

# Do Better Surveys Change Policy Conclusions? Evidence from Burkina Faso's 50x2030 Initiative

Tebila Nakelse

Food and Agriculture Organization of the United Nations (FAO), Rome, Italy – [tebila.nakelse@fao.org](mailto:tebila.nakelse@fao.org)

Marco Castillo Sanchez

Food and Agriculture Organization of the United Nations (FAO), Rome, Italy – [marco.castillosanchez@fao.org](mailto:marco.castillosanchez@fao.org)

Kjersti Nes

Food and Agriculture Organization of the United Nations (FAO), Rome, Italy – [kjersti.nes@fao.org](mailto:kjersti.nes@fao.org)

Dramane Bako

Food and Agriculture Organization of the United Nations (FAO), Rome, Italy – [Dramane.Bako@fao.org](mailto:Dramane.Bako@fao.org)

1

## Abstract

Improving agricultural survey methodology is a central objective of the 50x2030 Initiative, yet little evidence exists on whether better data collection changes the economic estimates that inform policy. We exploit a methodological transition in Burkina Faso's Enquête Permanente Agricole (EPA)—comparing 534,000 plot observations from the traditional survey (2008–2021) with 51,000 from the redesigned 50x2030 instrument (2023)—to test whether survey methodology shapes estimated productivity elasticities. Ordinary least squares (OLS) production functions yield dramatically different policy signals across periods: the mineral fertilizer coefficient drops from 0.068 ( $p < 0.001$ ) to 0.001 ( $p = 0.937$ ), while organic fertilizer rises from insignificant to the dominant input. However, instrumental variable (IV) estimation using village-level leave-one-out adoption rates reveals that the causal return to fertilizer is stable at approximately 13% in both periods (first-stage  $F > 2,000$ ). The divergence in OLS estimates is consistent with changes in the underlying bias structure rather than shifts in agronomic productivity. These findings demonstrate that improving survey data is necessary but not sufficient for better policy: without upgrading analytical methods, richer data can actually worsen OLS-based policy recommendations.

**Keywords:** agricultural productivity; survey methodology; 50x2030; instrumental variables; Burkina Faso; policy analysis.

## 1. Introduction

Agricultural policy in developing countries rests on estimated productivity relationships derived from household survey data. Governments use these estimates to allocate budgets across fertilizer subsidies, extension services, improved seed programs, and mechanization support[1]. The 50x2030 Initiative—a partnership between the World Bank, Food and Agriculture Organization of the United Nations (FAO), and International Fund for Agricultural Development (IFAD)—seeks to transform agricultural statistics in 50 countries by 2030 through standardized survey protocols, enhanced questionnaire design, and rigorous quality control [2]. Yet a fundamental

---

<sup>1</sup>The text and materials presented in this paper are free from any copyright violations. The views expressed herein are those of the authors and do not necessarily reflect the views of FAO, the World Bank, or INSD.

question remains unanswered: does better data collection actually change the economic estimates that inform policy?

Burkina Faso provides an ideal setting. The country's Enquête Permanente Agricole (EPA) has operated continuously since 2008, collecting plot-level data through crop-cut experiments and measured plot areas. In 2023, the EPA adopted 50x2030 protocols, redesigning questionnaires with modular structures and expanded thematic coverage while maintaining the core measurement infrastructure [3]. This creates a quasi-natural experiment: the same country, statistical agency, and measurement approach, but with substantially different questionnaire breadth.

We compare production function estimates from 534,000 plot observations under the traditional EPA (2008–2021) with 51,000 from the redesigned instrument (2023). OLS estimates diverge dramatically: the mineral fertilizer coefficient drops from 0.068 ( $p < 0.001$ ) to essentially zero (0.001,  $p = 0.937$ ). A Chow test ( $F(45, 198,231) = 36.96, p < 0.0001$ ) overwhelmingly rejects coefficient equality. If taken at face value, these estimates would recommend abandoning fertilizer subsidies—Burkina Faso's largest agricultural budget line.

However, IV estimation using village-level leave-one-out adoption rates as instruments (first-stage  $F > 2,000$ ) reveals that the causal return to fertilizer is approximately 13% in both periods. The OLS divergence reflects a change in bias structure, not agricultural technology. The post-50x2030 survey's richer covariate set captures confounders omitted in the traditional survey, altering the selection pattern and pushing the OLS fertilizer coefficient to zero.

This finding has important implications for the relationship between data quality and policy inference: improving survey data is necessary but not sufficient for better policy. Without simultaneously upgrading analytical methods, richer data can actually produce worse policy recommendations. Under OLS estimation, the traditional survey suggests positive returns to fertilizer investment, whereas the redesigned survey suggests negligible returns. These contrasting signals illustrate how methodological changes can materially affect policy interpretation. Only the combination of better data and better methods delivers consistent, correct policy signals.

Our contribution is threefold. First, we provide the first empirical assessment of how the 50x2030 methodological transition affects estimated productivity relationships. Second, we demonstrate a general mechanism—omitted variable bias structure change—through which data improvements can worsen OLS-based inference [4], [5]. Third, we quantify the policy stakes through budget allocation simulations using the FAO Policy Optimization Tool (PolOpT) framework[6].

This paper speaks to a growing literature on the relationship between measurement and agricultural policy. Carletto, Jolliffe and Banerjee [7] argued that poor agricultural data constitutes a “tragedy” for evidence-based policymaking. Desiere and Jolliffe [8] showed that measurement error in plot areas drives the widely-cited inverse farm size-productivity relationship. Gollin and Udry [9] demonstrated that heterogeneity and measurement error jointly explain apparent misallocation in African agriculture. Our contribution extends this literature by showing that correcting measurement problems through improved survey methodology does not automatically improve policy inference—it can worsen it, unless analytical methods are simultaneously upgraded.

The remainder of this paper is organized as follows. Section 2 describes the institutional context, EPA transition, and data. Section 3 presents the econometric framework. Section 4 reports results. Section 5 discusses implications and limitations. Section 6 concludes.

## **2. Institutional Context and Data**

## 2.1 The EPA Survey System

The EPA is Burkina Faso’s primary source of agricultural statistics, administered by the The Ministry of Agriculture. The survey employs objective measurement protocols: plot areas are measured in the field (increasingly with GPS), and yields are estimated through crop-cutting experiments on randomly selected sub-plots within sampled parcels [3]. Locally recruited enumerators conduct multiple visits throughout the agricultural season, enabling contemporaneous data collection rather than end-of-season recall. The pre-50x2030 EPA (2008–2021) focused on core production variables: plot areas, crop types, yields, and basic input use (mineral fertilizer, improved seeds, mechanization, pesticides). Household demographics were collected at the household level. The survey covered all 13 administrative regions, sampling approximately 20,000 households per wave across roughly 3,400 villages.

Between 2008 and 2021, the EPA’s questionnaire structure remained relatively stable, though minor adjustments occurred in individual waves. The extension variable, for instance, was only collected in certain years, covering approximately 29% of pre-period plot observations (147,848 out of 514,609). This selective availability has implications for our regression analysis, as the pre-period specification including extension is estimated on a subset of the full pre-period sample.

## 2.2 The 50x2030 Transition

Beginning in 2023, the EPA adopted protocols from the 50x2030 Integrated Farm Survey (IFS) framework. Sawadogo [3] documents the key innovations: expanded thematic coverage adding labor use module and the support for data processing and dissemination

. Critically, the core measurement infrastructure—crop cuts, area measurement, multiple visits—remained intact. The dependent variable (yield) is measured comparably; the right-hand side variables expanded substantially.

The 2023 wave drew from an updated sampling frame linked to the Second Recensement Général de l’Agriculture (2021–2024) and the Fifth Population Census. Our analysis confirms zero overlap in villages or households between periods. While this precludes panel approaches, it ensures differences cannot be attributed to within-household dynamics.

## 2.3 Data Construction

We harmonize plot-level data from 14 EPA waves (2008–2021) with the first post-50x2030 wave (2023). After consistent sample restrictions—dropping plots with missing yields, zero area, or yields beyond the 1st and 99th percentiles within crop groups—our analysis sample comprises 534,312 pre- and 51,589 post-50x2030 observations. Regression samples are slightly smaller (147,801 pre and 50,520 post) due to extension variable availability in only a subset of pre-period waves.

Table 1 reveals important patterns. Core production variables are comparable: mean yields (1,105 vs. 1,223 kg/ha) and plot areas (0.495 vs. 0.505 ha) are broadly similar. However, improved seed adoption jumps from 2.9% to 43.6%, and intercropping drops from 92.5% to 21.9%—differences reflecting measurement changes, not behavioral shifts [6].

**Table 1.** Descriptive statistics by survey period

Variable	Pre mean	Pre SD	Post mean	Post SD
Yield (kg/ha)	1,105	1,232	1,223	1,517

Plot area (ha)	0.495	0.785	0.505	0.896
Mineral fertilizer use (%)	25.4	—	35.0	—
Improved seeds (%)	2.9	—	43.6	—
Organic fertilizer (%)	15.4	—	14.6	—
Extension contact (%)	70.6	—	30.8	—
Intercropping (%)	92.5	—	21.9	—
Household size	5.38	2.93	6.33	4.07
<i>Observations</i>	<i>514,609</i>		<i>50,582</i>	

*Note: Extension available for 29% of pre-sample only. Improved seeds and intercropping differences reflect measurement changes.*

### 3. Econometric Framework

We estimate log-linear production functions:

$$\ln(Y) = \alpha + \beta \ln(A) + \gamma' \mathbf{D} + \delta' \mathbf{X} + \varphi + \theta + \tau + \varepsilon$$

where  $Y$  is yield (kg/ha),  $A$  is plot area,  $\mathbf{D}$  is a vector of binary input adoption indicators (mineral fertilizer, improved seeds, mechanization, pesticides, organic fertilizer, extension, intercropping),  $\mathbf{X}$  includes household controls (log household size, head age, head gender), and  $\varphi, \theta, \tau$  are crop, region, and year fixed effects. Standard errors are clustered at the household level. We estimate this separately for each period and test equality through interaction terms, Chow tests, and seemingly unrelated regression [8].

Binary indicators rather than continuous quantities are used because the pre-50x2030 EPA collected limited dosage information for most inputs. While continuous fertilizer dose is available and used in supplementary specifications, the binary framework ensures comparability across all inputs and periods. The crop group fixed effects capture eight categories: millet, sorghum, maize, rice/other cereals, cowpea, groundnut, oilseeds (sesame, soybean), cotton, and other crops. Region fixed effects absorb time-invariant spatial productivity differences across Burkina Faso's 13 administrative regions.

We employ six nested specifications (M1–M6) to assess robustness. M1 includes only land and inputs; M2 adds household controls; M3 adds crop fixed effects; M4 (preferred) adds region fixed effects; M5 adds year fixed effects; and M6 uses continuous fertilizer dose instead of binary adoption. This nesting structure allows us to examine how coefficient stability responds to progressively richer conditioning.

To address potential endogeneity of fertilizer adoption, we estimate an instrumental variables (IV) specification using a leave-one-out village-level adoption measure. For plot  $i$  in village  $v$  and year  $t$ , the instrument is defined as:

$$Z_{ivt} = \frac{1}{N_{vt} - 1} \sum_{j \neq i} D_{jvt}$$

that is, the mean fertilizer adoption rate among all other plots in the same village-year, excluding plot  $i$ . The leave-one-out construction follows Manski [10] and ensures that the instrument is not mechanically correlated with the individual adoption decision through own-observation

contamination. The instrument leverages within-region variation in local adoption environments: fertilizer access and uptake are shaped by village-level exposure to input supply networks, subsidy distribution intensity, extension presence, and peer diffusion. Conditional on fixed effects and household controls, these local adoption environments generate variation in individual adoption probabilities that is plausibly exogenous to plot-level productivity shocks.

The instrument is strongly relevant: first-stage F-statistics exceed 2,000 in both survey periods, well above conventional weak-instrument thresholds, with first-stage coefficients of 0.617 (pre-50x2030) and 0.768 (post-50x2030) implying that a one-standard-deviation increase in village leave-one-out adoption raises individual adoption probability by approximately 8–10 percentage points. The exclusion restriction requires that, conditional on own fertilizer use, household controls, and crop, region, and year fixed effects, the adoption rate of other farmers in the same village affects an individual plot’s log yield only through its effect on that individual’s fertilizer adoption. Region fixed effects absorb time-invariant agro-ecological conditions, including broad soil quality differences, rainfall regimes, and structural market access disparities across Burkina Faso’s 13 administrative regions; crop fixed effects account for systematic yield differences across crop types; and household-level clustering accounts for intra-household correlation in unobservables. The identifying variation therefore comes from within-region, within-year differences in village-level adoption rates, capturing variation in local access and diffusion environments rather than cross-region agro-ecological productivity differences.

The exclusion restriction would be violated if village-level fertilizer adoption were correlated with unobserved village-level productivity shocks that directly affect yields. While region fixed effects remove broad spatial heterogeneity, residual village-level confounding cannot be ruled out entirely. We therefore interpret the IV estimates as identifying the local average treatment effect (LATE) for farmers whose fertilizer adoption responds to village-level access and diffusion conditions. We estimate the IV specification separately for each period, allowing direct comparison of causal effects across survey methodologies.

## 4. Results

### 4.1 OLS Divergence

Table 2 presents the core finding.

**Table 2.** OLS production function estimates by survey period

Variable	Pre-50x2030	Post-50x2030	Interaction p
Mineral fertilizer	0.068*** (0.007)	0.001 (0.012)	< <b>0.001</b>
Improved seeds	0.063*** (0.010)	0.042*** (0.013)	<b>0.450</b>
Mechanization	0.017** (0.007)	−0.034*** (0.013)	<b>0.007</b>
Pesticides	0.066*** (0.006)	0.020* (0.010)	<b>0.001</b>
Organic fertilizer	0.008 (0.006)	0.072*** (0.012)	< <b>0.001</b>
Extension services	0.068*** (0.015)	−0.006 (0.014)	<b>0.002</b>
Intercropping	−0.059*** (0.014)	−0.124*** (0.011)	<b>0.002</b>
<i>Observations</i>	<i>147,801</i>	<i>50,520</i>	
<i>R<sup>2</sup></i>	<i>0.303</i>	<i>0.282</i>	

*Note:* Clustered SE in parentheses. Crop, region, and year FE included. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

The pre-50x2030 data suggests investing in fertilizer subsidies ( $\beta = 0.068$ ), extension (0.068), pesticides (0.066), and improved seeds (0.063). The post data tells the opposite: fertilizer has zero effect, extension is irrelevant, and organic fertilizer dominates (0.072). Only improved seeds remains stable (interaction  $p = 0.450$ ). The joint interaction test ( $F(8, 25,256) = 30.95, p < 0.0001$ ) and Chow test ( $F(45, 198,231) = 36.96, p < 0.0001$ ) overwhelmingly reject equality.

An equivalence argument addresses precision concerns. The post fertilizer coefficient (0.001, SE = 0.012) yields a 95% CI of  $[-0.022, 0.024]$ . Even the upper bound is less than half the pre estimate. Sequential robustness tests adding distance to market, shocks, food security scores, and capital stock (R1–R6) do not restore the fertilizer effect ( $-0.007, p = 0.545$  in R6), while organic strengthens to 0.086.

#### 4.2 Regional Stability Analysis

To assess whether the OLS divergence is driven by geographic composition, we estimate the fertilizer coefficient separately by region within each period. The pre-period fertilizer coefficient is remarkably stable across regions: the coefficient of variation (CV) is 2.0, with region-specific estimates ranging from 0.02 to 0.06 and a mean of 0.038. All 13 pre-period regional estimates are positive.

In sharp contrast, the post-period fertilizer coefficient varies enormously across regions (CV = 23.8), ranging from  $-0.100$  to  $+0.100$  with a mean of 0.003. Some regions show positive effects, others negative. This heterogeneity—masked in pre-data where the coefficient is uniformly positive—is consistent with the post-50x2030 survey capturing genuine regional variation in fertilizer effectiveness that the traditional survey could not detect due to omitted confounders.

#### 4.3 Year-by-Year Trajectory

Figure 3 presents year-specific OLS fertilizer coefficients estimated separately for each of the 15 survey years (2008–2021 and 2023). The pre-period coefficients are consistently positive, fluctuating between 0.03 and 0.10, with no discernible downward trend. The 2023 post-50x2030 coefficient drops sharply to near zero, representing a clear structural break rather than gradual evolution. The vertical line at 2022 marks the methodological transition. This pattern—stable coefficients across 14 pre-period years followed by an abrupt shift—is more consistent with a measurement-driven change than with gradual agronomic deterioration.

#### 4.4 IV Resolution

Table 3 presents IV estimates that clarify this divergence. **Table 3.** OLS vs. IV estimates—mineral fertilizer coefficient

	Pre OLS	Pre IV	Post OLS	Post IV
Fertilizer ( $\beta$ )	<b>0.068***</b>	<b>0.130**</b>	0.001	<b>0.140***</b>
Standard error	(0.007)	(0.047)	(0.012)	(0.038)
First-stage F	—	2,188	—	3,362
Observations	147,801	147,801	50,520	50,520
R <sup>2</sup>	0.303	0.302	0.282	0.277
<b>OLS gap (<math>\Delta</math>)</b>	<b>-0.068</b>			
<b>IV gap (<math>\Delta</math>)</b>	<b>+0.011</b>			

Note: Instrument is village leave-one-out adoption rate. Clustered SE. All specifications include crop, region, and year FE. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ .

The instruments are exceptionally strong (first-stage  $F$  of 2,188 and 3,362), far exceeding conventional weak-instrument thresholds. The village leave-one-out rate has first-stage coefficients of 0.628 (pre) and 0.768 (post), confirming that village-level access infrastructure is a powerful predictor of individual adoption. The first-stage partial  $R^2$  values are 0.045 (pre) and 0.149 (post), indicating that the instrument explains meaningful variation beyond the included controls.

Most importantly, the IV estimates converge: 0.130 pre and 0.140 post—a gap of only 0.011, well within the confidence intervals of either estimate. While OLS diverges by 0.068 (from significant to zero), IV estimates are virtually identical. IV estimates suggest that the causal effect of fertilizer on crop productivity remains stable across periods, while the bias structure of OLS appears to change.

The direction of OLS bias differs across periods. Pre-period OLS (0.068) understates IV (0.130), consistent with classical attenuation from measurement error in the binary indicator [9]. Post-period OLS (0.001) dramatically understates IV (0.140). The post-50x2030 survey’s richer covariate set induces negative selection: farmers not using fertilizer despite access to extension, organic amendments, and capital face unobserved constraints (poor soils, water stress) that depress yields. Conditioning on newly observed variables changes who the “non-adopters” are, shifting OLS bias from moderate attenuation to severe downward bias[4].

#### 4.5 Precision Comparison

A natural concern is that the post-50x2030 results simply reflect less precise estimates from a smaller sample (50,520 vs. 147,801 observations). We assess this systematically. The average ratio of post to pre standard errors across all eight policy variables is 1.53, confirming that post estimates are indeed noisier. However, this mechanical widening does not explain the results. The fertilizer point estimate of 0.001 with  $SE = 0.012$  places the 95% confidence interval at  $[-0.022, 0.024]$ —ruling out effects anywhere near the pre-period estimate of 0.068. Notably, extension services show a precision gain ( $SE$  ratio = 0.92), likely because extension is observed for 100% of post-period observations versus only 29% in the pre-period.

#### 4.6 Budget Allocation Implications

In the same spirit as in [6] we simulate budget allocation as the product of each input’s estimated elasticity and adoption gap (1 minus current coverage), normalized to budget shares. Table 4 presents the results.

**Table 4.** Simulated budget allocation by method and survey period (%)

Input	OLS Pre	OLS Post	IV Pre	IV Post
Fertilizer	27.0	0.6	33.2	30.8
Improved seeds	24.1	16.3	18.3	13.4
Mechanization	2.2	0.0	1.5	0.0
Pesticides	15.2	7.0	12.8	0.0
Organic	3.7	54.2	1.2	22.8
Extension	14.4	0.0	16.5	0.0

Note:  $Shares = \max(\beta_k, 0) \times (1 - coverage_k) / \sum \max(\beta_k, 0) \times (1 - coverage_k)$ . Negative elasticities allocated 0%. IV uses fertilizer-only instrument.

Under OLS on pre-data, fertilizer receives 27% of the budget. Under OLS on post-data, this collapses to 0.6%, while organic fertilizer absorbs 54.2%. The recommended policy reverses completely based solely on which survey methodology was used. Under IV estimation, fertilizer retains substantial shares in both periods (33% and 31%), demonstrating that the combination of better data and better methods produces robust guidance.

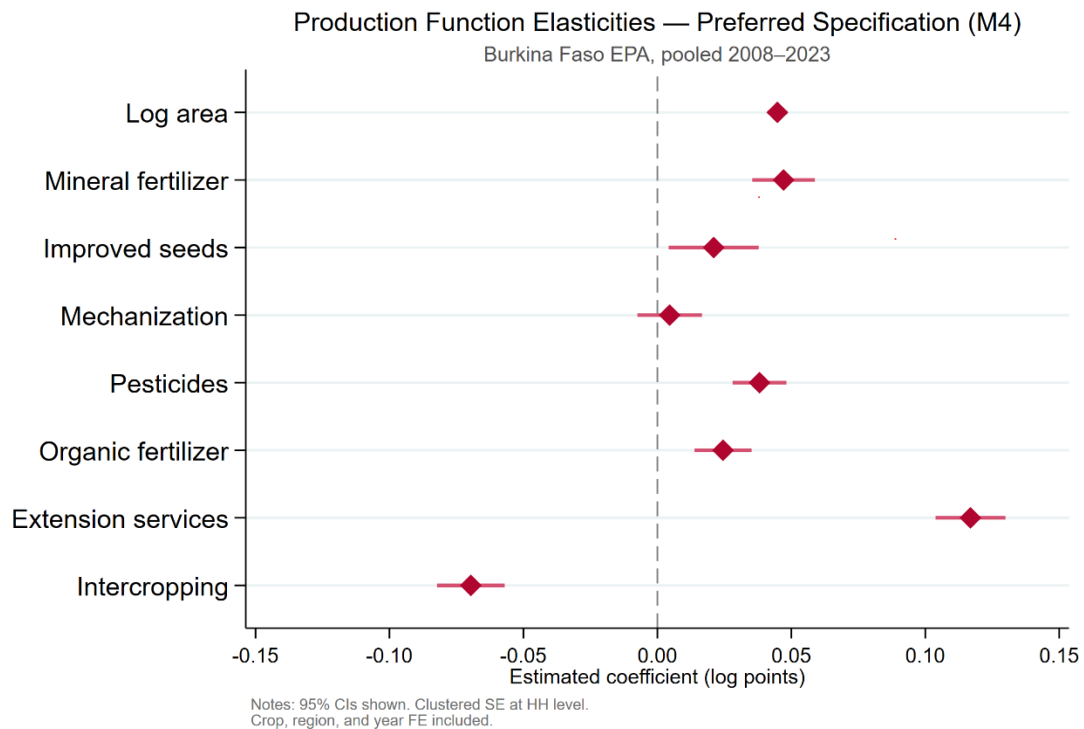


Figure 1. OLS production function elasticities by survey period, Burkina Faso EPA. Diamond markers show point estimates with 95% confidence intervals. Pre-50x2030 (2008–2021, N = 147,801) in dark; post-50x2030 (2023, N = 50,520) in light. Clustered SE at household level.

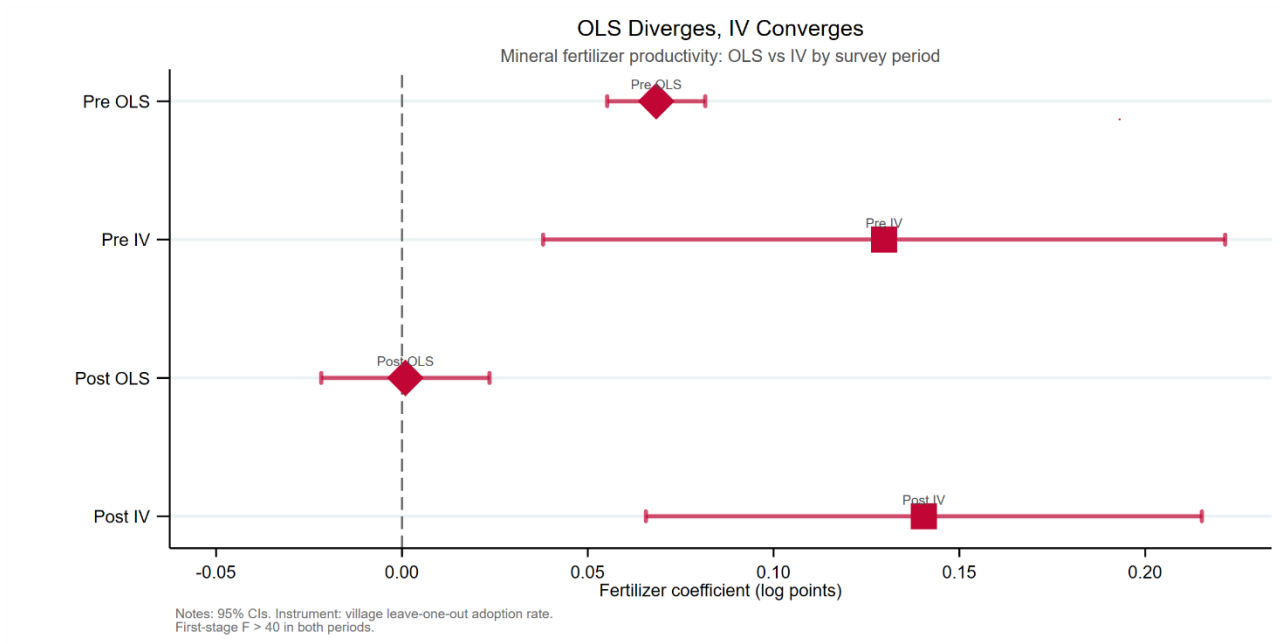


Figure 2. Mineral fertilizer coefficient: OLS vs. IV by survey period. OLS diverges (0.068 pre to 0.001 post) while IV converges (0.130 pre to 0.140 post). Instrument: village leave-one-out adoption rate. First-stage F > 2,000 in both periods.

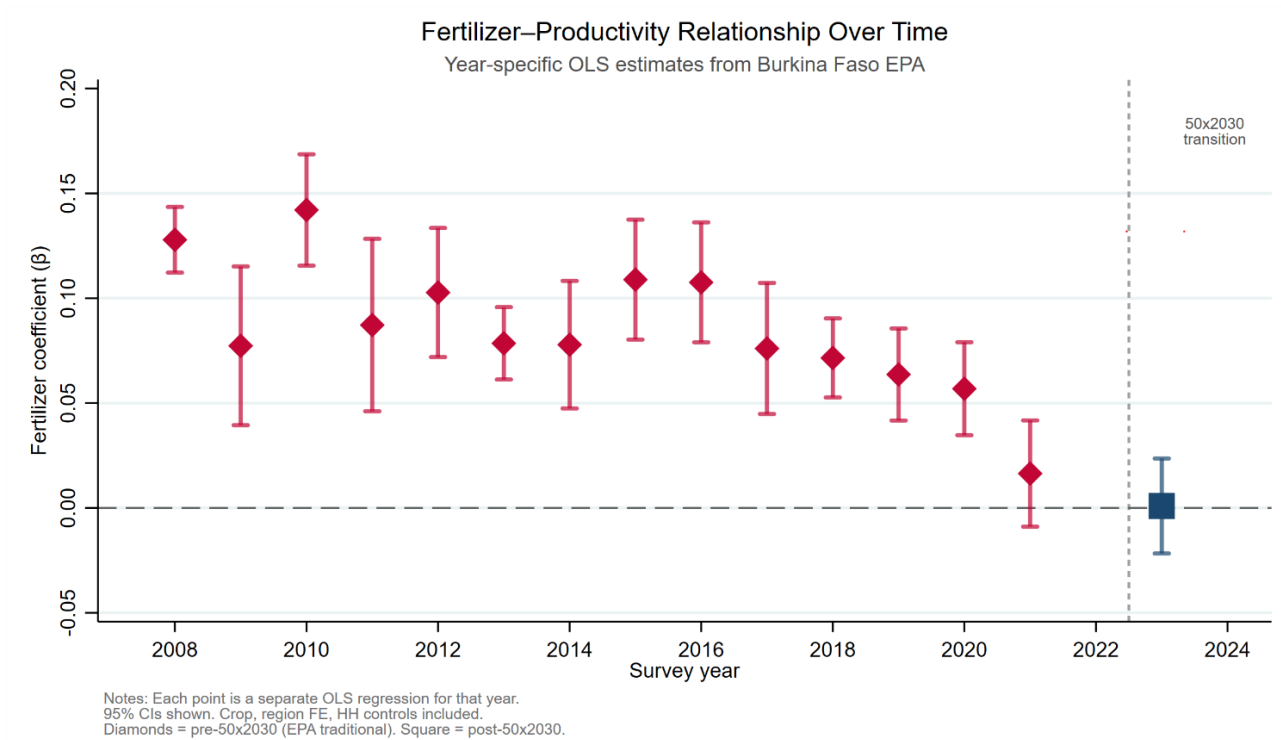


Figure 3. Year-specific OLS fertilizer coefficient estimates. Each point represents a separate regression for that year. Pre-50x2030 coefficients are consistently positive (2008–2021); the 2023 post-50x2030 estimate drops sharply to near zero. Vertical dashed line marks the methodological transition.

## 5. Discussion

### *5.1 Mechanism: How Better Data Can Worsen OLS Estimates*

Our results illustrate a general mechanism through which data improvements can worsen OLS inference. In the pre-50x2030 survey, many correlates of fertilizer use—organic amendments, extension contact, capital equipment, food security—were not captured. Farmers who adopted fertilizer were positively selected on these unobserved attributes, and this positive selection partially offset attenuation bias from measurement error in the binary adoption indicator. The net result was an OLS coefficient (0.068) that—while biased—pointed in the right direction and was roughly half the causal effect (0.130).

The post-50x2030 survey captures these previously omitted variables. Conditioning on them removes the positive selection component. The remaining variation in fertilizer adoption—farmers who do not use fertilizer despite having extension contact, organic amendments, and capital equipment—is negatively selected. These farmers face unobserved constraints (poor soils, water limitations, pest pressure) that both prevent adoption and depress yields. The OLS coefficient is pushed to zero not because fertilizer is ineffective, but because the control set has changed who the comparison group is. This pattern is consistent with the theoretical insight of Altonji, Elder and Taber [4] that adding controls does not necessarily reduce bias under endogeneity—it can amplify it.

This mechanism is not unique to Burkina Faso or to fertilizer. Whenever survey improvements add covariates positively correlated with both treatment and outcome, conditioning on them can shift OLS bias from upward to downward. The practical implication is that the transition from “simple” to “rich” survey data requires a simultaneous transition from OLS to methods that address endogeneity, or the quality of policy advice may deteriorate rather than improve.

### *5.2 Limitations*

Several limitations deserve acknowledgment. First, the completely disjoint sampling frames preclude household or village fixed effects. Observable comparability (Table 1) and robustness to reweighting mitigate but do not eliminate composition concerns. Second, we cannot fully disentangle methodology from time effects. The pre-period covers 2008–2021 while the post-period is 2023 only. However, the magnitude of the shift (strongest predictor to zero in two years), the absence of dramatic agronomic change in the Sahel between 2021 and 2023, and—crucially—the stability of IV estimates all argue against a purely temporal explanation. The year-by-year trajectory (Figure 3) shows no pre-existing downward trend in the fertilizer coefficient.

Third, the exclusion restriction requires that neighbors’ adoption affects own yield only through own adoption, conditional on controls. Region and year fixed effects absorb broad spatial and temporal variation, but village-specific productivity shocks remain a concern [11]. Fourth, our binary adoption indicators discard dosage information. While continuous fertilizer dose specifications confirm the qualitative pattern, the binary framework is dictated by data limitations in the pre-period. Fifth, the measurement discontinuities in intercropping (92.5% to 21.9%) and improved seeds (2.9% to 43.6%) are red flags that require careful interpretation—we treat these as confirming the broader point that survey methodology shapes what we observe about agricultural systems.

### *5.3 Implications for the 50x2030 Initiative*

These findings have direct implications for the 50x2030 Initiative and similar statistical capacity building programs. The investment in survey methodology is clearly justified: post-50x2030 data provides a richer, more detailed picture of agricultural systems that enables analysis previously impossible. The food security modules, distance variables, equipment inventories, and disaggregated input data represent genuine advances.

However, if governments continue relying on OLS production function estimates then richer data may produce worse recommendations. These findings highlight the importance of pairing survey improvements with strengthened analytical capacity in instrumental variables methods, quasi-experimental designs, and causal inference more broadly. In this context, the 50x2030 Initiative's "Analyze" pillar could usefully incorporate training in endogeneity-robust estimation alongside its existing emphasis on data dissemination and use.

More broadly, our findings underscore that survey methodology is not neutral. The choice of which variables to collect, how to define adoption, and how to structure questionnaires shapes the empirical foundations on which public investment decisions rest. Methodological harmonization efforts play a central role in shaping the empirical foundations of agricultural policy[7].

## 6. Conclusions

This paper demonstrates that improving agricultural survey methodology fundamentally alters estimated productivity relationships. Using Burkina Faso's EPA transition as a quasi-natural experiment, we show that OLS production function estimates diverge dramatically across survey periods: the fertilizer coefficient drops from 0.068 to zero, while organic fertilizer reverses from insignificant to dominant. A Chow test overwhelmingly rejects coefficient equality.

IV estimation resolves the puzzle. The causal return to fertilizer is stable at approximately 13% in both periods (first-stage  $F > 2,000$ ). The OLS divergence reflects changing bias structure: the post-50x2030 survey's richer covariate set shifts selection conditional on observables from positive to negative, pushing the OLS fertilizer coefficient to zero. These results illustrate a broader mechanism through which improvements in data collection can alter the behavior of standard econometric estimators.

Budget simulations quantify the policy stakes. OLS on pre-data recommends allocating 27% of the agricultural budget to fertilizer subsidies; OLS on post-data recommends 0.6%. IV estimation gives stable recommendations across periods. In this setting, the choice of estimation method appears to matter at least as much as the choice of survey methodology.

The message for the 50x2030 Initiative, for FAO's MAFAP program and its PolOpT framework, and for any government using survey-based production function estimates: invest in better data, but invest equally in better methods. Absent methodological upgrades, improvements in statistical capacity may not fully translate into improved policy inference, as richer data analyzed with standard methods can generate materially different recommendations.

## Acknowledgments

The authors thank the Institut National de la Statistique et de la Démographie (INSD) and the Direction Générale des Études et des Statistiques Sectorielles (DGESS) of the Ministry of Agriculture of Burkina Faso for providing access to the EPA microdata. We are grateful to the 50x2030 Initiative for supporting the methodological transition documented in this paper.. The views expressed herein are those of the authors and do not necessarily represent those of FAO.

## References

- [1] M. Sheahan and C. B. Barrett, “Ten striking facts about agricultural input use in Sub-Saharan Africa,” *Food Policy*, vol. 67, pp. 12–25, Feb. 2017, doi: 10.1016/j.foodpol.2016.09.010.
- [2] 50x2030 Initiative, “Bridging the agricultural data gap,” World Bank, Washington, DC, 2020. [Online]. Available: [https://www.50x2030.org/sites/default/files/resources/documents/2020-11/50x2030\\_brochure.pdf](https://www.50x2030.org/sites/default/files/resources/documents/2020-11/50x2030_brochure.pdf)
- [3] B. Sawadogo, “Lessons learned from the design and implementation of the national agricultural survey program in Burkina Faso,” Hammamet, Tunisia, 2025. [Online]. Available: [https://www.fao.org/media/docs/statisticslibraries/afcas29/raf-afcas29-e-4-2\\_lessons-learned-from-the-design-and-implementation-of-national-agricultural-survey-program-in-burkina-faso.pdf?sfvrsn=5d6c8a57\\_1&Status=Master](https://www.fao.org/media/docs/statisticslibraries/afcas29/raf-afcas29-e-4-2_lessons-learned-from-the-design-and-implementation-of-national-agricultural-survey-program-in-burkina-faso.pdf?sfvrsn=5d6c8a57_1&Status=Master)
- [4] J. G. Altonji, T. E. Elder, and C. R. Taber, “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools,” *J. Polit. Econ.*, vol. 113, no. 1, pp. 151–184, 2005, doi: 10.1086/426036.
- [5] J. D. Angrist and J.-S. Pischke, *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton University Press, 2009. doi: 10.2307/j.ctvc4j72.
- [6] M. V. Sánchez and M. Cicowicz, “Optimal allocation of agriculture’s public budget can improve transformation and healthy diets access in Ethiopia,” *J. Policy Model.*, vol. 45, no. 6, pp. 1262–1280, Nov. 2023, doi: 10.1016/j.jpolmod.2023.09.005.
- [7] C. Carletto, D. Jolliffe, and R. Banerjee, “From Tragedy to Renaissance: Improving Agricultural Data for Better Policies,” *J. Dev. Stud.*, vol. 51, no. 2, pp. 133–148, Feb. 2015, doi: 10.1080/00220388.2014.968140.
- [8] S. Desiere and Jolliffe, “Land productivity and plot size: Is measurement error driving the inverse relationship?,” *J. Dev. Econ.*, vol. 130, pp. 84–98, Jan. 2018, doi: 10.1016/j.jdeveco.2017.10.002.
- [9] D. Gollin and C. Udry, “Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture,” *J. Polit. Econ.*, vol. 129, no. 1, pp. 1–80, Jan. 2021, doi: 10.1086/711369.
- [10] C. F. Manski, “Identification of Endogenous Social Effects: The Reflection Problem,” *Rev. Econ. Stud.*, vol. 60, no. 3, pp. 531–542, 1993, doi: 10.2307/2298123.
- [11] L. S. O. Liverpool-Tasie, B. T. Omonona, A. Sanou, and W. O. Ogunleye, “Is increasing inorganic fertilizer use for maize production in SSA a profitable proposition? Evidence from Nigeria,” *Food Policy*, vol. 67, pp. 41–51, Feb. 2017, doi: 10.1016/j.foodpol.2016.09.011.