang, Jien, Peiyu Cao, and Chaoqun Lu. 2021. “Half-Century History of Crop Nitrogen Budget in the Conterminous United States: Variations Over Time, Space and Crop Types.” *Global Biogeochemical Cycles* 35 (10): e2020GB006876. <https://doi.org/10.1029/2020GB006876>.
- Zhang, Xin, Eric A. Davidson, Denise L. Mauzerall, Timothy D. Searchinger, Patrice Dumas, and Ye Shen. 2015. “Managing Nitrogen for Sustainable Development.” *Nature* 528 (7580): 51–59. <https://doi.org/10.1038/nature15743>.

5 Supplementary information

5.1 Decomposition method

This section gives the full specification of the index decomposition summarised in the main-text Methods. The method is additive **LMDI-I** (the logarithmic mean Divisia index, “Model 1”) with period-averaged endpoints, following Ang and Liu (2001) and Ang (2005) and the agricultural application of Blomqvist, Yates, and Brook (2020); a chained (rolling annual) decomposition is also computed and reported as a trajectory sensitivity. Zero- and negative-valued sub-categories are handled by the analytical-limit convention of Ang and Liu (2007a) and Ang and Liu (2007b).

5.1.1 The decomposition identity

For a given year, total US agricultural field nitrogen balance B (kg N yr⁻¹) is the sum over sub-categories indexed by county i and crop j ,

$$B = \sum_i \sum_j B_{ij},$$

where B_{ij} is the cell-level balance assembled in the data-processing section below ($B_{ij} = N_{\text{in},ij} - N_{\text{out},ij}$). Every quantity is year-specific; the year subscript t is suppressed throughout to reduce clutter and reinstated only where the exposition refers to specific years (e.g. the chained year-on-year decomposition and parts of the data construction). Each sub-category balance factors into five multiplicative components,

$$B_{ij} = H \cdot S_i \cdot S_{ij} \cdot Y_{ij} \cdot \text{NBI}_{ij},$$

which telescopes back to the identity on substitution of the definitions,

$$H \cdot \frac{H_i}{H} \cdot \frac{H_{ij}}{H_i} \cdot \frac{P_{ij}}{H_{ij}} \cdot \frac{B_{ij}}{P_{ij}} = B_{ij}.$$

The five factors are:

- $H = \sum_{ij} H_{ij}$ — total US harvested area across all in-scope crops and counties (ha);
- $S_i = H_i/H$ — county i 's share of national harvested area, where $H_i = \sum_j H_{ij}$ (dimensionless);
- $S_{ij} = H_{ij}/H_i$ — crop j 's share of county i 's harvested area (dimensionless);
- $Y_{ij} = P_{ij}/H_{ij}$ — yield, i.e. production per unit area (kg ha^{-1});
- $\text{NBI}_{ij} = B_{ij}/P_{ij}$ — the **nitrogen balance intensity**, balance per unit of crop output (kg N per kg output).

Here H_{ij} and P_{ij} are the harvested area and production of crop j in county i . Note that H does not vary with i or j , and S_i does not vary with j : they are constant within a year and within a county-year respectively.

5.1.2 Additive LMDI-I decomposition

The change in aggregate balance between a base year (superscript 0) and a terminal year (superscript T), $\Delta B = B^T - B^0$, decomposes additively into five factor effects that sum to it exactly:

$$\Delta B = \Delta B_H + \Delta B_{S_i} + \Delta B_{S_{ij}} + \Delta B_{Y_{ij}} + \Delta B_{\text{NBI}_{ij}}.$$

Each effect is a logarithmic-mean-weighted sum of the log-changes of that factor over the sub-categories,

$$\Delta B_k = \sum_{ij} L(B_{ij}^T, B_{ij}^0) \ln \left(\frac{x_{k,ij}^T}{x_{k,ij}^0} \right),$$

where $x_{k,ij}$ is factor k in sub-category (i, j) and $L(a, b)$ is the logarithmic mean,

$$L(a, b) = \frac{a - b}{\ln(a/b)} \quad (a \neq b), \quad L(a, a) = a.$$

The weight is written as a ratio of logarithms rather than the equivalent difference $\ln a - \ln b$ so that it stays well defined for the negative-balance cells discussed below: whenever B_{ij}^T and B_{ij}^0 share the same sign their ratio is positive and $\ln(B_{ij}^T/B_{ij}^0)$ is real. This is equivalent to the additive-inverse convention of Ang and Liu (2007b).

The decomposition is *perfect* — it leaves no residual term — because the factor weights are identical across the five effects and the factor log-changes sum to the log-change of the sub-category total, so that

$$\sum_k \Delta B_k = \sum_{ij} L(B_{ij}^T, B_{ij}^0) \ln \left(\frac{B_{ij}^T}{B_{ij}^0} \right) = \sum_{ij} (B_{ij}^T - B_{ij}^0) = B^T - B^0,$$

using the identity $L(a, b) \ln(a/b) = a - b$ (Ang and Liu 2001). The displayed sum is written for cells where every factor keeps a constant sign; the analytical-limit cells introduced below satisfy the same per-cell identity $\sum_k \Delta B_{k,ij} = B_{ij}^T - B_{ij}^0$ by the $1/m$ split, so the aggregate remains residual-free. LMDI-I is additionally *consistent in aggregation*: factor effects computed on any subset of sub-categories sum to the effect computed on their union, so county- or state-level

contributions to each factor effect aggregate exactly to the national total (Ang and Liu 2001; Ang 2005). This property underpins the sub-national attribution reported with the results.

5.1.3 Estimator: period-averaged single-step

The headline decomposition compares two *averaged states* rather than single endpoint years. For a base window (1990–1994) and a terminal window (2018–2022) we compute the cell-level mean of area, production, and B over each window, then run a single decomposition step between these averaged states. All five factor quantities (H , S_i , S_{ij} , Y_{ij} , NBI_{ij}) are re-derived from the averaged quantities so that the accounting identity holds at the averaged endpoints. Cells absent from an entire window are treated as births or deaths by the outer join in the single decomposition step and handled by the analytical-limit rule described below.

This estimator is chosen because the nitrogen balance is a signed quantity and a substantial share of cells lie near the break-even point where B changes sign between years. In a chained decomposition, each such sign change triggers the mechanical $1/m$ attribution, and these attributions accumulate along the chain into the cumulative NBI effect. The result is that the chained cumulative NBI is sensitive to the choice of terminal year. Over 1990–2022 the chained NBI runs from approximately -1.4 Tg N at end-2016 to $+0.6$ Tg N at end-2022, solely from the choice of stopping year. That is a swing of about 2.0 Tg N, driven by noise in the annual near-zero-balance cells rather than by genuine trend differences. A single comparison between period-averaged states encounters no sign changes between the averaged endpoints, so it never accumulates those mechanical attributions. The single-step comparison necessarily sets aside the intermediate annual trajectory, but for a long-run structural comparison this is appropriate: the quantity of interest is the net change between the two periods, and averaged endpoints are far more stable than any single pair of reference years. This approach is established practice for long-run structural comparisons using LMDI (Kastner et al. 2012).

The chained year-on-year decomposition,

$$\Delta B_k(t_0, t_n) = \sum_{t=t_0}^{t_n-1} \Delta B_k(t, t+1),$$

is also computed over 1990–2022 and reported alongside the headline as a trajectory sensitivity that reveals within-period dynamics the period comparison cannot resolve — following the presentation in Blomqvist, Yates, and Brook (2020). Because LMDI-I is perfect at every annual step, the chained effects aggregate exactly to the observed $B^{2022} - B^{1990}$ without remainder.

5.1.4 Zero and negative sub-categories

Two features of the panel break the plain logarithmic-mean formula. First, crop $\times$ county cells enter and leave the panel between years (a crop is “born” in or “dies” out of a county), so a factor can be zero in one year: $\ln(x^T/x^0)$ then diverges and the weight $L(0, c)$ collapses. Second, the balance B_{ij} can be negative (see the negative- B audit below) and can change sign between years. We handle both with the analytical-limit (AL) convention (Ang and Liu 2007a) and its extension to sign changes (Ang and Liu 2007b).

When a cell has no harvested area in a year ($H_{ij} = 0$, hence no production) we set $Y_{ij} = \text{NBI}_{ij} = 0$ by convention, consistent with the limiting behaviour of the formula (no yield is observed and no balance arises). More generally $\text{NBI}_{ij} = 0$ is set whenever $P_{ij} = 0$, which also covers the rare cell that records harvested area but zero production. Consider a sub-category in which m of the five factors undergo a zero-change or a sign-change between the two years, while every other factor keeps a constant nonzero sign. Each of the m affected factors is credited with an equal share $\frac{1}{m}(B_{ij}^T - B_{ij}^0)$ of the cell’s change, and the unaffected factors are credited with zero. The remaining cells, where all five factors keep a constant sign in both years, take the standard logarithmic-mean formula above, which — in the ratio form

given earlier — is well defined for two negative values of the same sign. Cells absent in both years contribute nothing and need no computation.

The dominant application is the born/died cell. For a crop appearing in a county that is itself present in both years ($H_i > 0$, $H > 0$), the factors S_{ij} , Y_{ij} , and NBI_{ij} all transition $0 \rightarrow +$ while H and S_i stay positive. So $m = 3$, and each of the three is credited $B_{ij}^T/3$. A dying cell is the mirror image, with $-B_{ij}^0/3$. If the county itself enters or exits (so S_i also changes), $m = 4$. The AL convention is exact rather than approximate (unlike the small-value substitution sometimes used), and resolves the overwhelmingly common case — a crop perpetually absent from a county — at zero computational cost.

The other route into the analytical-limit branch is a pure sign change of the balance. Because H , S_i , S_{ij} , and Y_{ij} are non-negative by construction, the sign of B_{ij} is carried entirely by $\text{NBI}_{ij} = B_{ij}/P_{ij}$. A cell with positive harvested area in both years whose balance changes sign between them ($+ \rightarrow -$ or $- \rightarrow +$) is therefore an $m = 1$ case, with NBI the sole sign-changing factor. The standard log-mean cannot be applied, because the weight's ratio B_{ij}^T/B_{ij}^0 would be negative. The entire cell change is therefore credited to the NBI effect, and the other four factors receive zero for that cell. This is the correct analytical-limit behaviour, but it does load balance sign changes wholly onto NBI; the prevalence of such changes, and the sensitivity of the headline NBI effect to them, is reported with the negative- B robustness checks below.

5.1.5 Verification

The decomposition is implemented in R. Every run is checked: the headline period-averaged step and each of the 32 annual steps of the chained trajectory. We verify the accounting identity at both endpoints ($\sum_{ij} B_{ij} = B$ for each state, and $B_{ij} = H S_i S_{ij} Y_{ij} \text{NBI}_{ij}$ for every cell with positive area), and confirm that the five factor effects sum to the observed change in B to machine precision. The chained effects likewise sum to the observed $B^{2022} - B^{1990}$. Factor

effects are reported in physical units (kg N), with a positive effect indicating that the change in that factor raised aggregate balance and a negative effect that it lowered it.

5.2 Data sources and processing

5.2.1 Nitrogen application data

County-level nitrogen application data are taken from the inputs developed for the US National Greenhouse Gas Inventory (Hoskovec et al. 2026), a national reconstruction of synthetic fertilizer and manure nitrogen applied to cropland at the county $\times$ crop $\times$ year level, covering 1990–2023. The inventory is built on the USDA National Resources Inventory (NRI), a longitudinal survey conducted at a large panel of sample points whose locations are fixed in space and revisited over time. Crop identity at each point is taken from the NRI through 2017 and from the USDA Cropland Data Layer thereafter. Because the sample points are fixed, the set of points falling in a given county does not vary from year to year, and neither does whether that county is represented in the inventory at all. That is a property of the county rather than of the county-year, and the panel construction below relies on it.

Application rates are estimated from the survey together with ancillary agronomic and economic data and delivered as six multiply-imputed realizations that carry the uncertainty in the estimation; each realization reports, for every county $\times$ crop $\times$ year cell, the fertilized (or manured) area and the total nitrogen mass applied. Application rates are modelled crop- and location-specifically with a gradient-boosting model for corn, cotton, soybeans, and wheat, with sorghum borrowing corn’s rates and barley and oats wheat’s. The remaining crops receive coarser average rates from survey data for lumped categories such as vegetables and horticulture. A top-down fertilizer-sales total is imposed as an over-application constraint across all crops. This coarser treatment of the non-major crops is the source of the elevated specialty-crop nitrogen-balance residuals noted later.

The two source files — synthetic fertilizer and manure — are processed in parallel. The inventory is delivered already aggregated to county scale, with NRI survey points area-weighted to their county. Fertilizer and manure nitrogen are summed, within each imputation, to an applied-nitrogen total per cell, and the six imputed totals are averaged to give the headline applied input,

$$N_{\text{appl},ij} = N_{\text{fert},ij} + N_{\text{man},ij},$$

used throughout; the six individual realizations are retained for a downstream imputation-sensitivity check. NRI crop labels are mapped to the harmonized NASS commodities of the analysis through the crop crosswalk (Table S1). The inventory contains a row only for cells in which the crop was recorded at the sampled points, so an applied amount of zero is represented by the absence of a row rather than an explicit zero. The consequences of that convention for the panel are taken up in the field-balance-panel subsection below. Biological fixation and atmospheric deposition are added to $N_{\text{appl},ij}$ in later subsections to form the total nitrogen input $N_{\text{in},ij}$.

5.2.2 NASS data acquisition

Harvested area and crop production data were obtained from the United States Department of Agriculture National Agricultural Statistics Service (USDA-NASS) Quick Statistics database, accessed through the NASS application programming interface using the `rnassqs` package (Potter 2019) in R. For each commodity we queried harvested area and production, at both the county and state geographies, over 1990–2023, from two sources: the annual sample-based survey and the every-farm Census of Agriculture (conducted in 1992, 1997, 2002, 2007, 2012, 2017, and 2022 within the study period).

Every query was restricted to commodity-level totals, excluding the management-practice domain subdivisions that NASS publishes alongside them. Those subdivisions would double-count if summed with their parent aggregate. Production-practice and utilization-practice breakdowns are not covered by this restriction and are reduced in a separate step below.

Crops covered by the NRI nitrogen application data were matched to NASS commodities, and to NASS classes within them, wherever county-level data are available. Two NRI catch-all categories and two commodities lacking usable county-level NASS coverage were excluded (Table S1).

Table S1: NRI crop categories and their inclusion status in the analysis.

Crop	Included	Notes
Barley	Yes	Winter and spring varieties aggregated
Corn	Yes	
Cotton	Yes	Upland and Pima varieties summed
Lentils	Yes	
Oats	Yes	Winter and spring varieties aggregated
Peanuts	Yes	
Peas	Yes	Production yield-imputed from state survey yield
Potatoes	Yes	
Rice	Yes	
Sorghum	Yes	
Soybeans	Yes	Winter and spring varieties aggregated
Sugar beets	Yes	
Sunflower	Yes	Oil-type and non-oil-type varieties summed
Sweet Potatoes	Yes	Production yield-imputed from state survey yield
Tobacco	Yes	
Tomatoes	Yes	census-only; 1990–1991 missing (no anchor)
Wheat	Yes	Restricted to processing tomatoes
Close-grown crops	No	Winter and spring retained as separate classes
Dry Beans	No	No NASS commodity equivalent
Onions	No	County-level NASS coverage begins only in 2019
truck crops	No	No county-level NASS survey data
	No	No NASS commodity equivalent

5.2.3 Parsing reported values and unit conversion

NASS marks two kinds of unavailable value: the disclosure code “(D)”, withheld when fewer than three operations report, and the threshold code “(Z)”, for quantities below half the reporting unit. The `rnassqs` parser converts both to NA, and we retain that treatment. Reading “(Z)” as missing rather than as zero is conservative: such quantities are below half a bushel or hundredweight, so a cell later filled by interpolation under this convention receives at most a negligible positive value. Non-physical reporting units were dropped.

Reported values were converted to kilograms using USDA Federal Grain Inspection Service test weights for bushel-denominated grains and standard factors for the hundredweight, short-ton, and cotton-bale units; areas were converted from acres to hectares. Unit labels that differ between survey and census for the same physical unit were reconciled first, so the coalescing step below cannot mix bases.

Two auxiliary series were retained alongside the county panel. State-level data support the suppression diagnostics, the survey-existence indicator, the yield-bound coupling, and the potato imputation described below. NASS also publishes district-level aggregates of the suppressed counties within an Agricultural Statistics District; these carry no FIPS code and are excluded from the county panel, but they are by construction the natural yield denominator for the counties they cover, which the yield-bound step exploits.

5.2.4 Aggregating across production and utilization practices

For several commodities NASS returns several rows per county–year, subdividing the crop by production practice (irrigated, non-irrigated, and for barley and wheat two further non-irrigated categories) or by utilization practice (grain, silage, processing). We reduce these to exactly one row per cell: the aggregate row is used where NASS publishes one, and where it does not, the

complementary sub-types are summed to recover the total. The same reduction is applied to the county, state, and district series.

Three choices within this step are worth stating. First, we accept an aggregate row even when it is suppressed, rather than falling through to whichever sub-types happen to be published. Falling through would recover a few hundred barley and wheat cells in the high-Plains states, but it undercounts area and inflates implied yields, because sub-type publication is itself often partial in the same county-years where the parent is suppressed; the yield-bound step below recovers these cells more reliably. Second, corn and sorghum have no aggregate utilization row, so we take grain only, consistent with the nitrogen application inventory, and exclude silage- and syrup-only counties so they cannot enter the panel without matching nitrogen input. Third, tomatoes are restricted to processing tomatoes: counties reporting only fresh-market production would carry the fresh-market area without its production, giving a spuriously low implied yield. This drops roughly 9% of US tomato production, which we note as a limitation.

5.2.5 Coalescing Survey and Census sources

For each cell the survey value is retained where published, and the census value fills the gap where it is not. The provenance of every cell is recorded. Two features of the source data make this hybrid necessary.

For several geographically concentrated specialty crops — potato, sugar beet, and non-oil sunflower — county-level survey publication has deteriorated over the study period as production consolidated into fewer counties and crossed NASS’s disclosure thresholds. Survey rows for potato area fell from about 4,200 in 1990 to six in 2018, and to none from 2019 onward. The Census of Agriculture enumerates every operation and is not subject to the same suppression, so it fills these gaps in census years. Tobacco is a stronger case still: it is published only through the census at the county level, in every year. The tobacco panel is therefore drawn entirely from

the census and interpolated between census-year anchors, and 1990 and 1991 have no anchor and remain missing — a documented limitation.

One consistency caveat attaches to the 2022 census. County-level census sums in that year exceed the state survey total by 10–33% for several major Corn Belt state–crop combinations, among them Kansas sorghum, Nebraska corn, and Missouri soybeans. The most plausible explanation is that the 2022 census enumerates small operations below the survey sampling threshold, so the two are not strictly reconcilable that year. This does not corrupt the zero-fill and missing-value assignment below, which triggers only on under-coverage. It does mean that the 2021 → 2022 step of the chained decomposition may register an area increase that is partly a change in enumeration rather than real agricultural change.

5.2.6 Handling suppressed values

The procedure below applies to the values NASS withholds under the “(D)” and “(Z)” codes, and equally to any county that simply lacks a row in the API response. The API returns that silent absence indistinguishably from genuine zero production, which is the central difficulty.

An absent county row has three possible explanations. The county genuinely grew none of that crop; production was withheld without an explicit marker; or the crop was not surveyed at the county level that year. The rule described here bounds the second case. The state total is an independent NASS estimate that includes all suppressed counties’ production, while the county sum excludes suppressed rows of either kind, so the gap between them captures suppression whether explicit or silent. Genuine zeros leave no trace in that gap, since a county with no production contributes nothing to either side. The third case is invisible to the rule, because an unsurveyed year has no state total either; it is handled in the next subsection.

For each state, commodity, and year we computed the state residual: the state total less the sum of known county values, as a fraction of the state total. Where the residual fell below 5%,

absent county observations in that state were assigned zero. Where it reached 5% or more, they were retained as missing, to avoid recording suppressed production as an absence of production. Those missing values then pass to the interpolation step below.

This treats genuine zeros only as a byproduct: correctly zero-filled in low-residual states, but assigned missing in high-residual ones, which may be wrong and is the origin of the phantom-production artefact discussed below. The obvious alternative — zero for every absent observation — avoids that artefact and handles genuine zeros correctly. We rejected it because where suppression is pervasive it systematically understates production in the suppressed counties, manufacturing crop-cell births and deaths that reflect NASS disclosure practice rather than agriculture. Those would bias the crop-mix, yield, and NBI contributions in exactly the years when suppression begins or lifts, and unlike phantom production they cannot be bounded from the data. The residual rule accepts a quantifiable error in place of an unquantifiable one.

The 5% threshold is a pragmatic choice. The full distribution of residuals across commodities and years is reported below (Figure S3) so that its consequences can be judged.

5.2.7 Distinguishing zero production from unsurveyed years

Several specialty crops — dry bean, pea, lentil, sweet potato, and tomato among them — are not surveyed annually at the county level, appearing only in the quinquennial census. Treating an unsurveyed year as a year of zero production would create spurious cell births and deaths, charging several years of accumulated change to a single year-to-year transition.

We use the state-level series as a survey existence indicator: a crop is treated as surveyed in a state and year if the state record exists. Where it does, the crop was grown somewhere in the state, so a county-level absence can only be a genuine zero or a suppressed value, and the residual rule above applies. Where neither the state nor the county record exists, the absence

carries no information at all about production, so the gap is retained as missing and filled by interpolation between census-year anchors.

Figure S1 summarises the full decision procedure for missing and absent county-level observations.

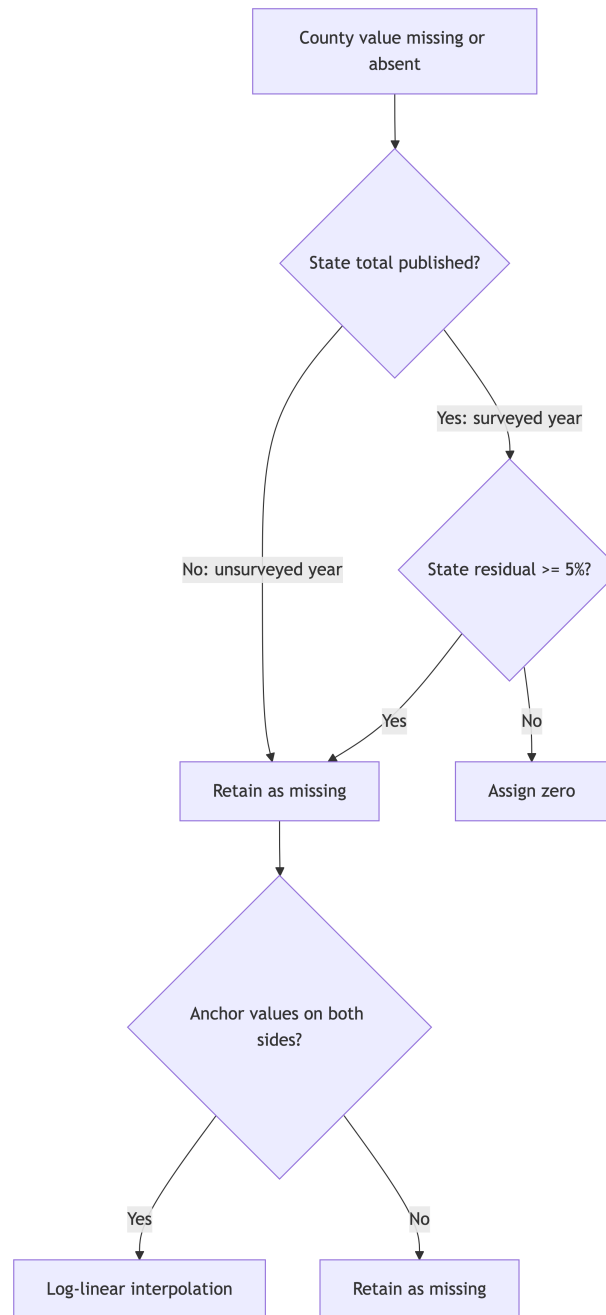


Figure S1: Decision procedure for handling missing and absent county-level NASS observations.

5.2.8 Log-linear interpolation

Missing values retained by the procedure above, whether from unsurveyed years or from county observations withheld under the residual rule, were filled by log-linear (geometric) interpolation within each county $\times$ commodity series. For a gap between two observed anchor values x_{t_0} and x_{t_1} , the interpolated value at year t is:

$$\hat{x}_t = x_{t_0} \left(\frac{x_{t_1}}{x_{t_0}} \right)^{\frac{t-t_0}{t_1-t_0}}$$

Area and production were interpolated separately within each county $\times$ commodity series. The geometric form suits quantities bounded below by zero that grow multiplicatively, making the implied year-to-year changes proportional rather than additive. Two restrictions apply. The procedure does not extrapolate beyond the first or last observed positive value, so a series stays missing outside its observed range. Nor does it bridge across a zero anchor: where either endpoint is zero the gap is left missing. A zero anchor means the crop was absent that year, and bridging one manufactures production for the years between at implied yields that can be an order of magnitude too high. The rule trades a few unfilled gaps, which the yield-bound step can still recover, against those artefacts. All interpolated values are flagged, so sensitivity analyses can be restricted to observed data.

The residual rule and interpolation interact to produce one artefact worth stating plainly. A county with genuinely zero production in a year returns no row. If the state residual that year reaches 5%, because other counties were suppressed, that absence is recorded as missing rather than zero; and if the county produced the crop in adjacent years, interpolation will fill the gap with a positive value that never existed.

Three things bound this. Interpolation needs positive anchors on both sides, so the artefact requires production to fall to zero and then return — an unusual pattern, and one that cannot arise

for counties that decline to zero permanently or have never grown the crop. Where it does arise, the interpolated value lies between the two anchors and cannot exceed the larger of them, so it is small wherever the surrounding values are small. And the yield-bound step below rescales interpolated areas so that their state-level sum matches the published survey total, which directly limits over-allocation. Table S2 reports the affected area by commodity, in hectares and as a share of each commodity’s panel area. It is an upper bound in both directions: not every one of these cells is a genuine zero, and the figures precede the yield-bound rescaling.

Table S2: Cells flagged as high-residual NA under the residual rule that were subsequently filled by log-linear interpolation, by commodity. Reported as an upper bound on the phantom-production artefact, before the yield-bound rescaling further constrains it.

Commodity	Cells (n)	Interpolated area (ha)	% of commodity area
Pea	401	1,166,963	19.20
Sunflower	1162	4,138,121	17.08
Sweet potato	405	128,362	15.28
Lentil	185	470,145	14.16
Sorghum	4337	9,171,871	9.58
Oats	8287	2,386,273	9.32
Peanut	966	1,717,211	8.92
Wheat	8014	50,389,349	7.44
Potato	3213	966,800	7.30
Barley	3422	2,884,449	5.22
Dry bean	5	46,304	4.78
Rice	247	1,257,168	3.19
Corn	6339	32,863,254	3.14
Soybean	4042	28,425,788	2.85
Cotton	1487	3,930,209	2.64
Tomato	79	14,576	0.77
Sugar beet	44	112,594	0.67
Tobacco	3	77	0.01

5.2.9 Yield-bound coupling and state-survey area scaling

Because area and production are interpolated independently, the two series can carry different observed anchors — a county whose area is published in years when its production is sup-

pressed, or the reverse. Interpolating them separately then yields implied yields several times the state survey value, and can distribute enough interpolated area across silently suppressed counties that the state sum exceeds the published survey total. Both pathologies occur in this data.

To bound implied yields and reconcile the county panel with the state survey, a coupling step after interpolation classifies each cell by the original, pre-interpolation observation status of its area and production:

- **Both observed:** retain both values unchanged.
- **Area observed, production imputed:** recompute production as $H_{ij,t} \cdot Y_{i,j,t}^{\text{ref}}$.
- **Production observed, area imputed:** recompute area as $P_{ij,t} / Y_{i,j,t}^{\text{ref}}$.
- **Both imputed:** rescale the interpolated areas within the (state, commodity, class, year) cell so that the panel sum matches the state survey total minus area already attributed to cells in the first three classes (the “fixed” portion); production is then recomputed from the rescaled area via $Y_{i,j,t}^{\text{ref}}$.

The reference yield $Y_{i,j,t}^{\text{ref}}$ is the district-level suppressed-county yield where NASS publishes one — the ratio of production to area over exactly the suppressed counties it aggregates, and so their natural denominator — falling back to the state survey yield otherwise. Cells with neither retain their interpolated values. The step is applied to every commodity except potato and sweet potato, which follow the dedicated path described next.

Rescaling requires at least one cell in each state–year with both quantities imputed and an interpolated area to absorb the residual. Where every suppressed county in a state–year has a zero anchor blocking interpolation, the residual stays unallocated and the panel falls short of the published state total. The unallocated quantity is small, and it is known-suppressed area that is simply not attributed to any named county, but state-level sums from the panel may not reconcile with published survey totals in those cells.

5.2.10 Yield-imputation for potatoes and sweet potatoes

County-level potato production is not reliably published over the study period. Survey county production deteriorates and ends in 2018, and the census stops publishing county-level potato production after 2002, enumerating only acreage thereafter. State-level survey data for both area and production are robust throughout. Sweet potato has similar gaps, plus a further pathology: census area anchors can be far smaller than adjacent survey years for the same county, which produces absurd implied yields when survey-scale production is interpolated across them.

For both crops, county production is therefore imputed as

$$P_{ij,t} = H_{ij,t} \cdot Y_{s(i),j,t}^{\text{state}}$$

where $Y_{s,j,t}^{\text{state}}$ is the state-level survey yield for state s , commodity j , and year t , and $s(i)$ is the state containing county i . It is applied across the whole study period, including years with observed county production, so that the panel carries no discontinuity at the boundary between observed and imputed years. The imputation assumes yield is roughly uniform within a state. That is defensible for these two crops because production is geographically concentrated in each state — potatoes in the Snake River Plain, the Columbia Basin, and Aroostook County, sweet potatoes in a four-state cluster of North Carolina, California, Louisiana, and Mississippi.

Potato and sugar beet also need one adjustment to the zero-fill rule: absent county cells in non-census years are forced to missing regardless of the state residual, so that interpolation can bridge between census anchors. Without it, the few commercial counties still in the survey after coverage collapsed would account for nearly all state area, and the general residual rule would zero-fill every other county in the major potato states. Potato coverage consequently ends in 2022, since the survey has ended, the next census falls outside the study period, and the procedure does not extrapolate.

Where NASS publishes no state-level yield because production falls below the reporting threshold, imputed production is missing, and we set the cell's area to missing too so an unmatched area–production pair cannot enter the decomposition. The affected states are all trivially small producers.

Cross-validation against the observed county production cells that remain gives a median imputed-to-observed ratio of 1.05 for potato ($n = 4,284$ overlapping cells) and 1.06 for sweet potato ($n = 940$), with interquartile ranges of $[0.93, 1.34]$ and $[0.94, 1.27]$. About 57% and 61% of cells respectively fall within $\pm 20\%$ of the observed value. The upper-tail bias is consistent with suppressed counties being marginal producers with below-average yields. Both crops are a minor share of the national nitrogen balance, and a sensitivity run excluding them bounds their influence on the decomposition.

5.2.11 Aggregating multi-class commodities

Several commodities are queried as multiple NASS classes and aggregated to one label per crop in the final panel: upland and Pima cotton are summed, as are oil-type and non-oil-type sunflower, and durum and non-durum spring wheat. Barley, oats, and sugar beet are published by NASS only as a single class. The NRI data instead distinguish winter and spring varieties of wheat, barley, oats, and sugar beet, so NRI nitrogen inputs for the spring and winter varieties of the latter three are summed to match the NASS aggregate.

Wheat alone is kept as two crops, winter and spring, because their grain nitrogen content differs materially — roughly 0.020 against $0.024 \text{ kg N kg}^{-1}$, following the difference in grain protein. Aggregating them would mis-estimate N_{out} , and hence NBI, in counties dominated by one type. No comparable difference exists between the seasonal varieties of barley, oats, and sugar beet, so aggregating those loses nothing. Within each aggregated class, area and production

are summed across the constituent classes, and a cell whose contributing classes are all missing remains missing.

5.2.12 Diagnostic checks

Four diagnostics show how much of the panel rests on imputation, and where. The phantom-area upper bound is reported above (Table S2), and the distribution of state residuals underlying the 5% rule in the subsection that follows (Figure S3). A third records the share of county-year rows carrying an explicit disclosure flag, fixing the upstream suppression burden before any imputation. The fourth is the reach of the yield-bound step (Table S3), which recomputes a few percent of cells for the major Corn Belt crops and a far larger share for tobacco and the specialty crops that depend heavily on census supplementation.

Table S3: Cells whose area or production was recomputed by the yield-bound coupling step, by commodity. Potato and sweet potato are zero by design — they are handled by their own yield-imputation paths.

Commodity	Cells (n)	Bounded (n)	% bounded	Bounded area (ha)
Tobacco	22,372	16,863	75.4	1,403
Peanut	20,604	12,447	60.4	1,723,014
Cotton	28,968	16,076	55.5	41,833,135
Lentil	4,862	2,445	50.3	485,513
Sugar beet	8,330	4,187	50.3	1,498,179
Oats	87,482	43,892	50.2	2,394,517
Barley	61,846	29,567	47.8	2,895,135
Sorghum	64,906	28,493	43.9	9,191,535
Rice	6,902	2,945	42.7	1,259,531
Wheat	130,220	52,645	40.4	88,765,688
Sunflower	46,750	16,732	35.8	8,337,083
Corn	99,246	32,744	33.0	32,892,975
Soybean	84,796	27,138	32.0	28,456,245
Pea	22,950	6,288	27.4	1,173,460
Tomato	57,324	5,925	10.3	230,603
Dry bean	918	52	5.7	46,304
Potato	94,282	0	0.0	0
Sweet potato	60,112	0	0.0	0

A coverage heatmap of county-level harvested-area publication rates by commodity and year (Figure S2) summarises where the panel relies on imputation.

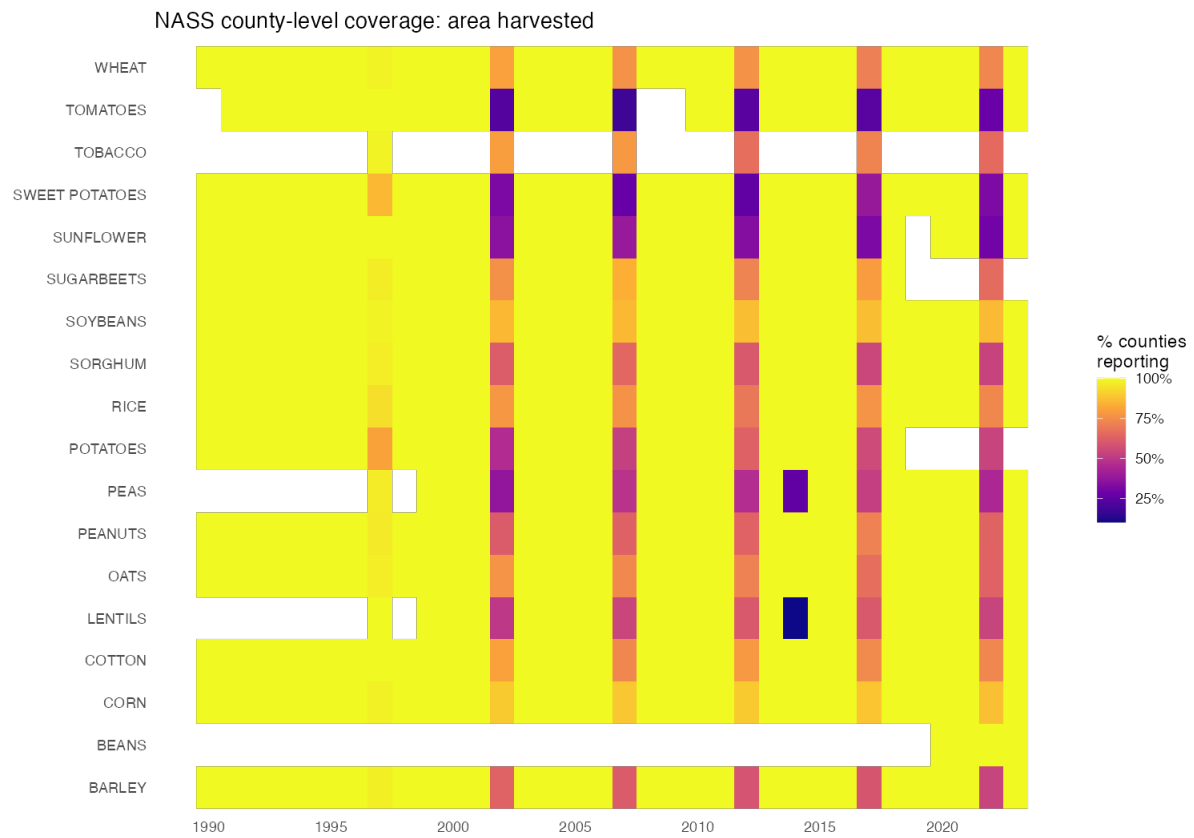


Figure S2: Share of (county $\times$ commodity) cells with a non-suppressed harvested-area value after coalescing survey and census sources, by commodity and year. White cells indicate no published value from either source. Vertical low-coverage bands at 1997, 2002, 2007, 2012, 2017, and 2022 are Census of Agriculture years (see text).

The vertical low-coverage bands at the seven census years are a structural feature of the source data rather than a coverage problem in the usual sense. The census enumerates a long tail of small marginal producers that fall below the survey sampling threshold. These counties enter the panel only in census years, but their small operation counts also trigger disclosure suppression at a high rate, so the denominator of the reporting share expands faster than the numerator. The effect is strongest for crops with a wide tail of small producers — tomato, sweet potato, potato, sunflower, and sugar beet — and is barely visible for crops grown at scale across most US counties. For tomato it is compounded by the processing-only restriction, which

drops fresh-market-only census counties. Tobacco is coloured only in census years, reflecting its census-only publication, and the potato and sugar beet gaps in non-census years from about 2018 mark the survey collapse described above, with the 2022 census filling them.

5.2.13 State-residual distribution

The 5% residual threshold used in the zero-fill / NA assignment rule (see the handling of suppressed values above) is a pragmatic choice that trades a small risk of phantom production (interpolating a true zero) for avoiding a larger risk of zero-filling silently suppressed values. Figure S3 reports the distribution of state residuals — $(\text{state total} - \text{county sum}) / \text{state total}$ — across all (state, year) observations in the study period, faceted by commodity. The dashed red line at +5% marks the rule cutoff: observations to the right trigger the high-residual NA assignment for absent counties; observations to the left receive zero-fill.

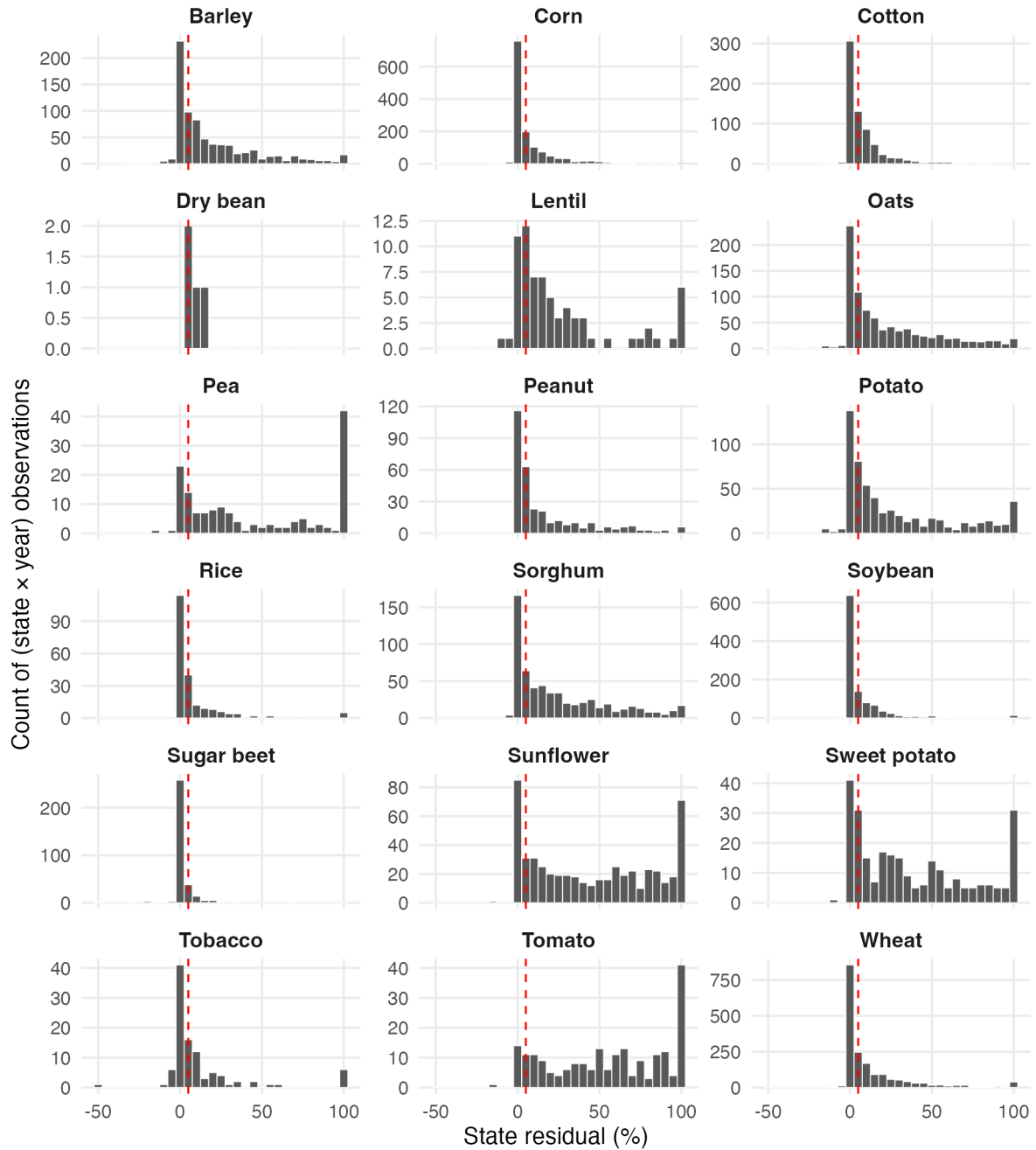


Figure S3: Distribution of state residuals by commodity, pooled across all states and years in the study period. The state residual is the gap between the NASS state-level total and the sum of published county-level values, expressed as a percentage of the state total. The dashed red line at +5% marks the threshold used to choose between zero-fill (residual < 5%) and NA assignment (residual $\geq$ 5%) for absent county observations. Residuals beyond $[-50, 100]\%$ are visually clipped to the panel boundary; the underlying counts include all observations.

For the major Corn Belt crops — corn, soybean, and wheat — the bulk of the distribution lies below the 5% threshold, so the simpler zero-fill rule would already be approximately correct for most state-year cells. The crops in which high residuals are common — potato, sugar beet, sunflower, tobacco, and the dry-legume specialty crops — coincide with those flagged above as requiring census supplementation, the potato and sugar beet non-census-year override, or the yield-bound coupling step. Negative residuals (county sum exceeding state total) are concentrated in census years. They reflect the documented inability of state survey estimates to reconcile with county-level census enumeration when the census exposes small operations below the survey sampling threshold. This is the artefact discussed under the coalescing of survey and census sources above. Because the rule triggers only on residuals at or above +5%, negative residuals do not lead to incorrect NA assignment.

5.2.14 Nitrogen content in crops

Nitrogen removed in the harvested product, $N_{\text{out},ij}$, is the production of crop j in county i multiplied by a crop-specific nitrogen content coefficient:

$$N_{\text{out},ij} = P_{ij} \cdot n_j$$

where n_j is the nitrogen mass per unit mass of harvested product (kg N kg^{-1}). A single coefficient is applied per crop across all counties and years, with the sole exception of wheat, which carries distinct coefficients for its winter and spring classes (see above). The coefficient therefore captures differences in nitrogen concentration between crops but not spatial or temporal variation in grain protein within a crop. Because production is converted to a common mass basis (kg) earlier in the pipeline, $N_{\text{out},ij}$ inherits those units directly.

The coefficients are United States agronomic nutrient-removal rates: the nitrogen contained in the harvested product per unit of yield, compiled from field-measurement-based extension and industry sources rather than from food-composition data. This places N_{out} on the same grain- and fiber-removal basis used throughout the agronomic nitrogen-balance literature. The published values, their units, the derived coefficients, and the source for each crop are listed in Table S4. Two of the coefficients need comment.

Spring wheat has no published removal estimate of its own, so it is derived here. We scale the winter-wheat coefficient by the ratio of spring to winter grain protein reported by USDA FoodData Central (<https://fdc.nal.usda.gov/>) — 15.4 against 12.6 g per 100 g — giving a factor of 1.22 and a coefficient of 0.024 kg N kg⁻¹. Because both classes share the same nitrogen-to-protein conversion, that protein ratio is exactly the ratio of their grain nitrogen concentrations.

Cotton also warrants a note. NASS reports cotton production as lint, but the removal coefficient is the nitrogen in the lint *plus the associated cottonseed*, both of which leave the field at ginning. This is the correct quantity for a field nitrogen balance. A lint-only coefficient would understate cotton’s harvested-nitrogen removal roughly thirtyfold, because the fiber is nearly nitrogen-free and the seed carries almost all of the harvested nitrogen.

Table S4: Crop nitrogen removal coefficients used to compute N_{out} , on a US agronomic harvested-product-removal basis. The published removal value is the nitrogen contained in the harvested product per unit of yield (per bushel, hundredweight, ton, or 480-lb cotton bale; or g N per kg); the N content is that value converted to kg N per kg of production using each crop’s standard test weight. Sources: Roberts et al. (n.d.) for the major field crops, Bouwman, Van Drecht, and Van der Hoek (2005) for the minor and specialty tail, with potato from Shober (2024) and tobacco from Murrell (2008). Wheat is retained as separate winter and spring classes; Barley, oats, and sugar beet take a single coefficient each, as their seasonal varieties do not differ in nitrogen content.

Crop	Removal (published)	N content (kg N/kg)	Source
Barley	17 g N/kg	0.0170	Bouwman et al. 2005
Corn	0.67 lb N/bu	0.0120	Roberts et al.
Cotton	32 lb N/bale	0.0667	Roberts et al.

(continued)

Crop	Removal (published)	N content (kg N/kg)	Source
Lentil	35 g N/kg	0.0350	Bouwman et al. 2005 (pulses)
Oats	16 g N/kg	0.0160	Bouwman et al. 2005 (other cereals)
Pea	35 g N/kg	0.0350	Bouwman et al. 2005 (pulses)
Peanut	40 g N/kg	0.0400	Bouwman et al. 2005 (groundnuts in shell)
Potato	0.37 lb N/cwt	0.0037	Shober 2024
Rice	0.57 lb N/bu	0.0127	Roberts et al.
Sorghum	0.66 lb N/bu	0.0118	Roberts et al.
Soybean	3.3 lb N/bu	0.0550	Roberts et al.
Spring wheat	1.47 lb N/bu	0.0244	Roberts et al., scaled by FoodData Central protein ratio
Sugar beet	2 g N/kg	0.0020	Bouwman et al. 2005
Sunflower	34 g N/kg	0.0340	Bouwman et al. 2005
Sweet potato	3 g N/kg	0.0030	Bouwman et al. 2005
Tobacco	3.6 lb N/cwt	0.0360	Murrell 2008
Tomato	2 g N/kg	0.0020	Bouwman et al. 2005 (vegetables)
Winter wheat	1.2 lb N/bu	0.0200	Roberts et al.

5.2.15 Biological nitrogen fixation in legumes

Legumes acquire much of their nitrogen from the atmosphere through symbiotic biological nitrogen fixation (BNF). That nitrogen enters the soil–plant system in the same accounting sense as applied fertilizer, so it is added to the inputs of legume crops, and the applied base defined above becomes

$$N_{\text{appl},ij} + \text{BNF}_{ij},$$

where $\text{BNF}_{ij} = 0$ for all non-legumes, and atmospheric deposition completes $N_{\text{in},ij}$ in the next subsection. The rationale is given in the Methods: including BNF aligns the input definition with X. Zhang et al. (2015) and prevents the spurious deficits that omitting a legume’s dominant nitrogen source would otherwise create. BNF is computed for the four legumes in the panel — soybean, peanut, lentil, and dry pea.

The fixation literature reports the fraction of nitrogen a legume derives from the atmosphere as β_j (%Ndfa), defined relative to *total plant nitrogen at maturity* rather than to the nitrogen in the harvested grain. The pipeline observes only grain nitrogen, so BNF is recovered in three steps that first reconstruct whole-plant nitrogen and then apply β_j :

$$\begin{aligned} N_{\text{plant},ij}^{\text{ag}} &= N_{\text{out},ij} / \text{NHI}_j, \\ N_{\text{plant},ij}^{\text{total}} &= N_{\text{plant},ij}^{\text{ag}} \cdot f_{\text{BGN},j}, \\ \text{BNF}_{ij} &= \beta_j \cdot N_{\text{plant},ij}^{\text{total}}. \end{aligned}$$

The first step rescales grain nitrogen to total *aboveground* plant nitrogen using the nitrogen harvest index NHI_j , the fraction of aboveground plant nitrogen held in the grain at maturity. The second scales aboveground nitrogen up to the whole plant with a belowground multiplier $f_{\text{BGN},j}$ covering roots, nodules, and rhizodeposition, which shoot sampling routinely omits and which ^{15}N -tracer studies show to be a large fraction of plant nitrogen (Herridge, Peoples, and Boddey 2008). Applying β_j to that total gives the fixed nitrogen. Combining the steps, $\text{BNF}_{ij} = (\beta_j f_{\text{BGN},j} / \text{NHI}_j) N_{\text{out},ij}$, so fixed nitrogen scales with grain nitrogen and hence with yield. The harvest-index step matters: applying β_j directly to grain nitrogen would understate BNF by roughly the value of NHI_j , about 30% for soybean, and this is a common pitfall.

The parameters are crop-specific and are listed with their sources in Table S5. Two choices within them are worth stating. For soybean we use the United States–specific %Ndfa of 0.60 rather than the global soybean–groundnut value, and take the nitrogen harvest index from the

directly measured meta-analytic mean of Salvagiotti et al. (2008). The other legumes have no equally authoritative direct measurement, so their NHI is derived as $\text{NHI}_j = \text{HI}_j n_j / w_j$, where HI_j is the dry-matter harvest index, n_j the grain nitrogen content already used for N_{out} (Table S4), and w_j the nitrogen concentration of aboveground biomass. That formula combines a dry-matter harvest index with an as-reported grain nitrogen concentration, whose moisture bases differ slightly. The soybean override sidesteps the approximation for the dominant legume, and for the minor legumes its effect on national totals is negligible.

Two variants of BNF are computed in parallel: an aboveground-only estimate, setting $f_{\text{BGN},j} = 1$, and the headline estimate including belowground nitrogen. The headline includes it because shoot-only measurements are now understood to understate fixation materially (Herridge, Peoples, and Boddey 2008), and because under it soybean's field nitrogen balance comes out approximately neutral, consistent with the long-standing finding for well-nodulated legume crops (Salvagiotti et al. 2008). The aboveground-only variant is retained for sensitivity analysis. For US soybean the headline implies fixation on the order of 200 kg N ha^{-1} , against the roughly 190 kg N ha^{-1} that Herridge, Peoples, and Boddey (2008) estimate for the United States.

Table S5: Crop-specific parameters for the biological nitrogen fixation calculation. The Ndfa fraction (β_j) is the fraction of plant nitrogen derived from the atmosphere; the harvest index (HI_j) is the dry-matter grain fraction; shoot N (w_j) is the nitrogen concentration of aboveground biomass; NHI is the nitrogen harvest index — the directly measured value for soybean and the derived $\text{HI}_j n_j / w_j$ for the others; the BGN multiplier is $f_{\text{BGN},j}$. All values are from Herridge et al. (2008) except the soybean NHI, which is from Salvagiotti et al. (2008).

Crop	Ndfa fraction	Harvest index	Shoot N (%)	NHI	BGN multiplier
Soybeans	0.60	0.40	3.0	0.73	1.5
Peanuts	0.68	0.40	2.3	0.70	1.4
Lentils	0.63	0.35	2.0	0.61	1.4
Dry peas	0.63	0.35	2.0	0.61	1.4

5.2.16 Atmospheric nitrogen deposition

Reactive nitrogen deposited from the atmosphere — wet plus dry, oxidized plus reduced — is a further input to every field regardless of crop. Following X. Zhang et al. (2015), it is added to all cells, completing the total nitrogen input:

$$N_{\text{in},ij} = N_{\text{appl},ij} + \text{BNF}_{ij} + N_{\text{dep},ij}, \quad N_{\text{dep},ij} = d_i \cdot H_{ij},$$

where d_i is the deposition rate ($\text{kg N ha}^{-1} \text{ yr}^{-1}$) in county i and H_{ij} the cell's harvested area. Deposition falls on the land independently of what is planted, so the rate varies by county and year but not by crop, and is scaled to each cell by its area.

No single fine-resolution product spans the study period, so two gridded total-nitrogen products are combined. For 2000–2022 we use the NADP Total Deposition (TDep) measurement–model fusion product (Schwede and Lear 2014; National Atmospheric Deposition Program 2025), which fuses monitored wet-deposition and air-concentration data with modelled dry-deposition velocities over the contiguous United States. For 1990–1999, which TDep does not cover, we use the NACID-NDEP1 reconstruction over North America (Hember 2018), anchored to ground monitoring from 1990 onward. Each annual raster is reduced to a county mean by area-weighted zonal extraction, each grid cell contributing in proportion to its overlap with the county polygon (Baston 2025), giving one value $d_{i,t}$ per county and year.

The two products are independent and built by different methods, so NACID is not used in absolute terms. Its 1990–1999 values are rescaled to the TDep level by a county-specific factor estimated over the 2000–2013 overlap. For each county i we compute

$$r_i = \frac{\overline{d_{i,t}^{\text{TDep}}}}{\overline{d_{i,t}^{\text{NACID}}}} \Big|_{t \in 2000-2013}, \quad d_{i,t}^{1990-99} = r_i \cdot d_{i,t}^{\text{NACID}},$$

with counties lacking overlap data, about 4%, assigned the national median ratio. The headline series is therefore TDep for 2000–2022 and harmonized NACID for 1990–1999. The two products agree closely over the overlap (median $r_i \approx 1.1$, interquartile range 1.0–1.25), and the splice introduces no discontinuity in the national mean at the 1999→2000 boundary (Figure S4). The final series covers 3,084 contiguous-US counties; Alaska, Hawaii, and the territories fall outside TDep’s domain and are dropped. The national cropland-county mean runs at about 9–10 kg N ha⁻¹ yr⁻¹ in the 1990s and declines to roughly 8 by the 2010s, consistent with the fall in oxidized-nitrogen deposition that followed Clean Air Act controls on NO_x emissions.

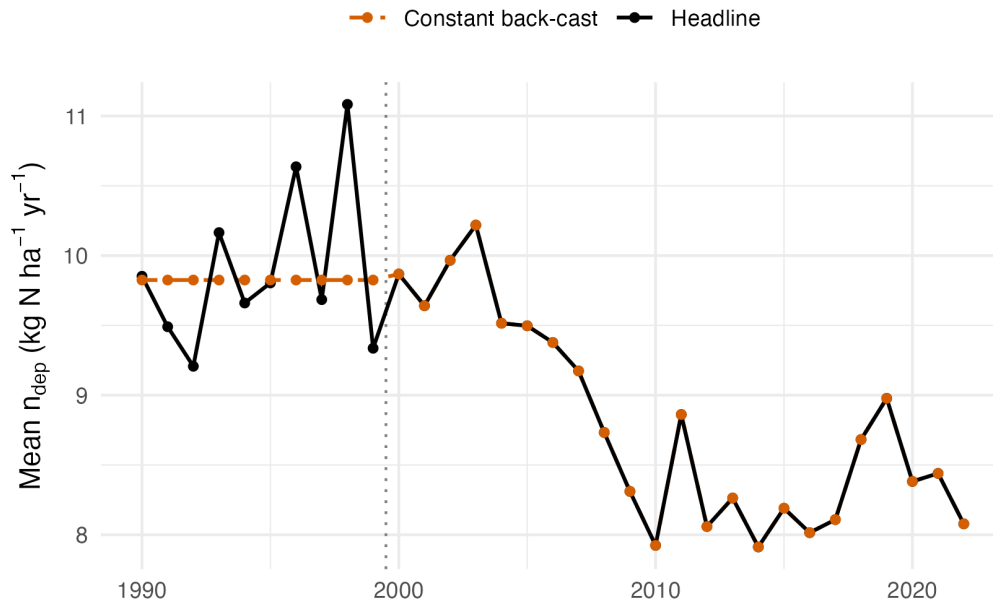


Figure S4: National mean atmospheric nitrogen deposition rate across contiguous-US counties, 1990–2022. The headline series uses NADP-TDep for 2000–2022 and harmonized NACID-NDEP1 for 1990–1999 (split at the dashed line); the constant back-cast sensitivity holds the 2000–2002 TDep county mean fixed across 1990–1999. The two back-casts agree to within about 1 kg N ha⁻¹, and neither introduces a discontinuity at the splice.

Four limitations attach to the series. The 1990s rest on a splice, and although the harmonization removes any systematic level offset, only NACID’s relative year-to-year variation informs those values, not its magnitude. A constant back-cast holding the 2000–2002 TDep mean fixed

agrees with the harmonized series to within about 1 kg N ha⁻¹ (Figure S4), so the pre-2000 treatment is not material. Deposition is assumed uniform across crops within a county, so sub-county gradients, such as elevated deposition downwind of concentrated livestock operations, are unresolved. Both products capture organic nitrogen only partially, so N_{in} is mildly underestimated. And nine counties in central Texas carry rates of 50–75 kg N ha⁻¹ yr⁻¹ in recent years, which are published in the source product rather than aggregation artefacts and are passed through unchanged. Separately, much of the deposition in the Corn Belt is driven by agricultural ammonia emissions, so the input is partly endogenous to agriculture — a feature of regional nitrogen budgets the balance framework does not disentangle.

5.2.17 Assembling the field-nitrogen-balance panel

The four input components above combine into the cell-level quantity that the decomposition explains. For each crop j in county i and year t , the field nitrogen balance is the total nitrogen input minus the nitrogen removed in the harvested product:

$$B_{ij} = N_{\text{in},ij} - N_{\text{out},ij} = (N_{\text{appl},ij} + \text{BNF}_{ij} + N_{\text{dep},ij}) - N_{\text{out},ij}, \quad N_{\text{appl},ij} = N_{\text{fert},ij} + N_{\text{man},ij}.$$

B_{ij} is the net nitrogen entering the field from human (fertilizer and manure), biological (fixation), and atmospheric (deposition) sources that is not exported in grain or fiber, identical in construction to the N_{sur} indicator of X. Zhang et al. (2015). A positive B_{ij} is nitrogen lost to the environment together with any net accumulation in the soil organic pool; a negative B_{ij} marks a field that is, over the year, a net nitrogen sink relative to these inputs. Soil organic matter mineralization is absorbed into B rather than modelled separately, following the standard

soil-surface convention (Oenema, Kros, and De Vries 2003; Van Beek, Brouwer, and Oenema 2003) and X. Zhang et al. (2015) (Methods).

The panel is keyed on (county, crop class, year) and is *sparse*: only cells with an observed positive harvested area are materialized. Absent cells need no rows, because the decomposition reconstructs births and deaths from each cell’s presence across the year-pair being compared.

Computing B_{ij} requires a nitrogen value from the inventory, so the usable universe is bounded by inventory coverage. The inventory is a fixed-point sample, its survey locations fixed in space and re-observed every year, so whether a county belongs to the sampled frame is a property of the county rather than of the county–year. We therefore gate the panel on county-level frame membership: a positive-area cell is retained only if its county appears somewhere in the inventory. This drops 2,266 cells, 0.0185% of in-scope harvested area, all in counties the inventory never sampled and for which B cannot be computed. It deliberately keeps cells in surveyed counties that show no crop record in a given year, which a county-by-year gate would wrongly discard.

The inventory encodes “no application” by the absence of a row rather than an explicit zero. A retained cell with no matching row is therefore assigned $N_{\text{appl}} = 0$, which applies to 34% of panel cells. For the low-input crops, chiefly legumes, a true zero of synthetic input is correct. For some fertilized-crop cells the crop was grown and almost certainly fertilized but missed at the fixed sample points, so the zero understates the input and produces a spurious small negative B . These are dominated by wheat and the other small grains; corn is among the best-covered crops, with under 1% of its area affected. The cells are numerous but spatially slight — 34% of the panel by count, but only about 4% of harvested area, roughly two-thirds of it wheat. They are retained for coverage, and the artefact is quantified in the negative-balance audit below.

Harvested area also carries three states rather than two: positive, explicit zero, and missing. The decomposition treats missing cells as absent, so explicit zero and missing are not distin-

guished, and a missing-to-positive transition enters as a born cell even though missing denoted “unknown” rather than “zero”. Most surviving missing values are immaterial here, being either adjacent to a zero anchor, where a birth is legitimate anyway, or on the leading edge of a series before its first record — about 4,500 cells, under 0.1% of area — where the area genuinely is unknown. We accept this treatment and note the leading-edge births as a minor caveat.

Only the headline B is materialized in the panel. The sensitivity specifications — excluding belowground fixation, excluding deposition, or both — and the alternative nitrogen-input imputations are produced downstream by re-evaluating B from the component columns the panel retains, so the scaffold is assembled once and only this final layer varies.

5.2.18 Negative-balance audit

Under the headline specification B is positive across the great majority of harvested area, as expected of a nitrogen balance. Negative B touches about 11% of harvested area and 12% of production (Table S6). These negatives fall into two populations that must be read differently, separated by whether the inventory records a positive input for the cell.

In **positive-input** cells, B is computed from a real recorded input, so a negative value is a potentially interpretable signal. These supply about two-thirds of the negative area and are the subject of what follows. In **genuine-zero** cells — surveyed counties where the crop carried no matching row, so N_{appl} is set to zero (above) — the input side collapses to $B = N_{\text{dep}} - N_{\text{out}}$, which is negative almost by construction for any producing cell. Over 90% of genuine-zero area is negative for that mechanical reason. These are a survey-coverage artefact rather than an agronomic signal, and are excluded from the headline interpretation. That the negative production share slightly exceeds the negative area share confirms the negatives are not confined to marginal low-yield land.

Within the positive-input population, two further groups separate (Table S7). Legume cells are essentially never negative, because fixation supplies enough nitrogen that even high-yielding soybean fields stay net-positive — confirming that the BNF reconstruction does the job that motivated it. Among non-legume cells, negative B covers about 12% of area, concentrated in the low-input small grains and a few specialty crops.

Two diagnostics establish where these negatives come from: whether they reflect overstated removal or understated input, and whether they are merely the inventory’s own estimation uncertainty.

First, they are an input-side phenomenon. Splitting each non-legume crop’s positive-input cells by the sign of B (Table S8), the negative cells carry a median applied rate three- to tenfold below the positive cells of the same crop. The medians run corn at 60 against 220, wheat at 39 against 136, and oats at 17 against 204 kg N ha⁻¹. Their removal rate is comparable or, being modestly higher-yielding, slightly higher. Deposition adds a further 4–14 kg N ha⁻¹, near-identical across the two groups. The gap that pushes B negative is an implausibly low recorded input for the yields observed, not an inflated N_{out} . The removal coefficients are exonerated.

Second, they are structural rather than an artefact of that uncertainty. Recomputing B under each of the six imputed nitrogen realizations, 82% of the headline-negative area stays negative under all six. These negatives are a property of the recorded input level relative to harvest removal, stable across the inventory’s own uncertainty band.

What remains is not a correctable data error but two features of the input set. The structurally negative area concentrates precisely in the crops the inventory treats with coarser rates — barley and oats borrow wheat’s, and the specialty crops take top-down sales-balance averages rather than modelled rates (above) — which under-allocate nitrogen to the high-yielding tail of those crops. Beyond that, the balance is missing inputs by construction. A high-yielding small grain or root crop recorded at 30–60 kg N ha⁻¹ is drawing the remainder from mineralization and

from rotational or residue nitrogen, which the framework does not model and absorbs into B as soil-nitrogen mining.

Neither diagnostic can fully separate a cell where the sample points missed applied fertilizer from one genuinely mining soil nitrogen: both look like low recorded input with normal removal. The distinction does not change the treatment. Because the decomposition explains changes in B rather than its level, and accommodates negative and sign-changing cells, these negatives are retained rather than dropped.

Table S6: Negative field-nitrogen-balance ($B < 0$) footprint under the headline specification, decomposed into the positive-input and genuine-zero populations. The first two columns give the share of each population's own harvested area and production that is negative; the last column gives each population's share of the total negative area. Positive-input and genuine-zero cells partition all cells, so their contributions sum to 100%. The genuine-zero rate is high by construction ($B = N_{\text{dep}} - N_{\text{out}}$ with no recorded input) and is treated as a survey-coverage artifact, not an agronomic signal.

Cell population	Negative-B area (% of group)	Negative-B production (% of group)	Share of total negative area (%)
All crops, all cells	11.2	12.1	100.0
Positive input	7.8	9.4	66.6
Genuine-zero input (artifact)	90.5	95.0	33.4

Table S7: Share of cells and of harvested area with a negative field nitrogen balance ($B < 0$) under the headline specification, among cells with a positive recorded nitrogen input, by crop. Legume cells are essentially never negative because biological fixation offsets harvest removal; the non-legume negatives concentrate in low-input small grains and specialty crops and are interpreted as genuine soil-nitrogen mining plus the inventory's coarser treatment of unmodelled crops (see text). Cells with a genuine-zero input are excluded here and discussed separately.

Class	Crop	Cells (n)	% cells with $B < 0$	Panel area (Mha)	% area with $B < 0$
Legume	Soybean	47,986	0.0	958	0.0
Legume	Peanut	4,538	0.0	17	0.0
Legume	Pea	559	0.0	3	0.0
Legume	Lentil	132	0.0	1	0.0
Non-legume	Corn	59,630	6.5	1006	3.9
Non-legume	Wheat	42,809	22.5	533	25.3
Non-legume	Cotton	13,284	11.4	144	7.6

(continued)

Class	Crop	Cells (n)	% cells with $B < 0$	Panel area (Mha)	% area with $B < 0$
Non-legume	Sorghum	13,278	5.0	89	4.2
Non-legume	Barley	7,683	29.8	44	37.1
Non-legume	Rice	2,916	9.0	38	4.4
Non-legume	Sunflower	2,409	36.5	17	44.5
Non-legume	Oats	13,923	5.6	16	11.3
Non-legume	Sugar beet	2,652	33.5	16	37.1
Non-legume	Potato	3,327	20.8	12	28.5
Non-legume	Tomato	272	23.5	2	38.6
Non-legume	Tobacco	815	12.8	1	20.4
Non-legume	Sweet potato	271	31.0	0	48.5

Table S8: Median applied-nitrogen and harvest-removal rates (kg N ha^{-1}) for non-legume cells with a positive recorded input, split by the sign of the field nitrogen balance B . In every crop the negative cells carry a far lower applied (fertilizer-plus-manure) rate than the positive cells, while their removal rate is comparable or slightly higher — showing that negative B is driven by low recorded input rather than overstated removal. Atmospheric deposition (not shown) adds a further $4\text{--}14 \text{ kg N ha}^{-1}$, near-identical across the two groups.

Crop	Applied N (neg B)	Applied N (pos B)	Removal (neg B)	Removal (pos B)
Barley	31	138	64	56
Corn	60	220	103	96
Cotton	33	131	64	54
Oats	17	204	38	34
Potato	90	407	160	138
Rice	55	218	98	92
Sorghum	29	216	58	47
Sugar beet	70	200	123	105
Sunflower	28	86	57	47
Sweet potato	21	117	55	50
Tobacco	49	236	87	82
Tomato	113	335	205	180
Wheat	39	136	76	61

The negatives trace to recorded inputs that are low relative to harvest removal, so they could distort the factor effects. Four facts bound their influence.

First, the nitrogen input reaches four of the five effects only through the shared logarithmic-mean weight, never through their ratios, which come entirely from NASS area and production (Methods). The crop-mix effect is the one materially exposed, because that weight carries the

sign of B_{ij} . A bounding sensitivity that floors the small-grain balances at zero confirms this: the yield and NBI effects are essentially unchanged, and the whole shift lands on crop mix, concentrated in wheat.

Second, the weight in the NBI effect is the magnitude of each cell's balance, and it sits overwhelmingly in the cleanest crops. Corn, 3.9% of whose area is negative, and soybean, which is never negative, together carry about 72% of the aggregate balance. The crops with elevated negative rates carry only about 15%, most of it wheat, with the high-rate specialty crops below half a percent each. Cells near the break-even point, where the sign of B is least certain, carry small weight by construction, so the noise falls where it counts least.

Third, the period comparison examines the change in B between two averaged states, not its level in either, so a time-stable under-recording offset partially cancels. Averaging also suppresses the sign changes that would otherwise trigger the analytical-limit rule, for the reasons set out under the period-averaged estimator above. The effect is large: wheat's NBI contribution falls from about +0.57 Tg N in the chained run to about +0.02 Tg N in the period-averaged headline, leaving the headline NBI effect corn-dominated.

Fourth, an unknown share of the negatives is not error but the balance correctly registering mineralization and rotational nitrogen the inventory does not carry.

Two residual risks are tested rather than assumed. The under-recording is mildly yield-correlated, so the offset does not cancel perfectly. And the 2017 switch from NRI to CDL crop identities in the inventory is a potential confound; the terminal window spans both sides of it, diluting any level shift into the five-year average, but it remains a tested sensitivity for the chained trajectory. The headline effects are therefore assessed against a robustness suite, reported with the results, that drops the elevated-negative crops, substitutes each of the six nitrogen imputations, and varies the fixation and deposition specifications.

Two of those variants answer different questions, and conflating them would overstate what either shows. Dropping the elevated-negative crops tests the crop-structure narrative, not data quality. Because LMDI-I is consistent in aggregation, removing a crop simply subtracts its own NBI contribution and leaves every other crop's untouched, so the shift in NBI is a mechanical identity — expected for any large crop whatever its data quality, and largest for wheat because wheat is large. What that variant does inform is the share factors, where removing a crop reshuffles the area denominators and shows whether the crop-mix story depends on it. Data quality is instead answered by the six imputations, which vary only N_{appl} while retaining every crop and cell, so their spread bounds the imputation uncertainty directly.

For the chained trajectory, where wheat's NBI is large, we additionally split wheat's contribution by computational route. Year-pairs in which B keeps its sign flow through the smooth logarithmic-mean formula and carry a genuine intensity signal; pairs that cross zero or are born or die take the mechanical $\frac{1}{m}$ attribution. About 30% of wheat's harvested-area pairs take the mechanical route, and they carry roughly 35% of its chained NBI contribution — proportional to their share of pair area, not less. That component nets a large positive value rather than cancelling along the chain. The chained wheat signal is therefore about two-thirds genuine and one-third mechanically attributed, which is why the chained NBI is read for its sign alone. It also explains the collapse to +0.02 Tg N under the period-averaged headline, where the accumulation cannot occur.

Table S9 reports the imputation robustness of the canonical decomposition directly. Recomputing the period-averaged factor effects under each of the six nitrogen imputations moves every effect by roughly 1% of its value or less. The NASS-driven scale and yield effects are essentially fixed, and even the input-built crop-mix and NBI effects vary by at most about 1%. This is the input-uncertainty band summarised in the caption of the headline decomposition figure (Figure 2).

Table S9: Sensitivity of the period-averaged factor effects to nitrogen-input imputation. Each effect is recomputed under the six multiply-imputed realizations of the applied-nitrogen inventory, holding all else at the headline specification. ‘Headline’ is the effect under the imputation-averaged input used throughout; ‘Min’ and ‘Max’ bound it across the six imputations; ‘Spread’ is the half-range as a percentage of the headline magnitude. Base 1990–1994, terminal 2018–2022; values in Tg N.

Factor	Headline	Min	Max	Spread ($\pm\%$)
Harvested area (H)	-0.112	-0.112	-0.112	0.0
Geographic distribution (S_i)	-0.154	-0.154	-0.151	1.1
Crop mix (S_{ij})	+0.504	+0.499	+0.511	1.2
Yield (Y_{ij})	+1.987	+1.985	+1.987	0.1
Balance intensity (NBI_{ij})	-1.397	-1.406	-1.391	0.5
Net change (ΔB)	+0.827	+0.825	+0.829	0.2

5.3 Attribution tables

Factor contributions from the period-averaged single-step decomposition (base 1990–1994, terminal 2018–2022) broken down by Land Resource Region and by crop. Units are Tg N; positive values increase the aggregate nitrogen balance and negative values reduce it. The H factor is omitted from both tables. Its log-change is the same scalar in every cell, so breaking its effect out by region or crop would only redistribute the national scale effect in proportion to each group’s share of the log-mean weight. The crop table additionally omits S_i , which carries no crop index. Both effects enter the national totals reported with the results.

5.3.1 Geographic attribution by Land Resource Region

County-level factor contributions are apportioned to Land Resource Regions (LRRs) using synthetic fertilizer N-application weights from the nitrogen application inventory: for counties that span multiple LRRs, each LRR receives the fraction of the county’s contribution proportional to its share of that county’s total nitrogen application across all years and crops. Ten of the 2,783 counties in the decomposition straddle two LRRs and receive split weights; the rest are assigned entirely to one LRR.

Table S10: Factor contributions to the change in aggregate N balance (Tg N) by Land Resource Region, period-averaged single-step decomposition. LRR codes and names follow the NRCS MLRA 52 (2022) delineation. The Total row reproduces the national factor totals.

LRR	Region name	ΔB_{S_i}	$\Delta B_{S_{ij}}$	$\Delta B_{Y_{ij}}$	$\Delta B_{NBI_{ij}}$	Total
A	Northwestern Forest, Forage, and Specialty Crop Region	-0.003	-0.003	0.000	0.000	-0.006
B	Northwestern Wheat and Range Region	-0.016	0.020	0.019	-0.026	-0.003
C	California Subtropical Fruit, Truck, and Specialty Crop Region	-0.091	-0.023	0.006	0.050	-0.057
D	Western Range and Irrigated Region	-0.057	-0.001	0.008	0.006	-0.043
E	Rocky Mountain Range and Forest Region	-0.003	0.005	0.003	0.007	0.012
F	Northern Great Plains Spring Wheat Region	0.012	0.218	0.083	0.060	0.372
G	Western Great Plains Range and Irrigated Region	-0.017	0.037	0.014	0.044	0.078
H	Central Great Plains Winter Wheat and Range Region	-0.117	0.125	0.124	0.025	0.157
I	Southwest Plateaus and Plains Range and Cotton Region	-0.034	-0.012	0.005	0.005	-0.035
J	Southwestern Prairies Cotton and Forage Region	-0.038	-0.013	0.004	-0.008	-0.054
K	Northern Lake States Forest and Forage Region	0.041	-0.015	0.056	-0.052	0.030
L	Lake States Fruit, Truck Crop, and Dairy Region	0.022	-0.016	0.140	-0.099	0.047
M	Central Feed Grains and Livestock Region	0.186	0.133	1.019	-0.845	0.492
N	East and Central Farming and Forest Region	0.046	0.003	0.096	-0.133	0.012

(continued)

LRR	Region name	ΔB_{S_i}	$\Delta B_{S_{ij}}$	$\Delta B_{Y_{ij}}$	$\Delta B_{NBI_{ij}}$	Total
O	Mississippi Delta Cotton and Feed Grains Region	-0.036	0.024	0.122	-0.082	0.029
P	South Atlantic and Gulf Slope Cash Crops, Forest, and Livestock Region	-0.028	0.021	0.165	-0.249	-0.091
R	Northeastern Forage and Forest Region	0.015	-0.002	0.030	-0.023	0.020
S	Northern Atlantic Slope Diversified Farming Region	-0.003	0.000	0.032	-0.006	0.022
T	Atlantic and Gulf Coast Lowland Forest and Crop Region	-0.035	-0.001	0.055	-0.074	-0.055
U	Florida Subtropical Fruit, Truck Crop, and Range Region	0.002	0.003	0.004	0.003	0.013
Total		-0.154	0.504	1.987	-1.397	0.939

5.3.2 Crop attribution

Table S11: Factor contributions to the change in aggregate N balance (Tg N) by crop, period-averaged single-step decomposition. Wheat is retained as separate winter and spring classes; all other crops are aggregated to a single class. Crops sorted by absolute yield contribution. The S_i and H factors are omitted, for the reasons given at the start of this section. The Total row reproduces the national factor totals for the three factors shown.

Crop	$\Delta B_{S_{ij}}$	$\Delta B_{Y_{ij}}$	$\Delta B_{NBI_{ij}}$	Total
Corn	0.284	1.137	-1.249	0.172
Soybean	0.538	0.638	0.017	1.193
Cotton	-0.046	0.104	0.011	0.069
Rice	0.015	0.042	-0.052	0.006
Winter wheat	-0.092	0.032	-0.106	-0.167
Peanut	-0.004	0.019	-0.012	0.003
Pea	0.007	0.007	0.007	0.021

(continued)

Crop	$\Delta B_{S_{ij}}$	$\Delta B_{Y_{ij}}$	$\Delta B_{\text{NBI}_{ij}}$	Total
Sorghum	-0.105	-0.006	-0.048	-0.159
Potato	-0.002	0.006	-0.011	-0.007
Tomato	0.008	0.004	-0.010	0.002
Sugar beet	-0.002	0.004	-0.026	-0.025
Lentil	0.002	0.002	0.002	0.007
Oats	-0.078	-0.001	-0.007	-0.087
Spring wheat	-0.009	0.001	0.104	0.096
Sweet potato	0.000	0.000	-0.001	-0.001
Tobacco	0.000	0.000	0.000	0.000
Sunflower	-0.001	0.000	-0.012	-0.013
Barley	-0.011	0.000	-0.003	-0.015
Total	0.504	1.987	-1.397	1.093

5.4 Conventional NUE versus within-field efficiency

The main text reports that the apparent change in aggregate nitrogen-use efficiency (NUE) over the study period overstates the genuine field-level efficiency gain, because a substantial part of it is compositional rather than a change in field management. This section gives the method and the full result.

For one crop in one county, NBI and NUE are the same measurement on different scales: $\text{NBI}_{ij} = n_j(1/\text{NUE}_{ij} - 1)$, one-to-one and strictly decreasing in NUE_{ij} (Methods). The scaling n_j is what separates them once cells are combined, and it varies widely across the panel (Table S4): more than fivefold among the grains and oilseeds, and over thirtyfold once the high-moisture crops, reported on a fresh-weight basis, are included. Two cells at the same NUE therefore do not leave the same surplus. At any common NUE, a kilogram of soybeans ($n_j = 0.055$) leaves 4.6 times the surplus of a kilogram of corn ($n_j = 0.012$), because producing it runs 4.6 times as much nitrogen through the field. NUE reports the conversion rate and is silent on the throughput.

Aggregate NUE is the harvest nitrogen removed per unit of nitrogen applied, summed over all cells, $\text{NUE} = \sum_{ij} N_{\text{out},ij} / \sum_{ij} N_{\text{in},ij}$. Writing $\phi_{ij} = N_{\text{in},ij} / \sum_{ij} N_{\text{in},ij}$ for a cell's share of total input and $\text{nue}_{ij} = N_{\text{out},ij} / N_{\text{in},ij}$ for its own efficiency, aggregate NUE is exactly an input-weighted mean of cell efficiencies,

$$\text{NUE} = \sum_{ij} \phi_{ij} \text{nue}_{ij},$$

structurally identical to the $B = \sum_{ij} P_{ij} \text{NBI}_{ij}$ form that underlies the balance decomposition. The same additive LMDI-I then splits the change ΔNUE into an **intensity effect** — the composition-free change in within-cell efficiency — and a **structural effect** carrying only the shift in input weights ϕ_{ij} across crops and counties:

$$\Delta \text{NUE} = \underbrace{\sum_{ij} L(v_{ij}^T, v_{ij}^0) \ln \frac{\text{nue}_{ij}^T}{\text{nue}_{ij}^0}}_{\text{intensity (within-cell)}} + \underbrace{\sum_{ij} L(v_{ij}^T, v_{ij}^0) \ln \frac{\phi_{ij}^T}{\phi_{ij}^0}}_{\text{structural (composition)}},$$

with $v_{ij} = \phi_{ij} \text{nue}_{ij}$ and $L(\cdot, \cdot)$ the logarithmic mean used throughout. The structural effect is the composition the main text refers to: the part of an apparent aggregate-NUE change that a naive analyst would book as efficiency but that reflects only which crops and regions grew. The endpoints and cell universe are identical to the canonical period-averaged single-step decomposition (base 1990–1994, terminal 2018–2022; finite balance, positive area). The exact split is computed on cells present and strictly positive in both windows; the small uncovered remainder (births, deaths, zero-output cells) is itself compositional, so the matched split is a conservative estimate of the structural share.

Table S12 reports the result for two input definitions and, for symmetry, the same split applied to aggregate NBI. Under the study's full nitrogen budget (the definition used for the balance, matching X. Zhang et al. (2015)), aggregate NUE rises over the period, but roughly a third of

that rise is structural — composition inflates the genuine within-field gain by nearly half. The composition effect is larger, not smaller, under the conventional farm-gate definition (synthetic plus manure nitrogen only). There the within-cell efficiency change is actually negative, and the entire apparent improvement is manufactured by the system’s shift toward biologically-fixing legumes, whose near-zero applied-nitrogen denominator gives them very high farm-gate NUE. That figure is partly artefactual, since fixation is excluded from the denominator. Applied-only NUE is in this sense mis-specified for a legume-bearing system, and is undefined for the legume cells that receive no synthetic or manure nitrogen (hence the slightly lower coverage). The symmetric NBI row is shown only to make the contrast explicit and should be read with care: treating aggregate NBI as a single output-weighted ratio forces the exclusion of all sign-changing and net-deficit cells (the lower coverage), so it is not the composition-free NBI effect of the decomposition. That effect — which handles signed and zero cells exactly through the analytical-limit rule and is the paper’s actual efficiency measure — is the -1.40 Tg N reported in the main text, compositional by construction in none of its parts.

Table S12: Decomposition of the change in aggregate efficiency into within-cell (intensity) and compositional (structural) effects, period-averaged single-step (base 1990–1994, terminal 2018–2022). NUE is reported for the study’s full nitrogen budget and for the conventional farm-gate definition (synthetic + manure only). The NBI row applies the same output-weighted ratio split for symmetry; it excludes sign-changing and net-deficit cells (note the lower coverage) and is therefore not the composition-free NBI effect of the main decomposition. ‘Comp.’ is the structural (compositional) effect and ‘Comp. %’ is that effect as a share of the within-cell + structural total; ‘Cov.’ is the share of the base-period denominator on matched cells. Within-cell and composition sum to the change over matched cells, which differs from the naive Δ only by the contribution of births and deaths (themselves compositional).

Metric	Base	Term.	Δ (naive)	Within	Comp.	Comp. %	Cells	Cov.
NUE, full N budget	0.5673	0.6418	0.0745	0.0505	0.0225	31	8,425	100%
NUE, applied N only	0.8664	1.1771	0.3107	-0.1157	0.4130	139	6,244	96.8%
NBI, balance/output	0.0144	0.0114	-0.0030	-0.0027	-0.0002	6	5,339	82.3%