DOES LAND INEQUALITY INCREASE MODERN AGRICULTURE'S VULNERABILITY TO CLIMATE CHANGE? EVIDENCE FROM SPACE[*]

Ignacio Flores[†]   Dylan Glover[‡]

This version: March 2026

Total word count: 16,151

Abstract

Agricultural land consolidation delivers economic efficiency gains through scale economies in costs, but adverse climate shocks primarily affect the biological production function, not the cost structure. This paper investigates whether the ownership distribution that maximizes aggregate output also maximizes vulnerability to extreme heat. Combining satellite spectrometry of crop biomass, farm-level cadastral records, and meteorological data at quasi-weekly frequencies from 2015 through 2021, we estimate dynamic causal effects of staggered heat stress exposure on gross primary production (GPP) – a measure of the biological base of agricultural output, stripped of factor costs and prices – during the prime maturation and harvesting season in France. Under extreme heat more unequally distributed farmland suffers significantly larger GPP losses, with effects translating directly to year-end yield declines. We present evidence that a key driver of this vulnerability is the proportion of arable land allocated to protective seminatural ecosystems, farm hectares that produce no yield but buffer adjacent crops against heat stress. We conjecture that under scale economies, the opportunity cost of maintaining these buffers rises with operational scale, so high-inequality landscapes systematically under-provide this ecological protection. This pattern does not reduce to average farm size: the land Gini captures the polarization of the scale distribution which shapes both the composition of production and, critically, the landscape-level vulnerability under climate change. We examine the incentives formally, with an application to the European Union's Common Agricultural Policy and conjecture that local land Ginis are informative statistics for the effectiveness of subsidies aimed at ecosystem maintenance.

JEL Classification codes: O13, Q15, Q54 (Candidates: Q57)

---

[*] We are grateful to Cameron Harries, Oswaldo Mena, and Selen Özdoğan for their outstanding research assistance and to seminar participants at College de France, CUNY, INSEAD, PSE, Paris 1, SMU and ZEW Mannheim. We have also benefited from helpful feedback from Thomas Blanchet, Raia Massad, Thomas Piketty, Esteban Rossi-Hansberg and Dimitrije Ruzic. This project benefited from funding from the INSEAD's Stone Centre for Wealth Inequality. A complete set of replication files, including all codes and necessary data to replicate our findings will be made available on a public GitHub repository.

[†] E-mail: ignacio.flores@psemail.eu

[‡] Corresponding author. Address: Boulevard de Constance, 77300 Fontainebleau, France. E-mail: dylan.glover@insead.edu

1 Introduction

A defining feature of the world economy is the pairing of economic growth with the consolidation of agricultural production, which has deepened agricultural land inequality in many countries (Lowder et al., 2021). Farmland concentration at the top of the farm-size distribution reflects profit incentives that arise from scale economies in mature, price-taking markets.[1] This concentration has contributed to increases in agricultural output sufficient to meet a sustained rise in food demand over the past century. Yet such expansion has often relied on intensive farming practices that jeopardize the long-run productivity of arable land and the stability of key ecosystem services (Tilman et al., 2002). These same production systems are also faced with significant challenges in adapting to climate risk and, in particular, more frequent and harsher heat stresses (Lüttger and Feike, 2018; Beillouin et al., 2020; Webber et al., 2020). Together, these pressures threaten the resilience of food production systems and raise broader concerns for global food security (Ortiz-Bobea et al., 2021). The existing economic literature has focused on forecasting damage and sustainability through technological adaptation and reallocation (Vogel et al., 2019; Moscona and Sastry, 2023, 2025; Bilal and Känzig, 2024; Bilal and Rossi-Hansberg, 2023). Yet the hypothesis that agricultural consolidation moderates climatic shock effects on food production and security remains untested.

The present study broaches the question. We examine how land inequality moderates plant productivity and climate resilience by linking yearly land ownership and utilization records with high-frequency plant growth observations in France. We obtain consolidation measures through geo-referenced cadastral ownership data, providing detailed information on agricultural plots, including their shape, size, location, crop composition and natural ecosystem exposure. The data allow us to construct a panel of yearly farmland Gini coefficients at an arbitrary resolution. We link these ownership distributions with satellite spectrometry estimates of quasi-weekly crop biomass production at the plot level. This measure –Gross Primary Production (GPP)– is the baseline biological metric used to measure biomass production per unit area and provides a proxy for crop productivity and yields (He et al., 2018). These spectrometry estimates from orbit, adjusted for multiple confounding factors, measure GPP flows at a 500x500m resolution and compiled at a quasi-weekly level (every eight days), allowing researchers to accurately track growth dynamics over the entire year. An advantage, and novel to economics, is that GPP captures the biological base of agricultural production rather than revenue or profit. This distinction matters: the inverse farm-size productivity literature (start with Foster and Rosenzweig (2022)) has primarily studied revenue or value-added per hectare, where factor prices, market access, and subsidies conflate physical output with economic returns. GPP strips these away, isolating how land organization shapes the physical capacity of cropland to convert sunlight and water into biomass—the layer on which all agricultural value ultimately rests. Because GPP is uncorrelated with short-run price effects, it allows us to cleanly measure crop productivity in terms of carbon content and calories. We construct panel observations at the second-most dis-

---

[1] Scale economies in agricultural production are well documented and derive from crop specialization, lumpiness or indivisibility of factor inputs and factor prices and opportunity costs that are scale dependent (MacDonald and McBride, 2009; MacDonald, 2012; Rada and Fuglie, 2019; Adamopoulos and Restuccia, 2014; Foster and Rosenzweig, 2022).

aggregated French administrative unit, the canton, which is roughly seven times more granular than the official agricultural unit in France. This provides substantial farm-scape granularity while remaining sufficiently aggregated to produce reliable inequality statistics. The full panel is comprised of 2,485 cantons observed over seven years with each year nesting 45 quasi-weeks of primary productivity and meteorological variation.[2]

Our results reveal a nuanced relationship between land inequality, farm productivity and temperatures. The literature relating land inequality to agricultural productivity has remained thin since the cross-country analysis of Vollrath (2007). Hence, we begin with a descriptive contribution. Conditional on crop composition, agricultural region fixed effects and weekly environmental controls, we find that crop GPP flows are negatively associated with inequality—more consolidated farmland produces less biomass per hectare. But this association reverses when we scale by cropland area to measure aggregate primary production. Total output combines two margins: per-hectare intensity and cropped area. Inequality is associated with lower intensive margins but larger cropped areas, so the net effect on totals can be positive even as per-hectare crop performance declines.

This sharpens the recent evidence that the farm-size productivity relationship is U-shaped rather than monotonically declining (Foster and Rosenzweig, 2022), a stylized fact that has birthed decades of literature since Sen (1962). Crucially, the U-shape documented by Foster and Rosenzweig (2022) is in *profits* per hectare: the economic gains from scale derive from cost reductions in labor productivity and increased capital efficiency, not from superior crop growing conditions. Our GPP data measure the biological base of production, stripped of factor costs and prices, and reveal that per-hectare primary production declines monotonically with land inequality even in a modern economy. This is consistent with a classic IFSP in physical output coexisting with a U-shape in economic returns. The divergence matters: the economic case for consolidation rests on cost efficiency, not crop biology. Higher land inequality reflects a polarization of the size distribution with fewer mid-sized operations, displaced by large and super farms that cultivate more total area; aggregate output can rise even as the biological productivity of each hectare falls.

This has direct implications for climate resilience. Heat shocks damage the biological production function, not the cost structure. The margin that makes large farms economically efficient offers no protection against the margin that makes their cropland climatically vulnerable. Food balance sheets, regional accounts, and carbon inventories are measured in totals, so "more output" under consolidation can mask both weaker per-hectare performance and weaker resilience to the extreme weather events that are becoming more frequent. The recent literature suggests that a planner maximizing aggregate output should favor high land inequality (Foster and Rosenzweig, 2022). Our findings add a critical caveat: the ownership structure that maximizes output may also be the most climate-vulnerable. Importantly, these results are confirmed in marketable year-end yields for which we estimate a per-hectare yield–GPP elasticity of approximately 0.6

---

[2] French Agricultural Regions (ARs) – over 400 contiguous units adjusted repeatedly over the last 70 years – are designed to maximize similarity in production specialization and shared environment for agricultural policy design. To our knowledge, this is the most granular temperature–crop productivity panel data constructed for a modern agricultural economy at national scale and additionally links it to ownership.

and document a highly comparable yield–Gini relationship.[3] These results on average crop performance and aggregate output remain descriptive, but an aspect of climate resilience can be causally identified.

Hence, our main contribution is to examine whether land consolidation moderates temperature effects on the intensive GPP margin. First, we find that more unequally distributed farmland has lower average GPP throughout the temperature gradient consistent with the descriptive per-area productivity evidence. We then examine GPP-temperature variation under heat stress. We find little evidence that GPP of more or less consolidated farmland trends differently over "normal" temperatures. Only when exposed to extreme heat, roughly 30°C and above, do we observe the emergence of differing trends. Cropland in the first Gini quartile suffers a 6-7% weekly GPP loss for every additional degree above 30°C on average, while cropland in the fourth quartile GPP drops by 10-11%.

The rich meteorological data and plot level crop measures enable us to capture farmland specific heat stress shocks. We estimate these granular temperature shocks by calculating the local crop weighted average of detrimental heat exposure documented in the agronomy literature to define farmland specific damage thresholds. We propose a modern empirical approach adapted to staggered heat shock exposure, documenting pre-trend support and provide a thorough description of exposure paths. This allows us to clarify target estimands when (1) shocks are non-absorbing but treatment effects vary dynamically and (2) reconcile differences in estimates obtained using classic two-way-fixed-effect (TWFE) and recent dynamic heterogeneity-robust estimators. We build a Wooldridge-Mundlak OLS estimator based on Wooldridge (2025) in which we explicitly model all possible cohort-time effects non-parametrically. This allows one to define a wide variety of dynamic ATE estimands that can be tested through weighted linear combinations of the estimated parameters on the cohort-time dummies that provide asymptotically valid standard errors.[4]

Our event study specification estimates pre-trend coefficients which are uninformative of future treatment exposure effects and, importantly, do not exhibit systematic Gini heterogeneity, but threshold crossing effects are stark. We estimate a week-of-the shock ATE on weekly GPP flows of 17%, with Gini inter-quartile heterogeneity of 8 percentage points. Flow effects persist in the second week but then vary considerably in magnitude and even in sign subsequently. We then document how flow effects translate into highly persistent negative effects on the GPP stock: At six weeks, crossing the crop-weighted stress threshold at least once reduces primary production by 3.5% on average with a Gini inter-quartile ATE of roughly 2 percentage points. And because we use farmland-specific thresholds, these can be considered generalized average treatment effects (ATE) with high external validity across the majority of farm landscapes.[5]

---

[3] Yearly yields are available at the French *département* which is an administrative unit larger than the AR.

[4] We illustrate how classic TWFE estimators capture fundamentally different estimands and why caution is warranted when modeling treatment exposure as absorbing when shocks are transitory but can be repeated and outcomes have memory.

[5] Wooldridge (2025) recently established that one can parameterized this linear model that can be combined to produce numerically identical results to the majority of the recent heterogeneity-robust DiD estimators that have emerged recently. We contribute by showing its ability to replicate results, but with tighter standard errors, of even the most advanced estimators such as those proposed by de Chaisemartin and D'Haultfoeuille (2024)

The empirical framework also provides a limpid interpretation of linear ATE heterogeneity with respect to fixed moderators and credible identification of dynamic effects to measure shock impacts with or without multiple shock exposure during the season.

We then test the hypothesis that arable land-use input decisions are correlated with consolidation and heat stress effect heterogeneity. The leading candidate agro-biology mechanisms center around two literatures: crop diversification strategies (Abson et al., 2013; Renard and Tilman, 2019) and seminatural ecosystem integration within farmland (Altieri et al., 2015; Rolo et al., 2023; Burgess et al., 2022). First, crop portfolio effects allow farmers to hedge climate risk as different crops have different heat thresholds and allow for protective interspecies symbiosis effects. We construct raw crop counts, crop Herfindahl–Hirschman indices (HHI) and a primary risk-reducing crop portfolio metric (the crop diversification ratio). We detect modest shock effect heterogeneity over canton HHIs and more significant moderation related to the portfolio metric, independently. Yet they have little predictive power for the land inequality effect gradient. The second candidate mechanism corresponds to arable land maintained in a seminatural state that provide a variety of ecosystem services beneficial to crops. They include small forests and woodlots, grasslands, pastures and hedgerows integrated in and around crop plots. Long standing theory and evidence highlight a myriad of biotic mechanisms such as providing pollinator habitat and refuge, hydrologic transfers through root systems, evapotranspiration and reduced erosion and increased pest predation. Recent plot-level evidence also demonstrates substantial additional abiotic cooling effects on adjacent crops under extreme heat conditions (Ghafarian et al., 2024). We show that seminatural allocation is a much stronger candidate in understanding the inequality shock ATE gradient. We first document how inequality remains highly correlated with lower levels of semi-natural allocation conditional on our control variables. We then illustrate substantial stress mitigation when modeling seminatural ATE heterogeneity directly. Indeed, it provides the strongest predictive power over values of the Gini, among all candidate biological mechanisms, in reproducing the baseline non-parametric Gini ATE heterogeneity. We then examine whether covariation affects the Gini ATE point estimates when estimating ATE heterogeneity jointly. We find that additively "controlling" for seminatural shock effect heterogeneity attenuates Gini-shock ATE estimates by roughly half. We conclude our empirical analysis by confirming that heat stress impacts on GPP and its inequality gradient translate directly to effects on actual yields even when the identifying variation is much coarser. We find that heat stress ATEs on yields over inequality and seminatural allocation are highly consistent when aggregated. The evidence that seminatural allocation itself strongly moderates Gini shock heterogeneity is also robust.[6] In the final third of the paper, we examine whether an agricultural production technology with scale economies not only incentivizes consolidation, as is well documented, but also shapes the private incentives that lead to lower levels of protective investment on aggregate in high inequality landscapes. Importantly, we show that the presence of scale effects alone can generate the observed empirical patterns, without making behavioral

---

which, for example, allow one to isolate highly parsimonious identifying variation to relax pre-trend assumptions and allow flexibility in defining control-group units that serve as controls in cell-level ATEs before weighted aggregation.

[6] In fact, average and heterogeneous yield effect estimates are actually larger than stock GPP effects at six weeks.

assumptions on individual preferences, risk aversion, or information asymmetry. Our analysis illustrates how such patterns reflect the heterogeneity of costs farmers face in allocating arable land to non-yield producing ecosystem buffers. With scale economies in the production function, the per-hectare opportunity cost of a discrete reallocation of land away from protection rises with farm size.

Giving up the "next" hectare is cheap for small farms but expensive for a large one because that hectare sits at a high marginal product. Two testable implications follow. First, aggregate protection is decreasing in land inequality when it comes from land consolidation. That is, holding total arable land fixed, regions where land is reallocated toward the upper-tail of the farm size distribution maintain less semi-natural area. Second, any flat per-hectare payment creates sharp size cutoffs: small and mid-sized farms respond, while large farms typically do not unless the payment is very high.[7] Adding heat shocks strengthens the case for protection but does not overturn the basic scale force: higher shock risk raises the incentive needed to induce protection across the board, yet the "bite" of a uniform payment still fades with size. These predictions map directly into subsidy requirement curves, removal frontiers, and decision maps that we provide through simulations using the estimated protective effects on yields, a scale parameter and existing subsidy payment thresholds.

We use the simple framework to emphasize the farmland Gini as a valuable policy metric beyond its ability to predict the most shock-vulnerable farmland. Farms are embedded in landscapes, so protection generates local ecological spillovers that individual producers do not internalize. Every farm therefore under-invests relative to the social optimum. Large farms account for a disproportionate share of this "missing" protected area and land inequality helps identify this: when local land is more concentrated in a few large operations, the “at-risk” hectares, i.e. the hectares most likely to flip to yield producing input, are disproportionately controlled by those with the strongest private opportunity costs. In such high-inequality landscapes, a euro of uniform payment will typically buy fewer protected hectares for society and will exhibit stronger levels of free-riding. In short, we ask the question whether policy must be both scale-aware and landscape-aware to be effective in maintaining and recovering protection where it is most effective.

We structure the remainder of the paper as follows. Section 2 situates our contribution in the literature on climate adaptation and returns to scale in agriculture. Section 3 describes the data, defines our primary productivity and inequality measures, and constructs threshold-based heat-shock indicators. Section 4 describes the association of land inequality with per-meter gross primary, aggregate yearly output and its mapping to marketable yields. Section 5 presents our causal results on productivity and temperature across the land-inequality distribution and

---

[7] A flat per-hectare payment, such as those available to French farmers under the EU’s Common Agricultural Policy (CAP), creates sharp size cutoffs: the adopted protective share is effectively decreasing in farm size. This policy is central to the European Green Deal and recent CAP reforms. In France, agro-ecology, established as a legal objective in 2014 and supported by *France Relance* and *Écophyto*, remains largely framed as a mitigation and biodiversity policy rather than an adaptation tool, even if the agrobiology literature is extensive. The United States’ *Conservation Reserve Program* (CRP) similarly protects semi-natural areas—grasslands, riparian buffers, and woodlots—through annual subsidies for land retirement, with payments determined proportionally to local land productivity rather than provided as a lump sum.

provides average treatment effect estimates of staggered heat-stress exposure and its dynamics. Section 6 studies the mechanisms behind the observed heterogeneity of shock impacts. Section 7 develops a simple framework that clarifies the opportunity cost of resilience and draws out policy implications. Section 8 concludes.

2 Related literature

This study combines insights from two literatures. The first is the increasingly active literature on adaptation to climate change. The second is the economics of inequality and allocational efficiency under scale economies.

2.1 Adaptation to climate change

Our findings intersect with the literature that examines how productive systems will fare and adapt to average rising temperatures and increasingly frequent extreme weather events. Over the last 60 years, Ortiz-Bobea et al. (2021) argue that anthropogenic climate change has already slowed global agricultural productivity growth, highlighting that climate change is not only a source of future damages but also a drag on realized long-run productivity. Burke and Emerick (2016) study historical data on U.S. agricultural production and find scarce evidence of increased resilience to extreme heat.[8] They find that long-run adjustments have done little to mitigate the short-run impacts of extreme heat on corn yields, despite the clear detrimental effects of heat-stress exposure. Expanding beyond the agricultural sector in the U.S., Burke et al. (2024) examine various outcomes—such as crop yields, mortality, and economic activity—in regions including the U.S., Europe, and Brazil. Similarly, they conclude that while some areas show reduced sensitivity to climate change, the majority of outcomes studied do not exhibit significant adaptation, suggesting that the net effect of adjustments to date has not substantially reduced climate-related damage. In this context, adaptation is inferred from changes over time in the relationship between weather shocks and outcomes, which is informative about overall resilience but less so about the precise mechanisms behind it.

One specific channel for adaptation is technological change. Moscona and Sastry (2023) investigate the potential for directed innovation and demonstrate the significant, though incomplete, role of technological change in adapting to global warming in U.S. agriculture. Their research shows that directed innovation has offset approximately 20% of potential losses in U.S. agriculture due to damaging climate trends since 1960, with projections suggesting that innovation could offset 13-16% of the potential damage by 2100. These innovations include a range of strategies such as the development of drought- and heat-tolerant crop varieties, advances in hybrid breeding, and in some cases genetically modified crops that increase stress resistance. Whether innovation translates into adaptation also depends on institutions and regulation. In the European Union, including France, the cultivation of genetically modified crops for commercial purposes is prohibited, and field experimentation is subject to strict approval processes.

---

[8] Their reduced-form analysis summarizes the net change in the temperature–yield relationship over several decades, reflecting the combined effect of many adjustments and coincident changes, whether these ultimately dampen or amplify damages.

These rules are grounded in precautionary principles and reflect societal preferences regarding risk management, environmental protection, and consumer choice. Consequently, the contribution of directed innovation to climate adaptation depends not only on technological feasibility and ecological fit but also on the regulatory and institutional context in which innovation takes place. Furthermore, Moscona and Sastry (2025) show that technological developments are often tailored to the ecological environments of the countries where they are created. As a result, transferring these technologies to other regions can be challenging, as they may perform suboptimally or fail to address the most pressing local stresses. This underscores the need for innovation and policy that consider local integration to make global agriculture more resilient.

A complementary margin operates under this consideration. Climate adaptation can run through ecological and management-based adjustments, in which farmers' decisions interact with biological processes to shape resilience to heat. The biology literature robustly links biodiversity and landscape structure to productivity stability and stress tolerance through ecosystem-service channels (Cadotte et al., 2008; Kremen and Miles, 2012; Duffy et al., 2017; Renard and Tilman, 2019). A more specific agroecology and agroforestry literature focuses on plot- and landscape-level configurations designed to make yields more resilient under climate stress (start with Altieri et al. 2015; Rolo et al. 2023). A recurring theme is that more frequent and more intense heat shocks can be more damaging than gradual shifts in mean temperatures, and that buffering strategies often operate primarily through two mechanisms: diversification and seminatural ecosystem integration within farmland (Burgess et al., 2022).

While these mechanisms are well documented at the plot level, they remain weakly connected to the economics literature on climate damages in two key respects. First, we know much less about how these buffering strategies translate into aggregate output losses during extreme events, as opposed to average warming. Second, we have limited evidence on how land consolidation shapes the feasibility and effectiveness of ecological buffering, given that many strategies depend on spatial configuration, coordination, or land set-asides that mean forgone yields. Our paper speaks to these gaps by documenting how heat-shock impacts vary across the land-inequality distribution and by linking this heterogeneity to diversification and seminatural buffering patterns that relate directly to the biological sciences (Tamburini et al., 2020; Beillouin et al., 2021).

A first mechanism is diversification through crop portfolio choice. Varied portfolios can hedge climate risk because crops and varieties differ in heat thresholds and growth responses to temperatures. Furthermore symbiosis effects, both within and between crop species can further stabilize outcomes (Abson et al., 2013; Renard and Tilman, 2019). Consistent with this view, several studies show that diversification can protect production exposed to climatic shocks, particularly for subsistence farmers (Valdivia et al., 1996; Di Falco and Chavas, 2009; Birthal and Hazrana, 2019; Seo, 2010).

Complementary literature highlights a series of, often overlapping, biotic and abiotic mechanisms which mitigate heat stress through integration of natural ecosystems. This allocation affects solar radiation, wind speed, humidity, evaporation and water availability (Bramer et al., 2018). Recently, Ghafarian et al. (2024) provide granular evidence pointing to significant direct cooling effects of these seminatural areas under extreme heat. These cooling effects pass through direct solar radiation absorption (Donat et al., 2023) and evapotranspiration of the seminatural areas themselves: substantial amounts of water taken from deep root systems evaporates from the canopy leading to significant cooling effects at surprising large radii (Rolo et al., 2023).[9] Tree roots also move water passively from deeper and wetter layers to shallower and dryer soil layers along a gradient of soil water potential, a process called "hydraulic lift" (Liste and White, 2008). The water becomes available in crop topsoil to evaporate and absorbed by crops shallow roots allowing more direct transpiration. They also create wind breaks which reduce wind speed and preserve microclimates by reducing air mixing (Brandle et al., 2004) and prevent erosion (Donjadee et al., 2010). Both of which can significantly increase water retention and soil moisture, reducing hydric stress within the plot. The evidence thus strongly points to seminatural ecosystem integration as a strong candidate in adapting to climate change. But increased resilience through this channel is not a free margin because it means less arable land allocated for revenue producing yields. How this opportunity cost relates to level of agricultural consolidation, and therefore operational scale remains an open and important question.[10]

2.2 Returns to scale in agriculture and the cost of resilience

The observed relationship between farm size and economic productivity in developing countries has long intrigued economists: smaller farms often report higher yields per square meter than larger ones within a given country. This inverse relationship, first noted by Chayanov (1926) and later expanded upon by Sen (1962) and surveyed in Berry et al. (1979), has been a challenging question in economics. Known as the Inverse Farm Size-Productivity (IFSP) relationship, the stylized fact has fueled significant research and fueled debates over the benefits of land redistribution since breaking up larger farms should lead to higher aggregate productivity (Cornia, 1985; Ellis, 1993; Lipton, 2009). Importantly, such a monotonic relationship implies that higher land inequality is directly related to lower aggregate productivity, a prediction first tested by Vollrath (2007).[11]

The main theoretical explanations that emerged to justify an IFSP relationship revolve around market failure, missing markets (Sen, 1962, 1966; Feder, 1985) and internalization in production risk (Srinivasan, 1972; Barrett, 1996). Sen argued that developing economies are characterized by family labor that is cheap relative to hiring in the external market and this leads small farms that are more intensely labored.[12] But Feder showed that this alone is insufficient. One would need at

---

[9] A mature oak tree can release over a 14,000 L. of water per day into the atmosphere while smaller birch trees release 75 L (see this source).

[10] There is also a small literature exploring direct biological costs of natural ecosystem integration. These costs relate to inter-species competition for water, nutrients, and light; see Mustafa et al. (2022) and Viaud and Thomas (2019).

[11] Vollrath finds a negative relationship between a country's agricultural land Gini and per-area revenue, on average. Since, the literature has remained silent on the direct relationship between inequality and productivity with the recent exception of Borrero (2025). He structurally estimates yearly TFP in Peru and also finds a negative relationship.

[12] The effective wage of family labor falls below that of hired labor due to the low opportunity cost of household members' time and the absence of supervision costs inherent in employing from the labor market. This wage differential generates a systematic relationship between farm size and labor intensity: small family farms allocate household workers at rates that exceed what market wages would dictate because they face lower effective marginal labor costs. Thus they have higher levels of labor input per hectare and, consequently, yields. Larger operations, in contrast, must recruit from the external labor market at higher effective wages. Thus they employ fewer

least one additional market failure for the IFSP relationship to be sustained.[13] Variation in risk exposure over scale could also be at play if farmers face a subsistence constraint and yields are stochastic. If so, farm size will systematically affect optimal effort with small farms investing more intensive labor into plots to self-insure. Large farmers, by contrast, generate sufficient output even in poor years and thus face no comparable incentive to over-invest in effort.[14]

A separate branch in the IFSP literature emerged, arguing that its existence was artifact of poor data quality or suspect identification. But more recent studies cast doubt on whether empirical flaws explain the pattern.[15]

What is certain is that a monotonic IFSP curve is extremely puzzling when confronted with the stylized fact that agricultural productivity of developed agriculture – where farm scale is much higher on average – is substantially higher than it is in low-income countries – where farm scale is low. And that the most credible empirical estimates of the modern production technology (using farm panel data in the US) find returns-to-scale parameters above one (Paul et al., 2004b,a).

Recent work by Foster and Rosenzweig (2022) provides a compelling reconciliation consistent with MacDonald and McBride (2009); MacDonald (2012); Rada and Fuglie (2019).[16] Using plot-level data from India, they present evidence that the relationship between farm size and "productivity" is actually U-shaped. They propose that two forces drive this pattern. At smaller scales, labor costs rise with farm size, pulling down output per hectare as farms grow from very small to mid-sized. But capital goods, though highly productive, require a minimum threshold of land to justify their use. In their data this threshold hits around 4 hectares, beyond which mechanization and other capital inputs begin to push yields back up. The result is a dip across the middle of the size distribution, with higher output per hectare at both extremes.

At face value, a U-shaped distribution implies that a social planner who wants to maximize output should *maximize* land inequality with a mass of micro farms (low labor costs, high yields) and super farms (big enough to exploit capital at minimum efficient scale).

But crucially, the U-shape productivity-scale curve documented by Foster and Rosenzweig (2022) is in profits per hectare, not physical output. The classic IFSP in yields per hectare may well remain monotonically declining even when the profit relationship is U-shaped because the economic gains from scale come from average cost reductions rather than superior growing conditions. GPP data – which measure the biological base of production, that abstract from

---

workers per unit of cultivated area.

[13] If only the labor market is distorted, transactions in other factor markets can arbitrage away the productivity differential. If the large farm could simply rent land to the small farmer until both operations were of a scale worked entirely by family labor, this would equalize yields across farms and raise aggregate output.

[14] The result is higher observed average yields on small farms — not because small farms are inherently more productive, but because their operators are optimally responding to the interaction of risk and a binding subsistence constraint.

[15] Omitted variable bias like soil quality might inversely correlate with farm size (Bhalla and Roy, 1988; Benjamin, 1995). Measurement error, particularly in how farm size and output are recorded, could be negatively correlated with farm size, thus producing the observed inverse relationship (Lamb, 2003). Barrett et al. (2010) refute the claim that omitted soil quality creates a spurious IFSP and Carletto et al. (2013), using GPS tracking of farm size actually reinforces the evidence for an IFSP curve.

[16] They explore the presence of factor lumpiness and indivisibility, labor specialization and relative factor prices in incentivizing farmland consolidation through scale effects.

short-run factor costs and prices – allow us to examine this directly. We find that per-hectare primary production declines monotonically with land inequality, consistent with a classic IFSP in physical output even in a modern economy. The economic case for consolidation thus rests on cost efficiency, not crop biology. This distinction becomes consequential under climate stress: heat shocks damage the biological production function, not the cost structure, so the margin that makes large farms economically efficient offers no protection against the margin that makes them climatically vulnerable.

But this prescription is silent on climate resilience: if scale economies also raise the private cost of ecological buffering, the output-maximizing ownership structure may systematically underinvest in the biological assets that protect food production under heat stress.

Our contribution is to ask what happens to climate resilience under a U-shaped profit-scale relationship. When scale economies raise the opportunity cost of non-yield-producing land, does the resulting distribution of incentives across the farm-size distribution shape aggregate provision of biological buffers, whether through risk-reducing crop portfolios or forgone hectares allocated to protective ecosystems?[17]

Farmers have traditionally cultivated interspersed natural ecosystems for their intrinsic benefits and cultural value. Yet in highly consolidated landscapes, the opportunity costs of maintaining these buffers sit disproportionately on the few large operations that control the most land. When seminatural areas also serve as heat-stress buffers that increase food security, the case for maximizing land inequality requires careful reconsideration.

3 Data and methods

We combine three high-resolution sources for mainland France over 2015–2021: (i) quasi-weekly satellite measures of plant production, (ii) farm-level administrative records that locate cropland and describe land use, and (iii) weather-station temperatures interpolated to local agricultural areas. The core unit of observation in the analysis is a canton-by-date panel, where dates follow the satellite measurement frequency (8-day intervals) and cantons are defined using pre-2013 boundaries; we exclude urban cantons. Appendix A provides more details on data inputs and treatment, including validation exercises.

3.1 Agricultural productivity from repeated satellite spectrometry

Our primary outcome is gross primary production (GPP), a biophysical measure of photosynthetic carbon uptake that closely tracks biomass accumulation and, in agricultural settings, is tightly related to marketable yields (see section A.1). We use the MODIS-based GPP layer produced from the NASA Terra satellite, which provides cloud-adjusted measures at an 8-day frequency. The baseline outcome is expressed in carbon mass per unit area and observed on 321 dates over 2015–2021. Figure A.1a provides a sample visualization at the national level.

---

[17] Beyond forgone hectares, the costs are also exacerbated when capital-intensive technologies are not optimized for crop diversity nor for the spatial arrangement of plots required for natural ecosystem exposure benefits.

Because MODIS pixels cover all land cover types, we restrict the outcome to productive cropland by overlaying the satellite grid with administrative plot boundaries and retaining only pixel areas that intersect declared cropland (see figure A.2 for an illustration of such overlap). This masking step is central: it minimizes forest or permanent natural vegetation from contaminating measured agricultural production in a canton. We validate and quantify primary productivity's relationship to economic value using marketable yield declarations aggregated to the yearly level.

3.2 Land use and inequality measures from cadastral data

We observe land organization using the French Graphic Parcel Register (RPG), an administrative dataset in which nearly all farmers report the precise location and land use of agricultural plots. The RPG provides georeferenced polygons and crop allocation information by year; we use 2015–2021 because earlier years record only the dominant crop within plots rather than an exhaustive declaration (figure A.1b showcases national coverage and figure A.2 illustrates its granularity).

A key conceptual distinction is between the *exploitation*, the administrative farm unit typically used in census tabulations, and the land block (*îlot*), a contiguous area of farmland managed by one administrative entity that may contain multiple parcels. Our analysis focuses on land blocks, which aligns the inequality measure with physical land configuration rather than administrative consolidation of dispersed holdings (Section A.2 provides more details on the French cadastre and its definition of land units).

Our main inequality measure is the Gini coefficient of cropland area within canton-year, computed from the distribution of land holdings observed in the RPG. To limit concerns that local productivity shocks affect land organization, we measure land inequality using the pre-period distribution of cropland holdings (2015) and hold it fixed over time in the main specifications.

3.3 Panel construction

To capture land consolidation, we aggregate data at the level of French administrative divisions. The most obvious candidate is the official Agricultural Region (AR) which partitions France to maximize the homogeneity of agricultural production in terms of specialization and shared environment under political and bureaucratic constraints related to agricultural policy. Currently, a total of 432 agricultural regions partition mainland France.[18]

By conditioning on AR fixed effects and its yearly interaction, we isolate within year variation of similar agricultural units to estimate associations of land inequality with GPP, crop diversity and biological mechanisms, characteristics of farmland that vary little over our seven year panel. Yet two more granular administrative units are nested within the administrative agricultural region. The most granular unit is the *commune* or "municipality" which number more than 35k. This level of granularity is highly attractive, but (1) even if farms are administratively

---

[18] The statistical zoning, developed jointly by Insee and the Scees (SSP), led to a first joint Insee/Scees publication in 1956 containing both the zoning reference framework and the associated map and has been updated roughly every 20 years. They have an intermediate size between the commune and the department. Significant zoning updates were carried out in 1974 and 2007. See this source.

attached to a single commune, a significant proportion span multiple communes and (2) even more concerning is the danger of small sample bias in computing a Gini coefficient for both its point estimate (Deltas, 2003) and its variance (Langel and Tillé, 2013). Indeed, Deltas (2003) shows that the Gini can be significantly biased downward, by as much as 15% for samples between 5 and 10 observations, while the bias becomes negligible for samples approaching 30. For communes, this is extremely problematic because the median number of farms per commune in France is 7 with a mean of 12.4(23.5), min of 0 and max of 974.[19] We thus use the next most granular administrative unit to build our Gini coefficients, the *canton* and the canton-year serves as the panel unit for causal impacts. This panel comprises 2,485 cantons. The median number of farms per canton is 850 with a mean of 1045(sd=918) and min and max of 31 and 8846, respectively. We rely on the cantonal delimitation that preceded a 2013 reform, when there were 4,032 cantons of which we exclude urban cantons. Our sample thus represents 406 Agricultural Regions which nest 6.5 cantons, on average.[20] To calculate the Gini coefficient, farms' crop fields within a canton are ranked by their surface size. From this ranking, we construct a Lorenz curve and apply the standard formula. This provides the best balance between an adequate level of aggregation to compute reliable cropland ownership concentration metrics while conserving a high level of granularity in agricultural production that even goes beyond the nesting agricultural region used for policy. The spatial distribution of our cropland Gini coefficients is presented in Figure A.1c.

We estimate a mean cropland Gini of 0.52 over the seven years with a standard deviation of 0.061 and an inter-quartile range of 0.066. Importantly, it varies little over this time span (Figure A.3), exhibiting a within-canton standard deviation of 0.0068. We thus exploit between canton variation to study the Gini's association with outcomes while conditioning on the most relevant canton-nesting geographic units, such as the AR (e.g. the within-AR Gini standard deviation is 0.035) or the *département* (e.g. for which marketable yield outcomes are available; within-*département* Gini standard deviation of 0.042.) to make the correlational results as informative as possible. For our causal analysis, we model canton-level heterogeneity as fixed (at base year), by which conditionally exogenous temperature exposure moderates primary production and yields, to keep interpretation as clean as possible.[21]

3.4 Crops, semi-natural ecosystems, and local production types

The RPG also reports crop allocation within plots, which we use for two purposes. First, to define crop-types that enter the construction of our heat thresholds (below). Second, it allows us to measure features of the agricultural landscape that are central to the mechanism analysis.

We measure semi-natural ecosystems as the area of woody and grassland features embedded

---

[19] Distribution statistics from the 2020 French Agricultural Census publicly available here.

[20] They nest within a single AR at a rate of over 90%. For cantons that do span ARs, we assign each canton to the AR containing the largest share of its area.

[21] The official agricultural land Gini includes all agricultural production, not just crops, and has evolved from 0.57 in 1970 to 0.6 in 2020. This puts it close to estimates of the European average. In contrast, the North American Gini is significantly higher and trending upwards at a significantly higher rate. For example the USDA estimated a Gini of 0.78 in 2007. See the French Statistics Agency report, "Un panorama de agriculture en France", (2024) and the recent joint report "Research findings from the land inequality initiative", (2024).

in agricultural land, e.g., hedgerows, thickets, groves, woodlots, gallery forests, and prairies, expressed as a share of cropland in the canton-year (with farmland defined net of semi-natural area). These structures provide documented buffering services in hot periods; in our setting they are also an input whose opportunity cost plausibly rises with farm scale. We measure crop diversity using standard concentration and variety metrics, including crop counts and the Herfindahl–Hirschman Index (HHI). For portfolio-style diversification, we additionally compute a standard diversification ratio calibrated from historical yield covariance. See Sections A.4 and A.5 for exact details.

3.5 Temperature data and generalized heat stress exposure

Temperature data are collected from France's national meteorology institute, Météo France. The raw data are recorded at weather stations every three hours. Because stations are not canton-specific, we interpolate temperatures to canton centroids using kriging, with latitude, longitude, and altitude as spatial predictors (see figure A.1d). Kriging substantially improves upon simpler inverse-distance methods in this setting. See section A.6 for more details on our temperature data and its comparison with alternatives.

We summarize heat exposure using a threshold-based shock indicator tailored to local production types. For each canton-year, we construct a crop-weighted critical temperature threshold using the local crop mix and agronomy estimates of heat sensitivity. We then define a binary heat shock on each 8-day date as an indicator that average daytime temperature exceeds the canton-year threshold. This produces a treatment variable that varies across space, time, and crop composition, and is directly interpretable as an episode of heat stress relative to locally relevant crop physiology. Appendix A.7 provides formulae, details underlying crop thresholds and explains how we treat crops for which French-specific estimates are not available. Figure A.4 illustrates the geographical spread of this "locally relevant" shock indicator, as well as raw differences in cumulative GPP between shock and non-shock years.

Importantly, since we observe the same canton each season, this allows us to isolate very parsimonious identifying variation, even up to comparing the canton's farmland GPP with itself with week times year variation, i.e. compare a canton's weekly GPP in shocked vs non-shocked weeks and test whether the size of that within-canton effect varies with the Gini. This means that the fixed heterogeneity over which we examine heat stress impacts might be very close to a benchmark of random assignment to canton farmland.

4 Land inequality and productivity: lower per-area growth, higher aggregate output

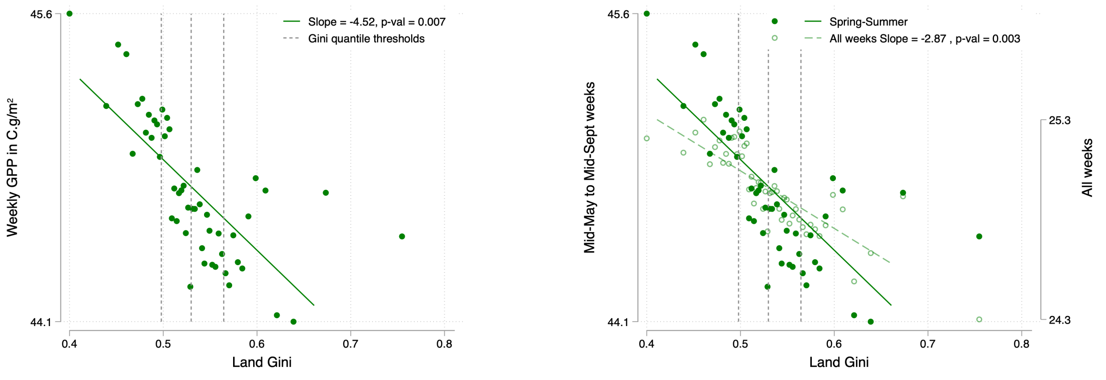
The literature has remained thin on the association of inequality to agricultural productivity since Vollrath (2007) tested the implication that an inverse farm-size productivity (IFSP) implies falling per-area yields in inequality using cross-sectional variation across countries in revenue per acre. To our knowledge there exists no evidence that links high frequency measures of primary production to granular measures of land inequality in either a developing or mature agricultural economy. We thus begin our empirical analysis with a descriptive analysis of land inequality and productivity in terms of GPP and marketable yields.

Table 1: Gross primary productivity (GPP C.g/m^2 per week) and farmland inequality

	(1)	(2)	(3)	(4)	(5)
Panel (a): Linear Gini					
Ginist	-1.698***	-0.308***	-0.274***	-0.348***	-0.318***
	(0.160)	(0.112)	(0.101)	(0.106)	(0.104)
Mean GPP	44.7				
Panel (b): Gini quantiles					
q2	-3.399***	-0.799***	-0.344***	-0.376***	-0.387***
	(0.277)	(0.153)	(0.126)	(0.132)	(0.131)
q3	-5.453***	-1.194***	-0.608***	-0.685***	-0.688***
	(0.284)	(0.182)	(0.154)	(0.162)	(0.160)
q4	-5.892***	-1.257***	-0.687***	-0.825***	-0.781***
	(0.316)	(0.224)	(0.191)	(0.202)	(0.198)
Mean GPP 1st quantile (ref.)	48.4				
p-val equal	0.0000	0.0074	0.0226	0.0066	0.0114
R-squared	0.030	0.230	0.244	0.280	0.845
N	355024	355024	355024	355024	355024
(AR x Year) FEs		✓	✓	✓	✓
Crop compo.			✓	✓	✓
Meteorological				✓	✓
(AR, Crop compo., Meteo.) x Week and Year FEs					✓

Notes. Dependent variable is weekly gross primary production in grams of carbon per square meter of canton cropland (C.g/m^2) from mid-June through mid-September. Panel (a) Gini centered and standardized. Panel (b) quartile differences in reference to bottom Gini quartile. Meteorological controls are average weekly temperature, precipitation and wind speed. Standard errors are clustered at the canton. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure 1: Weekly crop primary production over land inequality



(a) Weekly GPP C.g/m^2 mid-May to mid-September

(b) Comparing with all weeks

Notes. 50 quantile spaced bins. (a) Weekly GPP flows of gross primary production (GPP) during weeks from mid-May to mid-September. (b) Overlays bins and slope for all weeks. GPP expressed in carbon grams per square meter, C.g/m^2. Residualization and linear fits correspond to model (3) in Table 1.

Panel (a) of Table 1 presents evidence of a significant negative association between land inequality and average weekly crop biomass production during the prime maturation and harvesting season (mid-May to mid-Sept.). The cropland Gini coefficient is standardized, $Gini^{st}$. Panel (b) shows estimates of the relationship non-parametrically and the p-value tests whether conditional means are equal over quantiles of the Gini. The raw association indicates that a one standard deviation increase in the Gini correlates with a decrease in weekly GPP of about 1.7 carbon grams per square meter (C.g/m^2) and a Gini inter-quartile difference of roughly 5.9 grams. The results in the following columns isolate variation between GPP and the canton Gini within the AR-Year nest, providing more informative estimates of land inequality's association with crop productivity.[22] We see this attenuates the estimates significantly but does not eliminate the negative association. We then directly control for crop inputs that vary at the canton-year level. We observe minimal further attenuation, suggesting that crop inputs across cantons vary relatively little within agricultural regions over years. We then add meteorological covariates that vary at the canton-week: average temperatures, the amount of liquid precipitation and wind speed. We note that the negative GPP-Gini association grows slightly stronger. Finally, we saturate the regression with time effects that allow agricultural regions to trend non-parametrically by week within a year and for continuous control slopes to vary by week and year. The model explains 84.5% of weekly GPP variation and the negative Gini association with crop GPP flows remains stable and robust to sampling both on average and across the Gini quantiles.[23]

Figure 1a illustrates the estimated GPP-Gini CEF non-parametrically during the prime growth season using quantile spaced bins, residualized using model (3) of Table 1, and partitions the graph with vertical lines that denote Gini quantiles. We then scale the axis to compare the spring and summer CEF with that of GPP week flows for all weeks during the year in Figure 1b. We note that GPP flows are significantly higher during the spring and summer weeks, but also exhibit a significantly steeper Gini slope (p-val < 0.001) during the spring and summer weeks. On average, a point increase in the Gini is associated with roughly -0.027 less grams of weekly carbon flows per square meter over the whole year while the association is about -0.051 during these spring and summer weeks.

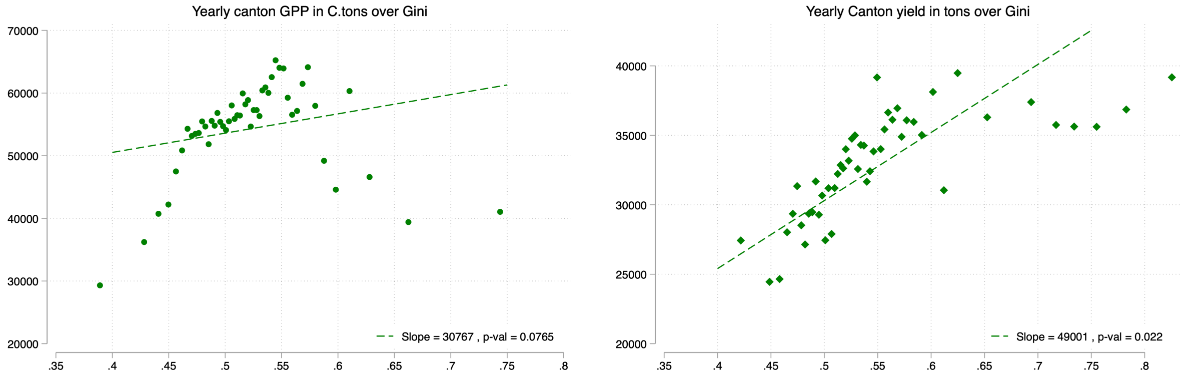
If more unequally distributed cropland ownership is associated with lower average per-area productivity, a natural question is whether this translates into lower food supply over inequality. This distinction matters because regional accounting of agricultural output, carbon inventories, food security, and comparative productivity metrics for policy are typically functions of aggregate annual production. Strikingly, when we shift the outcome to yearly aggregate GPP, by simply scaling the per-area GPP measure by the canton's crop surface area, the descriptive relationship with land inequality becomes positive. Figure 2a displays a clear, positive, and potentially concave association between the land Gini and annual crop biomass measured in carbon tons per canton.

---

[22] This allows us to control for differences in all fixed or slow moving unobservable inputs like soil quality, variation in local factor markets and features of mesoclimate exposure that vary outside the nesting and compares farmland within "administratively homogeneous" farmland.

[23] See Table A.3 for estimates by model for average flows over the whole year not just the prime growing season.

Figure 2: Yearly agricultural GPP and declared crop yields



(a) Yearly GPP per canton C.ton

(b) Yearly yield per canton C.tons

Notes. 50 quantile spaced bins residualized on AR x year FEs. Cumulative yearly GPP on 31 December and yearly declared yield expressed in tons of carbon by canton. Declared yield expressed in tons of carbon by canton by dividing the *département* measure by the number of nested cantons. Residualization and linear fits correspond to model (3) in Table 1 with standard errors clustered at the canton (a) or *département* (b) level. N=22,189

To benchmark our satellite-based quantities against administrative outcomes, we use publicly available yield declarations compiled by the French Ministry of Agriculture at the *département*-year level. While the *département* is a broader unit of aggregation than the canton or AR, these records provide an independent, outcome-based counterpart to our GPP measures.[24] We merge the *département*-year yield declarations to our canton-year observations and pair them with cumulative annual GPP. This allows us to estimate the relationship between declared yields and nesting canton Gini's weighted by the share of its crop surface area to the *département*'s. In Figure 2b we observe a Gini relationship CEF that strongly resembles that of its yearly GPP-Gini counterpart.[25]

Table 2 provides level correlations and yield-GPP elasticity estimates. A one gram increase cumulative year-end GPP is conditionally associated with a 0.56 gram increase in declared marketable yield. In the next column we see that the level effect translates into a constant elasticity estimate of about 0.7. We observe that controlling for yearly variation in crop compositions affects the estimate little but weighting the regression by the share of the canton's farmland hectares within total department farmland hectares increases the elasticity estimate. The final column provides an elasticity estimate of around 0.44 of total declared *département* yields with total canton-level crop carbon production.

---

[24] Cantons are nested within the 94 *départements* of metropolitan France, with both the median and mean number of cantons per *département* at roughly 38.

[25] A natural concern is that the Gini coefficient is not additive across nested units: the *département*-level land Gini cannot be recovered from the canton Ginis because it also reflects between-canton variation in farm sizes. Our analysis does not require this aggregation. The canton Gini enters as a moderator of canton-level shock exposure, not as a *département*-level statistic. Canton fixed effects absorb all time-invariant canton characteristics, and identification comes from within-canton changes in shock exposure across years. The area-weighted merging ensures that each canton's contribution to *département* yields is scaled by its share of total farmland, so the estimates capture how differential shock exposure across cantons with different ownership structures maps to the *département*'s aggregate yield outcome. Non weighted results are presented in Figure A.9.

Table 2: Yield-Cumulative GPP elasticities

	Yield g/m^2	log(Yield g/m^2)			log(Yield ton per Canton)
	(1)	(2)	(3)	(4)	(5)
Cum. GPP C.g/m^2	0.563***				
	[0.081]				
log(Cum. GPP)		0.685***	0.670***	0.719***	
		[0.079]	[0.076]	[0.104]	
log(Cum. GPP ton per Canton)					0.442***
					[0.046]
Mean Yield (levels)	913				46385
Mean GPP (levels)	1145				54226
N	22189	22189	22189	22189	22189
Canton and Year FEs	✓	✓	✓	✓	✓
Crop compo.	✓	✓	✓	✓	✓
Area share weighted				✓	✓

Notes. Yearly declared yields in g/m^2 regressed on cantons' end-of-year cumulative gross primary production measured in grams of carbon per square (C.g/m^2). Weights correspond to the canton's farmland area share of the nesting *département*'s total farmland surface area. Standard errors are clustered at the *département*. * $p < .1$, ** $p < .05$, *** $p < .01$.

A pattern of results that are consistent across productivity metrics suggests that the relationship between aggregate agricultural productivity and inequality is a composite of two margins – an intensive margin (per-area productivity) and an extensive margin (crop area scale). The evidence further suggests that the extensive margin dominates because yields rise in inequality even if per–ha. crop productivity falls in it.

So what is the land Gini capturing, really? Figure 3a shows that lower-inequality farmland tends to have larger small and mid-sized operations that hold more of the total land. In the bottom Gini quartile, farms with cropland of 2–10 ha control over half of total cropland, while farms above 50 ha control a negligible share. Moving up the Gini distribution, we observe a clear reallocation of land ownership away from small and mid-sized operations toward large (20–50 ha), very large (50–100 ha), and super farms ($\geq$ 100 ha). By contrast, the land share held by micro operations remains relatively stable. Higher land inequality reflects an allocation of productive resources towards larger operations but also captures the level of polarization in the farm-scape, on average. In Section C, we show the Gini maps to the share of control over operation type and clarifies how ownership polarization underpins land inequality.

Figure 3b illustrates consistent results with these raw means when conditioning on our controls. Land inequality's partial elasticities with land share holdings across scale bins exhibit a strong positive association with bigger operations and an inverted U-shaped CEF due to its smaller, positive association with micro farms.[26] We observe that the land Gini can capture a rich summary statistic for the allocation of productive resources by operational scale. Figure C.1 and Table C.1 show results at more granular size bins.

---

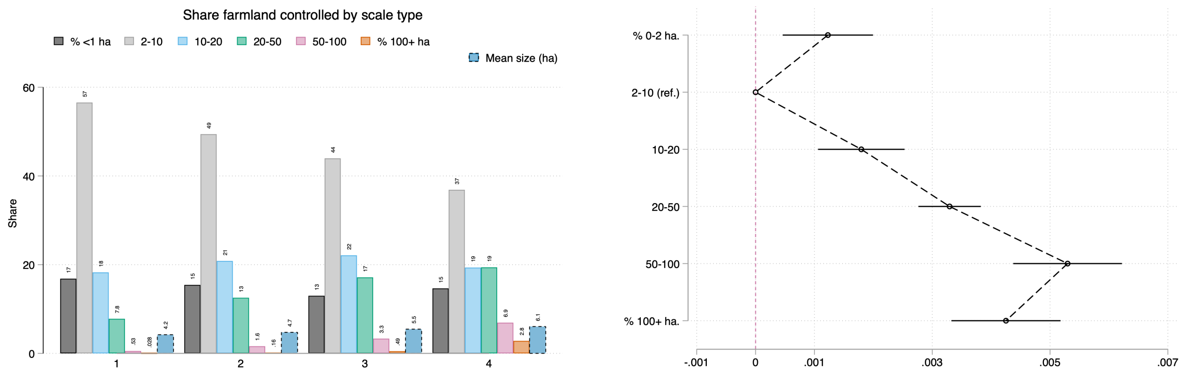
[26] Share holding bins sum to 100% and are constant within year thus one interprets estimates as follows: a 1%-point increase in the type's controlled share – that displaces 1%-point of the omitted share – affects the outcome by 100% $\times \beta$ See Morawetz and Klaiber (2025) for a formalized interpretation of linear regression estimates with this type of clustered aggregation.

Figures 3c and 3d then map these associations onto the intensive and aggregate margins. The partial elasticities indicate that a one percentage point increase in the share of land managed by very large farms (100+ ha) is associated with about a 0.1% decline in weekly mean biomass production (relative to a comparable change in the share of land controlled by 2–10 ha. farms). We then observe the same symmetry as before. When regressing total annual biomass in the canton on share bins, the association flips.

A higher Gini equates to a more polarized farm size distributions, with (1) significantly higher shares of operations in the curve's upper right tail, and (2) modestly more micro operations in the left tail.

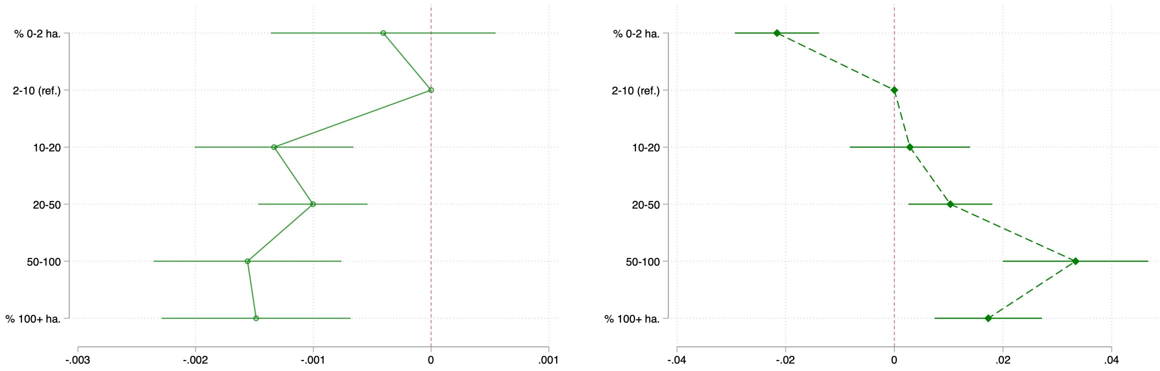
Overall the evidence is consistent with an economy's agricultural productive allocation trending towards the most economically productive farms. Yet the fact that average plant productivity is negatively correlated with the Gini presents a wrinkle: In the IFSP curve's upper right tail, it may be that optimal crop conditions do not necessarily equate with efficiency when production decisions maximize profit. From this perspective, the Gini may also summarize the distribution of both scale complementarities in the production technology, and private opportunity costs that vary over scale.

Figure 3: Farmland control over land inequality and relation with productivity



(a) Land shares by farm-size bin over Gini quantiles

(b) Gini - land share elasticities



(c) Weekly GPP flows per m^2 elasticities

(d) Year-end GPP per canton elasticities

Notes. (a) Share of total farmland held by each farm-size type across Gini quantiles. (b) Estimated partial elasticities of the land Gini with respect to farmland scale share control (c) Estimated partial elasticities of weekly GPP flows with respect to holdings by size from mid-May (d) Homologous estimated partial elasticities on year-end canton GPP stocks. All estimates condition on agricultural region x year FEs and yearly crop compositions. Standard errors clustered at the canton level. 95% confidence intervals. N=22,189

5 Primary productivity responses to temperature variation over land inequality

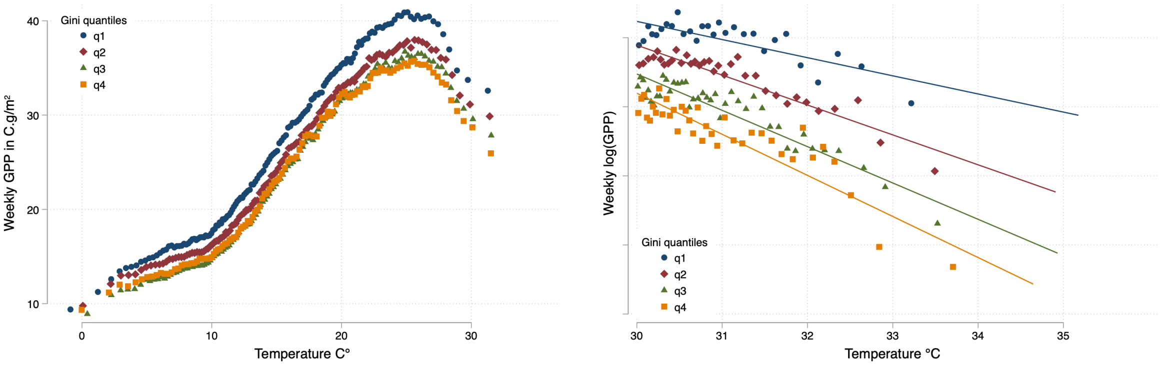
We begin by examining how primary productivity varies with temperature and whether the relationship exhibits heterogeneity with respect to the level of farmland consolidation. Growing seasons repeat each year and we have GPP measurements for every quasi-week for seven years. This allows us to isolate variation within the canton-year unit to provide an analysis of the GPP-temperature function when isolating identifying variation to within the most granular panel unit.

5.1 The temperature–production gradient

The residualized scatter plots presented in Figure 4 provide non-parametric estimates that help one visualize the entire GPP-temperature gradient. We observe in (a) that increasing temperatures are strongly associated with higher GPP and that more consolidated land exhibits lower productivity across the entire temperature gradient, on average; the estimated CEFs of Gini quartiles three and four are dominated by those of the first and second quartiles over almost all temperature bins. The nonlinearity in the right tail indicates the presence of heat-stress effects on crop carbon production. Figure 4b zooms in, restricting to weekly observations in which mean temperature meets or exceeds 30°C with linear fits on log GPP. The visual evidence suggests that, under heat-stress, more unequally distributed cropland sees steeper declines. For the remainder of the causal analysis we focus on the prime maturation and harvesting season, mid-May to mid-September (quasi-weeks 17-32) and where all stress events are contained. Section B.1) provides the specifications and assumptions for our raw temperature models.

Table 3 provides estimates and inference from the spline regression equation B.2. Columns correspond to the temperature subsample for the predicted marginal effects. Panel (a) provides the average one-degree log-linear effect on weekly GPP flows and its heterogeneity from a standard deviation increase in the Gini coefficient. Panel (b) presents the GPP-temperature elasticity heterogeneity non-parametrically by Gini quartile and provides p-values of the Wald test that temperature effects are equal across the quartiles. On average, there is a 1.4% return in GPP for a one-degree temperature increase. We also note significant Gini heterogeneity. Across the columns we observe point estimates in-line with the non-linear GPP-Temperature CEFs observed previously. Returns are increasing and then strongly decreasing. Land inequality strongly moderates effects, especially when crops are likely to face thermal stress. At 30°C or above, every degree increase reduces weekly GPP flows by about 8% while a one-standard deviation increase in the Gini exacerbates the GPP losses by roughly 1.3%. The final three columns provide robustness tests. Identifying variation becomes more and more restricted by adding non-parametric week trends for nesting geographic units up to the canton itself. Indeed, we can condition on canton time trends and still identify extreme heat exposure variation because seasons repeat. Linear Gini effect heterogeneity estimates are attenuated using these more demanding models, yet we still robustly reject the null for the average effect and Gini effect heterogeneity with non-parametric heterogeneity estimates remaining statistically identical.

Figure 4: Weekly GPP v. temperature over land Gini quantiles



(a) All temperatures (b) ≥30ºC

Notes. Bins residualized on canton-year and week fixed effects. (a) N=1,016,929 corresponding the canton-week-year observations for the full year. (b) N=16,712. Restricted to weeks with average temperatures 30ºC or above. Bin selection and partitioning follows the integrated mean squared error minimization algorithm proposed by Cattaneo et al. (2024).

Table 3: GPP-temperature impacts over land inequality

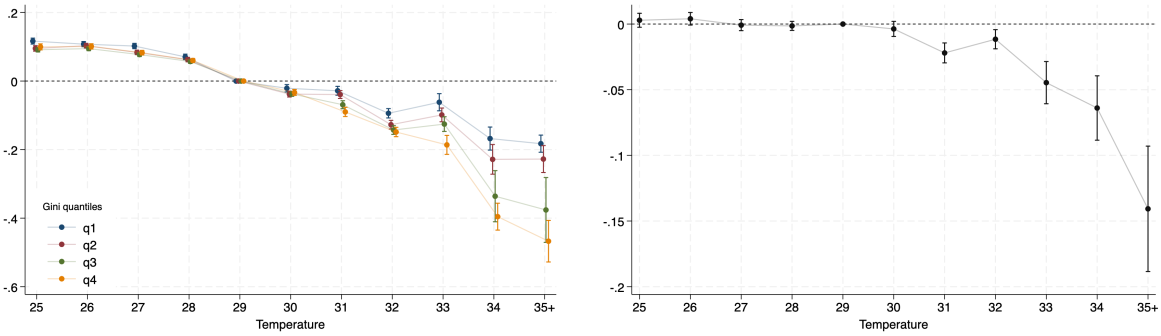
	All T	T <20°	20 ≤ T < 25	25 ≤ T < 30	30°≤ T°			
Dep. Variable: logGPP	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(a): TºC effects with linear Gini								
Temp.C°	0.014***	0.056***	-0.002***	-0.044***	-0.083***	-0.079***	-0.080***	-0.081***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Temp.C° x Gini^{st}	-0.001**	0.000	-0.001*	-0.003***	-0.013***	-0.008***	-0.007***	-0.007***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
(b): TºC effects by Gini quantile								
q1	0.018***	0.053***	0.004***	-0.043***	-0.062***	-0.068***	-0.071***	-0.072***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
q2	0.015***	0.057***	-0.001	-0.045***	-0.073***	-0.072***	-0.073***	-0.074***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
q3	0.013***	0.058***	-0.002***	-0.043***	-0.086***	-0.080***	-0.080***	-0.081***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)	(0.004)
q4	0.008***	0.057***	-0.008***	-0.047***	-0.109***	-0.096***	-0.096***	-0.097***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)
p-val equal	0.000	0.011	0.000	0.202	0.000	0.000	0.000	0.000
N	355024	355024	355024	355024	355024	355024	355024	354928
Canton x Year and Week FEs	✓	✓	✓	✓	✓	✓	✓	✓
Gini x Week FEs	✓	✓	✓	✓	✓	✓	✓	
Dept. x Week FEs						✓		
AR x Week FEs							✓	
Canton x Week FEs								✓

Notes. Predictions of temperature effects and its Gini heterogeneity over different temperature thresholds derived from estimates of equation B.2. Panel (a) presents the partial elasticities of GPP with temperature in degrees Celsius and its interaction with sample centered and standardized land Gini. Panel (b) presents the partial GPP temperature elasticity within each Gini quartile. Baseline estimates used to construct the splines are presented in Table B.1. Standard errors clustered at the canton. * $p < .1$, ** $p < .05$, *** $p < .01$.

We further examine Gini effect heterogeneity with a "temperature event study." Specification B.3 models dummies for all (rounded) temperature integers and their interaction with the Gini. This allows for a fully non-parametric examination of the GPP-temperature CEF and whether inequality heterogeneity is correlated with how weekly GPP flows trend over "normal" to extreme temperatures.

Figure 5a presents the effect by quartile for bins of 25°C and above using 29°C as reference. We note that GPP falls gradually leading to this reference temperature, on average, but that trends are highly comparable over Gini quartiles. We then see the emergence of significantly larger negative effects of temperature increases for more unequally distributed cropland. At 31°C, we estimate that the least consolidated cropland (Gini quantile 1) suffers a 4.3% relative loss while the most consolidated land (Gini quantile 4) experiences a 9.7% drop. The Gini-effect gradient continues to "fan out" as temperatures become more and more extreme. Though standard errors increase due to decreasing exposure support, inference is still possible even at these extremes. For example, at 35°C or higher, the inter-quartile effect gap is more than 25 percentage points in reference to 29°C. Figure 5b presents the GPP-temperature elasticity estimates using continuous Gini heterogeneity.[27] Compared to the reference temperature bin, at 31°C a standard deviation increase in the Gini associates with 2% higher GPP losses and this differential over inequality only grows stronger as stress increases. At 35°C or above, the heat stress differential is 14% for a standard deviation increase in the Gini. We formalize how to interpret support for the partial identification requirement of the raw temperature models in Section B.1.1.

Figure 5: Temperature "event study" estimates



(a) Gini quantiles

(b) $\text{Gini}^{st} \times 1\{\text{Temp}^\circ\text{C} = j\}$

Notes. N=355,024. Estimates from "temperature event" model B.3. (a) of the GPP-temperature partial elasticity by Gini quantiles using integer temperature bin dummies. (b) Coefficient estimates on the standardized Gini interacted with temperature bin dummies. Reference temperature is 29°C. Illustration restricted to estimates for 25-35+°C. Standard errors are clustered at the canton. Vertical lines correspond to 95% confidence intervals.

5.2 Average treatment effects under staggered heat stress shocks

We now estimate average treatment effects (ATEs) using the crop-weighted temperature shock indicator. As opposed to raw temperature increases, this indicator allows us to capture exposure to heat stress as a function of the yearly crop composition in the canton. It thus provides a generalized treatment effect for all crop farmland for our most granular panel unit and test a full identification assumption under dynamic heat stress exposure.

Our first target estimands are the "week-of-shock" ATE on GPP and its heterogeneity over

---

[27] Figure B.1 reproduces both figures when additionally conditioning on non-parametric canton-week time trends.

land inequality: the effect on crop productivity when average weekly temperature crosses the crop-specific heat stress threshold for the canton. We start with a two-way-fixed-effects (TWFE) estimator and provide a framework that helps clarify what this estimator captures in our setting and why it might differ from our target estimand. We then propose a more parsimonious empirical estimation strategy adapted to staggered heat-stress exposure and potential dynamic effect heterogeneity. We show that this allows one to cleanly isolate the week-of-shock ATE and dynamic effects on GPP flows and stocks in subsequent weeks.

All canton-year specific shocks, $D_{iwt} \in \{0, 1\}$, occur in weeks 20 through 27 for each year $\mathcal{T} = \{2015, \ldots, 2021\}$. Weeks are denoted $\mathcal{W} = \{17, \ldots, 32\}$ (mid-May through mid-September) which ensures that all treated canton-year units have at least three pre-shock, and six post-shock periods.[28]

In specification B.4, D_{iwt} is a treatment indicator for canton cropland i equal to one if temperature equals or exceeds the canton-specific stress threshold in week w in year t, and zero if not. Linear land inequality ATE heterogeneity is modeled by interacting the shock indicator with the cropland Gini coefficient in 2015. It is centered and standardized using the sample mean and standard deviation, $Gini_i^{st}$. We condition on canton×year and Gini×week fixed effects to isolate variation within canton-year pairs and again allow land inequality to vary non-parametrically by week within each year. As above, we also model ATE-Gini relationship non-parametrically by substituting the $Gini_i^{st}$ terms with a vector of Gini quantile q indicators to estimate ATEs by inequality quantile. Further description and identifying assumptions are presented in Section B.2.

Table 4 presents shock ATE and linear heterogeneity estimates in panel (a) while panel (b) presents shock ATE estimates by Gini quartiles. In column 1 of panel (a) we observe that, on average, weekly crop GPP falls by roughly 19.6% when temperatures meet or exceed the cropland specific heat stress threshold. The coefficient on the interaction implies that a one-standard deviation increase in the Gini exacerbates the negative effect by 2.1%, on average. The ATE estimates by Gini quantile presented in Panel (b) are consistent with this linear ATE heterogeneity. While crop GPP in the first Gini quartile experiences a 16.7% drop, the most consolidated farmland experiences a 22.7% drop.

Though we have strong evidence in support of "parallel temperature trends," we noted that this supports a partial identification assumption. Cropland GPP can experience different dynamic trends over the temperature gradient and trends may depend on underlying Gini heterogeneity. We now augment the baseline TWFE model B.4 (column 1 of Table 4) with leads of D_{iwt} and their interactions with the Gini to test the no anticipation (NA) and conditional parallel trends (CPT) assumptions necessary for causal inference.[29] Simply, does future heat stress exposure predict contemporaneous effects on GPP? Figure 6 illustrates. The top plot, Figure 6a, presents

---

[28] See Figure A.8 for the distribution of weighted shocks. A small number of first shocks occur in weeks 28-30 but are isolated to very few cantons and thus exhibit insufficient Gini variation.

[29] The two lead coefficients are identified for all treated cohorts in our summer window. The earliest shocks occur in week 20 (cohort 20), with weeks 17 and 18 available as leads in reference to week 19. We therefore bin all earlier available weeks for cohorts 21-27 into a single "far" pre-shock indicator to avoid composition differences driven by late-treated cohorts.

the estimated lead coefficients – using lead 1 as the omitted reference, for the ATE and Gini interaction term – respectively. The lead estimates shows that cropland GPP flows tend to be slightly higher and slightly decreasing leading up to the shock, on average, consistent with the temperature event study results. We then observe a sharp discontinuity. Crop GPP drops by roughly 17% in reference to the previous week.

We observe little evidence that future heat stress effects are predicted by land inequality. We observe small and statistically insignificant estimates for the $D_{iwt} \times Gini_i^{st}$ leads. We then observe that a one standard deviation increase in the Gini is associated with over 2.5% larger GPP flow losses when shocked in reference to the first lead 1. Bottom Figure 6b plots the dynamics by Gini quantiles. We observe similar pre-trends and then the clear emergence of a monotonic Gini-effect gradient where farmland in the first Gini quartile suffers roughly 13% GPP losses while the top Gini quartile loses over 20%.

Columns 2 and 3 in Table 4 display estimates when adding non-parametric – agricultural region (AR)- and canton-week (our preferred model) trends to the baseline model. We observe a modest compression of effects in Gini quartiles 2, 3 and 4, but overall, ATE and Gini heterogeneity estimates that are statistically indistinguishable across the 3 models.

Column 4 presents estimates when we drop post-shock week observations for treated units. This sample restriction gives an, a priori, closer estimate to the target week-of-shock estimand. By construction, this restriction prevents "forbidden comparisons" by ensuring that previously shocked units are never included as controls.[30] It also requires no assumptions about ATE dynamic heterogeneity. Interestingly, we observe a similar average effect but stronger linear Gini ATE heterogeneity. Turning to panel (b), it appears that the change in the linear Gini moderating gradient is driven by an attenuated estimated shock ATE in the bottom Gini quartile.

It is informative to see what happens when we redefine the shock indicator to be absorbing: It turns "on" (=1) in the initial crossing-threshold week and remains on for the remaining weeks of the canton-year unit. Estimates are presented in the remaining column. This model implicitly uses weighted averages over all post exposure periods to determine treated mean GPP. We observe significant attenuation for all parameters. Comparing across models 3-5 suggests that there may be significant dynamic ATE heterogeneity post initial exposure and that its Gini heterogeneity may also be dynamic. A standard TWFE model abstracts from this dynamic heterogeneity and or multiple shock exposures.

---

[30] Formally, this set is

$$\mathcal{S} = \big\{(iwt):\ w \in \mathcal{W} \text{ and } w \leq w_i^{\star}(D_{iwt}=1)\big\} = \big\{(iwt): w \in \mathcal{W},\ D_{iwt}=0\big\} \cup \big\{(iwt): w \in \mathcal{W},\ w = w_i^{\star}(D_{iwt}=1)\big\}.$$

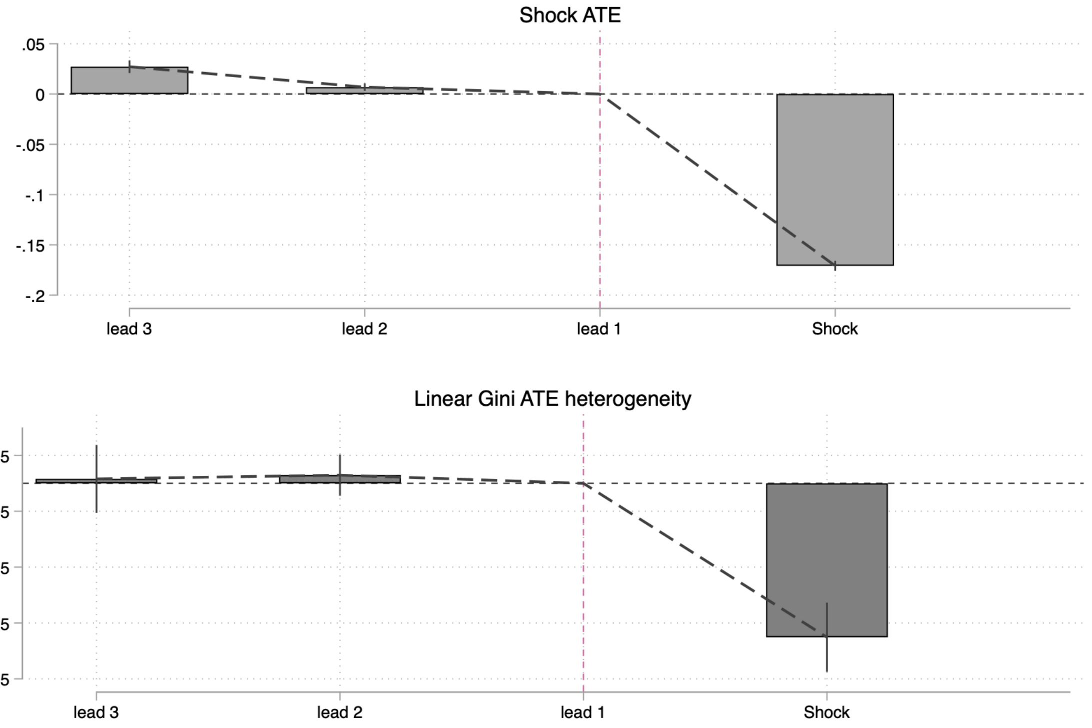
, where $w_i^{\star}(D_{iwt}=1) := \inf\{\, w \in \mathcal{W}:\ D_{iwt}=1\,\}$.

Table 4: Generalized heat shock TWFE impact estimates

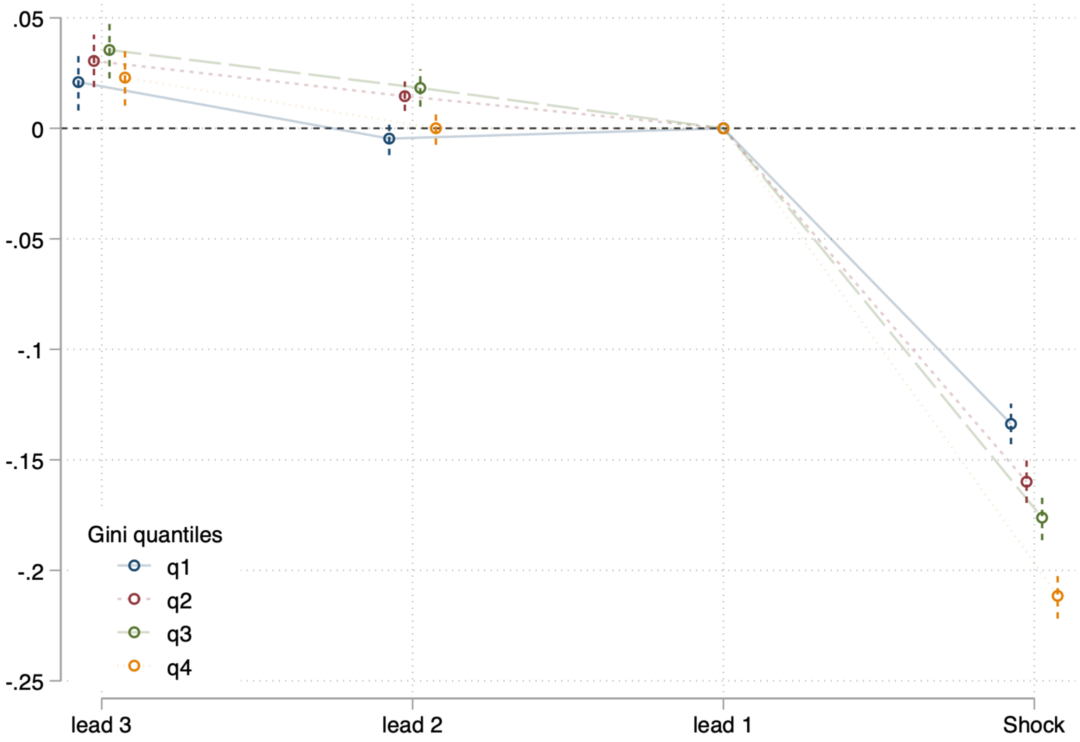
Dep. Variable: logGPP	(1)	(2)	(3)	No post-shock weeks (4)	Absorbing shock (5)
Panel (a): ATE and linear Gini					
Shock	-0.196***	-0.183***	-0.187***	-0.184***	-0.088***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Shock x Ginist	-0.021***	-0.018***	-0.019***	-0.030***	-0.015***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.002)
Panel (b): ATE by Gini quantile					
q1	-0.167***	-0.164***	-0.167***	-0.134***	-0.044***
	(0.006)	(0.005)	(0.005)	(0.006)	(0.007)
q2	-0.188***	-0.180***	-0.182***	-0.182***	-0.092***
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)
q3	-0.199***	-0.181***	-0.186***	-0.189***	-0.102***
	(0.006)	(0.005)	(0.005)	(0.006)	(0.004)
q4	-0.227***	-0.208***	-0.215***	-0.226***	-0.104***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.004)
p-val equal	0.0000	0.0000	0.0000	0.0000	0.0000
R-sq.	0.723	0.812	0.825	0.819	0.824
N	355024	355024	355024	330067	355024
Canton x Year FEs	✓	✓	✓	✓	✓
Gini x Week FEs	✓				
Gini x AR x Week FEs		✓			
Canton x Week FEs			✓	✓	✓

Notes. Estimated heat shock effects and inequality heterogeneity by log of land Gini (a) and by quantile of the land Gini (b) from baseline models B.4 and B.5, respectively. (1-3) include all weeks in prime season in which shock indicator can turn on and off. (4) drops shocked cantons post initial shock week. (5) Absorbing shock. Standard errors are clustered at the canton. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure 6: TWFE estimates with shock leads



(a) Average Shock Effect and Linear Gini heterogeneity



(b) By Gini quantile

Notes. Estimates of lead terms and their interactions added to B.4. Vertical lines correspond to 95% confidence intervals. Standard errors are clustered at the canton. N=355,024

5.3 Dynamic heterogeneity robust shock ATEs

Because initial shock exposure happens during different weeks within a year (staggered exposure) interpreting the TWFE shock estimates thus far exhibited requires caution. First, TWFE estimates are a weighted average of all available 2x2 difference in differences (DiD) estimates for each period w that exhibit exposure variation (de Chaisemartin and D'Haultfoeuille, 2020; Goodman-Bacon, 2021). It therefore includes iwt outcomes that do not correspond to the target estimand. These "forbidden comparisons" exist because TWFEs exploits treatment variation within each period w thus "control group" outcomes are averaged over both never-treated units (if they exist), not-yet-treated and previously/future treated units since their contemporaneous treatment status does change for later switching cohorts.

Second, staggered treatment exposure timing may lead to different OLS weights applied to each DiD_w estimate, with weights a function of a unit's treatment exposure timing and duration. If the shock ATE is (1) constant after exposure, (2) treated units are shocked only once, and (3) DiD_w TWFE weights are independent of exposure timing, we would expect to see all parameter estimates of models 3-5 in Table 4 to converge. The pattern of results suggests that some or all of these assumptions may not hold requiring us to enrich the empirical strategy to pin down the immediate week-of-shock ATE and dynamics beyond the first week.[31]

We adopt a Mundlak-Wooldridge (MW) OLS estimator recently proposed by Wooldridge (2025) to explicitly model dynamic heat stress ATEs and its heterogeneity. It is now established that this estimator can provide numerically equivalent estimates to many popular estimators aimed at overcoming the failings of the classic TWFE estimator in the presence of staggered exposure and dynamic effect heterogeneity (Borusyak and Jaravel, 2018; de Chaisemartin and D'Haultfoeuille, 2020; Sun and Abraham, 2021; Callaway and Sant'Anna, 2021; Borusyak et al., 2024). Perhaps less well known is that it can produce numerically equivalent estimates to those obtained using the de Chaisemartin and D'Haultfoeuille (2024) estimator and its useful extensions: we show that this equivalence holds when including not-yet-treated units as controls and when one conditions on non-parametric time trends for dimensions that nest canton-year units. Because the MW estimator is OLS, it is straightforward to model ATE heterogeneity over fixed continuous covariates, i.e. the land Gini and additional modeling assumptions underlying the standard DiD robustness checks remain transparent. We incorporate the insights of de Chaisemartin and D'Haultfoeuille (2024) who show that ATE heterogeneity is a function of exposure paths when treatment is non-absorbing and rely on their estimator when non-parametric time trends have high dimensionality.[32]

To begin, define a series of treated cohorts that denote the period in which the canton-year unit is first shocked within year: dg where $g \in \{20, 21, \ldots, 27\} \equiv \mathcal{G}$ corresponds to the week w. For

---

[31] We describe the pitfalls in more detail and examine the OLS weights applied to the cell-level DiDs on which TWFE estimator is constructed in Section B.2.1 using the diagnostic tools of de Chaisemartin and D'Haultfoeuille (2024). Reassuringly, only 7.9% of the DiD cell estimates receive a negative weight, thus estimates in columns (1-3) in Table 4 are unlikely to be wrongly signed but may still not correspond to the target estimand in the presence of dynamic heterogeneity and multiple shock exposures.

[32] Their estimator can be much more efficient when either the number of staggered treated cohorts is large and/or the dimensionality of the nesting unit in which cell-level ATEs are calculated is high (such as the AR or canton). Clustered standard errors are slightly larger but still highly comparable to our OLS estimator's.

treated cohort g, we define treatment effects in week $w \in \mathcal{W} \setminus \{17, 18, 19\}$,

$$ATE_w(g) = GPP_w(g) - GPP_w(\infty), \quad g = w, \ldots, W, \;\; g \leq w$$

where $GPP_w(\infty)$ represents the potential outcome of the treated unit if not treated. The treatment effects we are searching to identify are the ATEs in periods where the cohorts are actually subjected to the shock:

$$\tau_{gw} = E[ATE_w(g) \mid dg = 1], \quad g = w, \ldots, G \;; \;\; w = g, \ldots, W \;; \; g \leq w.$$

Further defining a series of period indicators denoted by fw, $w \in \{17, \ldots, 32 = W\}$, our baseline dynamic-heterogeneity-robust event-study estimator writes as,

$$logGPP_{iwt} = \sum_{g}^{G} \left(\sum_{w=17}^{g-2} \theta_{gw} \quad \cdot dg_{it} \cdot fw_t + \sum_{w \geq g}^{W} \tau_{gw} \cdot dg_{it} \cdot fw_t + dg_{it} \right) + \sum_{w \in \mathcal{W} \setminus \{17\}}^{W} fw_t + u_{iwt}. \tag{1}$$

and can be estimated with pooled OLS (POLS) or a numerically and inferentially equivalent "Extended TWFEs" estimator where the individual dg and fw terms are replaced with canton x year and week dummies respectively. The model is saturated with all cohort x period, treated cohort and period dummies, but excludes reference terms $dg_{it} \cdot f(g-1)_t$ and period indicator $f17_t$, and, importantly, the never-treated cohort $d\infty_{it}$. This heterogeneity robust event study estimator simultaneously estimates all possible DiD estimates in parameter vector τ and pre-trend parameter vector θ using the never-treated units' contemporaneous evolution as the control outcome. The week-of-shock estimand for cohort g in this model is,

$$\begin{aligned}
\tau_{gw} = {} & E(logGPP \mid dg = 1, \; fw = 1) - E(logGPP \mid dg = 1, \; f(g-1) = 1) \\
& - \big(E(logGPP \mid d\infty = 1, \; fw = 1) - E(logGPP \mid d\infty = 1, \; f(g-1) = 1)\big), \quad g = w.
\end{aligned}$$

It represents the contemporaneous effect when first crossing the crop-specific threshold during the year. Modeling Gini heterogeneity is achieved by interacting all observed terms with our continuous measure or quantile dummies to create triple differences as formalized in equation B.8.

With estimates of the τ_{gw}'s (θ_{gw}'s) in hand, all that remains is to compute the treatment effect (pre-trend placebo) of interest as a convex combination of the estimates with the appropriate sub-population weights to obtain the average partial effects that corresponds to the ATE estimand of interest. See section B.3.2 for further details.

Table 5, column 1, presents week-of-shock effect estimates using the event study estimator with never-treated units serving as controls. On average, the shock causes a 16.4% drop in GPP in the first week. This is roughly 3 percentage points lower than the TWFE estimates. The estimated coefficient on the Gini interaction term (demeaned and standardized at the cohort dg level, "$d.st$"), indicates that a one standard deviation increase in land inequality exacerbates the immediate effect by about 1.9%, similar to the TWFE estimates.[33] Turning to non-parametric ATE heterogeneity in the lower panel, the ATE inter-quartile range is roughly 6.4 percentage points with the top quartile suffering a 19.1% GPP loss during the shock week.

In column 2 of Table 5, we can additionally allow not-yet-treated units to serve as controls by stacking the individual treated cohort-time DiD estimates and covariances with the appropriate sample restrictions for each treated cohort to reproduce this standard de Chaisemartin and D'Haultfoeuille (2024) estimator. We observe statistically identical results for the average shock effect and its Gini heterogeneity (see table notes and Section B.3.4 for precise details and why the equivalence does not hold with the way in which de Chaisemartin and D'Haultfoeuille (2024) model fixed linear heterogeneity). Column 3 shows estimates from our preferred event study model in which we condition on non-parametric canton-week trends (equation B.9). This will be our primary estimator for the remainder of the paper as it controls for how our most granular unit trends naturally week to week each season. Inference remains unchanged and point estimates remain consistent. On average, we estimate a week-of-shock ATE of 16.7% lower GPP flows with a one standard deviation increase in a cohort's Gini exacerbating the loss by 2.4% and an ATE Gini inter-quartile range of 8 percentage points.[34]

The Mundlak-Wooldridge model allows further flexibility in how one defines the control group. If one omits all pre-trend terms in equation B.8 we obtain a "lags only" estimator where the control mean is pooled over both treated cohorts in all pre-treatment periods and never-treated cohorts. We observe that the shock effect and linear Gini heterogeneity remain comparable but are larger in magnitude.[35]

Now that shock effect dynamics are explicitly modeled, we examine how the effects on GPP flows and stocks evolve before and after the initial shock using the event study estimator. In Figure 7a, we observe strong pre-trend assumption support. We then see that negative flow effects are slightly attenuated in week 2 ($w + 1$) and then highly attenuated in week 3 after the initial crossing. Interestingly, we see that GPP flows drop significantly again in week 4 and that the Gini ATE gradient inverts after which effects attenuate and Gini heterogeneity shrinks and becomes inconsistently signed.

Overall, we detect significant and rather complex dynamic ATE heterogeneity on weekly GPP flows. There is a strong detrimental effect over the first two weeks with additional negative

---

[33] Demeaning at the cohort is required for the average effect parameter to correspond to that when no heterogeneity is modeled.

[34] In Section B.3.4 we further document the MW parameterization that provides numerically equivalent estimates to the de Chaisemartin and D'Haultfoeuille (2024) estimator when conditioning on non-parametric unit time trends interacted with cohort-time dummies. While the interpretation of the estimand is straightforward over subsamples of fixed heterogeneity, i.e. Gini quartiles, the appropriate way to model linear heterogeneity that varies within the coarser nesting unit or not at all, i.e. the canton, is still an active area of applied econometrics.

[35] See Section B.3.3 for a discussion about the trade-offs in terms of bias versus efficiency compared to the event-study estimator. We also note that these estimates are closest in magnitude to those obtained using the classic TWFE estimator which also pools pre-treated and never-treated units into the control group.

effects in some subsequent weeks that vary in magnitude. It is important to note that the week-of-shock effect is most likely to best approximate the pure biological heat-stress response. ATEs beyond the shock week continue to capture dynamic biological heat-stress effects, but may be more and more likely to comprise unobserved farm-level behavioral responses endogenous to initial shock exposure. That said, an "all encompassing" estimand remains highly relevant.

We now turn to how impacts on GPP flows translate to effects on cumulative GPP, i.e. crop biomass stocks. Figures 7b and 7d provide estimates when substituting weekly flows with cumulative (log) GPP during the season as the dependent variable. We observe that heat stress shocks lead to persistent negative effects on the GPP stock that stabilize around -3.5% with a standard deviation increase in the Gini exacerbating the effect by roughly 0.4%. At week six post initial shock (w+5) – the last week with full cohort-week support – cropland in the fourth quartile suffers a cumulative loss of about 4% while the first quartile only a 2% drop.

Figure B.2b reproduces the main results but uses a single raw temperature threshold indicator in place of the generalized weighted stress threshold indicator. One might be concerned that the yearly readjusted cropland specific threshold is endogenous. We set this at 31.83°C which corresponds to the median temperature observed in the data (mean of 31.82°C) for all French cropland that crossed the weighted threshold at least once.[36] We observe a highly comparable pattern of results with somewhat noisier pre-trends. And we note that effects sizes on flows and stocks are also attenuated. This would be consistent with increased measurement error in treatment variation, particularly for crop compositions that have relatively low damage thresholds or vise versa.

We now provide a detailed exposition of shock exposure dynamics. Table B.3 displays frequencies and proportions of the 18 treatment paths observed in the data to clarify the primary source of exposure variation in the presence of possible repeated stress events. There are a total of 3,366 shocked cropland units and a total of 4,454 shocks. A large majority (74.6%) of treated units follow a treatment path where they are shocked only once during the year. The next most frequent path (14.9%) are shocked twice, with this second shock happening three weeks after first exposure. By column, we note that roughly 9% (298 units) receive two weeks of consecutive exposure after the first shock. The remaining rows/columns highlight that there exists a wide variety of exposure paths but that these more exotic paths provide relatively little identifying information due to their infrequency. Figure B.3 lets us further visualize the path's association with the land Gini. We observe evidence that average land inequality appears to be almost uniformly distributed across the paths.

Plots 7e and 7f reproduce the flow and stock impact estimates by Gini quartiles, but now the estimating sample iteratively restricts to treated units that do not receive a second shock. Under the assumption that additional heat stress shocks are exogenous to the initial shock this allows us to examine whether repeated shock exposure is determinant. The evidence does not strongly support this hypothesis. We see ensuing effects (post week-of-shock) on flows and stocks that are very similar to the baseline results. We do observe some attenuation and compression starting

---

[36] See Figure A.6 for the density estimate.

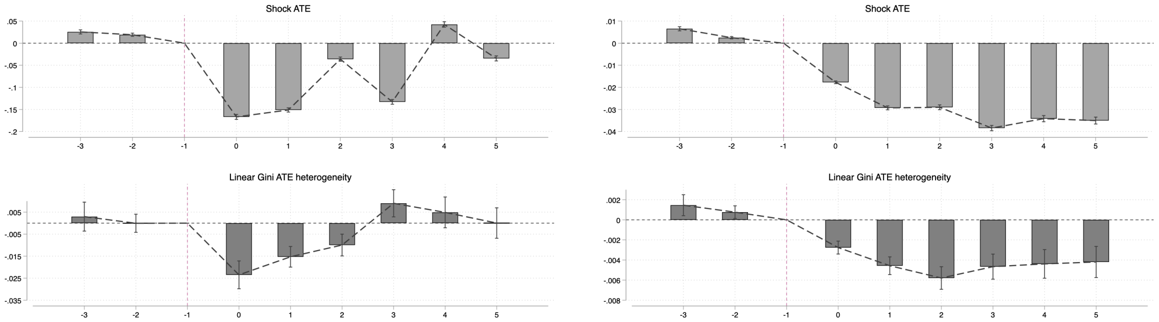
at week 4 for stock effects, but clearly the initial shock is the primary driver of effect dynamics on GPP stocks.

Table 5: Dynamic heterogeneity-robust week-of-shock ATE estimates

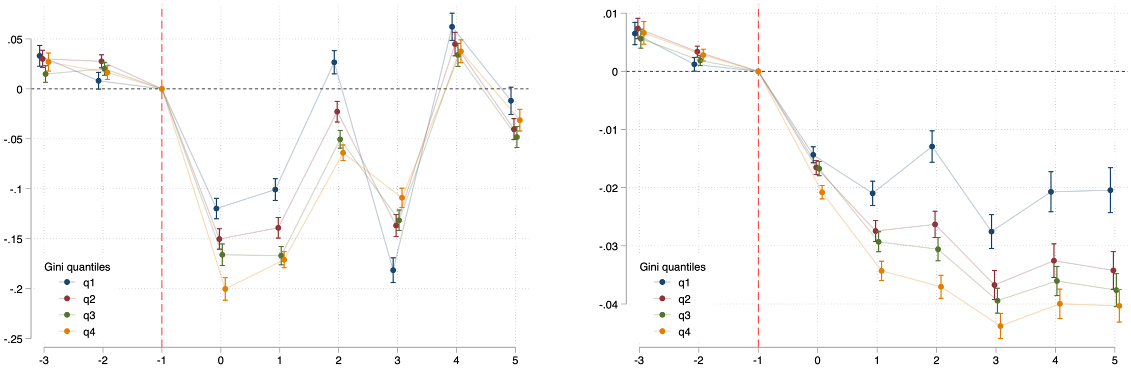
Dep. Variable: logGPP	Event study			Lags only
	(1)	(2)	(3)	(4)
Panel (a): ATE with linear Gini				
Shock	-0.1644***	-0.1619***	-0.1668***	-0.1905***
	(0.0027)	(0.0027)	(0.0028)	(0.0029)
Shock x Gini$^{d.st}$	-0.0194***	-0.0197***	-0.0235***	-0.0259***
	(0.0029)	(0.0029)	(0.0032)	(0.0031)
N	355024	191574	355024	355024
Control units: Never shocked	✓		✓	✓
Control units: Not yet shocked		✓		
Canton x Week FEs			✓	✓
Panel (b): ATE by Gini quantile				
q1	-0.1268***	-0.1253***	-0.1198***	-0.1312***
	(0.0053)	(0.0052)	(0.0053)	(0.0056)
q2	-0.1490***	-0.1466***	-0.1503***	-0.1794***
	(0.0051)	(0.0051)	(0.0052)	(0.0057)
q3	-0.1586***	-0.1564***	-0.1661***	-0.1903***
	(0.0052)	(0.0052)	(0.0055)	(0.0059)
q4	-0.1908***	-0.1885***	-0.2003***	-0.2256***
	(0.0052)	(0.0052)	(0.0058)	(0.0056)
p-val. equal ATEs	0.0000	0.0000	0.0000	0.0000
N	355024	191574	355024	355024
Control units: Never shocked	✓		✓	✓
Control units: Not yet shocked		✓		
Canton x Week FEs			✓	✓

Notes. OLS estimates of week-of-shock treatment effects. (a) All terms interacted with the the continuous Gini standardized and centered at treated cohort values. (b) All terms interacted with baseline Gini quartile indicators. (1-3) cohort treatment effects relative to the period prior to the cohort's shock week using the event study estimator. (2) Uses stacked coefficient and covariance matrices from separate OLS regressions for each cohort $g = w, \ldots, G$, using not-yet-treated cohorts $k > g$ and the never-treated panel as controls, restricted to periods fg_t and $f(g-1)_t$, with their Gini interactions. (3) conditions on non-parametric canton-week time trends (equation B.9). (4) the models only $\sum_g^G \sum_{w \geq g}^W dg_{it} \cdot fw_t$ terms and Gini interactions. See section B.3.3. Standard errors are clustered at the canton. * $p < .1$, ** $p < .05$, *** $p < .01$.

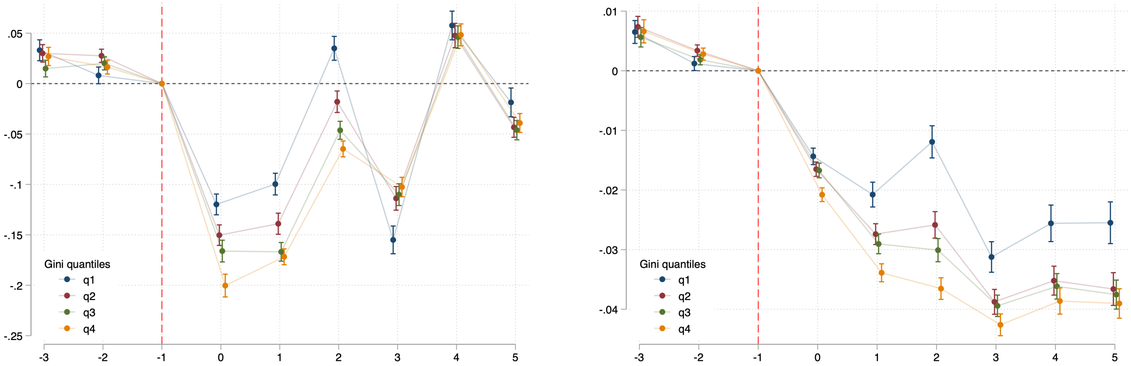
Figure 7: Dynamic heterogeneity-robust shock effect estimates on GPP flows and stocks



(a) Effects on flows (b) Effects on stock (year to date cumulative GPP)



(c) Effects on flows by Gini quantile (d) Effects on stock (year to date cumulative GPP)



(e) Shocked once flow effects (f) Shocked once stock effects

Notes. Placebo and treatment estimates on log GPP using event study estimator B.9 with in reference to treated cohorts' initial pre-shock week using linear Gini and Gini quartiles up to six weeks. Left column are flow effects, right cumulative GPP starting week 17.(e-f) effects isolating treated units that have not yet received second shock over w+1 to w+6 in Table B.3). Standard errors clustered at the canton. 95% confidence intervals. N=355,024

6 Land-use patterns and climate resilience

6.1 Diversification strategies and biological buffers

As highlighted in Section 2, the leading biological candidates that mitigate crop heat stress revolve around two channels: (1) modifying crop inputs in the production function to hedge against heat stress risk on yields and (2) natural ecosystem proximity and integration within plots. The yearly cadastral data on crop compositions and the proportion of arable land allocated to seminatural ecosystems allows test for biological moderating effects.

Figure 8a substitutes Gini quantiles with quantiles of the biological candidate mechanisms (again, fixed at 2015 values) in our preferred event study specification to estimate direct heat stress week-of-shock effect heterogeneity on GPP flows. We observe that ATEs by crop species diversity quantiles (measured in numbers of different species) provide no clear pattern. ATEs over the Herfindahl–Hirschman crop index (HHI) present a clearer pattern, with losses moderately declining in the crop HHI, on average. Extrapolating, this indicates that shock effects are stronger as we approach monoculture. The ATEs over the crop diversification ratio are increasing and exhibit a concave form, indicating that strategic crop portfolio selection does moderate heat stress losses but with decreasing effectiveness. Finally, natural ecosystem allocation clearly has the strongest positive moderating effect ceteris paribus. We see a monotonic ATE distribution over the quantiles of proportion allocated as a share of arable land. The bottom half of seminatural allocation exhibits roughly 20% week-of-shock losses on GPP flows while the top half suffers 12%, on average.

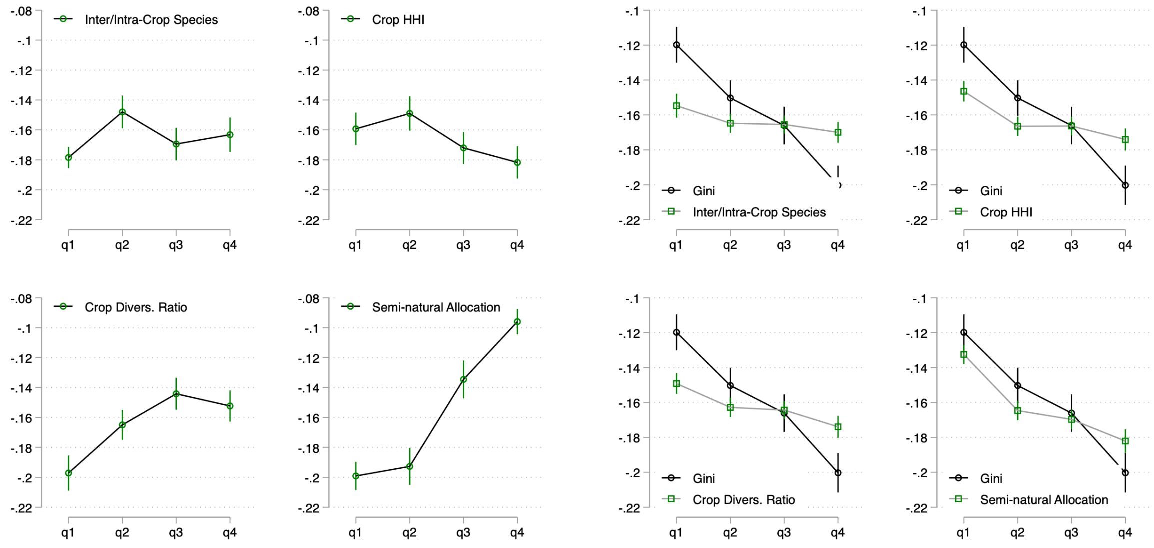
Figure 8b overlays the baseline Gini week-of-shock quantile ATEs with a composition-weighted prediction from each biological candidate model: for each Gini quantile, we compute the average predicted ATE using the empirical distribution of the candidate's quantiles among treated observations within that Gini quantile. Essentially, we ask whether the joint distribution of the biological candidate and the Gini can explain the gradient (see Appendix B.6 for details).[37]

We see that the crop diversity, concentration and even the risk-reducing portfolio candidates provide relatively little predictive power for the Gini impact gradient. But this is not the case for seminatural allocation. Its predicted ATEs at Gini quantiles maps much more tightly to the baseline Gini quantile effects. The binscatter illustrates the correlation between the level of seminatural allocation and the land Gini, conditional on AR x year and crop composition controls. In sum, the composition-weighted prediction that best reproduces the Gini gradient uses seminatural allocation: high-Gini cantons happen to cluster in low-seminatural quantiles, and low-seminatural quantiles have large negative ATEs.

---

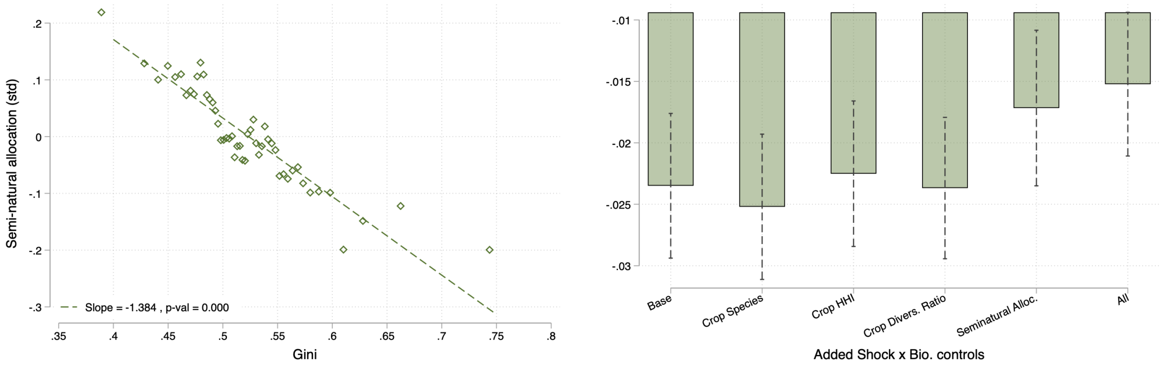
[37] Table B.6 also presents elasticity point estimates of different measures of crop diversity and the level of seminatural allocation with the Gini in panel (a) and partial elasticities with the proportion of land held farm scale in panel (b).

Figure 8: Impact heterogeneity through candidate biological mechanisms



(a) ATE by biological candidate quantiles

(b) Predicted impacts over Gini quantiles



(c) Seminatural - Gini CEF

(d) Linear Gini het. controlling for bio moderators

Notes. (a) Estimated shock ATEs by quantile of biological candidate mechanism corresponding to column 3 of Table 5. See Appendix E.1 for a spatial distribution of these variables. (b) Baseline estimated impacts by Gini quantiles overlaid with the predicted impacts using models in (a) at mean values for each Gini quantile. c) Seminatural allocation (standardized) over land Gini with a linear fit. (d) Comparing baseline week-of-shock ATE linear Gini heterogeneity with models that additionally condition on candidate mechanism shock heterogeneity. (a,b,d) N=355,034. (c) N=22,183. Standard errors clustered at the canton level. 95% confidence intervals.

Figure 8d shows how estimated linear Gini ATE heterogeneity changes when we add the shock interactions of the candidate biological mechanisms directly to the Gini event study estimator. Only seminatural allocation significantly attenuates the Gini ATE gradient – by 30% for the week-of-shock Gini heterogeneity – suggesting that a significant proportion of reduced shock resilience in high inequality farmland passes through its negative correlation with exposure to natural ecosystem services.

6.2 Service provision of seminatural ecosystems

Though reduced form, the evidence strongly points to the proportion of integrated seminatural ecosystems as a key mechanism behind land inequality's role in reducing heat stress resilience. These seminatural areas directly mitigate heat shocks through microclimate cooling effects on adjacent crops. Indeed, a very active literature highlights a series of, often overlapping, biotic and abiotic mechanisms underpinning the buffering of heat exposure on crops. Bramer et al. (2018) summarize that effects pass primarily through chnages in solar radiation, wind speed, humidity, evaporation and hydric stress.

Direct cooling effects pass through solar radiation absorption, wind breaks and the evapotranspiration of the seminatural areas themselves which contribute to differentiating plot microclimates from the surrounding meso- and macroclimate.[38]

The most compelling reduced form evidence on cooling effects come from recent work by Ghafarian et al. (2024). Using highly granular land surface temperature measurements (LST) in Germany during the hottest hours of the day over four years, they show that adjacent and integrated woodlots, groves and hedgerows on arable land have a detectable and significant cooling effects up to 75 meter radii into the plot during extreme heat shocks. Figure B.6 and accompanying Table B.7 reproduce their key findings: the cooling gradient and temperature variance at different distances from the seminatural area. Interestingly, they find that the strongest cooling effects are found on winter crop plots such as wheat. In their preferred specification they find cooling effects averaging -2.2 C°.

The main French crop is indeed winter wheat and its association with the Gini is the most elastic of all cereal crops, 0.74 in our context (Table C.2). Shock exposure happens in June, July and August. July shocks are by far the most frequent (Figure B.4) and present the strongest GPP week-of-shock losses (Figure B.5). Incidentally, July is when wheat yields are particularly vulnerable to heat stress in France because it is undergoing the final grain filling stage before August harvest.[39]

---

[38] Substantial amounts of water taken from deep root systems evaporates from the canopy leading to significant temperature decreases through evapotranspiration (Rolo et al., 2023). For example, an oak tree can release over a 1000 L. of water per day into the atmosphere while smaller birch trees release 75 L (see this source). Tree roots also move water passively from deeper and wetter layers to shallower and dryer soil layers along a gradient of soil water potential, a process called "hydraulic lift" (Liste and White, 2008). The water becomes available in crop topsoil to evaporate and be absorbed by crops shallow roots for crop transpiration. Their wind breaking effects reduce wind speed and preserve microclimates by reducing air mixing Brandle et al. (2004) and prevent top soil loss and erosion Donjadee et al. (2010). This can significantly increase water retention and soil moisture reducing hydric stress within the plot. Beyond these direct effects on heat stress mitigation, a plethora of indirect biological mechanisms related to nutrient cycling and transportation, the soil microbiome, pest moderation and pollinator health have been shown to be beneficial to yields and increase resilience to abiotic (such as heat shocks) and biotic stressors (Rolo et al., 2023). Although the evidence strongly points to seminatural allocation as positive in the face of climate change there can be yield costs related to interspecies competition for water, nutrients and light. See Mustafa et al. (2022) and Viaud and Thomas (2019) for discussion.

[39] Winter wheat's optimal photosynthesis temperature is approximately 25°C (Nagai and Makino, 2009; Lepekhov, 2022) and during the filling stage Wardlaw et al. (1989) estimates that a 1°C increase above the optimal temperature decreases wheat yield by 3–4%.

6.3 Validation against yearly crop yield declarations

We observe that (1) the negative effects of heat shocks on weekly cropland GPP are increasing in land inequality and (2) land inequality's association with amount of plot allocation to natural and semi-natural ecosystems is the strongest candidate sub-mechanism for which we have data. We now examine how impacts on GPP stocks map to impacts on actual year-end yields and whether this candidate biological mechanism is a factor in protecting marketable output.

The first three columns of Table 6 present estimates from a regression of (log) yearly declared yields at the *département* level on the yearly number of shock exposures of its nested cantons along with interactions with the centered and standardized land Gini and seminatural allocation share.

$$logYield_{k(i),t} = a+b\cdot Shocks_{ikt}+b_1\cdot(Shocks_{ikt}\cdot Gini^{st}_{ik})+b_2\cdot(Shocks_{ikt}\cdot Seminat^{st}_{ik})+c_{ik}+year_t+e_{k(i),t}$$

Because observations are at the canton-year level but yields only vary at the *département*-year level, b_1 captures how differential exposure to shocks to cantons within the same *département*-year maps to *département*-level yield declarations. It is thus an extrapolated "exposure-weighted aggregate" effect. With the fixed effects, identification comes from within-canton changes in shock exposure over years. Since,

$$\overline{Shock}_{kt} = N_k^{-1}\sum_{i\in k} Shocks_{it}$$

where N_k is the number of cantons in the *département*, b_1 measures the effect of a *département* going from having 0 shocked cantons to all (N_k) cantons shocked. We observe that this extrapolated effect is associated with a yearly yield loss of roughly 6.4% per hectare. Furthermore, we observe that land inequality and seminatural allocation have significant moderating effects on heat stress related yield losses: A one standard deviation increase in a *département*'s canton-level Gini exacerbates the yield loss by roughly 3.5% while the same change in seminatural allocation reduces the loss by roughly 3.3%. As with the granular weekly estimates, we see that controlling for the biological mechanism of seminatural allocation attenuates the inequality impact heterogeneity. Table B.4 tests the robustness to collapsing at the *département*-year level so that variation is harmonized for all covariates. We find highly comparable estimates and inference on effects.

Columns (4-6) presents estimates when scaling the predicted shock effect by the département-level share of shocked cantons: $\overline{Shock}_{kt} \cdot (\hat{b} \cdot Shock_{ikt} + \hat{b}_1 \cdot Shock_{ikt} \cdot Gini^{st}_{ik} + \hat{b}_2 \cdot Shock_{ikt} \cdot Seminat^{st}_{ik})$. This proxies for the average "dosage" effect, i.e. the estimated effect on *département*-level crop yields for an additional shocked canton. On average, an additional shocked canton associates with a 1.2% drop in *départemental* yields per hectare at mean levels of the covariates.

Columns 7-9 display stock ATE estimates that correspond to week six estimates in Figure 7b: effects on (log) cumulative GPP estimated using our dynamic heterogeneity robust OLS equation B.9. Model (7) thus corresponds to the last bars in Figure 7b. We observe that stock effects on weekly GPP and its heterogeneity are strongly predictive of yield losses but clearly do not map one-to-one. This is consistent with GPP capturing all crop biomass carbon production, not just the marketable part nor its water content at harvest.

Finally, Table B.5 reproduces results but regresses declared yields in levels. On average, the extrapolated department-level shock is associated with a 0.65 ton per ha. loss off a baseline yield of 8.4 t/ha. We also find strong support for nonlinearity in the moderating effects of both the Gini and seminatural allocation, an important point that we return to in Section 7 below.

These results also offer an important sanity check. Heat stress ATE heterogeneity on crop productivity is not driven by the GPP of the seminatural areas themselves. Indeed, the satellite data are granular, but not granular enough to exhaustively exclude all seminatural areas within plots. If seminatural area GPP is simply more resilient to heat shocks compared to adjacent cropland, then evidence for its protective biological effect could be biased due to measurement error, with bias inversely proportional to farm size. But the evidence using crop yields strongly rejects this concern. Rather, it supports the biological mechanism that proximity to natural ecosystem services are effective at mitigating heat stress effects. In sum, heat stress effects on crop GPP, its inequality gradient and, importantly, protective ecosystem moderation correlate consistently with effects on marketable yields. Furthermore, impacts on GPP may capture a lower bound on yield shock effects.

Table 6: Comparing with aggregate yields

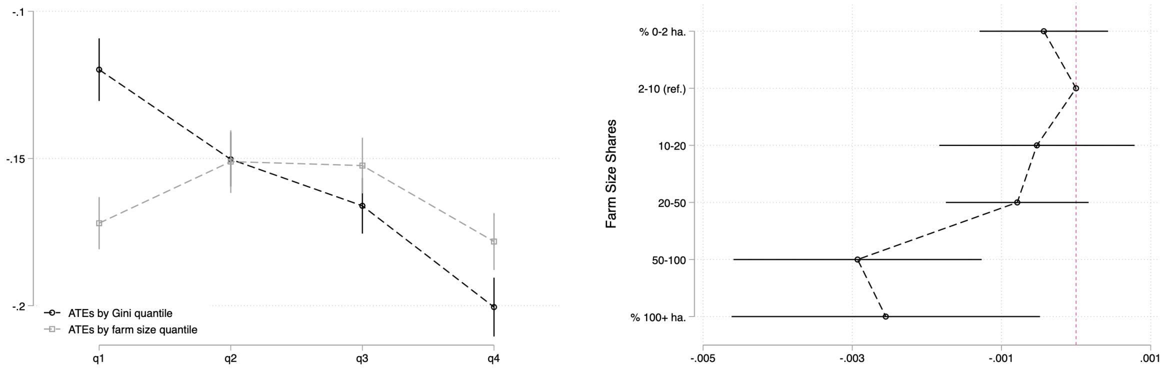
	Yearly log(Crop Yield (C.t/ha.))						log(Cum. GPP (C.t/ha.))		
	Department-level ATE			Dosage ATE per nested canton			Week 6 of exposure Het. Rob. ATE		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Shock	-0.0637***	-0.0519***	-0.0523***	-0.0121***	-0.0124***	-0.0126***	-0.0350***	-0.0350***	-0.0351***
	[0.0076]	[0.0065]	[0.0065]	[0.0016]	[0.0017]	[0.0015]	(0.0008)	(0.0008)	(0.0008)
Shock x Ginist	-0.0162***		-0.0142***	-0.0018**		-0.0012*	-0.0042***		-0.0030***
	[0.0060]		[0.0050]	[0.0009]		[0.0006]	(0.0008)		(0.0009)
Shock x Seminaturalst		0.0347***	0.0333***		0.0044***	0.0041***		0.0032***	0.0018**
		[0.0079]	[0.0084]		[0.0015]	[0.0016]		(0.0007)	(0.0008)
N	22183	22183	22183	22183	22183	22183	355024	355024	355024
Canton FEs	✓	✓	✓	✓	✓	✓			
Year FEs	✓	✓	✓	✓	✓	✓			

Notes. (1-3) Crop yields correspond to marketable food crops transformed as the log of metric tons per hectare for each French *département* for years 2015-2021 by the French Agricultural Ministry. (4-6) Yield canton dosage effects per *département* by scaling (1-3) by the proportion of shocked cantons for each year. (7-9) Dynamic heterogeneity robust event study estimates on (log) cumulative GPP at treatment week six (w+5) using our preferred specification (model 3 of Table 5. Model (7) estimates correspond to the last bars in Figure 7b. Standard errors in brackets indicate clustering at the *département* while those in parentheses indicate clustering at the canton. * $p < .1$, ** $p < .05$, *** $p < .01$..

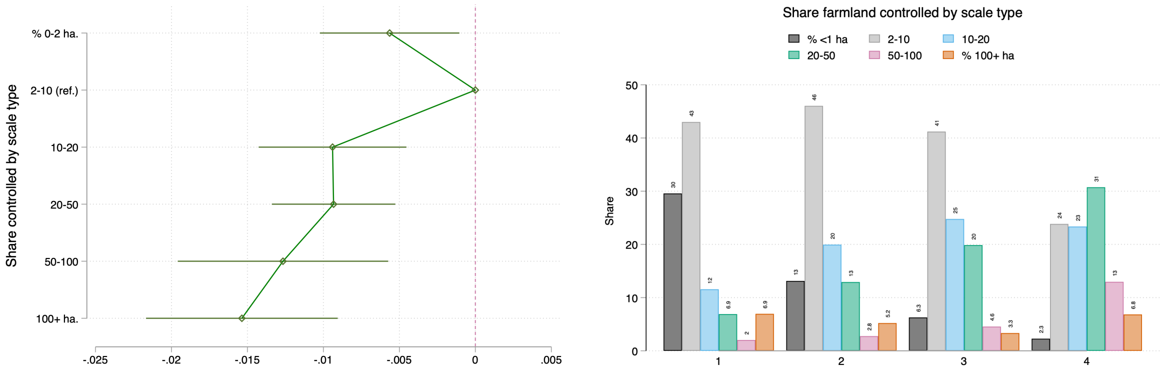
7 Production technology and biological buffers

We now provide a conceptual framework that can rationalize the empirical results with relatively few assumptions. Figure 9 illustrates the key patterns motivating the mechanism: the Gini ATE gradient reflects the distribution of farm-scale shares rather than average farm size (panels a-b), and this maps closely onto heterogeneity in seminatural allocation (panel c). Farm scale is key to understanding incentives that drive consolidation and farm-level decisions, but average scale may be an insufficient statistic when the ownership distribution is highly polarized. Average scale can mask multimodal distributions in which large farms exist alongside small operations (panel d).

Figure 9: Comparing shock impact heterogeneity over seminatural allocation and farm size types.



(a) Week-of-shock impacts quantiles of mean size (b) Week-of-shock impacts by size shares



(c) Seminatural - size share elasticities (d) Size share control over mean scale quantiles

Notes. N=355,034 (a) Dynamic heterogeneity-robust week-of-shock ATEs by Gini quartiles from equation overlaid with those using farm size quantiles. (b) Week-of-shock estimates using holdings share by size of canton cropland in reference to share held farms of 2-10 ha. (c) Partial elasticity of yearly seminatural allocation with respect to size shares conditioning on agricultural region x year FEs and crop composition. N=22,189. (d) Raw distribution of land holdings by scale over mean farm size quantiles with at least one farm at 100+ ha. N=3,176. Standard errors clustered at the canton. 95% confidence intervals.

If farm technology features scale economies, the per-hectare opportunity cost of a discrete reallocation of land away from natural ecosystem protection rises with farm scale. To our knowledge, the only parameter estimates of a developed agricultural production function come from Paul et al. (2004b,a) who find support for significant increasing returns to scale in the US.[40] Multiple studies in the development literature highlight that factor lumpiness and indivisibility, labor

---

[40] Paul et al. (2004b,a) use farm panel data in the US and show that plot size and other input elasticities are consistently above 1.

specialization, relative factor prices and machine scale effects create a transition point where scale economies "kick-in" (MacDonald and McBride, 2009; MacDonald, 2012; Rada and Fuglie, 2019; Foster and Rosenzweig, 2022). Indeed, Foster and Rosenzweig (2022) provide a plot-level analysis of how increasing operational size enables discrete shifts to scale economies with minimal assumptions on production function parameters since they are captured in the average cost curve. Our data do not allow for equivalent tests of returns to scale as in Paul et al. (2004b,a) or Foster and Rosenzweig (2022), but we build on their insights to argue that a modern agricultural production technology not only underpins the incentives for consolidation, but can also lead to lower climate resilience at the local agricultural landscape.

7.1 Incentives for seminatural allocation

Let $\beta \in [0,1]$ denote the share of arable land S a farmer allocates to protective ecosystem services which can mitigate yield losses under heat stress. $(1-\beta)S$ is then the amount of land devoted to revenue-generating crops. Output from cultivated land is given by

$$Y(S,\beta) = f((1-\beta)S). \tag{2}$$

Here, $f(x)$ denotes the farm's reduced-form output when x hectares are allocated to crop production. Detrimental heat stress arrives with probability p, so expected output is given by

$$Y(S,\beta,p) = \begin{cases} f((1-\beta)S), & \text{with probability } 1-p, \\ \alpha(\beta) f((1-\beta)S), & \text{with probability } p, \end{cases} \tag{3}$$

where $\alpha(\beta) \in (0,1]$ is an increasing and concave function of β that captures how protective allocation mitigates yield penalties under thermal extremes. The representative (risk-neutral) farmer chooses β to maximize expected marketable output Y, which we assume is proportional to revenue under a competitive market price:

$$\max_{\beta \in [0,1]} E[Y(S,\beta,p)]. \tag{4}$$

The expected yield is

$$E[Y(S,\beta,p)] = [(1-p) + p\alpha(\beta)] f((1-\beta)S). \tag{5}$$

The first-order condition (FOC) is

$$\begin{aligned} \frac{\partial}{\partial \beta}\mathbb{E}[Y(\cdot)] = (1-p)\Big[-f'\big((1-\beta)S\big)S\Big] \\ + p\Big[\alpha'(\beta) f\big((1-\beta)S\big) - \alpha(\beta) f'\big((1-\beta)S\big)S\Big] = 0, \end{aligned} \tag{6}$$

which rearranges to

$$\underbrace{f'\big((1-\beta)S\big)S\big[(1-p)+p\alpha(\beta)\big]}_{\text{Marginal cost}} = \underbrace{p\,f\big((1-\beta)S\big)\alpha'(\beta)}_{\text{Marginal benefit}}. \tag{7}$$

Here, $f'((1-\beta)S)$ is the marginal product of land at $(1-\beta)S$. The left-hand side is the marginal opportunity cost of allocating additional land to protection, while the right-hand side is the marginal private benefit from reducing shock losses on the remaining cultivated land. This condition determines the farmer's privately optimal protection share β^*. The level of β^* depends on both the production side, through $f(\cdot)$, and the protection side, through $\alpha(\cdot)$.

A positive allocation β arises when the initial mitigation gain from the first protected hectare is large enough relative to the initial production loss. More generally, an interior solution equates these two margins. A higher heat-shock probability p raises the value of protection, while a steeper $\alpha'(\beta)$, especially near $\beta = 0$, supports a larger privately optimal protected share. Concavity of α is useful because it makes the marginal benefit of protection decline with β, which in turn guarantees a unique interior solution when one exists.

Suppose that f exhibits increasing returns (IRS) over the relevant range of cultivated hectares. For discrete shifts in β, the true opportunity cost is the total foregone output $f(S)-f\big((1-\beta)S\big)$, which strictly exceeds the linear approximation $f'\big((1-\beta)S\big)\,\beta S$ under any IRS:

$$f(S)-f\big((1-\beta)S\big) \;>\; f'\big((1-\beta)S\big)\,\beta S.$$

Intuitively, $f'\big((1-\beta)S\big)\,\beta S$ underestimates the area under the marginal product curve from $(1-\beta)S$ to S (see Appendix D.1). For small β or nearly constant returns the gap is negligible, but for stronger IRS and/or larger discrete shifts it can be substantial.

The discrete opportunity cost $f(S)-f((1-\beta)S)$ is always increasing in S under IRS. But does the smooth privately optimal share β^* also depend on farm size? Rewriting the first-order condition in terms of the scale elasticity of production, $\varepsilon(x) \equiv xf'(x)/f(x)$, clarifies when it does (see Section D.3).

$$\varepsilon\big((1-\beta)S\big)\big[(1-p)+p\alpha(\beta)\big] = p\,\alpha'(\beta)(1-\beta). \tag{8}$$

This representation makes clear that, in its most general form, scale enters the optimality condition through the elasticity term $\varepsilon((1-\beta)S)$. If $\varepsilon(x)$ increases with cultivated area, then larger farms face a higher marginal opportunity cost of protection at a given β, implying a lower optimal β^*. Indeed, this is consistent with settings in which scale gains reflect threshold effects, lumpy technologies, or mechanization.

Figure 10a illustrates that larger operations incur increasingly larger discrete opportunity costs for any given allocation β, even when the scale elasticity is constant. This distinction is key: under a homogeneous technology, the smooth interior β^* does not depend on S, but the total foregone output from a discrete reallocation does. It is this discrete margin that interacts with the stepwise European Common Agricultural Policy (CAP) subsidy structure.[41]

This becomes clear in the next section in which we define the required per-hectare incentive to compensate the opportunity cost, $[f(S)-f((1-\beta)S)]/S$: it is strictly increasing in S under any IRS technology, thus flat per-hectare payments create a sharp operational scale cutoffs above which protective allocation is privately unprofitable.

7.2 Heat stress buffers and policy implications

The farmer's problem is most likely a discrete decision on how much β to maintain. Furthermore, the specific CAP subsidies available to French farmers aimed at increasing or maintaining the proportion of arable land allocated to natural ecosystems have payouts that are linear in S.

We begin by defining the required per-hectare incentive to induce positive allocation $\beta = \bar{\beta}$:

$$R(S,\bar{\beta}) \;\equiv\; \frac{f(S) - f\big((1-\bar{\beta})S\big)}{S}.$$

A threshold payment τ_{rate} (per hectare) yields adoption iff $\tau_{\text{rate}} \geq R(S,\bar{\beta})$. We can write

$$R(S,\bar{\beta}) \;=\; \frac{1}{S}\int_{(1-\bar{\beta})S}^{S} f'(x)\,dx.$$

Because f' is increasing under IRS, $R(S,\bar{\beta})$ is strictly increasing in S: for any uniform per-hectare payment, adoption declines with farm size.

The adoption frontier is the set

$$\mathcal{A}(\bar{\beta}) \;=\; \big\{(S,\tau_{\text{rate}}) \;:\; \tau_{\text{rate}} \geq R(S,\bar{\beta})\big\},$$

and the cutoff size $S^*(\tau_{\text{rate}})$ solves $R(S^*,\bar{\beta}) = \tau_{\text{rate}}$.

Adding heat stress (shock probability p, protection $\alpha(\beta)$), the per-hectare payment that makes the farmer indifferent between $\beta = 0$ and $\beta = \bar{\beta}$ is

$$\tau_{\text{rate}} \;\geq\; \underbrace{(1-p)\left[\frac{f(S)-f((1-\bar{\beta})S)}{S}\right]}_{R_{1-p}(S,\bar{\beta})\text{: foregone output, normal times}} \;+ \underbrace{p\left[\alpha(0)\frac{f(S)}{S} - \alpha(\bar{\beta})\frac{f((1-\bar{\beta})S)}{S}\right]}_{R_p(S,\bar{\beta})\text{: expected shock differential}}$$

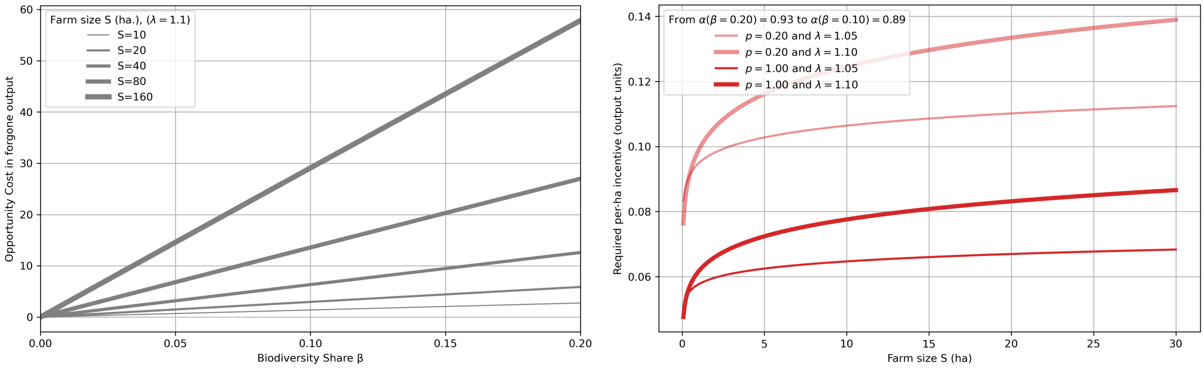
with $1 \geq \alpha(\bar{\beta}) > \alpha(0) > 0$. The first term is the forgone output per hectare in normal times weighted by $1-p$; under IRS it is increasing in S. The second term captures how shock mitigation affects the required payment: its sign is ambiguous and depends on the relative strength of baseline resilience and protection effectiveness.[42]

---

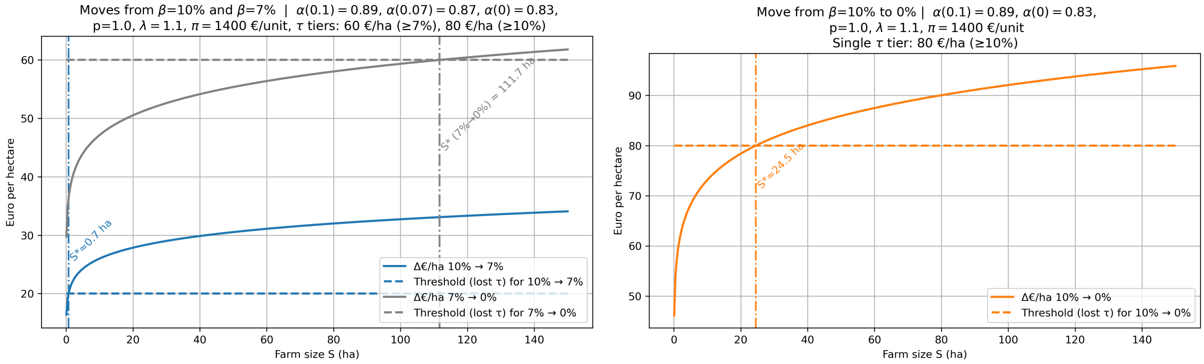
[41] If $\varepsilon(x)$ is constant, as under a homogeneous technology $f(x) = Ax^{\lambda}$, then farm size drops out and the optimal interior protection share is independent of S. It determines levels but not the share of arable land the farmer allocates. See Section D.3 and specifically D.3.1 for a detailed discussion.

[42] $R_p > 0$ (shocks raise the required payment) iff $\frac{\alpha(\bar{\beta})}{\alpha(0)} < \frac{AP(S)}{(1-\bar{\beta})AP((1-\bar{\beta})S)}$, where $AP(x) \equiv f(x)/x$. Intuitively, if the baseline is already quite resilient (high $\alpha(0)$), protection is only modestly better, so the payment must additionally cover the output shortfall on the remaining $(1-\bar{\beta})S$ hectares under heat stress. If instead the

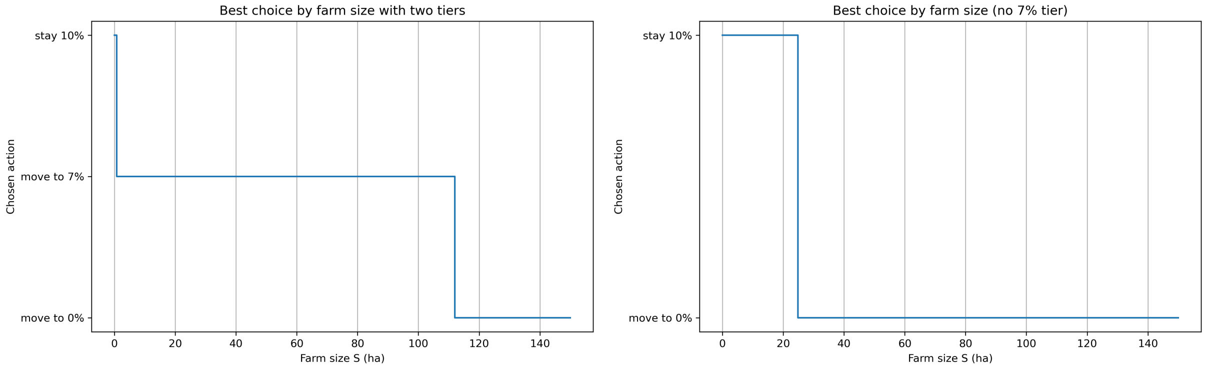
Figure 10: Opportunity costs, buffers and subsidy effects



(a) β allocation opportunity costs of foregone output (b) Required subsidies in output units



(c) Removal frontiers two tiers (d) Removal frontier one tier



(e) Decision map (f) Decision map one tier

Notes. (a) Simulation of $f(S) - f\big((1-\beta)S\big)$ with scaling parameter of 1.1 over allocation proportion β. (b-c) Removal thresholds over farm scale using estimates from B.5, model (3). (b) simulates removal thresholds with the two tiers, (c) simulates removal with only the 10% tier. (d-e) provide the corresponding decision maps.

The specific CAP subsidies available to French farmers aimed at incentivizing plot-level semi-natural allocation are a step function with two jumps, between and after which the payout is a simple linear function of farms total size S. Farmers receive 60€/ha if at least 7% of arable land

---

baseline is fragile (low $\alpha(0)$) and protection is highly effective (high $\alpha(\bar{\beta})$), the farmer would suffer large shock losses without protection, reducing the required subsidy.

is allocated to semi-natural ecosystems or 80 €/ha if 10% or above. Formally, let the subsidy be

$$\tau(\beta, S) \;=\; \begin{cases} 0, & \text{if } \beta < 0.07, \\ 60 \cdot S, & \text{if } 0.07 \leq \beta < 0.10, \\ 80 \cdot S, & \text{if } \beta \geq 0.10. \end{cases}$$

It is perhaps more intuitive and realistic to take the perspective that the majority of farmers' decisions are not whether to increase β but to reduce it. Imagine a farmer contemplating the removal of small woody features in the plot or shrinking a bordering hedgerow. This would increase the area of yield-producing crops and potentially increase labor productivity through machine scale effects.[43]

A farm that currently operates with a seminatural share $\bar{\beta}$ considers the discrete decision of reallocating some or all of $\bar{\beta}S$ to crops under the policy regime. Let risk and protection be summarized by

$$g(\beta) \equiv (1-p) + p\,\alpha(\beta).$$

Since the per-hectare payment $\tau_{\text{rate}}(\bar{\beta})$ is paid only if $\beta \geq \bar{\beta}$, conversion is optimal iff

$$\frac{g(0)f(S) - g(\bar{\beta})f((1-\bar{\beta})S)}{S} \;\geq\; \tau_{\text{rate}}(\bar{\beta}).$$

Here, the left-hand side equals the required per-hectare incentive derived above for adopting $\bar{\beta}$.[44]

With this set-up we can illustrate how land allocation decisions are jointly determined by shock frequency and protective effects captured in $g(\cdot)$ and subsidy payments $\tau(\cdot)$. In the following, farms of different sizes decide what to do with a baseline $\bar{\beta} = 10\%$ of seminatural allocation. There are three discrete decisions that derive from the policy payment thresholds: maintain, partial and total removal of $\bar{\beta}S$. We calibrate using impact and seminatural effect heterogeneity estimates on declared yearly yields presented in Table B.5, column 3. We use estimates for the estimated break-even price (€1400) for the representative farmer over the last five years using the report from the French agricultural consulting firm *Piloter sa ferme*.[45] The estimated protective effects at the policy thresholds are[46],

$$\hat{\alpha}(\beta) \;=\; \begin{cases} 0.83, & \text{if } \beta = 0, \\ 0.87 & \text{if } \beta = 0.07, \\ 0.89 & \text{if } \beta = 0.10. \end{cases}$$

---

[43] For example, a bigger harvester may now be optimal because of increased scale, broader passageways or increased plot contiguity.

[44] One could model conversion costs like a one-off $K(S)$ cost, but this does not change the comparative statics vis-à-vis increasing β, so we omit for clarity.

[45] Publicly available here.

[46] Concavity of α (i.e. $\alpha''(\beta) < 0$, consistent with the empirical estimates in Table B.5), ensures that the marginal benefit curve is downward-sloping in β while the marginal cost curve is upward sloping, guaranteeing a unique interior solution β^*. From Table B.5, $\hat{\alpha}(\beta) = (8.22 - 1.40 \cdot \mathit{Shock} + 5.03 \cdot \mathit{Shock} \cdot \mathit{Seminatural} - 4.81 \cdot \mathit{Shock} \cdot \mathit{Seminatural}^2)/8.22$

We start by simulating the price-free required payment (measured in output units) for the discrete choices $\bar{\beta} \in \{0.07, 0.10\}$ versus $\bar{\beta} = 0$ as a function of farm size S. We then illustrate how the CAP subsidies, given a market price, determine which farm sizes stay at $\bar{\beta} = 10\%$, move to $\bar{\beta} = 7\%$ or $\bar{\beta} = 0\%$. We assume shocks arrive with $p = 1$ and a scale parameter of $\lambda = 1.1$. Figure 10b shows that $R(S; \bar{\beta})$ is increasing in farm size for both $\bar{\beta} \in \{0.07, 0.10\}$. Under the calibrated power technology with $\lambda = 1.1$, the per-hectare measure $R \propto S^{\lambda-1} = S^{0.1}$ is concave; however, the total foregone output $R \cdot S = f(S) - f((1 - \bar{\beta})S) \propto S^{\lambda}$ grows super-linearly in S, consistent with the IRS technology. The shock-mitigation term,

$$g(\beta) = (1 - p) + p\,\alpha(\beta), \qquad \alpha(0) \approx 0.83,\ \alpha(0.07) \approx 0.87,\ \alpha(0.10) \approx 0.89,$$

shifts the requirement curves downward, increasing the set of farm sizes under the CAP payment thresholds. For each $\bar{\beta}$ the intersections with the horizontal CAP lines are economically meaningful thresholds: there exists a unique $S^*(\bar{\beta}, \tau)$ such that

$$\begin{cases} R(S; \bar{\beta}) \leq \tau & \Rightarrow \text{ policy is sufficient (no removal)} \quad \text{for } S \leq S^*(\bar{\beta}, \tau), \\ R(S; \bar{\beta}) > \tau & \Rightarrow \text{ policy is insufficient (removal pays)} \quad \text{for } S > S^*(\bar{\beta}, \tau). \end{cases}$$

Because $R(\cdot)$ rises in S while τ is size-invariant, the policy "bites" for small and mid-sized farms but loses traction for larger ones. Moreover, the $\bar{\beta} = 0.10$ curve lies far above the $\bar{\beta} = 0.07$ curve: given the common scale technology and shock frequency, the extra protection at 10% ($\alpha(0.10) > \alpha(0.07)$) is not large enough to offset the additional land diverted from crops. The right vertical axis maps output units to euros via the price calibrator π. Figure 10c illustrates the predicted size-removal thresholds in terms of profitability. It plots, for each discrete change in protection, the euro gain per ha. against the corresponding threshold loss per hectare. The shading shows the S-regions where each move is privately profitable.[47] A move is privately optimal iff $\Delta_{\text{€}/ha}(S; \beta_{\text{from}} \to \beta_{\text{to}}) \geq \mathcal{T}_{\text{€}/ha}(S; \beta_{\text{from}} \to \beta_{\text{to}})$. The gain term is proportional to $S^{\lambda-1}$, hence increasing and concave in S, while the threshold is flat in S. Each pair therefore has at most one crossing S^*, which we mark with a vertical line.

The two-tier system produces the ordering $S^*_{10\to7} < S^*_{7\to0}$. The step from 10% → 7% faces the smallest maintenance subsidy loss (€ 20/ha), so it becomes privately optimal at smaller S; the step from 7% → 0% has a much larger hurdle (€ 60/ha) and becomes profitable only for larger operations. Within the two-tier system, the full three-way ordering is $S^*_{10\to7} < S^*_{10\to0} < S^*_{7\to0}$: even though a direct jump 10% → 0% faces the largest loss hurdle (€ 80/ha), it delivers a larger output gain through the bigger recovered area, so $S^*_{10\to0}$ lies between the two partial thresholds. Figure 10d illustrates the removal threshold when there is no lower 7% tier: the only relevant

---

[47] This compares

$$\Delta_{\text{€}/ha}(S; \beta_{\text{from}} \to \beta_{\text{to}}) \;=\; \pi\big[g(\beta_{\text{to}})(1 - \beta_{\text{to}})^{\lambda} - g(\beta_{\text{from}})(1 - \beta_{\text{from}})^{\lambda}\big]\; S^{\lambda-1},$$

to the corresponding threshold loss per hectare,

$$\mathcal{T}_{\text{€}/ha}(S; \beta_{\text{from}} \to \beta_{\text{to}}) \;=\; \underbrace{\tau(\beta_{\text{from}}) - \tau(\beta_{\text{to}})}_{\text{lost maintenance}}.$$

margin is then 10% → 0%, showing how the lower tier creates option value that prevents full removal for a wide range of farm sizes. We discuss this in more detail in Section D.2 and its relation to fixed removal costs.

Three comparative statics are straightforward: (1) Raising the shock frequency p lowers $R(S;\bar{\beta})$ because the mitigation term $p\big[\alpha(0)\,f(S)/S - \alpha(\bar{\beta})\,f((1-\bar{\beta})S)/S\big]$ becomes more salient; the policy then bites at larger sizes; (2) steeper increasing returns (higher λ) shift $R(S;\bar{\beta})$ up and make it more curved, shrinking the set of farm sizes for which fixed, linear payments suffice; and (3) higher commodity prices (larger π) make a fixed per-hectare payment less generous in relative terms and shift crossings left; lower prices shift crossings right.

Figures 10e and 10f present the "best-action" map over farm size under the current two-tiered subsidies and when there is no lower tier. For each S, we compare the value of staying at the current tier versus moving to lower tiers and assign the action with the highest value (requiring a strictly positive gain to move). The breakpoints in the decision map align with the first and third S^* lines in Figure 10c derived from the inequality $\Delta_{\text{€}/ha} \geq T_{\text{€}/ha}$. The resulting piecewise-constant profile features three regions: those that stay at 10%; those that move to 7%; and those that go to 0%.

Two policy lessons emerge. First, stepwise, size-invariant payments inevitably create discrete margins in S: partial removal appears at smaller sizes than full removal, and the span of the "stay" region shrinks as either (1) increasing returns intensify (higher λ), (2) crop prices rise (larger π), or (3) shock frequency p falls (lower value of protection). Second, CAP subsidies trade off extensive and intensive margins: the 7% tier preserves protection for a larger set of farms (higher S^*), while the 10% tier delivers stronger mitigation α conditional on participation but loses traction sooner in S. The stepwise nature of $\tau(\cdot)$ also implies bunching at eligibility cutoffs: with the current parametrization no farm finds it optimal to set $\beta > 10\%$, small farms below $S^*(0.07, 60)$ face a relatively small loss from maintaining 7%, and farms above that threshold face growing incentives to drop to zero.

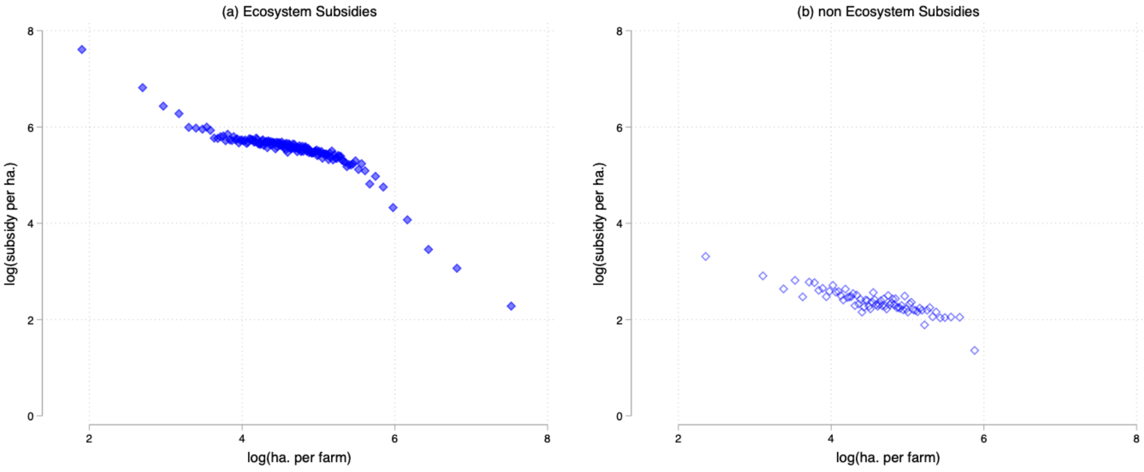
Figure 11 illustrates the relationship between local eco-subsidies received and mean farm scale in 2019. Though incomplete—only 16,220 communes are present in the data set for which one can isolate amounts for ecosystem preservation incentives—to our knowledge this 2019 cross-section is unique in linking ecological subsidy amounts received to the number and mean size of farms in a local area.[48] We observe an aggregate CEF with two kinks consistent with the predicted decision path of the farmer over operational scale: eco-subsidies initially fall in size, then flatten over a large continuum of mid-sized farms, and rise sharply in the upper tail of the size distribution. The second panel compares to all other subsidies received, which show a negative but much smaller and unkinked elasticity.[49]

---

[48] Overall, eco-subsidies account for over 83% of total PAC subsidies (10.3 billion euros) and 63% of the eco-subsidies correspond to eco-system maintenance and climate robust farming practices. The tiered surface subsidies account for 34%, specific "nature enhancing" bonuses for 19.6% and specific "climate support" farming practices for 4.7%.

[49] These data allow one to isolate more parsimonious variation at the département level and even the post-code level in which communes can be nested. When conditioning on these fixed effects, we find highly comparable empirical CEFs. See Figure E.2.

Figure 11: PAC subsidies disbursed to municipalities.



N bins: 152 and 75

Notes. We observe an aggregate CEF with two kinks consistent with the predicted decision path of the farmer over operational scale. N=16,220. Log subsidy amount received by communes over log mean farm size (SAU) of the commune in 2019. Optimal bin selection and partitioning follows the integrated mean squared error minimization algorithm of Cattaneo et al. (2024). Excludes subsidies for organic practices. Source: SAU et PAC par commune, 2019. Available here.

Because $R(\cdot)$ rises in S, this bunching is more pronounced in high-Gini landscapes where land acquisition opportunities for small farms are scarce because ownership is concentrated among relatively few large farms who have little incentive to scale down. This suggests that there may be alternative designs that scale payments with protected hectares (rather than total area) or index rates to market conditions π that might mechanically flatten the wedge between increasing payment requirements and linear subsidies, extending policy bite further up the size distribution. For example, the U.S. Conservation Reserve Program (CRP) compensates farmers according to the local rental value of the land they withdraw from production, estimated from county-level data on soil productivity and prevailing cash rents. In this policy setting, the subsidies may indeed better internalize the scale-dependent opportunity cost of foregone production. And like the CAP, eligibility depends on environmental criteria but contracts typically last 10 to 15 years rather than year to year.

7.3 Land inequality measures the level of market failure for climate resilience

The empirical evidence shows that regions with more unequal farmland exhibit lower average plant productivity and markedly larger output losses under heat stress. The model provides a simple organizing mechanism: under increasing returns to scale, the per-hectare opportunity cost of maintaining any positive protective share —$R(S, \bar{\beta}) = [f(S) - f((1-\bar{\beta})S)]/S$ – rises strictly with farm size. Under a flat per-hectare payment this creates a size cutoff S^*: farms with $S \leq S^*$ maintain $\beta = \bar{\beta}$ while larger farms optimally set $\beta = 0$, so the optimal protective share is effectively decreasing in S. In high-Gini farmland the size distribution shifts mass from small and mid-sized farms to large operations with a fringe of micro farms. And as we discuss in Section D.4, aggregate provision depends on how much land mass lies in the part of the scale distribution receiving low protection. If protection is decreasing in S, the area-weighted protection $B = \sum_i \beta_i S_i$ falls in inequality: large farms contribute little to ecological buffers despite large S_i, while small farms contribute proportionally more but the sum of their $\beta_i S_i$ may contribute relatively little to B. From this perspective, the "missing middle" lowers the landscape's protective base and amplifies vulnerability in hot years.

With protective spillovers, every farm under-invests in β relative to the social optimum because each internalizes only a subset of the spillover benefits. Even under a homogeneous technology where large farms need not have a larger gap in proportional protection, they can still account for a disproportionate share of missing protected hectares because each percentage-point shortfall applies to more of the farm-scape.

It follows that land inequality is a useful targeting statistic. It is a rich proxy for who controls the marginal hectares that are affected when conditions (p) or incentives (τ) change, as we highlight in Section D.4.1. High land inequality increases the probability that this margin is concentrated within a few large-scale decision-makers who face high private costs of β and benefit disproportionately from buffer spillovers. Whether the reduced form impacts over mean size and by share control illustrated in Figure 9 accurately capture the spillover effects of underinvestment by large farms remains an open empirical question. Yet a lack of exogenous variation should not deter further study of how inequality shapes incentives and policy impacts when farms are clustered in space.

Our data show that higher land inequality means more mass in the tails, especially in the upper tail. Local land Ginis summarize the margin concentration of at-risk hectares on high-cost decision-makers. In turn, it proxies for policy leverage. Flat per-hectare subsidies may have low leverage precisely where the marginal hectares are most in play. In high-Gini farmland, a few large farms currently sit on much of the margin where protective ecosystems are at high risk of conversion. It thus provides information on whether subsidies are effective at internalizing local spillovers to overcome the convex private costs of β and curb free-riding in the upper tail.

8 Conclusion

This paper documents that agricultural land inequality is associated with lower average crop productivity, and provides causal evidence that inequality amplifies the negative effects of heat stress on agricultural productivity. Our satellite-based GPP measure reveals that per-hectare biological productivity declines monotonically with land inequality, even in a modern economy where the economic case for consolidation rests on cost efficiency. Heat shocks damage the biological production function, not the cost structure: the margin that makes large farms economically efficient also makes their cropland climatically vulnerable.

We have shown that the presence of natural ecosystem integration plays an important role in mitigating the adverse effects of temperature shocks, highlighting the relevance of maintaining biodiversity in agricultural landscapes. Our results thus extend the traditional literature on farm size and productivity by testing key agrobiology mechanisms and the underlying incentives that shape farmland. A simple production technology with scale economies rationalizes these patterns: the per-hectare opportunity cost of maintaining protective ecosystems rises with operational scale, so flat per-hectare subsidies create sharp size cutoffs above which farms optimally abandon protection. These cutoffs shift with shock probability and subsidy levels, generating removal frontiers and decision maps that we calibrate to the EU's Common Agricultural Policy. The framework implies that policy must be both scale-aware and landscape-aware to recover protection where it is most effective.

Author Affiliations

Ignacio Flores: Paris School of Economics; Roma Tre (ignacio.flores@psemail.eu)
Dylan Glover: INSEAD (dylan.glover@insead.edu)

References

Abson, D. J., Fraser, E. D., and Benton, T. G. (2013). Landscape diversity and the resilience of agricultural returns: a portfolio analysis of land-use patterns and economic returns from lowland agriculture. *Agriculture & food security*, 2:1–15.

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6):1667–1697.

Altieri, M. A., Nicholls, C. I., Henao, A., and Lana, M. A. (2015). Agroecology and the design of climate change-resilient farming systems. *Agronomy for sustainable development*, 35(3):869–890.

Barrett, C. B. (1996). On price risk and the inverse farm size-productivity relationship. *Journal of Development Economics*, 51(2):193–215.

Barrett, C. B., Bellemare, M. F., and Hou, J. Y. (2010). Reconsidering conventional explanations of the inverse productivity–size relationship. *World development*, 38(1):88–97.

Beillouin, D., Ben-Ari, T., Malézieux, E., Seufert, V., and Makowski, D. (2021). Positive but variable effects of crop diversification on biodiversity and ecosystem services. *Global Change Biology*, 27(19):4697–4710.

Beillouin, D., Schauberger, B., Bastos, A., Ciais, P., and Makowski, D. (2020). Impact of extreme weather conditions on european crop production in 2018. *Philosophical Transactions of the Royal Society B*, 375(1810):20190510.

Benjamin, D. (1995). Can unobserved land quality explain the inverse productivity relationship? *Journal of Development Economics*, 46(1):51–84.

Berry, R. A., Cline, W. R., et al. (1979). *Agrarian structure and productivity in developing countries: a study prepared for the International Labour Office within the framework of the World Employment Programme.* Johns Hopkins Univ. Press.

Bhalla, S. S. and Roy, P. (1988). Mis-specification in farm productivity analysis: the role of land quality. *Oxford Economic Papers*, 40(1):55–73.

Bilal, A. and Känzig, D. R. (2024). The macroeconomic impact of climate change: Global vs. local temperature. Technical report, National Bureau of Economic Research.

Bilal, A. and Rossi-Hansberg, E. (2023). Anticipating climate change across the united states. Technical report, National Bureau of Economic Research.

Birthal, P. S. and Hazrana, J. (2019). Crop diversification and resilience of agriculture to climatic shocks: Evidence from india. *Agricultural systems*, 173:345–354.

Borrero, H. (2025). Land inequality, farm size, and productivity: Insights from peruvian agriculture. *Agricultural Economics*.

Borusyak, K. and Jaravel, X. (2018). *Revisiting event study designs*, volume 2826228. SSRN.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285.

Bramer, I., Anderson, B. J., Bennie, J., Bladon, A. J., De Frenne, P., Hemming, D., Hill, R. A., Kearney, M. R., Körner, C., Korstjens, A. H., Lenoir, J., Maclean, I. M., Marsh, C. D., Morecroft, M. D., Ohlemüller, R., Slater, H. D., Suggitt, A. J., Zellweger, F., and Gillingham, P. K. (2018). Chapter three - advances in monitoring and modelling climate at ecologically relevant scales. In Bohan, D. A., Dumbrell, A. J., Woodward, G., and Jackson, M., editors, *Next Generation Biomonitoring: Part 1*, volume 58 of *Advances in Ecological Research*, pages 101–161. Academic Press.

Brandle, J. R., Hodges, L., and Zhou, X. H. (2004). Windbreaks in north american agricultural systems. *Agroforestry systems*, 61(1):65–78.

Burgess, A. J., Cano, M. E. C., and Parkes, B. (2022). The deployment of intercropping and agroforestry as adaptation to climate change. *Crop and Environment*, 1(2):145–160.

Burke, M. and Emerick, K. (2016). Adaptation to climate change: Evidence from us agriculture. *American Economic Journal: Economic Policy*, 8(3):106–140.

Burke, M., Zahid, M., Martins, M. C. M., Callahan, C. W., Lee, R., Avirmed, T., Heft-Neal, S., Kiang, M., Hsiang, S. M., and Lobell, D. (2024). Are we adapting to climate change? Working Paper 32985, National Bureau of Economic Research.

Cadotte, M. W., Cardinale, B. J., and Oakley, T. H. (2008). Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences (PNAS)*, 105(44):17012–17017.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

Carletto, C., Savastano, S., and Zezza, A. (2013). Fact or artifact: The impact of measurement errors on the farm size–productivity relationship. *Journal of Development Economics*, 103:254–261.

Cattaneo, M. D., Crump, R. K., Farrell, M. H., and Feng, Y. (2024). On binscatter. *American Economic Review*, 114(5):1488–1514.

Chayanov, A. V. (1926). *AV Chayanov on the theory of peasant economy.* Manchester University Press.

Choueifaty, Y. and Coignard, Y. (2008). Toward maximum diversification. *The Journal of Portfolio Management*, 35(1):40–51.

Cornia, G. A. (1985). Farm size, land yields and the agricultural production function: An analysis for fifteen developing countries. *World development*, 13(4):513–534.

de Chaisemartin, C. and D'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, pages 1–45.

de Chaisemartin, C. and D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

Deltas, G. (2003). The small-sample bias of the gini coefficient: Results and implications for empirical research. *The Review of Economics and Statistics*, 85(1):226–234.

Di Falco, S. and Chavas, J.-P. (2009). On crop biodiversity, risk exposure, and food security in the highlands of ethiopia. *American Journal of Agricultural Economics*, 91(3):599–611.

Donat, M., Geistert, J., Grahmann, K., and Bellingrath-Kimura, S. D. (2023). Orientation of tree rows in alley cropping systems matters – the "shadot" modelling tool for tree growth and shading effects. *MethodsX*, 11:102282.

Donjadee, S., Clemente, R., Tingsanchali, T., and Chinnarasri, C. (2010). Effects of vertical hedge interval of vetiver grass on erosion on steep agricultural lands. *Land Degradation & Development*, 21(3):219–227.

Duffy, J. E., Godwin, C. M., and Cardinale, B. J. (2017). Biodiversity effects in the wild are common and as strong as key drivers of productivity. *Nature*, 549(7671):261–264.

Ellis, F. (1993). *Peasant economics: Farm households in agrarian development*, volume 23. Cambridge University Press.

Feder, G. (1985). The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints. *Journal of development economics*, 18(2-3):297–313.

Foster, A. D. and Rosenzweig, M. R. (2022). Are there too many farms in the world? labor market transaction costs, machine capacities, and optimal farm size. *Journal of Political Economy*, 130(3):636–680.

Ghafarian, F., Ghazaryan, G., Wieland, R., and Nendel, C. (2024). The impact of small woody features on the land surface temperature in an agricultural landscape. *Agricultural and Forest Meteorology*, 349:109949.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of econometrics*, 225(2):254–277.

He, M., Kimball, J. S., Maneta, M. P., Maxwell, B. D., Moreno, A., Beguería, S., and Wu, X. (2018). Regional crop gross primary productivity and yield estimation using fused landsat-modis data. *Remote Sensing*, 10(3):372.

Kremen, C. and Miles, A. (2012). Ecosystem services in biologically diversified versus conventional farming systems: benefits, externalities, and trade-offs. *Ecology and society*, 17(4).

Lamb, R. L. (2003). Inverse productivity: Land quality, labor markets, and measurement error. *Journal of Development Economics*, 71(1):71–95.

Langel, M. and Tillé, Y. (2013). Variance estimation of the gini index: revisiting a result several times published. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 176(2):521–540.

Lepekhov, S. (2022). Canopy temperature depression for drought-and heat stress tolerance in wheat breeding. *Vavilov Journal of Genetics and Breeding*, 26(2):196.

Lipton, M. (2009). *Land reform in developing countries: Property rights and property wrongs.* Routledge.

Liste, H.-H. and White, J. C. (2008). Plant hydraulic lift of soil water–implications for crop production and land restoration. *Plant and Soil*, 313(1):1–17.

Lowder, S. K., Bhalla, G., and Davis, B. (2025). Decreasing farm sizes and the viability of smallholder farmers: Implications for resilient and inclusive rural transformation. *Global Food Security*, 45:100854.

Lowder, S. K., Sánchez, M. V., and Bertini, R. (2021). Which farms feed the world and has farmland become more concentrated? *World Development*, 142:105455.

Lowder, S. K., Skoet, J., and Raney, T. (2016). The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Development*, 87:16–29.

Lüttger, A. B. and Feike, T. (2018). Development of heat and drought related extreme weather events and their effect on winter wheat yields in germany. *Theoretical and Applied Climatology*, 132(1):15–29.

MacDonald, J. (2012). Why are farms getting larger? the case of the us. *Proceedings "Schriften der Gesellschaft für Wirtschafts-und Sozialwissenschaften des Landbaues eV"*, 47(874-2017-1206):25–46.

MacDonald, J. M. and McBride, W. D. (2009). The transformation of us livestock agriculture scale, efficiency, and risks. *Economic Information Bulletin*, 1(43).

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77—-91.

Morawetz, U. B. and Klaiber, H. A. (2025). Regression analysis with independent variables in shares: a guide and an empirical example. *Empirica*, 52(1):63–98.

Moscona, J. and Sastry, K. A. (2023). Does directed innovation mitigate climate damage? evidence from us agriculture. *The Quarterly Journal of Economics*, 138(2):637–701.

Moscona, J. and Sastry, K. A. (2025). Inappropriate technology: Evidence from global agriculture. Technical report, National Bureau of Economic Research.

Mustafa, M., Szalai, Z., Divéky-Ertsey, A., Gál, I., and Csambalik, L. (2022). Conceptualizing multiple stressors and their consequences in agroforestry systems. *Stresses*, 2(3).

Nagai, T. and Makino, A. (2009). Differences between rice and wheat in temperature responses of photosynthesis and plant growth. *Plant and Cell Physiology*, 50(4):744–755.

Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., and Lobell, D. B. (2021). Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change*, 11(4):306–312.

Paul, C., Nehring, R., Banker, D., and Somwaru, A. (2004a). Scale economies and efficiency in us agriculture: are traditional farms history? *Journal of Productivity Analysis*, 22(3):185–205.

Paul, C. J. M., Nehring, R., and Banker, D. (2004b). Productivity, economies, and efficiency in us agriculture: a look at contracts. *American Journal of Agricultural Economics*, pages 1308–1314.

Piet, L. and Cariou, S. (2014). Le morcellement des exploitations agricoles francaises. *Economie rurale*, 342(4):107–120.

Rada, N. E. and Fuglie, K. O. (2019). New perspectives on farm size and productivity. *Food Policy*, 84:147–152.

Renard, D. and Tilman, D. (2019). National food production stabilized by crop diversity. *Nature*, 571(7764):257–260.

Rolo, V., Rivest, D., Maillard, E., and Moreno, G. (2023). Agroforestry potential for adaptation to climate change: A soil-based perspective. *Soil Use and Management*, 39(3):1006–1032.

Running, S. and Zhao, M. (2019). Mod17a2hgf modis/terra gross primary productivity gap-filled 8-day l4 global 500m sin grid v006 (data set). *NASA EOSDIS Land Processes DAAC.*

Schlenker, W. and Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to us crop yields under climate change. *Proceedings of the National Academy of sciences*, 106(37):15594–15598.

Sen, A. K. (1962). An aspect of Indian agriculture. *Economic Weekly*, 14(4-6):243–246.

Sen, A. K. (1966). Peasants and dualism with or without surplus labor. *Journal of political Economy*, 74(5):425–450.

Seo, S. N. (2010). A microeconometric analysis of adapting portfolios to climate change: adoption of agricultural systems in latin america. *Applied Economic Perspectives and Policy*, 32(3):489–514.

Srinivasan, T. (1972). Farm size and productivity implications of choice under uncertainty. *Sankhyā: The Indian Journal of Statistics, Series B*, pages 409–420.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics*, 225(2):175–199.

Tamburini, G., Bommarco, R., Wanger, T. C., Kremen, C., Van Der Heijden, M. G., Liebman, M., and Hallin, S. (2020). Agricultural diversification promotes multiple ecosystem services without compromising yield. *Science advances*, 6(45):eaba1715.

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, 418(6898):671–677.

Valdivia, C., Dunn, E. G., and Jetté, C. (1996). Diversification as a risk management strategy in an andean agropastoral community. *American Journal of Agricultural Economics*, 78(5):1329–1334.

Viaud, V. and Thomas, Z. (2019). Une réflexion sur l'état des connaissances des fonctions du bocage pour l'eau dans une perspective de mobilisation pour l'action. *Sciences Eaux & Territoires*, 30(4):32–37.

Vogel, E., Donat, M. G., Alexander, L. V., Meinshausen, M., Ray, D. K., Karoly, D., Meinshausen, N., and Frieler, K. (2019). The effects of climate extremes on global agricultural yields. *Environmental Research Letters*, 14(5).

Vollrath, D. (2007). Land distribution and international agricultural productivity. *American Journal of Agricultural Economics*, 89(1):202–216.

Wardlaw, I., Dawson, I., and Munibi, P. (1989). The tolerance of wheat to hight temperatures during reproductive growth. 2. grain development. *Australian Journal of Agricultural Research*, 40(1):15–24.

Webber, H., Lischeid, G., Sommer, M., Finger, R., Nendel, C., Gaiser, T., and Ewert, F. (2020). No perfect storm for crop yield failure in germany. *Environmental Research Letters*, 15(10):104012.

Wooldridge, J. M. (2025). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators: Jm wooldridre. *Empirical Economics*, pages 1–43.

Online Appendix

A Data and measurement

This section describes the three major data sources used in our empirical analysis in more detail: satellite imagery, agricultural cadastral ownership records, and weather estimates. We also quickly review methods used to construct concentration measures of both farmland and biological diversity, as well as the definition of temperature shocks used in our empirical assessment. Figure A.1 displays data samples, the following subsections elaborate on each source individually.

A.1 Productivity data

Satellites can detect photosynthesis, the fundamental process of plant growth. Some plant cells –usually leaf-cells– use the sun's energy to split CO_2 molecules around them in two parts. They keep carbon (C), which is mixed with other matter to build their own mass; and they release oxygen (O_2) back into the air as a byproduct. Such processes leave an invisible fluorescent signature track, which some satellites can measure. That is the case of the NASA Terra satellite, thanks to the MODIS remote sensor device launched in 2000. It measures the gross primary productivity (GPP) of plants, which is a generic measure of biomass production, gross of plant respiration, on a 500m x 500m grid for the entire globe. Cloud-coverage adjusted measures of cumulative production are provided by Running and Zhao (2019) at a quasi-weekly (8-day) frequency. These provide our main crop growth estimates with out baseline expressed as kilos of carbon per square meter (C.kg/m^2) for 321 dates over seven years. Figure A.1a showcases total cumulated production for a given year, with each pixel representing a data point.

Table A.1: GPP to Yield conversion factors, examples

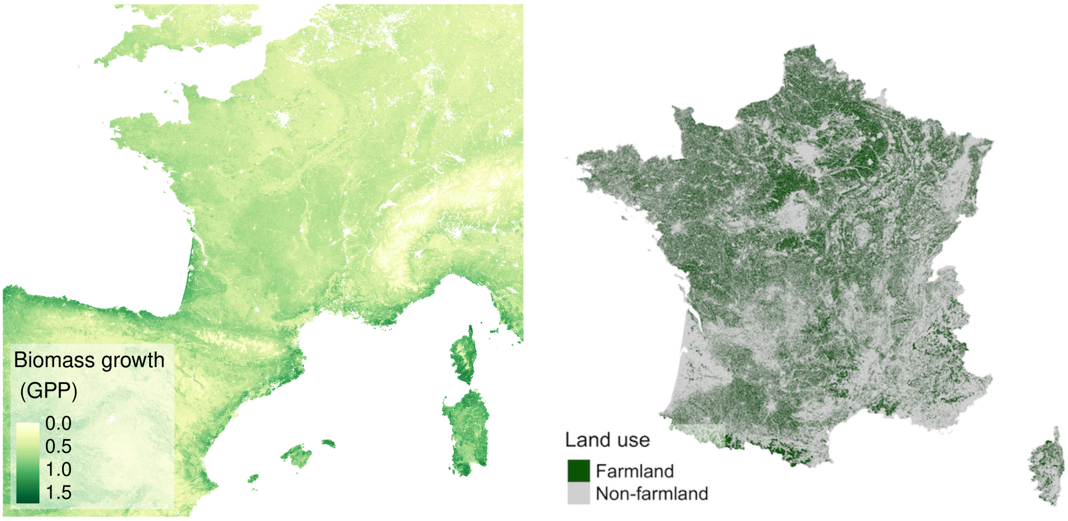
Crop	Factor
Alfalfa	0.55
Barley	0.42
Maize	0.44
Durum wheat	0.22
Peas	0.28
Spring wheat	0.24
Winter wheat	0.35

Notes. By He et al. (2018) for annual yield of staple crops in Montana, USA

A strong advantage of using the primary measure of crop biomass production is that it lets one abstract from price effects. Furthermore, the recent agro-biology literature shows that these remote sensor GPP measurements are highly correlated with marketable yields. Table A.1 displays conversion factors between satellite estimates of GPP and actual agricultural yield of crops from the US state of Montana by combining ground-level observations with the satellite estimates. Unfortunately, this literature is nascent and comparable conversion estimates over larger crop diversity lacking. We thus contribute to this measurement validation literature by providing yield-GPP elasticity estimates at the yearly level. This allows us to provide some clarity on how granular impacts on GPP map to impacts on economic value.

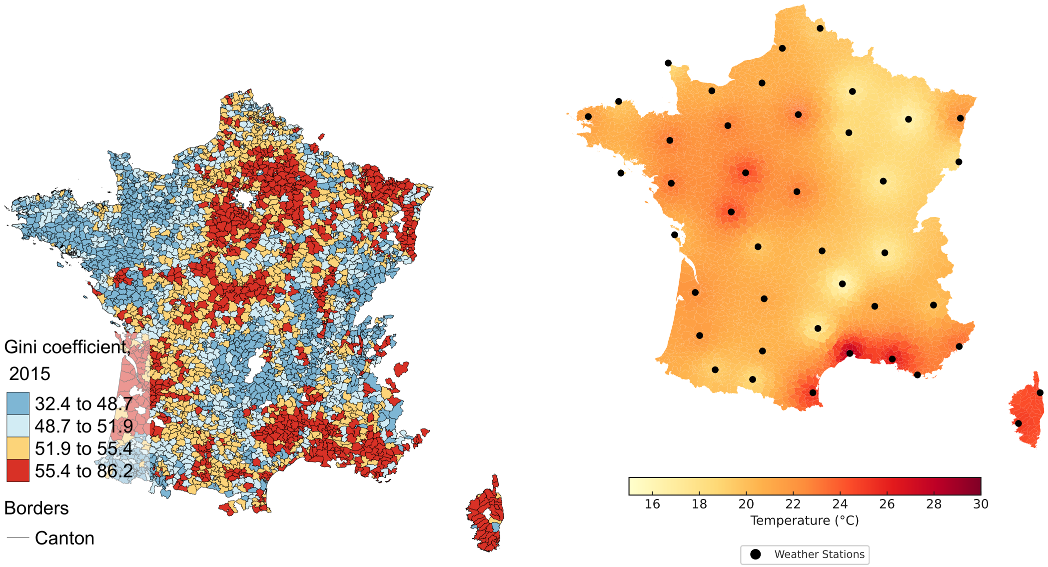
To isolate the part of the GPP layer that is relevant to our subject of study, we use exhaustive cadastral information to extract the slices of data that overlap with productive cropland, avoiding the contamination of our estimates with forest productivity, for instance. Figure A.2 provides

Figure A.1: Main inputs: high definition geo-located information



(a) Yearly land productivity, 2021

(b) Exhaustive farm information



(c) Land Gini coefficient

(d) Midsummer temperatures, 2021 (ºC)

Notes. A.1a: satellite data cover the whole national territory, allowing for high-frequency measurements of land productivity. A.1b: cadastral information precisely locates all farms in the French territory with exhaustive information on crop allocation. A.1c: showcases one of our benchmark farmland concentration measures, based on the cadastral data and aggregated to the canton level. A.1d: displays average temperatures over August as measured by official French weather stations, interpolated with the kriging method (more details in section A.6).

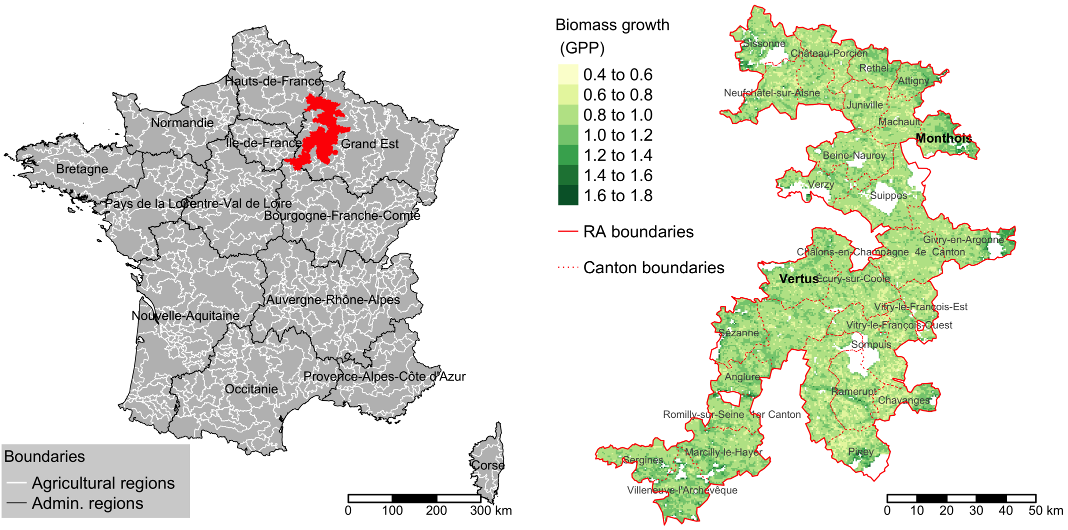
a visual representation of such overlap. The cadastral data is described in more detail in the following subsection.

A.2 Cadastral data

Almost all French farmers must declare information on land use to the national authority, including the exact crop allocation within agricultural plots (figure A.1b). Such declaration is necessary for them to receive annual subsidies from the European Union's Common Agricultural Policy. We obtain the administrative data from the Graphic Parcel Register dataset (RPG, for its acronym in French) provided jointly by the National Geographic Institute (IGN) and the Services and Payment Agency (ASP) of France. For every year the dataset reports the precise location of each exploitation, with detailed information of land use. We use the data between 2015-2021 for our analysis, despite the series starting five years earlier, because previous years contain information only on the dominant crop within plots instead of an exhaustive declaration. The only farmers that are excluded from the database are subject to some selective denominations of winemaking which don't benefit from subsidies.

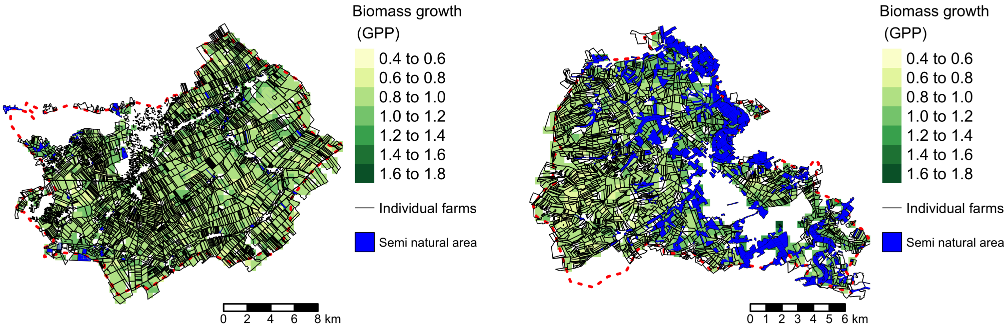
An important conceptual distinction arises when interpreting land organization through administrative data. In the French RPG, each land block (or *ilot* in French) represents a contiguous area of farmland managed by a single administrative entity –a legal or natural person– and may itself contain several parcels cultivated with different crops. An *exploitation*, in contrast, refers to the entire set of such land blocks managed under one administrative unit. The exploitation is the level at which the average farm size is typically measured in census and survey data, even though each may consist of several land blocks that are spatially dispersed, sometimes across large distances. Analyses of the RPG, such as Piet and Cariou (2014), indicate that French exploitations contain on average about 14 land blocks, each covering roughly 5 hectares, for a total mean size of around 70 hectares. This distinction highlights the difference between administrative consolidation, where dispersed plots are managed under a single entity, and physical land consolidation, which refers to the actual spatial regrouping of farmland. Our empirical analysis focuses on land blocks, thereby addressing issues of physical land consolidation rather than administrative concentration.

Figure A.2: Illustrative relation between land concentration, semi natural areas, and productivity



(a) Champagne-Crayeuse (Agricultural region)

(b) Yearly production in 31 cantons



(c) **Vertus:** Canton with high Gini, and low semi-natural areas

(d) **Monthois:** Canton with low Gini, and high semi-natural areas

Notes. (a) Highlighted in red is Champagne-Crayeuse, one of 432 agricultural regions nationwide, located within the broader and better known Champagne area. (b) The region is divided into 31 cantons, the most granular unit of observation in our panel, overlaid with satellite-measurement of yearly gross primary productivity cropped to farmland area (c, d) Two illustrative cantons are shown in detail, with overlaid parcel configurations (black lines) and presence of semi-natural areas (blue): (c) Vertus: Gini (62.2), Average farm size (6.2 ha.), Semi natural areas (1.3% of farmland) (d) Monthois: Gini (44.8), Average farm size (5.7 ha.), Semi natural areas (21.5% of farmland). These panels are illustrative of the tendency of higher land inequality (Gini) being strongly and negatively related to the presence of semi natural areas, which are also related to higher land productivity and resilience.

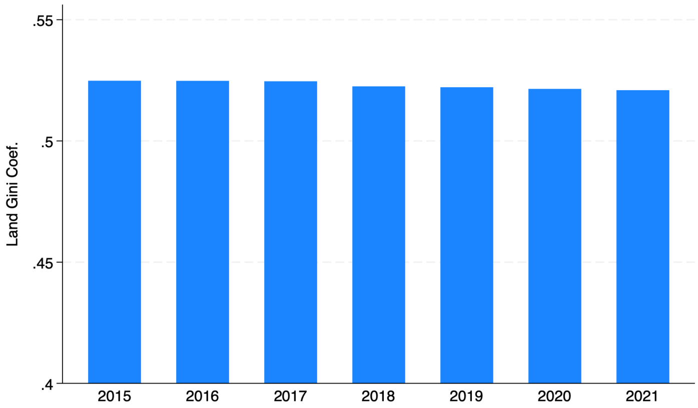
Figure A.2 provides an appreciation of the granularity of the cadastral data. In the background layer, one can notice productivity measurements coming at a coarser resolution, which explains why individual farm analysis is not possible with this data. The unit of observation of the cadastral data are productive agricultural plots controlled by a single person or entity, regardless of whether they control more than one, since they are recorded with a different id. As a consequence, our land inequality estimates probably represent a lower bound of land consolidation.

We use cadastral information to estimate yearly farmland concentration in both terms of land inequality, crop compositions and biological ecosystem diversity.

A.3 Average Gini

We estimate a mean Gini of .523 for our seven year panel with a standard deviation of sd=0.061 and an inter-quartile range of 0.066. It varies little over each year as can be seen in Figure A.3. The within canton standard deviation is 0.0068. We thus exploit between canton variation to study the Gini's association with GPP and yields while conditioning on relevant canton-nesting geographic units, such as the AR (e.g. the within AR Gini sd is 0.035) to make the correlational results as informative as possible. For our causal analysis, we use it as a fixed dimension of heterogeneity of the canton by which (conditionally) exogenous temperature exposure can moderate crop productivity. For these temperature impacts, we use the 2015 Gini to keep interpretation as clean as possible.

Figure A.3: Average Gini



Note. Mean canton cropland Gini by year built using farm-level cadastral data.

A.4 Crops and seminatural ecosystems

Inter- and intra-crop diversity are assessed by treating each crop as a distinct unit within the canton. For this, we use a simple count of crop types as well as the Herfindahl-Hirschman Index (HHI), a commonly used measure of market concentration.

In addition to examining crop diversity through simple counts and concentration indices, we explore the concept of crop diversity as portfolio diversification. For this purpose, we adopt the diversification ratio (D) as defined by Choueifaty and Coignard (2008). This ratio is calibrated using two decades of historical data on national crop yields. The diversification ratio is the ratio of the weighted average variance (WAV) to the overall portfolio variance (OPV). The WAV captures the variance within individual crop yields, while the OPV accounts for how these variances synchronize, incorporating their covariances. A higher diversification ratio indicates that the crop composition planted minimizes covariance among crops, signifying greater diversification. Further technical details on the calculation and calibration of the diversification ratio are provided in Appendix A.5.

To account for the presence of semi-natural areas, we measure their size as a percentage of the total farmland within each canton, where farmland excludes all semi-natural areas. These systems are patchy or linear structures of woody vegetation, such as solitaires, groves, woodlots, thickets, hedgerows, bocages or gallery forests and prairies.

In rare instances where semi-natural areas exceed the farmland area, this percentage can exceed 100%, highlighting regions where semi-natural landscapes –excluding fully natural areas like forests– are particularly prevalent.

A.5 The diversification of a crop portfolio

In our financial analogy, we draw upon the foundational principles of modern portfolio theory, pioneered by Markowitz (1952). This theory operates on a simple premise: for a given level of risk, higher yields are always preferable. Here, expected yields are represented by the expected value of a variable, while risk is quantified through its variance. A diversified portfolio is a means to mitigate risk. By diversifying our assets, or in this case, crops, we aim to reduce the overall level of risk. This is achieved by selecting assets with uncorrelated yields, or even better, yields that are negatively correlated.

To quantify diversification we use the diversification ratio (D) defined in Choueifaty and Coignard (2008). It is the ratio of the weighted average variance (WAV) divided by the overall portfolio variance (OPV), which considers not just the variances of the individual components but also how they move together, which is captured by their covariances. The numerator and the denominator of this ratio can differ due to the diversification effect. When assets in a portfolio are uncorrelated or negatively correlated, total portfolio variance can be lower than what one would expect from simply adding up the individual risks (after weighting them). In such case, the ratio would be above unity. Hence, a large ratio denotes a large diversification effect.

$$D = \frac{WAV}{OPV} \tag{A.1}$$

Weighted Average Variance refers to the sum of the individual variances σ_i of the portfolio components (crops, in our case) multiplied by their weights squared w_i. It represents what the portfolio's variance would be if the components were not correlated at all (i.e., if their performances were completely independent of each other). In a case with n crops, it is defined as follows.

$$WAV = \sum_{i=1}^{n}(w_i^2 * \sigma_i^2) \tag{A.2}$$

Overall Portfolio Variance is calculated by considering both the variances of the individual assets and the covariances between each pair of components i and j.

$$OPV = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j Cov(i, j) \tag{A.3}$$

A negative covariance between two assets means that when the productivity of one asset goes up, the return of the other tends to go down, and vice versa. This inverse relationship can significantly reduce the overall volatility of a portfolio because the negative performance of one asset can be offset by the positive performance of another. Essentially, when one asset is experiencing a downturn, another may be performing well, stabilizing the portfolio's overall performance. Very low covariance indicates that the returns on two assets have very little or no predictable relationship. While not as impactful as negative covariance in terms of reducing volatility, low covariance still contributes to diversification because the assets' returns do not move in tandem. This means that fluctuations in one asset will have a minimal impact on another, leading to more stable overall portfolio performance compared to if the assets were highly correlated.

A.6 Temperature data

We use temperature data published by *Météo-France*, the main public source of meteorological information in France. The dataset covers all 62 public weather stations distributed across the French territory. Each station records air temperature every three hours, providing detailed coverage at high temporal resolution. To estimate temperatures at the canton level, where direct measurements are often sparse, we interpolate the station data using a geostatistical method known as kriging. This method uses latitude, longitude, and altitude as spatial predictors, yielding reliable local temperature estimates. Compared with simpler techniques such as inverse distance weighting, kriging offers a more robust spatial interpolation. More sophisticated interpolation approaches such as those based on physical weather-models could in principle produce finer estimates, but high-frequency outputs (e.g., every three hours) are not publicly accessible.

In our empirical work, we focus on average afternoon temperatures, recorded between 12 p.m. and 3 p.m., and aggregate them into eight-day averages to match the frequency of our crop productivity data. For a sample of this data, see figure A.1d.

We also prefer these temperature estimates over satellite-based measures, such as laser readings from the Terra satellite, which can be affected by cloud cover. Because covered days are typically missing from raw satellite datasets, temperature averages tend to be upward biased. Weather stations are unaffected by this limitation, providing a more reliable basis for our analysis.

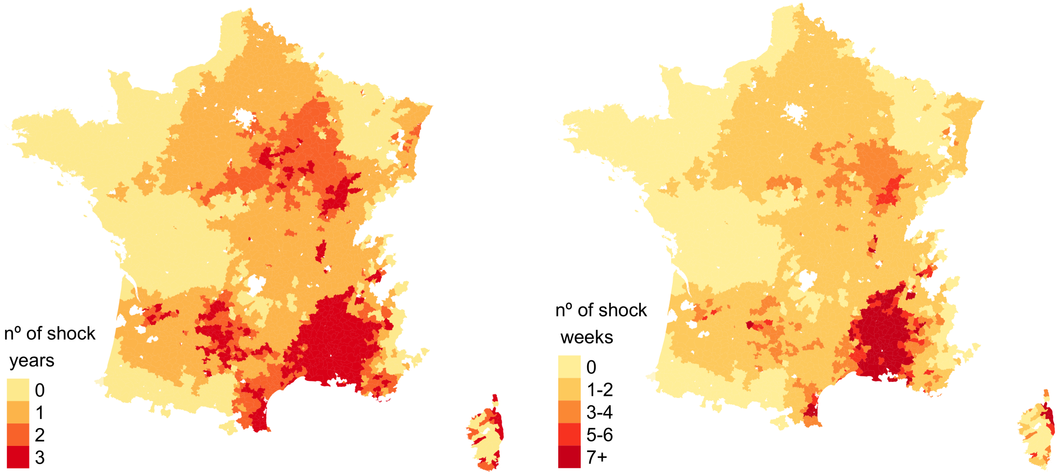
A.7 Constructing heat shock treatments indicators

Different species of plants have different climatic requirements, and they vary with the course of development phases. Spring and summer wheat, two major staple crops in French agriculture, are a good example. Winter wheat has very specific requirements at early stages of growth. During a phase called vernalization, the plant requires a period between one and two months of cold temperatures (0-5ºC) that are absolutely necessary to reach the next phase. During that period, temperatures as low as 6ºC could arguably be called thermal shocks, but not for spring wheat, which does not endure vernalization. Since we expect the impact of thermal shocks to be more visible at the extremes, we focus on climatic requirements during summer-spring seasons. During that part of the year, both spring and winter wheat go through phases known as flowering (anthesis) and grain-filling, which are directly relevant to the ultimate cereal yield, which is our object of study. For example, during grain-filling during June and July before the August harvests, wheat's optimal temperature is 25ºC and can suffer yield losses of 3–4% for every 1°C increase above the optimal temperature (Wardlaw et al., 1989; Nagai and Makino, 2009; Lepekhov, 2022). We define shock weighted thresholds as conservatively as possible: the upper limit of temperature stress demonstrated in the literature is presented in Table A.2. See Figure A.5 for the raw relationship between weekly GPP flows and temperature.

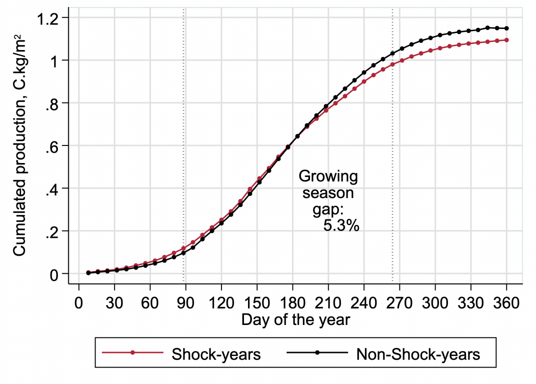
To take into account the presence of a variety of crops ($c = 1, 2, 3, \ldots, N$), we define the threshold T_A for treatment in area A as the mean value of crop-specific temperature thresholds T_c, during spring-summer, weighted by the area's crop composition $w_{A,c}$, as a share of surface.

$$T_A = \sum_{c=1}^{N} T_c * w_{A,c} \tag{A.4}$$

Figure A.4: The spatial distribution of heat shocks



(a) Years with at least one heat shock (b) Total number of heat shocks



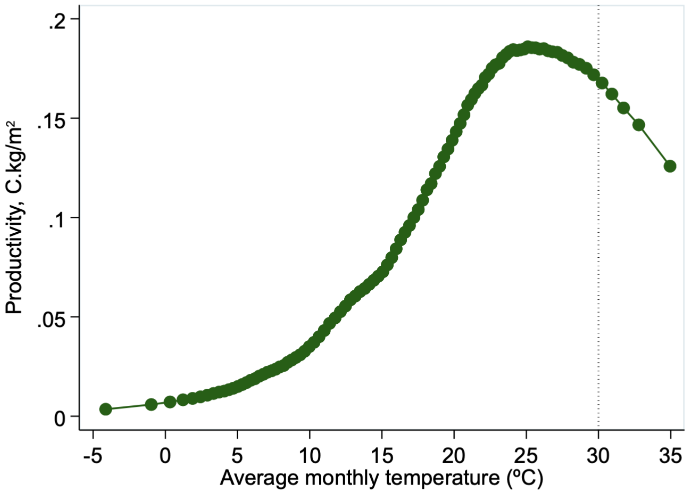
(c) Average loss under heat-shocks

Note. Figures A.4a and A.4b present the distribution of heat shocks across France from 2015 to 2021, based on weather data from Météo France. Heat shocks, or treatments, are defined as periods where the maximum daily temperature averaged over eight days exceeds the threshold outlined in Equation A.4. This threshold varies by region, accounting for local crop composition (see Figure A.7). Figure A.4c includes canton fixed effects. Shock years contain at least one occurrence within the year, as defined in section A.7. Own estimates based on data from Running and Zhao (2019) and Météo France.

In practice, finding crop-specific estimates of critical temperatures at that level of precision is not an easy task. Not only the actual estimation of such thresholds is subject to variations in methods, including issues such as controlling for other weather variables or the choice of functions for modeling, but they can also vary in their geographical and chronological dimension. We thus use estimates from French specific data or lab experiments in first priority when they are available, which is the case for many staple crops like wheat, corn, barley, rapeseed (colza), alfalfa and sunflower. Otherwise, we use data from research papers produced using data from other countries, which is the case for soybeans, for instance, where we use estimates from the United States (Schlenker and Roberts, 2009). For areas that are not covered by any of the options above we define the threshold at 30ºC.

For heat stress treatment effect estimation we define a binary shock indicator for the 321 dates of the seven-year panel that identifies when average daytime temperatures surpass the weighted threshold defined in equation A.4. The treatment variable is thus highly specific to local farmland production types for each year.

Figure A.5: Land productivity vs. monthly temperatures based on satellite data, 2000-2020



Notes. Own estimates based on data from MODIS sensor (Running and Zhao, 2019). Binned scatterplot, with canton-level observations. Displays average monthly farmland productivity in 100 equally sized groups, at corresponding temperature ranges.

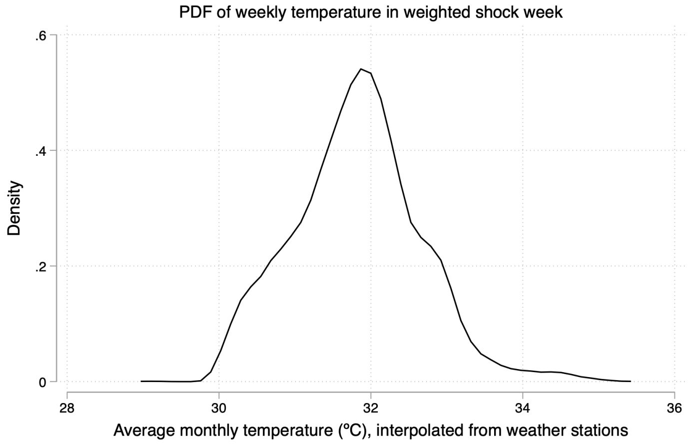
Table A.2: Crop-specific maximum temperatures

Crop	Max. temp (ºC)	Land share	Cumulative	Reference
Winter wheat	32	34.5	34.5	Gammans et al. (2017)
Corn/Maize	32	17.4	51.9	Hawkins et al. (2013)
Winter barley	33	7.4	59.3	Gammans et al. (2017)
Rapeseed	27	6.1	65.4	Pollowick and Sawhney (1988)
Sunflower	35	4.3	69.8	Rondanini et al. (2003)
Grapevine	30	3.6	73.3	Imputed
Spring barley	32	3.3	76.6	Gammans et al. (2017)
Alfalfa	30	2.8	79.5	Murata et al. (1965)
Beetroot	30	2.6	82.1	Imputed
Potato	30	1.1	83.2	Imputed
Soybean	30	1.0	84.1	Schlenker and Roberts (2009)
Spring wheat	33	0.2	84.3	Gammans et al. (2017)
Other (<1%)	30	15.6	100.0	Imputed

Notes. Compiled by the authors based on Cadastral data and references.

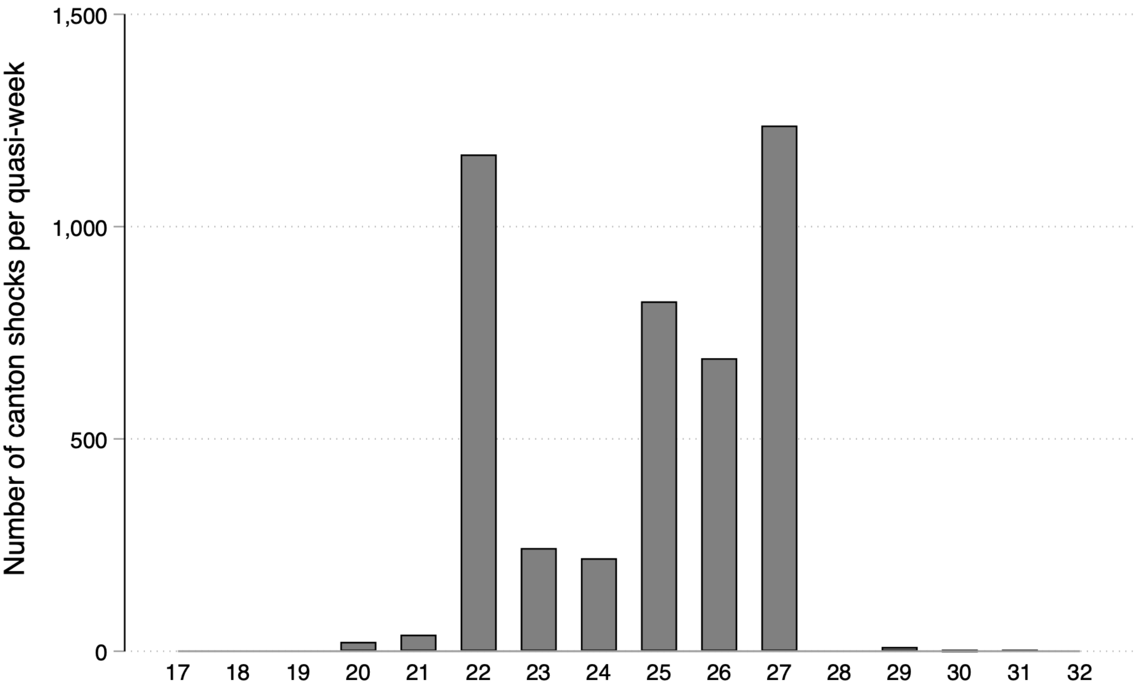
Figures A.4a and A.4b illustrate the distribution of heat-shocks across the French territory between 2015 and 2021. The shocks are broadly dispersed across most regions, though certain areas, particularly along the coastline and in mountainous regions, are never impacted due to their more temperate climates. Over the seven-year study period, most regions experience at least one heat-shock, with a maximum of three shock-years recorded for any given canton. Notably, some areas, especially along the southern Mediterranean coast, can experience multiple shocks within a single year, signaling more intense and prolonged exposure to heat. On average, land productivity is lower during shock-years, as 5.3% of the yield is lost by the end of the growth season. Figure A.4c shows that the effects are long-lasting and determinant to yearly production.

Figure A.6: Temperature distribution in shock week



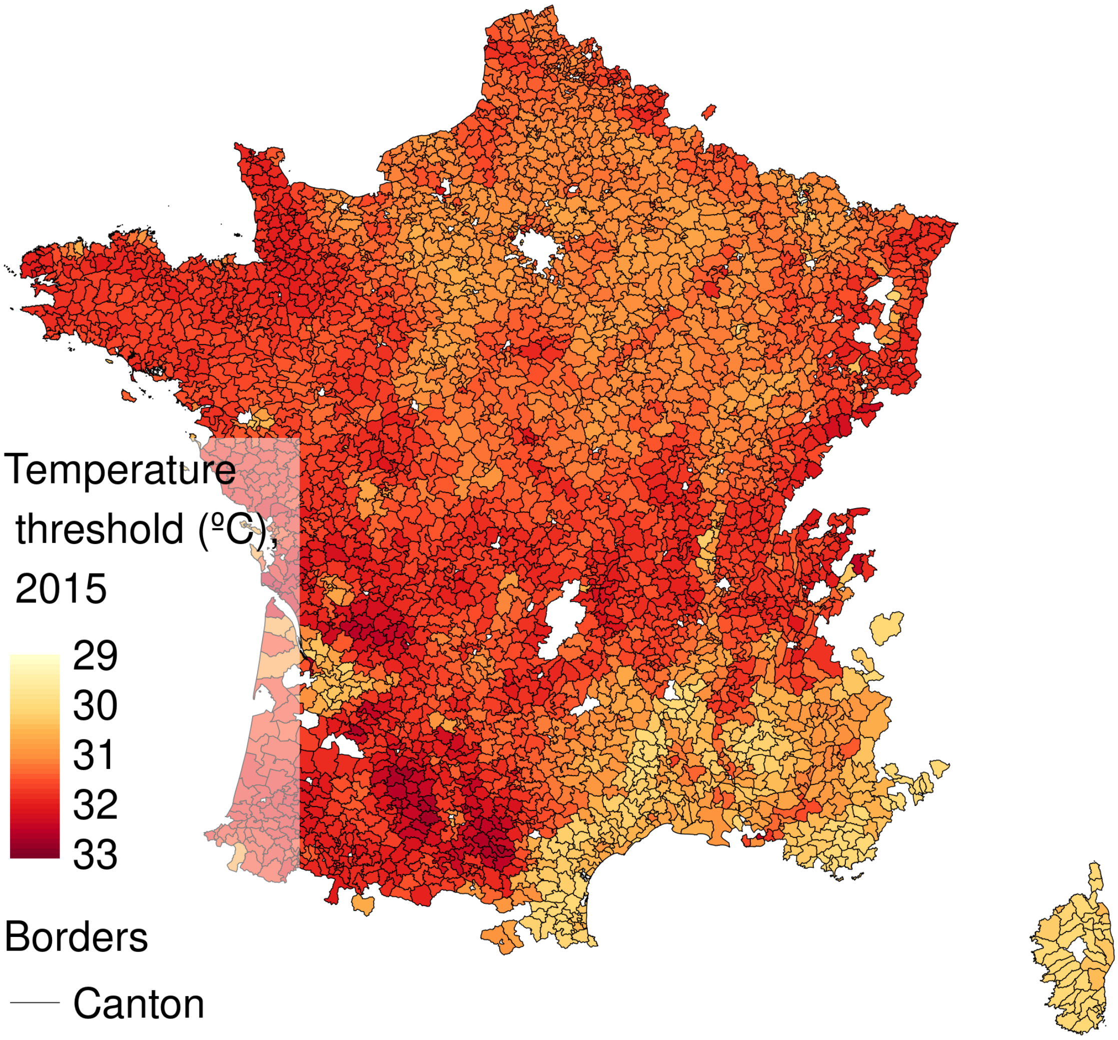
Notes. Kernel density estimates of the week's average temperature weighted shock weeks.

Figure A.8: Weighted shocks by quasi-week



Notes. Total number of canton-shocks per week.

Figure A.7: Treatment thresholds



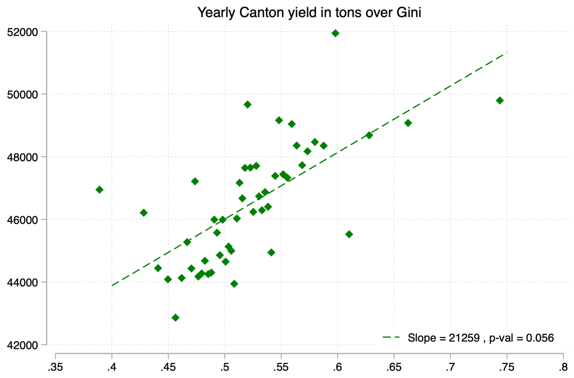
Note. The map presents the spatial distribution of treatment thresholds across France, reflecting region-specific temperature limits based on the local crop composition. Treatment thresholds are calculated as weighted averages of the maximum temperature tolerances for different crops, as outlined in Equation A.4. These thresholds account for the distinct thermal requirements in spring/summer. Crops for which temperature data is unavailable, such as pasture zones or other uncommon crops, a default threshold of 30ºC is applied (see figure A.5).

Table A.3: Weekly gross primary productivity (GPP C.g/m^2) on average over all weeks

	(1)	(2)	(3)	(4)	(5)
Panel (a): Linear Gini					
Ginist	-0.507***	-0.163**	-0.174***	-0.196***	-0.198***
	(0.123)	(0.064)	(0.058)	(0.069)	(0.061)
Mean GPP	25.0				
Panel (b): Gini quantiles					
q2	-1.745***	-0.333***	-0.135*	-0.299***	-0.198***
	(0.148)	(0.080)	(0.069)	(0.091)	(0.074)
q3	-2.829***	-0.565***	-0.283***	-0.496***	-0.376***
	(0.151)	(0.095)	(0.085)	(0.107)	(0.090)
q4	-2.596***	-0.664***	-0.401***	-0.557***	-0.485***
	(0.199)	(0.120)	(0.107)	(0.132)	(0.112)
Mean GPP 1st quantile (ref.)	26.8				
p-val equal	0.0000	0.0004	0.0027	0.0156	0.0017
R-squared	0.003	0.036	0.037	0.686	0.960
N	1016929	1016929	1016929	1016929	1016929
(AR x Year) FEs		✓	✓	✓	✓
Crop compo.			✓	✓	✓
Meteorological				✓	✓
(AR, Crop compo., Meteo.) x Week and Year FEs					✓

Notes. Dependent variable is weekly gross primary production in grams of carbon per square meter of canton cropland (C.g/m^2). Panel (a) Gini centered and standardized. Panel (b) quartile differences in reference to first Gini quartile. Meteorological controls are average weekly temperature, precipitation and wind speed. Standard errors are clustered at the canton. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure A.9: Yearly yield per canton C.tons (unweighted)



Notes. 50 quantile spaced bins residualized on AR x year FEs and yearly crop composition. Declared yield expressed in tons of carbon by canton by dividing the *département* measure by the number of nested cantons. Residualization and linear fits correspond to model (3) in Table 1 with standard errors clustered *département* level.

B Empirical models

B.1 Raw temperature models

Let temperature Temp.C° vary by week $\mathcal{W} = \{17, \ldots, 32\}$ (weeks mid-May through mid-September) each year $\mathcal{T} = \{2015, \ldots, 2021\}$.

$$\begin{aligned} \log(\mathrm{GPP})_{iwt} = b \ + \ & \big(\tau + \tau^g \times Gini_i^{st}\big) \cdot \mathrm{Temp_{iwt}C^\circ} \\ & + \underbrace{\sum_{t \in \mathcal{T} \setminus \{2015\}} \big(1 + c_i\big) \cdot year_t}_{\text{canton–year FEs}} \\ & + \underbrace{\sum_{w \in \mathcal{W} \setminus \{17\}} \big(1 + Gini_i^{st}\big) \cdot week_w \, + e_{iwt},}_{\text{Gini–week slopes}} \end{aligned} \tag{B.1}$$

The variable $Gini_i^{st}$ is the centered and standardized land Gini in 2015. The terms $week$, c and $year$ are week, canton and year fixed effects, respectively, and e is the error term. Though not dynamic, models allow land inequality to trend non-parametrically by week within each year. It also isolates variation within canton-year pairs, effectively making these pairs the panel unit of interest.

The coefficient τ represents the partial elasticity of GPP with respect a one degree temperature increase, on average, while τ^g allows this elasticity to vary with respect to a standard deviation increase in the Gini. We can provide further flexibility in this relationship by modeling land inequality heterogeneity non-parametrically, substituting the $Gini_i^{st}$ terms with Gini quantile q indicators. Their interactions with Temp.C° lead to straightforward predictions of the GPP-temperature partial elasticity within each Gini quantile.

We can isolate GPP-temperature partial elasticities and its Gini heterogeneity within different temperature splines by interacting terms with temperature threshold dummies.

$$\begin{aligned} \log(\mathrm{GPP})_{iwt} = c \ + \ & \big(\tau_o + \tau_o^g \cdot Gini_i^{st}\big) \cdot \mathrm{Temp_{iwt}C^\circ} \\ + & \sum_{s \in \{20,25,30,35\}} \mathbf{1}\{s \le \mathrm{Temp_{iwt}C^\circ} < s + 5\} \big(\tau_s + \tau_s^g \times Gini_i^{st}\big) \cdot \mathrm{Temp_{iwt}C^\circ} \\ + & \textstyle\sum_{t \in \mathcal{T} \setminus \{2015\}} \big(1 + c_i\big) \cdot year_t \quad + \displaystyle\sum_{w \in \mathcal{W} \setminus \{17\}} \big(1 + Gini_i^{st}\big) \cdot week_w + \nu_{iwt}. \end{aligned} \tag{B.2}$$

Table B.1: GPP-temperatures and land inequality spline regressions

Dep. Variable: logGPP	(1)		(2)		(3)		(4)	
(a): T°C effects with linear Gini								
Temp.°C	0.0561***	(0.0007)	0.0515***	(0.0005)	0.0524***	(0.0005)	0.0529***	(0.0005)
Temp.°C x Gini^{st}	0.0003	(0.0009)	-0.0002	(0.0007)	-0.0004	(0.0007)	-0.0003	(0.0007)
$1(20 \leq T < 25)$	0.5700***	(0.0157)	0.5231***	(0.0099)	0.5424***	(0.0092)	0.5513***	(0.0088)
$1(25 \leq T < 30)$	1.9084***	(0.0182)	1.8386***	(0.0144)	1.8697***	(0.0141)	1.8884***	(0.0136)
$1(30 \leq T°)$	2.4105***	(0.0565)	2.2751***	(0.0526)	2.2895***	(0.0538)	2.3280***	(0.0538)
Temp.°C × $1(20 \leq T < 25)$	-0.0297***	(0.0008)	-0.0262***	(0.0005)	-0.0271***	(0.0005)	-0.0276***	(0.0005)
Temp.°C × $1(25 \leq T < 30)$	-0.0831***	(0.0009)	-0.0789***	(0.0007)	-0.0803***	(0.0006)	-0.0811***	(0.0006)
Temp.°C × $1(30 \leq T°)$	-0.1005***	(0.0019)	-0.0940***	(0.0017)	-0.0949***	(0.0018)	-0.0963***	(0.0018)
Temp.°C x Gini^{st} × $1(20 \leq T < 25)$	-0.0003	(0.0010)	-0.0004	(0.0006)	0.0000	(0.0006)	0.0003	(0.0006)
Temp.°C x Gini^{st} × $1(25 \leq T < 30)$	-0.0023**	(0.0011)	-0.0014*	(0.0008)	-0.0012	(0.0008)	-0.0015**	(0.0008)
Temp.°C x Gini^{st} × $1(30 \leq T°)$	-0.0128***	(0.0019)	-0.0072***	(0.0017)	-0.0064***	(0.0017)	-0.0062***	(0.0017)
Gini^{st} × $1(20 \leq T < 25)$	-0.0103	(0.0190)	0.0026	(0.0121)	-0.0058	(0.0121)	-0.0107	(0.0118)
Gini^{st} × $1(25 \leq T < 30)$	0.0432*	(0.0240)	0.0266	(0.0182)	0.0240	(0.0179)	0.0334*	(0.0172)
Gini^{st} × $1(30 \leq T°)$	0.3597***	(0.0561)	0.1971***	(0.0517)	0.1749***	(0.0524)	0.1693***	(0.0534)
Constant	-4.2916***	(0.0140)	-4.2137***	(0.0105)	-4.2316***	(0.0100)	-4.2420***	(0.0095)
(b): T°C effects by Gini quantile								
Temp.°C	0.0528***	(0.0011)	0.0487***	(0.0008)	0.0503***	(0.0007)	0.0505***	(0.0007)
Gini quantile=1	0.0000	(.)	0.0000	(.)	0.0000	(.)	0.0000	(.)
Gini quantile=2 × Temp.°C	0.0038**	(0.0018)	0.0039***	(0.0013)	0.0030**	(0.0012)	0.0037***	(0.0011)
Gini quantile=3 × Temp.°C	0.0048***	(0.0015)	0.0073***	(0.0011)	0.0057***	(0.0011)	0.0060***	(0.0010)
Gini quantile=4 × Temp.°C	0.0039*	(0.0022)	0.0011	(0.0015)	0.0010	(0.0015)	0.0011	(0.0014)
$1(20 \leq T < 25)$	0.5301***	(0.0269)	0.4846***	(0.0160)	0.5245***	(0.0148)	0.5323***	(0.0136)
$1(25 \leq T < 30)$	1.8891***	(0.0299)	1.8279***	(0.0268)	1.8746***	(0.0267)	1.8905***	(0.0260)
$1(30 \leq T°)$	1.9133***	(0.1086)	2.0794***	(0.1039)	2.1539***	(0.1072)	2.1893***	(0.1077)
Temp.°C × $1(20 \leq T < 25)$	-0.0268***	(0.0013)	-0.0238***	(0.0008)	-0.0259***	(0.0008)	-0.0262***	(0.0007)
Temp.°C × $1(25 \leq T < 30)$	-0.0811***	(0.0013)	-0.0775***	(0.0012)	-0.0798***	(0.0011)	-0.0804***	(0.0011)
Temp.°C × $1(30 \leq T°)$	-0.0823***	(0.0036)	-0.0859***	(0.0034)	-0.0891***	(0.0034)	-0.0903***	(0.0034)
Gini quantile=2 × $1(20 \leq T < 25)$	0.0683*	(0.0411)	0.0623**	(0.0249)	0.0430*	(0.0232)	0.0518**	(0.0220)
Gini quantile=3 × $1(20 \leq T < 25)$	0.0317	(0.0367)	0.1117***	(0.0223)	0.0707***	(0.0210)	0.0712***	(0.0197)
Gini quantile=4 × $1(20 \leq T < 25)$	0.0350	(0.0463)	0.0003	(0.0276)	-0.0125	(0.0270)	-0.0244	(0.0256)
Gini quantile=2 × $1(25 \leq T < 30)$	0.0407	(0.0466)	0.0501	(0.0396)	0.0333	(0.0391)	0.0388	(0.0379)
Gini quantile=3 × $1(25 \leq T < 30)$	0.0117	(0.0417)	0.0769**	(0.0361)	0.0388	(0.0362)	0.0375	(0.0353)
Gini quantile=4 × $1(25 \leq T < 30)$	0.0358	(0.0567)	-0.0236	(0.0423)	-0.0224	(0.0416)	-0.0193	(0.0400)
Gini quantile=2 × $1(30 \leq T°)$	0.2020	(0.1506)	0.0166	(0.1425)	-0.0442	(0.1461)	-0.0349	(0.1451)
Gini quantile=3 × $1(30 \leq T°)$	0.5499***	(0.1595)	0.1898	(0.1535)	0.1095	(0.1585)	0.1079	(0.1579)
Gini quantile=4 × $1(30 \leq T°)$	1.1135***	(0.1567)	0.5845***	(0.1462)	0.5143***	(0.1495)	0.5079***	(0.1497)
Gini quantile=2 × Temp.°C × $1(20 \leq T < 25)$	-0.0040*	(0.0021)	-0.0035***	(0.0013)	-0.0024**	(0.0012)	-0.0030***	(0.0011)
Gini quantile=3 × Temp.°C × $1(20 \leq T < 25)$	-0.0027	(0.0018)	-0.0062***	(0.0011)	-0.0040***	(0.0011)	-0.0041***	(0.0010)
Gini quantile=4 × Temp.°C × $1(20 \leq T < 25)$	-0.0034	(0.0024)	-0.0007	(0.0014)	0.0000	(0.0014)	0.0006	(0.0013)
Gini quantile=2 × Temp.°C × $1(25 \leq T < 30)$	-0.0025	(0.0021)	-0.0029*	(0.0017)	-0.0020	(0.0017)	-0.0024	(0.0016)
Gini quantile=3 × Temp.°C × $1(25 \leq T < 30)$	-0.0019	(0.0019)	-0.0050***	(0.0016)	-0.0031**	(0.0015)	-0.0031**	(0.0015)
Gini quantile=4 × Temp.°C × $1(25 \leq T < 30)$	-0.0035	(0.0027)	-0.0000	(0.0019)	0.0001	(0.0019)	-0.0001	(0.0018)
Gini quantile=2 × Temp.°C × $1(30 \leq T°)$	-0.0084*	(0.0050)	-0.0024	(0.0046)	-0.0000	(0.0047)	-0.0006	(0.0047)
Gini quantile=3 × Temp.°C × $1(30 \leq T°)$	-0.0205***	(0.0053)	-0.0096*	(0.0050)	-0.0063	(0.0051)	-0.0064	(0.0051)
Gini quantile=4 × Temp.°C × $1(30 \leq T°)$	-0.0396***	(0.0053)	-0.0210***	(0.0048)	-0.0186***	(0.0049)	-0.0184***	(0.0049)
Constant	-4.2878***	(0.0136)	-4.2205***	(0.0097)	-4.2398***	(0.0094)	-4.2492***	(0.0090)
R-squared	0.758		0.833		0.845		0.858	
N	355024		355024		355024		354928	
Canton x Year and Week FEs	✓		✓		✓		✓	
Gini x Week FEs	✓		✓		✓			
Dept. x Week FEs			✓					
AR x Week FEs					✓			
Canton x Week FEs							✓	

Notes. OLS estimates using equation B.2. Panel (a) The standardized and centered land Gini is used. Panel (b) Gini quartile indicators are used. Observations with temperatures below 20°C serve as reference threshold. Standard errors, in parentheses, are clustered at the canton. * $p < .1$, ** $p < .05$, *** $p < .01$.

Now $\tau_{o/s}$'s and $\tau^g_{o/s}$'s capture the GPP-temperature partial elasticities and its linear Gini heterogeneity within temperature thresholds: min. to 20°C, from 20-25°C, 25-30°C and 30+°C allowing for non-nonlinearities in average and moderated temperature effects.

B.1.1 Temperature event study specification

GPP reactivity temperature and its heterogeneity over land inequality can also be modeled by saturating the regression with temperature bins to create a fully non-parametric "temperature event study" estimator. Let $\mathcal{J} = \{min, \ldots, 25, 26 \ldots, 35^+\}$ denote rounded integer temperature bins where temperatures at or above 35 are binned together. Then non-parametric temperature effects and Gini heterogeneity, this time using Gini quantiles, is modeled as,

$$\begin{aligned}\log(\mathrm{GPP})_{iwt} = a &+ \sum_{j\in\mathcal{J}\setminus\{29\}} \tau_j\, \mathbf{1}\{\mathrm{Temp°C} = j\}_{iwt} \\ &+ \sum_{q\in\{2,3,4\}} \sum_{j\in\mathcal{J}\setminus\{29\}} \tau_{qj}\Big(\mathbf{1}\{\mathrm{Temp°C} = j\}_{iwt} \cdot \mathbf{1}\{q_i = q\}\Big) \\ &+ \sum_{t\in\mathcal{T}\setminus\{2015\}} (1 + c_i)\cdot year_t \\ &+ \sum_{q\in\{2,3,4\}} \sum_{w\in\mathcal{W}\setminus\{17\}} (1 + \mathbf{1}\{q_i = q\})\cdot week_w + e_{iwt}. \end{aligned} \tag{B.3}$$

where $\mathbf{1}\{\mathrm{Temp} = j\}_{iwt}$ are temperature–bin dummies (29°C omitted in the first line), and the interaction coefficients τ_{qj} capture the *differential* response at temperature j for quartiles $q \in \{2, 3, 4\}$ relative to 29°C.

The event–study plot shows that interaction terms are flat and statistically indistinguishable across 25–29°C for higher–inequality quartiles relative to the first quartile. But parallel temperature trends is only a partial identification assumption since more or less equally distributed land may experience different dynamic trends over GPP x temperature bins. We formalize this as a local "normal parallel temperature trend":

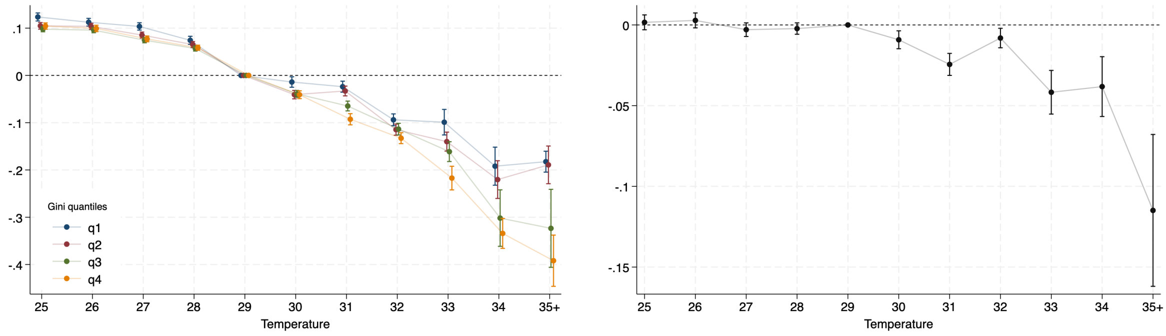
$$\textit{(LPTR)} \qquad \tau_{qj} = \tau_{q,29} \quad \text{for all } q \in \{1, 2, 3, 4\} \text{ and } j \in \{25, 26, 27, 28, 29\}.$$

Define $\Delta_{qj} := \tau_{qj} - \tau_{q,29}$. Under *LPTR*, $\Delta_{qj} = 0$ for all $j \leq 29$°C; hence for $j \geq 30$°C the set $\{\Delta_{qj}\}$ *identifies* the differential effect of extreme heat by inequality: excess heat sensitivity that emerges only beyond a certain temperature.

In our estimates, once weekly temperatures exceed 30°C, the responses diverge sharply: cropland in the first Gini quartile loses about 6.9% of weekly GPP per additional degree, while fourth–quartile cropland loses about 11.3% per degree, on average.

The assumption is a "partial" identification condition because we do not impose dynamic parallel paths before or after heat events; thus the "trends" show that conditional on the modeled fixed effects, more- and less-unequal cantons have the same GPP–temperature relationship in the non-extreme range (25–29°C). Under this local invariance, the post-29°C interaction coefficients identify the differential effect of extreme heat by inequality group.

Figure B.1: Temperature "event study" estimates with non-parametric canton-time trends



(a) Gini quantiles

(b) $\text{Gini}^{st} \times 1\{\text{Temp}°\text{C} = j\}$

Notes. N=355024. Non-parametric canton-week dummies added to equation B.3. (a) Estimates of the GPP-temperature partial elasticity by Gini quantiles using integer temperature bin dummies. (b) Coefficient estimates on the standardized Gini interacted with temperature bin dummies. Reference temperature is 29°C. Illustration restricted to estimates for 25-35+°C. Standard errors are clustered at the canton. Vertical lines correspond to 95% confidence intervals.

B.2 A two-way-fixed-effects estimator

A TWFE model to estimate heat stress treatment effects allowing for linear heterogeneity over the Gini can be formulated with,

$$\begin{aligned} log(GPP)_{iwt} =& \mu + \alpha \cdot D_{iwt} + \alpha^g \big(D_{iwt} \cdot Gini_i^{st}\big) + \sum_{w \in \mathcal{W} \setminus \{17\}} \big(1 + Gini_i^{st}\big) \cdot week_w \\ &+ \sum_{t \in \mathcal{T} \setminus \{2015\}} \big(1 + c_i\big) \cdot year_t + u_{iwt} \end{aligned} \tag{B.4}$$

where D_{iwt} is the treatment indicator for canton cropland i equal to one if temperatures equals or exceeds the weighted stress threshold in week w in year t and zero if not. The variable $Gini_i^{st}$ is the land Gini coefficient in 2015 centered and standardized using the sample mean and standard deviation. The terms $week$, c and $year$ are week, canton and year fixed effects, respectively, and u is the error term. We observe that the model allows canton i's land inequality to trend non-parametrically by week within each year. It also isolates variation within canton-year pairs, effectively making these pairs the panel unit.

With the classic assumptions of (1) no anticipation (NA), (2) conditional parallel trends (CPT) and, importantly that (3) heat shock effects are constant over time (CT), the average treatment effect (ATE) of the shock on percentage changes in weekly GPP can be consistently estimated with $\alpha = E[logGPP(D)_{w,t} - logGPP(0)_{w,t} | D_{w,t} = 1]$.

Again, we can provide further flexibility in modeling the ATE-Gini relationship non-parametrically by substituting the $Gini_i^{st}$ terms with a vector of Gini quantile q_i indicators to estimate ATEs within each quantile. With categorical Gini $q_i \in \{1, 2, 3, 4\}$ that denotes the Gini quartile we model ATE heterogeneity as,

$$
\begin{aligned}
logGPP_{iwt} =& a + \beta_1 \cdot D_{iwt} + \sum_{q \in \{2,3,4\}} \beta_q \cdot \mathbf{1}\{q_i = q\} \cdot D_{iwt} \\
&+ \sum_{w \in \mathcal{W} \setminus \{17\}} \left(1 + \sum_{q \in \{2,3,4\}} \mathbf{1}\{q_i = q\}\right) \cdot week_w \\
&+ \sum_{t \in \mathcal{T} \setminus \{2015\}} (1 + c_i) \cdot year_t + u_{iwt}
\end{aligned}
\tag{B.5}
$$

B.2.1 Assumptions and OLS weighting

OLS weights are determined by OLS residuals from a regression on the period-level treatment indicator D_{iw} on panel unit and period fixed effects using the subsample of "ever-treated" units. Hence the OLS weights applied to each period-level DiD estimate depend on the length of treatment exposure and total exposure variation within period when first treated. If the shock ATE and/or its Gini heterogeneity evolves dynamically post exposure, the implicit OLS weighting issue can lead to biased, and in extreme cases, an improperly signed TWFE estimate in the presence of negative weights. In the baseline model, roughly 7.9% of the DiD effect estimates receive a negative weight thus estimates in columns (1-3) in Table 4 are unlikely to be wrongly signed. Using the diagnostic tools of de Chaisemartin and D'Haultfoeuille (2024) we find that out of a total of 8058 DiD effect estimates, 634 (7.9%) receive a negative weight. The standard deviation across all $(w|D = 1, w - 1|D = 0)$-level DiD estimates would have to be at least 1.45 for the TWFE ATE estimate to be the opposite sign of the true ATE.

Restricting the sample to shocked-once treated units eliminates the presence of negative OLS weights in our data and indeed ATE estimates are highly comparable. Yet they may still not represent the week-of-shock effect if parallel trends do not hold after initial exposure. In addition, the remaining positive OLS weights are still a function of relative exposure timing. Modeling pre- and post-exposure trends with shock leads and lags can be a solution (Sun and Abraham, 2021). But in this context the D_{iw} shock is itself dynamic – it can switch on and off multiple times – and our main hypotheses requires conditioning on additional covariates.

B.3 Dynamic heterogeneity robust models

B.3.1 Event study estimators

Target estimand and identifying assumptions

Define a series of cohort indicators that denote the period in which a treated unit is first shocked within year t: dg at $g = w$ where $g \in \{20, 21, ..., 27 = G\}$ and the quasi weeks observed each year are $w \in \mathcal{W}\{17, \ldots, 32 = W\}$.

For treated cohort g, we define treatment effects,

$$ATE_w(g) = GPP_w(g) - GPP_w(\infty), \quad g = w, \ldots, W, \quad g \leq w$$

where $GPP_w(\infty)$ represents the potential outcome of the treated unit if not treated. The treatment effects we are searching to identify are the ATEs in periods where the cohorts are actually subjected to the shock:

$$\tau_{gw} = E[ATE_w(g) \mid dg = 1], \quad g = w, \ldots, G \ ; \quad w = g, \ldots, W \ ; \ g \leq w.$$

Identification follows the classic DiD assumptions adapted to staggered treatment exposure. Here we summarize using Wooldridge (2025) (section 4)

- Stable Unit Value Treatment Assumption (SUTVA): Potential outcomes in the population do not depend on the treatment status of other units in the population.
- No anticipation (NA): For the vector of treated cohorts $\mathbf{d} = (dg, ..., dG)$

  $$E(GPP_w(g) - GPP_w(\infty)|\mathbf{d}) = 0, \quad w < g$$

  , meaning that regardless of when the shock happens for a given unit the potential outcomes are the same prior to the shock.
- Parallel Trends (PT): We allow shock treatment exposure to differ by fixed levels for each cohort, but exposure cannot be systematically related to the trend in the never-treated-state.

  $$E(GPP_w(\infty) - GPP_{17}(\infty)|\mathbf{d}) = E(GPP_w(\infty) - GPP_{17}(\infty)), \ w \in \mathcal{W} \setminus \{17\}.$$

  This assumption can be relaxed by conditioning on fixed covariates. Then the assumption is Conditional Parallel Trends (CPT):

  $$E(GPP_w(\infty) - GPP_{17}(\infty)|\mathbf{d}, \mathbf{x}) = E(GPP_w(\infty) - GPP_{17}(\infty)|\mathbf{x}), \ w \in \mathcal{W} \setminus \{17\}$$

- Linearity: Treatment effects are captured by comparing conditional means and therefore it is appropriate to use linear regression. See Wooldridge (2025).

Pooled OLS and the Extended TWFE estimators

When cohort $d\infty$ corresponds to units that are never treated (NT) during year t, a POLS event study estimator writes as,

$$\begin{aligned} logGPP_{iwt} = \sum_{g}^{G} \left(\sum_{w}^{g-2} \theta_{gw} \cdot dg_{it} \cdot fw_t) + \sum_{w \geq g}^{W} \tau_{gw} \cdot dg_{it} \cdot fw_t + dg_{it} \right) \\ + \sum_{w \in \mathcal{W} \setminus \{17\}}^{W} fw_t + u_{iwt}. \end{aligned} \tag{B.6}$$

As Wooldridge (2025) highlights, this ES estimator is fully saturated and thus contains all the dg sample weighted 2x2 difference-in-differences (DiD) regressions one requires to compute an aggregate ATE as convex combinations of the underlying DiD estimates. For example, the week-of-shock effect corresponds to a weighted average of all $g = w$ cohort 2x2 DiDs using $g - 1$ as the pre-period and $g = w$ as the post-period for the treated and $d\infty$ never treated units. The weights that determine the ATE depend on the share of cohort g among all treated cohorts.

For example, the week-of-shock ATE for cohort $d20$ is captured by τ_{2020} where $d20 \cdot f20 = 1$. This coefficient is the cross-sectional DiD parameter

$$\begin{aligned} \tau_{2020} =& E(logGPP \mid d20 = 1, \ f20 = 1) - E(logGPP \mid d20 = 1, \ f19 = 1) \\ & - \big(E(logGPP \mid d\infty = 1, \ f20 = 1) - E(logGPP \mid d\infty = 1, \ f19 = 1)\big) \end{aligned}$$

Persistence in the second week is captured by τ_{2021} where $d20 \cdot f21 = 1$, etc. ; the contemporaneous effect for units first shocked in week 21 is τ_{2121} where $d21 \cdot f21 = 1$ with week 2 persistence captured by τ_{2122} where $d21 \cdot f22 = 1$, etc.

$$
\begin{aligned}
\tau_{gw} =& E(logGPP \mid dg=1,\ fw=1) - E(logGPP \mid dg=1,\ f(g-1)=1) \\
&- \big(E(logGPP \mid d\infty=1,\ fw=1) - E(logGPP \mid d\infty=1,\ f(g-1)=1)\big)
\end{aligned}
$$

Cohort pre-trend are captured by the θ_{gw}'s. An event study version using two-way-fixed-effects writes as,

$$
\begin{aligned}
logGPP_{iwt} = \sum_{g}^{G} \left(\sum_{w}^{g-2} \theta_{gw} \cdot dg_{it} \cdot fw_t + \sum_{w \geq g}^{W} \tau_{gw} \cdot dg_{it} \cdot fw_t \right) \\
+ \sum_{\substack{c\in \\ \mathcal{C}\setminus\{1\}}}^{C} \ \sum_{\substack{t\in \\ \mathcal{T}\setminus\{2015\}}}^{T} c_i \cdot year_t + \sum_{\substack{w\in \\ \mathcal{W}\setminus\{17\}}}^{W} week_w + e_{iwt}.
\end{aligned}
\tag{B.7}
$$

It provides identical estimates and inference since cohort and time periods are absorbed by the canton x year and week fixed effects.

Modeling effect heterogeneity over time invariant moderators, such as Gini quantile indicators q is straightforward. One simply interacts all terms with these subsample indicators which is clearly equivalent to estimating separate regression for each subsample of q. But using OLS on the saturated specification provides flexibility on ATE aggregation that corresponds to the estimand of interest and has the added advantage of simplifying inference for linear tests of treatment effect heterogeneity over time and baseline heterogeneity.

Modeling ATE heterogeneity with continuous baseline variables such as the Gini coefficient is also straightforward. Using the POLS specification, we have,

$$
\begin{aligned}
logGPP_{iwt} = \sum_{g}^{G} \Bigg(\sum_{w}^{g-2} \theta_{gw} \cdot dg_{it} \cdot fw_t + \sum_{w \geq g}^{W} \tau_{gw} \cdot dg_{it} \cdot fw_t \\
+ \sum_{w}^{g-2} \phi_{gw} \cdot dg_{it} \cdot fw_t \cdot Gini_i^{st,d} + \sum_{w \geq g}^{W} \gamma_{gw} \cdot dg_{it} \cdot fw_t \cdot Gini_i^{st,d} \\
+ dg_{it} \Bigg) + \sum_{w \in \mathcal{W}\setminus\{17\}}^{W} fw_t \cdot Gini_i^{st,d} + Gini_i + u_{iwt}.
\end{aligned}
\tag{B.8}
$$

For the θ's and τ's to capture the average effect we demean the heterogeneity covariate of interest specifically for each treatment cohort dg. Regardless if or how one centers, the γ_{gw}'s capture the linear moderating effect of the Gini for each cohort's post treatment period, while the ϕ_{gw}'s capture whether a cohort's future treatment exposure can be predicted by baseline moderators. This holds whether one models the covariate's interactions as linear using the continuous measure or non-parametrically using arbitrary quantiles q. An attractive feature of this specification, as highlighted by Wooldridge (2025), is that it can be thought of as a covariate-flexible version of the Sun and Abraham (2021) leads and lags ES estimator. Also, the coefficients obtained on $dg_{it} \cdot fw_t$ and $dg_{it} \cdot fw_t \cdot Gini_i^{st,g}$ will be equivalent to the corresponding cross-sectional DiD regression using the "long differences" adjustment estimator proposed by Callaway and Sant'Anna (2021) when using never-treated units as controls. Again, this is because it encapsulates all the cross-sectional regression estimated differences anchored to each cohort's $g-1$ pre-shock period. But it also has the added value of explicitly modeling the treatment selection and/or moderating effects of the fixed covariates on the ATE like the land Gini.

It also lends itself to easily conditioning on nesting unit time trends non-parametrically. For example ATEs by Gini quantile q conditioning on the most granular unit, i.e. the canton c itself nested within canton-year writes as,

$$\begin{aligned} logGPP_{iwt} = \sum_{q\in\{1,2,3,4\}} \mathbf{1}\{q_i = q\} \sum_{g}^{G} \left(\sum_{w}^{g-2} \theta^{q}_{gw} \cdot dg_{it} \cdot fw_t + \sum_{w\geq g}^{W} \tau^{q}_{gw} \cdot dg_{it} \cdot fw_t \right) \\ + \sum_{\substack{c\in \\ \mathcal{C}\setminus\{1\}}}^{C} c_i \left[\sum_{\substack{t\in \\ \mathcal{T}\setminus\{2015\}}}^{T} year_t + \sum_{\substack{w\in \\ \mathcal{W}\setminus\{17\}}}^{W} week_w \right] + e_{iwt}. \end{aligned} \tag{B.9}$$

B.3.2 Cohort specific effects and aggregation

Abstracting from effect heterogeneity for simplicity, we can see that, with $w \in \mathcal{W}\{17, \ldots, 32 = W\}$ and the first cohort shocked at $w = 20$ and last at $w = 27$, we have $\sum_{g=20}^{G}(32 - g + 1) = 76$ treatment effects to estimate.

These treatment effect parameters are the τ_{gw}'s which are all referenced to week $w = g - 1$ and use the corresponding period means of never-treated units, $d\infty_{it}$. To obtain the average effect over all cohorts for a given exposure period one must calculate a convex combination of the estimates using appropriate weights.

The ATE over all cohorts and treatment periods is captured by the weighted treatment effect,

$$\bar{\tau}_{\omega} = \sum_{g}^{G} \sum_{w\geq g}^{W} \omega_g \tau_{gw}.$$

Weights applied are

$$\omega_g \equiv \frac{N_g}{\sum_{k=g_{\min}}^{G-1} (W - k + 1)\, N_k}$$

with $w = g, \ldots, W \quad ; \quad g = w, \ldots, G$. For the instantaneous, week-of-shock effect, the natural weights are simple cohort shares $\omega_g^{\text{WoS}} \equiv N_g / \sum_k N_k$:

$$\bar{\tau}_{\omega,g=w} = \sum_{g}^{G} \omega_g^{\text{WoS}} \tau_{g,g=w}$$

We have full cohort support for all treatment effects up to and including $g+5$ since the last cohort is shocked at week $w = 27$ and the window of observation ends at $W = 32$ (mid-September). Thus weights up to and including lag+5 (week six post initial exposure) are equivalent to each cohort's share among all treated cohorts. Beyond this point, persistence of effects at exposure time e rely on the restricted cohort subset of dg_i's where $g \leq W - 5$.

B.3.3 The lags-only, non-event study estimator

Including no lead interaction terms leads to a lags-only estimator,

$$y_{iwt} = \sum_{g}^{G} \left(\sum_{w\geq g}^{W} \tau_{gw} \cdot dg_{it} \cdot fw_t + dg_{it} \right) + \sum_{w\in\mathcal{W}\setminus\{17\}}^{W} fw_t + u_{iwt}$$

.

Here, one uses the average of all pre-treatment periods when implicitly constructing the control as opposed to the ES estimators in which we modeled period $g-1$ as the reference for treated and control units (arbitrarily). Both de Chaisemartin and D'Haultfoeuille (2024) and Wooldridge (2025) have highlighted this specification, in which one averages by cohort and time-periods as above, produces numerically identical estimates to those of the imputation estimator proposed by Borusyak et al. (2024). Wooldridge (2025) argues that there is an added advantage to using OLS in that it can provide less computationally intensive and less conservative standard errors. Wooldridge (2025) (Section 6) further highlights that the choice of whether one "should" use the event study or the lags only model boils down to a trade-off between robustness and efficiency.

B.3.4 Estimator equivalencies and non-parametric time trends

Table B.2 demonstrates the numerical equivalence of estimates with those obtained with the de Chaisemartin and D'Haultfoeuille (2024) (dCDH) estimator. One difference is how the estimators model linear ATE heterogeneity. Those obtained using the de Chaisemartin and D'Haultfoeuille (2024) algorithm are larger in magnitude. dCDH proposes regressing the predicted cohort ATEs directly on the heterogeneity variable which is equivalent to the OLS estimator when one does not explicitly model Gini heterogeneity of the never-treated units i.e. omitting the Gini interaction with week dummies.[50]

To include not-yet-treated observations as control units in addition to the never-treated cohort observations, one can stack separate "DiD" OLS regressions for each cohort specific treatment effect using not-yet-treated cohort observations and the never-treated panel cohort observations as control units.

For example, focusing on the week-of-shock effect, this means stacking regressions of y_{it} on $1, dg_{it}, fg_t$ and $dg_{it} \cdot fg_t$ for each cohort $g = w, \ldots, G$, using as controls the observations from not-yet-treated cohorts $k > g$ together with the never-treated panel ($d\infty_{it} = 1$), restricted to periods fg_t and $f(g-1)_t$. The corresponding stacked covariance matrix allows for a straightforward estimation of weighted ATEs for arbitrary aggregation and the modeling of ATE heterogeneity as above.

$$\begin{aligned}\tau_{gw} =& E(logGPP \mid dg = 1,\ fw = 1) - E(logGPP \mid dg = 1,\ f(g-1) = 1) \\ &- \big(E(logGPP \mid dq = 1,\ fw = 1) - E(logGPP \mid dq = 1,\ f(g-1) = 1)\big), \quad g \leq w,\ \forall\ q > g\end{aligned}$$

What is perhaps less well known is that both ES estimators, using not-yet-treated and/or never-treated units, as controls with the above weighting scheme will provide numerically equivalent lead and lag ATE estimates as the de Chaisemartin and D'Haultfoeuille (2024) estimators built on cell-level demeaning, but with slightly different standard errors.

Table B.2 also provides estimates when cell level ATE estimation is restricted with more and more granular geographical units in which our canton-year observations are nested. To compare treated units that switch into shock exposure to control units whose exposure status does not (or has not yet) change(d) – within a region, departement, agricultural region or even the canton itself.

This entails interacting all terms in equation B.8 with $R-1$ fixed effects that correspond to the nesting unit R, i.e. region, departement, agricultural region or even the canton.

---

[50] See de Chaisemartin and D'Haultfoeuille (2024) for their proposed two-step procedure to predict linear ATE heterogeneity. Estimates using their did_multiplegt_dyn STATA package do not vary when further allowing not-yet-treated units to act as controls.

$$
\begin{aligned}
logGPP_{iwt} = \sum_{k=2}^{R} \mathbf{1}\{\mathrm{r_i} = k\} \left[\sum_{g}^{G} \left(\sum_{w}^{g-2} \theta_{gw}^{k} \cdot dg_{it} \cdot fw_t + \sum_{w \geq g}^{W} \tau_{gw}^{k} \cdot dg_{it} \cdot fw_t + dg_{it} \right) + \sum_{w \in \mathcal{W} \setminus \{17\}}^{W} fw_t + \kappa_k \right] \\
+ \mu_{iwt}.
\end{aligned}
\tag{B.10}
$$

Now the model contains $R \cdot (W-1)$ additional cross-sectional difference in differences parameters that aggregate to the τ_{gw}. Hence, treated cohort-time effects (DiDs) are calculated within nesting unit r and weighted by treated cohort share of$r$ before aggregation. The seasonal canton cohort-period estimands represent:

$$
\begin{aligned}
\tau_{gw} = &\sum_{k=1}^{R} \Pr(\mathrm{r} = k \mid dg = 1) \big(E(\log GPP \mid \mathrm{r} = k,\, dg = 1,\, fw = 1) - E(\log GPP \mid \mathrm{r} = k,\, dg = 1,\, f(g-1) = 1) \big) \\
&- \sum_{k=1}^{R} \Pr(\mathrm{r} = k \mid d\infty = 1) \cdot \big(E(\log GPP \mid \mathrm{r} = k,\, d\infty = 1,\, fw = 1) \\
&- E(\log GPP \mid \mathrm{r} = k,\, d\infty = 1,\, f(g-1) = 1) \big).
\end{aligned}
$$

With the corresponding weighted aggregation, all dynamic pre-trend placebo and treatment effect estimates will be numerically identical to those obtained with the de Chaisemartin and D'Haultfoeuille (2024) estimator with never-treated units used as controls and specifying fixed nesting category R as the non parametric trends option. See their Section 3.3 for a detailed discussion for how this can relax the CPT assumption.

Column 3 of Table B.2 provides an example. We present estimates when identifying variation is restricted to cell level DiDs in which units are nested within the broadest administrative delineation: 22 French regions. Comparing this to column 4 we again note the numerical equivalence of the OLS estimates with those obtained using the de Chaisemartin and D'Haultfoeuille (2024) estimator. As de Chaisemartin and D'Haultfoeuille (2024) prove, these event study estimators are unbiased even if cohorts experience differential trends overall, but that all cohorts within the same nesting unit would, counterfactually, experience parallel trends.

The advantage of using the OLS version is that linear combinations of treatment effect parameters and ensuing inference is easy to obtain using the single covariance matrix. But when the dimensionality of the nesting category is large, the de Chaisemartin and D'Haultfoeuille (2024) estimator can be much more efficient. The remaining columns of Table B.2 test the robustness of the immediate effect to non-parametric time trends for the departement, the agricultural region and the canton. We find that estimates depart little from the baseline ES estimates and are most pronounced when isolating to within canton comparisons. The equivalence is straightforward for subsamples defined by fixed heterogeneity (e.g., Gini quartiles). However, the appropriate way to model continuous linear heterogeneity that varies within the nesting unit – or does not vary at all when the nesting unit is the canton – remains an active area of applied econometrics and is currently infeasible with the de Chaisemartin and D'Haultfoeuille (2024) estimator.

Table B.2: Dynamic-heterogeneity-robust week-of-shock ATE estimates with non-parametric time trends

Dep. Variable: logGPP	Event study			... restricted to within nesting unit				
				Region		Department	AR	Canton
	OLS	dCDH		OLS	dCDH			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel (a): ATE with linear Gini								
Shock	-0.1644***	-0.1644***	-0.1619***	-0.1620***	-0.1620***	-0.1623***	-0.1637***	-0.1726***
	(0.00269)	(0.00279)	(0.00278)	(0.00217)	(0.00220)	(0.00161)	(0.00163)	(0.00602)
Shock x Gini$^{d.st}$	-0.0256***							
	(0.00287)							
Predicted Gini het.		-0.0256***	-0.0256***					
		(0.00319)	(0.00319)					
N	355024	153950	169385	355024	79605	38918	32341	17694
Control units: Never shocked	✓	✓	✓	✓	✓	✓	✓	✓
Control units: Not yet shocked			✓					
Panel (b): ATE by Gini quantile								
q1	-0.1268***	-0.1268***	-0.1253***	-0.1182***	-0.1182***	-0.1171***	-0.1225***	-0.1260***
	(0.00526)	(0.00530)	(0.00528)	(0.00425)	(0.00442)	(0.00502)	(0.00642)	(0.01068)
N		40825	42806		15332	6550	4591	3107
q2	-0.1490***	-0.1490***	-0.1466***	-0.1465***	-0.1465***	-0.1483***	-0.1501***	-0.1547***
	(0.00512)	(0.00514)	(0.00513)	(0.00401)	(0.00408)	(0.00408)	(0.00470)	(0.01052)
N		38786	42525		16174	7212	6168	4326
q3	-0.1586***	-0.1586***	-0.1564***	-0.1601***	-0.1601***	-0.1654***	-0.1659***	-0.1713***
	(0.00525)	(0.00528)	(0.00527)	(0.00377)	(0.00391)	(0.00310)	(0.00567)	(0.01144)
N		38146	42464		18580	8574	7042	4785
q4	-0.1908***	-0.1908***	-0.1885***	-0.1915***	-0.1915***	-0.1929***	-0.1932***	-0.2077***
	(0.00522)	(0.00523)	(0.00522)	(0.00397)	(0.00408)	(0.00321)	(0.00403)	(0.01251)
N	355024	36193	41590	355024	19525	10811	9258	5476
p-val equal ATEs	0.0000			0.0000				

Notes. Columns 1 reproduces the baseline results of column 1 in Table 5 except that linear Gini heterogeneity is modeled as in de Chaisemartin and D'Haultfoeuille (2024) (dCDH). Estimates using estimator dCDH follow using never-treated (2) or additionally not-yet-treated (3) units as controls. Columns (4-8) present estimated ATEs when variation is restricted within the entitled geographic region. Columns (4) and (5) demonstrate the numerical equivalence of ATE estimates obtained through OLS when saturating model B.6 with nesting dummy interactions and the dCDH estimator. Standard errors clustered at the canton. * $p < .1$, ** $p < .05$, *** $p < .01$.

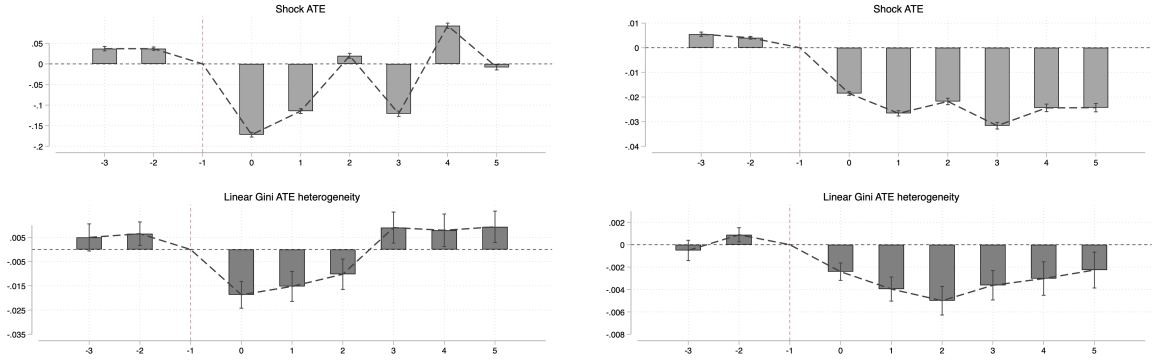
B.4 Exposure paths

Table B.3: Shock exposure paths

Path	N Treated	%	w-1	w	w+1	w+2	w+3	w+4	w+5	w+6	N Shocks
1	2511	74.6	0	1	0	0	0	0	0	0	2511
2	500	14.85	0	1	0	0	1	0	0	0	1000
3	175	5.2	0	1	1	0	0	0	0	0	350
4	40	1.19	0	1	1	1	0	0	0	0	120
5	27	.8	0	1	1	1	0	1	0	0	108
6	21	.62	0	1	0	1	0	0	0	0	42
7	21	.62	0	1	1	1	1	0	1	0	105
8	16	.48	0	1	0	0	1	0	1	0	48
9	13	.39	0	1	0	0	0	0	1	0	26
10	10	.3	0	1	1	0	0	1	0	0	30
11	8	.24	0	1	1	0	1	0	1	0	32
12	6	.18	0	1	1	1	0	0	1	0	24
13	5	.15	0	1	1	0	1	0	0	0	15
14	4	.12	0	1	0	0	0	0	0	1	8
15	4	.12	0	1	1	1	1	0	0	0	16
16	3	.09	0	1	0	1	1	0	0	0	9
17	1	.03	0	1	1	0	0	1	0	1	4
18	1	.03	0	1	1	1	1	1	1	0	6
	3366	100		3366	298	123	558	39	65	5	4454

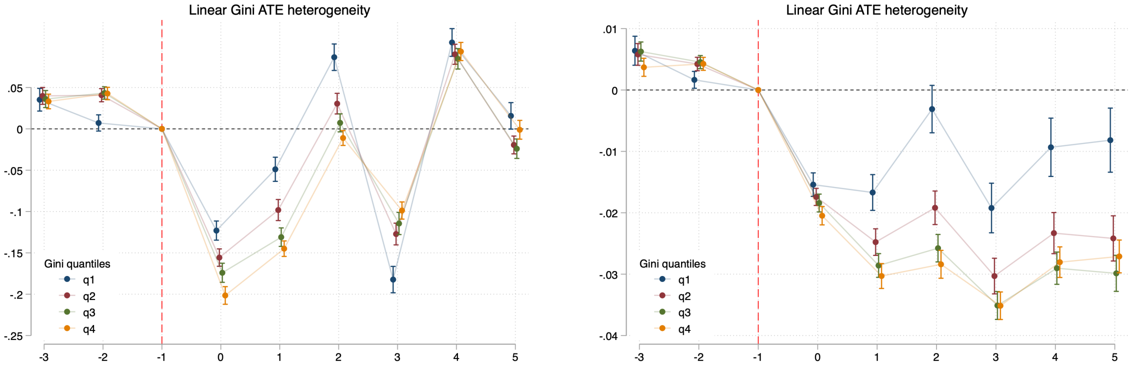
Notes. Treated cohorts exhibit 18 unique shock exposure paths that are sorted in descending order by proportion of the total 3,366 treated units. w corresponds to the week of unit's first shock.

Figure B.2: Dynamic heterogeneity-robust shock effect estimates on GPP flows and stocks using raw temperature threshold



(a) >31.8°C effects on flow over 8 weeks

(b) >31.8°C effects on stock (mid-May to date cumulative GPP)

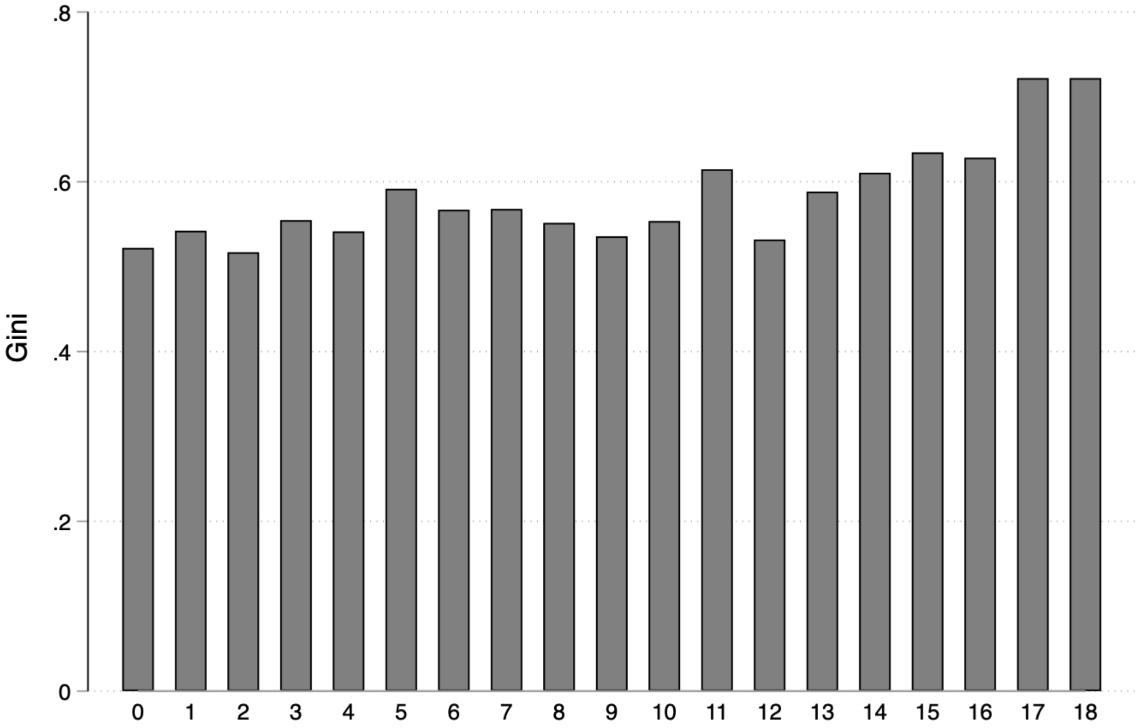


(c) >31.8°C effects on flows

(d) >31.8°C effects on stock (mid-May to date cumulative GPP)

Notes. Placebo and raw temperature threshold treatment estimates on log GPP using event study estimator B.9 with in reference to treated cohorts' initial pre-shock week using linear Gini and Gini quartiles up to six weeks. Threshold is the median stress temperature observed in the data, 31.83°C. Left column are flow effects, right cumulative GPP starting week 17. Standard errors clustered at the canton. 95% confidence intervals. N=355,024

Figure B.3: Land Gini distribution by shock exposure path



Notes. Mean land Gini by shock exposure paths corresponding to rows of Table B.3. Never-treated canton-year units take value 0.

Table B.4: Impacts on declared yields collapsed at the departement level

	Yearly log(Crop Yield (C.t/ha.))			
	(1)	(2)	(3)	(4)
Shock(k)	-0.089***	-0.090***	-0.086***	-0.088***
	[0.014]	[0.013]	[0.013]	[0.013]
Shock(k) x Gini(k)st		-0.031**		-0.015
		[0.014]		[0.014]
Shock(k) x Seminatural(k)st			0.029**	0.021
			[0.013]	[0.013]
N	658	658	658	658
Departement FEs	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓

Notes. Shock(k) is the (unweighted) mean of the canton-specific shock indicator over its nested cantons in departement k by year. Gini(k)st and Seminatural(k)st are the fixed mean Gini and seminatural allocation share of the nested cantons centered and standardized after collapsing. Crop yields correspond to marketable food crops transformed as the log of metric tons per hectare for each French *département* for years 2015-2021 by the French Agricultural Ministry. Standard errors in brackets indicate clustering at the *département*. * $p < .1$, ** $p < .05$, *** $p < .01$..

B.5 Level impacts on yields

Table B.5: Level impacts on declared yields

	Crop yield (t/ha.)			
	(1)	(2)	(3)	(4)
Shock	-0.65***	6.81**	-1.40***	1.03
	[0.11]	[3.14]	[0.25]	[2.27]
Shock x Gini		-24.51**		-7.46
		[10.47]		[7.67]
Shock x Gini^2		19.56**		5.48
		[8.34]		[6.04]
Shock x Seminatural			5.03***	5.08***
			[1.13]	[1.10]
Shock x Seminatural^2			-4.81***	-5.03***
			[1.13]	[1.16]
Cons.	8.36***	9.96***	8.22***	8.24***
	[0.02]	[0.51]	[0.03]	[0.35]
N	22189	22189	22189	22189
Canton FEs	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓

Notes. Crop yields correspond to marketable food crops for each French departement for years 2015-2021 by the French Agricultural Ministry. Yield expressed in metric tons per hectare. Gini and seminatural allocation are non-standardized. Standard errors clustered at the departement. * $p < .1$, ** $p < .05$, *** $p < .01$.

B.6 Direct biological moderating effects and their predictive power of Gini shock heterogeneity

Recall the event study estimator for arbitrary quantiles q of a baseline heterogeneity variable such as the Gini,

$$y_{iwt} = \sum_{q\in\{1,2,3,4\}} \mathbf{1}\{q_i = q\} \sum_{g}^{G} \left(\sum_{w}^{g-2} \theta^{q}_{gw} \cdot (1 - D_{iwt}) \cdot dg_{it} \cdot fw_t + \sum_{w\geq g}^{W} \tau^{q}_{gw} \cdot D_{iwt} \cdot dg_{it} \cdot fw_t \right)$$
$$+ \sum_{\substack{c\in \\ \mathcal{C}\setminus\{1\}}}^{C} c_i \left[\sum_{\substack{t\in \\ \mathcal{T}\setminus\{2015\}}}^{T} year_t + \sum_{\substack{w\in \\ \mathcal{W}\setminus\{17\}}}^{W} week_w \right] + e_{iwt}. \tag{B.11}$$

Direct effects by quantiles of the biological candidates is obtained by switching 2015 Gini quartile indicators with candidate quartile indicators.

The predictive power of a candidate biological mechanism in explaining week-of-shock Gini ATE heterogeneity is assessed as follows. Having estimated bio-quantile-specific ATEs $\hat{\tau}^{Bio_q}_{\omega,g=w}$ from the model above, we compute the *composition-weighted prediction* for each Gini quantile q':

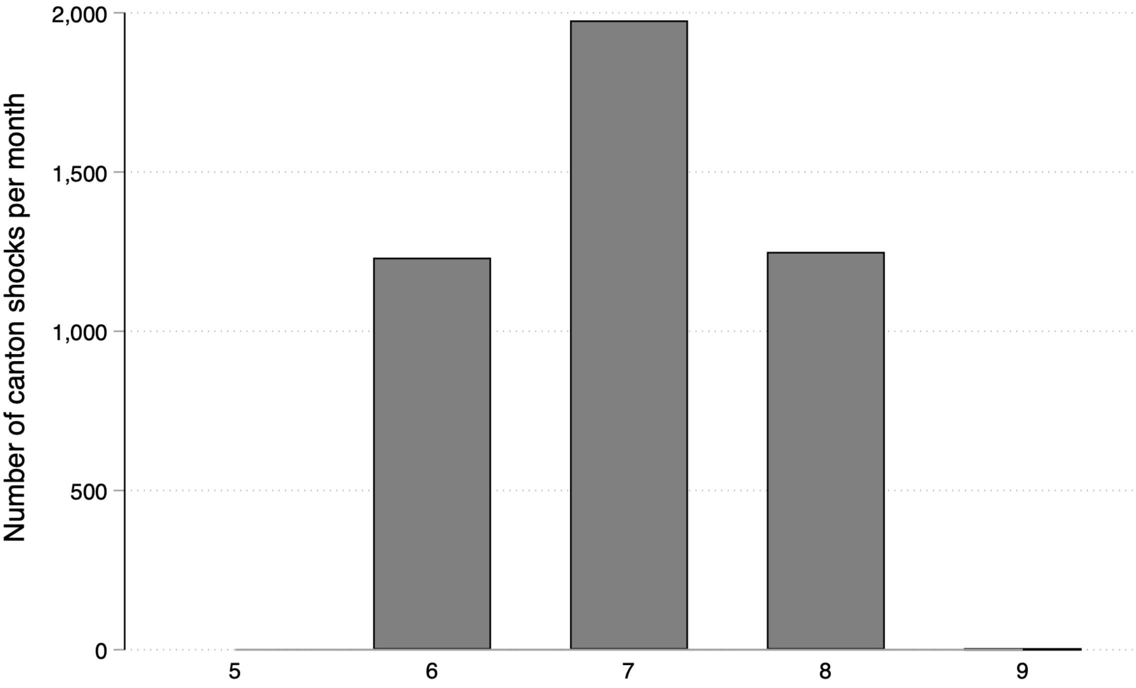
$$\hat{\tau}^{Bio}_{Gini=q'} = \sum_{q=1}^{4} \Pr\big(Bio = q \mid Gini = q', \text{exposed}\big) \cdot \hat{\tau}^{Bio_q}_{\omega,\, g=w}, \tag{B.12}$$

where $\Pr(Bio = q \mid Gini = q', \text{exposed})$ is the empirical share of bio quantile q among week-of-shock observations in Gini quantile q', and $\hat{\tau}^{Bio_q}_{\omega,g=w} = \sum_g \omega_g \hat{\tau}^{Bio_q}_{g,g=w}$ is the cohort-weighted week-of-shock ATE for bio quantile q.

Expression (B.12) asks: if this biological candidate fully explained shock heterogeneity, what Gini gradient would we predict given the empirical joint distribution of candidate and Gini quantiles among treated units? If $\hat{\tau}^{Bio}_{Gini=q'} \approx \hat{\tau}^{Gini_{q'}}_{\omega,g=w}$ for all q', the candidate accounts for the observed inequality–vulnerability gradient.

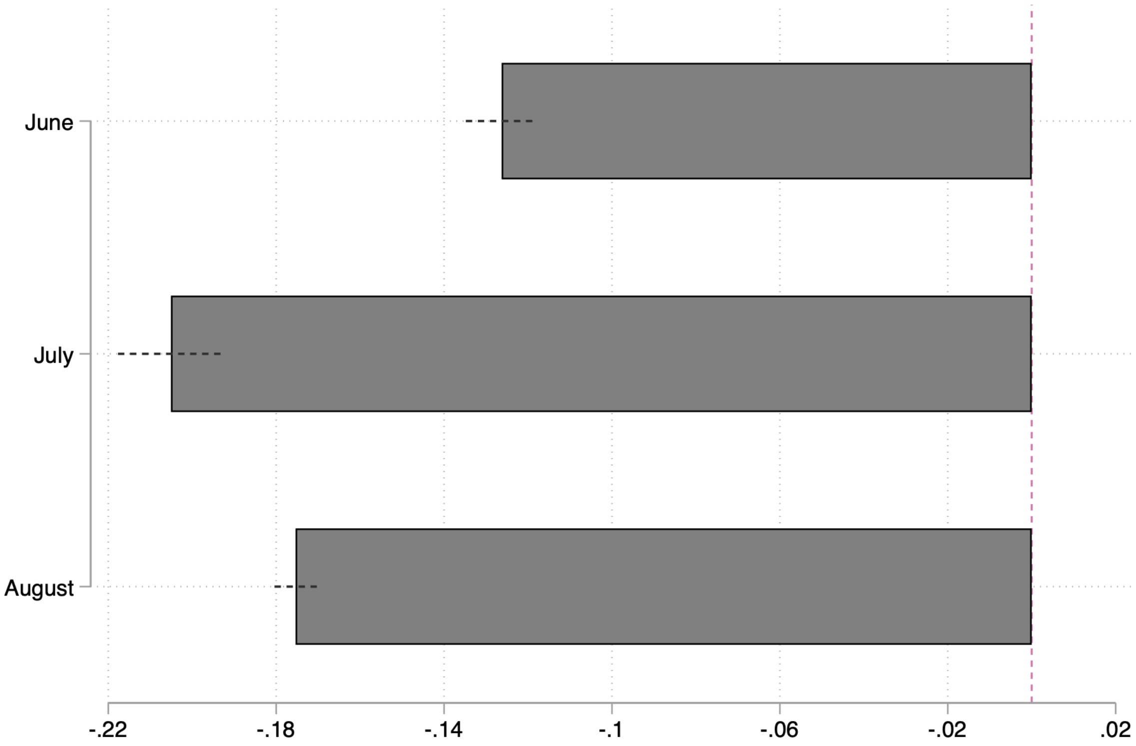
Table B.6 presents elasticity point estimates of different measures of crop diversity and the level of seminatural allocation with the Gini in panel (a) and partial elasticities with the proportion of farm-size shares. in panel (b). Also see Appendix E.1 for a spatial distribution of all these variables).

Figure B.4: Week-of-shock ATE on GPP flows by month



Notes. Total number of canton-shocks per month.

Figure B.5: Weighted shocks by month



Notes. Linear combinations of estimated parameters from the dynamic heterogeneity-robust week-of-shock MW specification B.6 with sample weighted aggregation by month.

B.7 Ecological buffers in the biology literature

Figure B.6: Seminatural distance-temperature gradients from Ghafarian et al. (2024)

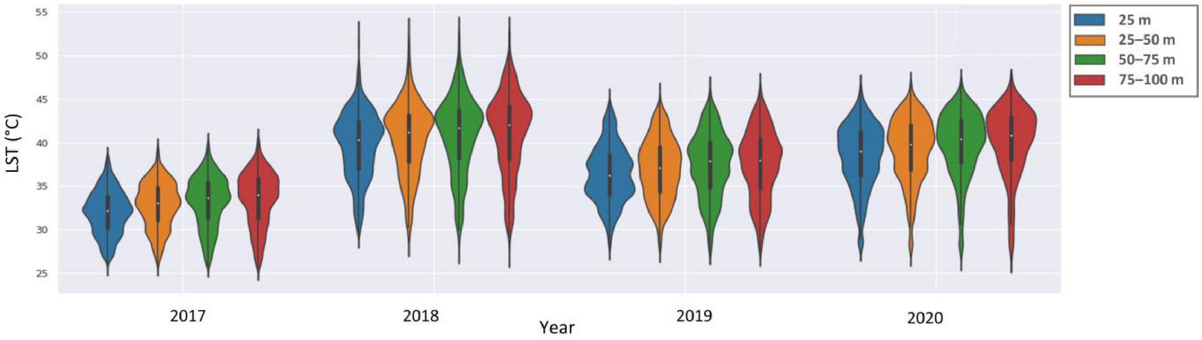


Table B.7: Mean and SD of distance-temperature gradient

	2017		2018		2019		2020	
Zone	Mean	SD	Mean	SD	Mean	SD	Mean	SD
25 m	32.7	2.1	40.9	3.4	37.3	2.8	39.3	2.9
25-50 m	33.6	2.2	41.8	3.5	38.2	2.8	40.1	3
50-75 m	34.2	2.3	42.5	3.6	38.8	2.9	40.8	2.9
75-100 m	34.5	2.3	42.8	3.7	39.1	3	41.2	2.9

Notes. Crop temperature-distance gradients from seminatural ecosystems from Ghafarian et al. (2024).

Table B.6: Elasticities of biodiversity with land Gini controlling share by size

	Log of			
	Inter/intra-crop species (1)	Crop HHI (2)	Crop Divers. ratio (3)	Seminatural allocation (4)
Panel (a): Gini				
log(Gini)	-1.175***	-0.080	0.007	-0.692***
	(0.188)	(0.092)	(0.051)	(0.153)
N	22189	22189	22087	22179
(AR x Year) FEs	✓	✓	✓	✓
Crop composition				✓
Panel (b): Farm size shares				
0-2 ha.	0.011***	0.011***	-0.002***	-0.006**
	(0.003)	(0.001)	(0.001)	(0.002)
2-10 (ref.)	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)
10-20	-0.012***	0.003**	-0.000	-0.009***
	(0.004)	(0.001)	(0.001)	(0.002)
20-50	-0.014***	0.003***	-0.001	-0.009***
	(0.003)	(0.001)	(0.001)	(0.002)
50-100	-0.027***	0.002	-0.000	-0.013***
	(0.004)	(0.001)	(0.001)	(0.004)
>100	-0.016***	0.007***	-0.000	-0.015***
	(0.004)	(0.002)	(0.002)	(0.003)
N	22189	22189	22087	22179
(AR x Year) FEs	✓	✓	✓	✓
Crop composition				✓

Notes. Elasticity estimates of crop diversity measures and seminatural allocation of farms in a canton with the Gini (a). Partial elasticity estimates of biological variables with control shares by scale relative to land held by operation of 2-10 ha. Standard errors are clustered at the canton level. * $p < .1$, ** $p < .05$, *** $p < .01$.

C Formalizing the relationship between the land Gini and holdings by farm size types

For a discrete distribution of farm sizes S, the Gini coefficient G can be expressed as

$$G = \frac{1}{2\,\bar{S}} \sum_i \sum_j \left| S_i - S_j \right| p_i\, p_j,$$

where p_i is the proportion of *farms* of type i and $\bar{S} = \sum_i p_i S_i$ is the mean farm size. For simplicity, partition farms into three ordered size categories $S_\mathrm{S} < S_\mathrm{M} < S_\mathrm{L}$ with farm count shares $p_s + p_m + p_l = 1$. The Gini reduces to

$$G = \frac{p_s\, p_m\, (S_\mathrm{M} - S_\mathrm{S}) + p_s\, p_l\, (S_\mathrm{L} - S_\mathrm{S}) + p_m\, p_l\, (S_\mathrm{L} - S_\mathrm{M})}{\bar{S}}.$$

However, we primarily use *land* shares: the proportion of total cropland controlled by each operational scale type. Let ω_k denote the land share of type k, with $\omega_s + \omega_m + \omega_l = 1$. Land shares and count shares are linked by the identity

$$\omega_k = \frac{p_k\, S_k}{\bar{S}}, \qquad \text{equivalently} \qquad p_k = \frac{\omega_k\, \bar{S}}{S_k}.$$

Substituting into the Gini formula and defining the inverse scale $\phi_k \equiv 1/S_k$ for each type, the Gini can be rewritten entirely in terms of land shares and type-specific inverse sizes:

$$G = \frac{\omega_s\, \omega_m\, (\phi_s - \phi_m) \; + \; \omega_s\, \omega_l\, (\phi_s - \phi_l) \; + \; \omega_m\, \omega_l\, (\phi_m - \phi_l)}{\omega_s\, \phi_s + \omega_m\, \phi_m + \omega_l\, \phi_l}. \tag{C.1}$$

The derivation follows from substituting $p_k = \omega_k \bar{S}/S_k$ into the three-type Gini formula and noting that $\bar{S} = (\sum_k \omega_k / S_k)^{-1}$, which cancels to yield the ratio above.[51]

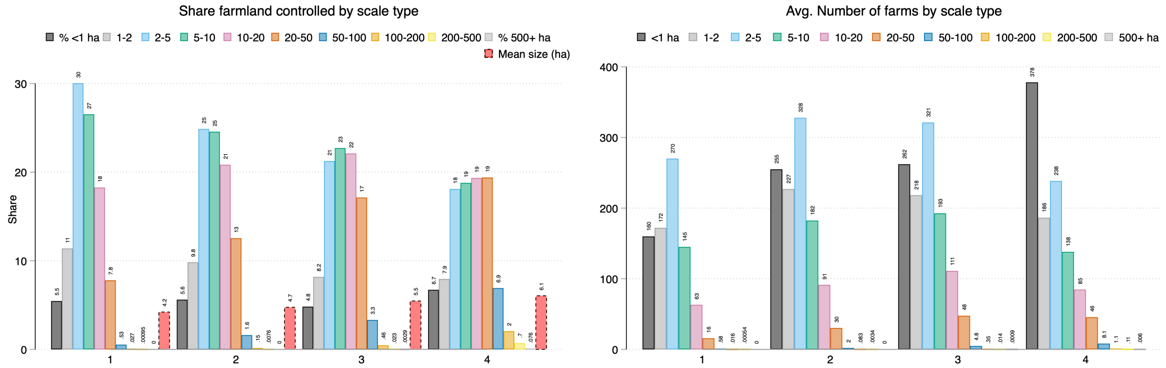
Expression (C.1) makes the Gini a function of land shares ω_k and type-specific inverse sizes ϕ_k. The ϕ_k values are approximately constant across cantons for a given bin, a "50–100 ha" farm is roughly 50–100 ha everywhere so cross-cantonal variation in the Gini is driven primarily by variation in land shares. This directly motivates regressing the Gini on the observed land shares in Table C.1.

The expression also clarifies how ownership polarization drives inequality. The numerator is a sum of pairwise land-share interactions $\omega_i\, \omega_j$ weighted by inverse-scale distances $(\phi_i - \phi_j)$. Since $\phi_s > \phi_l$ (small farms have much higher inverse size), the $\omega_s\, \omega_l\, (\phi_s - \phi_l)$ term dominates. Land consolidation that displaces mid-sized operations into the extremes. It raises $\omega_s\, \omega_l$ at the expense of ω_m. This amplifies this dominant term; moving up the Gini distribution, the land share of mid-sized farms falls while shares at both tails rise but much more so in the upper tail (Figure C.1).

---

[51] In detail: $|S_i - S_j|/(S_i S_j) = |\phi_j - \phi_i|$ for the numerator terms, and $\bar{S} = 1/(\sum_k \omega_k \phi_k)$ enters both as the pre-factor and in the denominator, producing the land-share-weighted harmonic mean in the denominator of (C.1).

Figure C.1: Share controlled and number by operational scale over Gini quartiles



(a) Share of total land

(b) Mean number

Notes. (a) Share of total farmland held by each farm-size bin across Gini quantiles. (b) Mean number of size types per canton.

Table C.1 provides the partial elasticity estimates and inference of the Gini with respect to total canton cropland by operational scale types in reference the share of land controlled by farms of 5-10 ha. The first column is unconditional while the second and third condition on our dynamic Agricultural Region fixed effects and canton level crop compositions. In our preferred model (column 3) we observe that a one percentage point increase in the share of land held by micro farms is associated with a 0.2% increase in land inequality while the share of crop operations of 50+ ha are associated with a 0.3-0.5% increase in the Gini in reference to a homologous change in the reference size type. As noted in the main text, share holding bins sum to 100% and are constant within year thus one interprets estimates as follows: a 1%-point increase in the type's controlled share – that displaces 1%-point of the omitted share – affects the outcome by 100%β. See Morawetz and Klaiber (2025).

C.1 Land inequality and crop specific effects

We now examine the association of land inequality and crop production. Table C.2 presents estimated elasticities of crop types with the Gini using separate regressions controlling for AR×Year fixed effects. Estimates are ordered by the proportion of surface area allocated to each type of crop. We observe that the production of wheat, corn, barley and colza (rapeseed) represents roughly 70% of crop land allocation in the cantons, on average: 32% wheat, 18% corn, 11.8% barley and 8.5% colza.

We also observe that elasticity estimates of the land allocated to these majority crops with respect to the land Gini is positive and varies between 1.09 for colza to 0.6 for corn.

Table C.1: Elasticity of Gini with respect to farm size types

Dep. var. log(Gini)	(1)	(2)	(3)
Share farmland controlled by scale type			
<1 ha.	0.004***	0.003***	0.002***
	(0.001)	(0.001)	(0.001)
1-2	0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
2-5	-0.002***	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
5-10 (ref.)	0.000	0.000	0.000
	(.)	(.)	(.)
10-20	0.000	0.001*	0.001*
	(0.001)	(0.001)	(0.001)
20-50	0.002***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)
50-100	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)
100-200	0.006***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)
200-500	0.004***	0.004***	0.003***
	(0.001)	(0.001)	(0.001)
	(0.003)	(0.001)	(0.001)
R-squared	0.36	0.68	0.69
N	22189	22189	22189
(AR x Year) FEs		✓	✓
Crop composition			✓

Notes. Gini elasticity estimates from an OLS regression of log(Gini) on shares of farmland controlled by each type of operational scale. The reference proportion is farms with 5-10 hectares of crop fields. Standard errors are clustered at the canton. * $p < .1$, ** $p < .05$, *** $p < .01$.

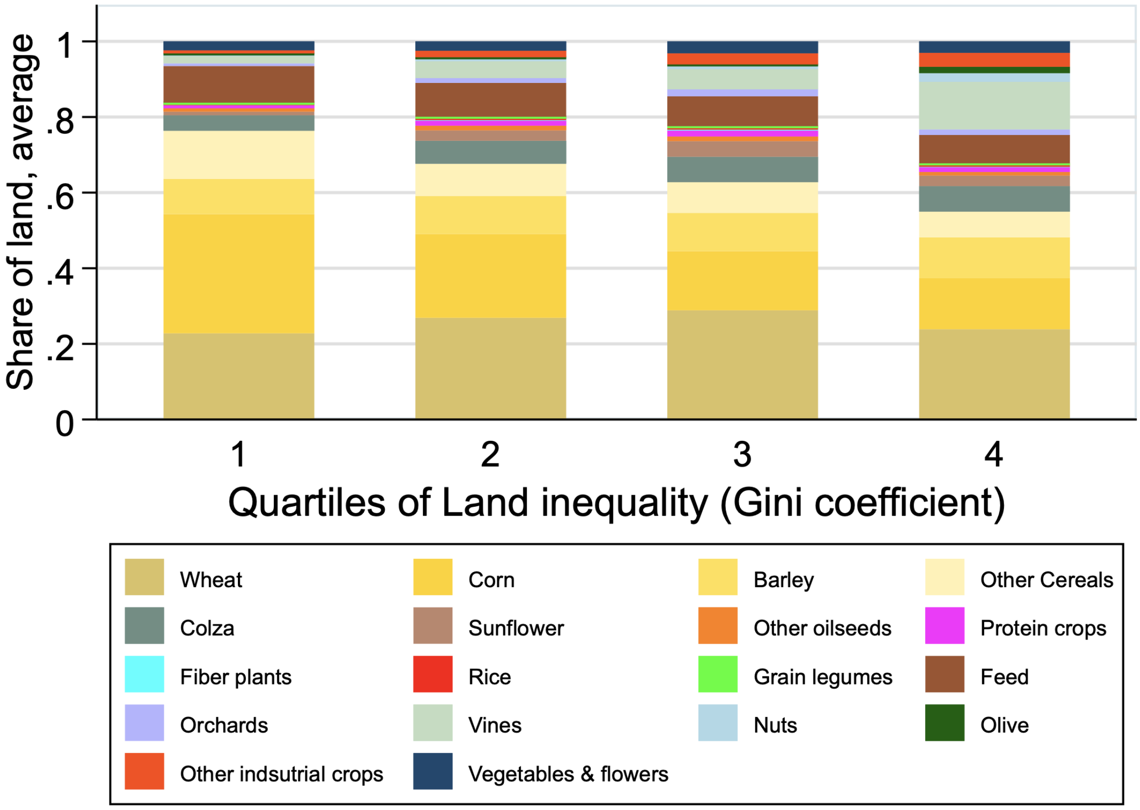
.

Table C.2: Elasticities of crop surface allocation with respect to Gini

	Mean ha. (1)	Pct. (2)	Elasticity ha. wrt Gini (3)	N (4)
Wheat	1980.2	0.320	0.735**	796128
			(0.359)	
Corn	1115.2	0.180	0.604	796128
			(0.463)	
Barley	729.0	0.118	0.710*	796128
			(0.366)	
Colza	527.2	0.0853	1.092***	796128
			(0.412)	
Other cereals	357.9	0.0579	0.498	796128
			(0.434)	
Feed	307.6	0.0498	0.0820	796128
			(0.418)	
Sunflower	236.7	0.0383	1.009**	796128
			(0.425)	
Vines	210.7	0.0341	4.885***	796128
			(0.464)	
Other indsutrial crops	202.6	0.0328	1.641***	796128
			(0.434)	
Vegetables and Flowers	156.2	0.0253	2.320***	796128
			(0.432)	
Protein crops	112.3	0.0182	0.737*	796128
			(0.406)	
Other oilseeds	70.23	0.0114	0.583	796128
			(0.400)	
Fiber plants	47.36	0.00766	0.0536	796128
			(0.388)	
Others	45.84	0.00741	-0.00240	796128
			(0.392)	
Orchards	40.61	0.00657	1.315***	796128
			(0.444)	
Grain legumes	21.50	0.00348	-0.0373	796128
			(0.314)	
Nuts	12.39	0.00200	-0.180	796128
			(0.236)	
Rice	4.809	0.000778	0.207**	796128
			(0.0948)	
Olive	3.524	0.000570	-0.240**	796128
			(0.114)	

Notes. Elasticity estimates from a OLS regressions of log of hectares of each crop type in a canton on the log of the land Gini. Standard errors are clustered at the canton level. * $p < .1$, ** $p < .05$, *** $p < .01$.

Figure C.2: Crop composition by Gini quantile



Note. Raw canton crop compositions by Gini quantile. 2015-2021.

D Proofs and further formalization

D.1 Formalizing returns to scale and non-linear increasing opportunity costs

The inequality

$$f(S) - f\big((1-\beta)S\big) \; > \; f'\big((1-\beta)S\big)\,\beta S$$

follows from f' being strictly increasing under IRS.

Define,

$$h(x) \;=\; f(S) \;-\; f(x).$$

When $x = (1-\beta)S$, the loss in output from not using the entire S is $h\big((1-\beta)S\big)$. By integrating $f'(t)$ from $t = x$ to $t = S$, we see

$$f(S) - f(x) \;=\; \int_x^S f'(t)\,\mathrm{d}t.$$

Because $f'(t)$ is increasing in t, we know

$$f'(x) \;<\; f'(t) \quad \text{for all } t > x.$$

Hence,

$$\int_x^S f'(t)\,\mathrm{d}t \;>\; \int_x^S f'(x)\,\mathrm{d}t \;=\; \big(S-x\big)\,f'(x).$$

Setting $x=(1-\beta)S$, we get

$$f(S)-f\big((1-\beta)S\big) \;=\; \int_{(1-\beta)S}^{S} f'(t)\,\mathrm{d}t \;>\; \big(S-(1-\beta)S\big)\,f'\big((1-\beta)S\big) \;=\; \beta S\,f'\big((1-\beta)S\big).$$

Thus, the actual output difference exceeds the linear (tangent-line) estimate creating the non-linearity.

D.2 A direct jump like $10\%\to 0\%$

Because the CAP schedule is stepwise in β, the policy-relevant margins are the moves across tier cutoffs, i.e. $10\%\to 7\%$ and $7\%\to 0\%$. In our baseline with per-hectare gains, the direct jump $10\%\to 0\%$ does not add an independent margin: both the gain and the subsidy loss decompose additively,

$$\Delta_{€/ha}(10\to 0)=\Delta_{€/ha}(10\to 7)+\Delta_{€/ha}(7\to 0), \qquad 80=20+60,$$

so a two-step adjustment weakly dominates a one-step jump by preserving option value. Hence,

$$S^*_{10\to 7} \;<\; S^*_{10\to 0} \;<\; S^*_{7\to 0}.$$

Although $10\to 0$ loses more maintenance than $7\to 0$ (€80 vs. €60 per ha), it also delivers a much larger output gain (a bigger recovered area and only a modest reduction in $g(\cdot)$ at $p=1$). The larger numerator is offset by the larger denominator in the ratio, so the $10\to 0$ frontier appears before $7\to 0$. Finally, $10\to 7$ has both the smallest hurdle and the smallest gain.

Accordingly, we present the move frontiers for $10\%\to 7\%$ and $7\%\to 0\%$ in Figure 10c, and the implied best-action regions in Figure 10e. The first frontier ($10\%\to 7\%$) appears at smaller sizes because it faces the smallest maintenance loss (€ 20/ha); the $7\%\to 0\%$ frontier appears later because it faces a larger hurdle (€ 60/ha). The breakpoints in the decision map align with these frontiers, partitioning farm sizes into stay at 10%, move to 7%, and, for sufficiently large S, move to 0%.

Fixed removal costs can play a role. When $K>0$ the frontiers all shift right by the term K/S in the discrete decision, but the ordering above is robust unless fixed costs are so large (or p / α so high) the inequalities are reversed. Imagine, fixed costs are per move (incurred twice in a two-step path) or if conversion costs are strictly convex (economies of doing it all at once), or if there are lumpy scale/contiguity gains (e.g. a discrete harvester upgrade attainable only below a β threshold), then $10\%\to 0\%$ can be privately optimal at sizes where neither step alone would be.

D.3 First-order condition with elasticity that varies over scale

The FOC in equation 7 can be written in terms of the scale elasticity of production, which makes transparent when farm size affects the privately optimal protection share. Let

$$x \equiv (1-\beta)S \tag{D.1}$$

denote cultivated land after the farmer allocates the share β of total land to protection. Using (D.1), (7) becomes

$$f'(x)\,S\,\big[(1-p)+p\alpha(\beta)\big]=p\,f(x)\,\alpha'(\beta).$$

Dividing both sides by $f(x)$ gives

$$\frac{xf'(x)}{f(x)}\,\frac{\big[(1-p)+p\alpha(\beta)\big]}{1-\beta}=p\,\alpha'(\beta).$$

Multiplying both sides by $(1-\beta)$, and defining the scale elasticity of production as

$$\varepsilon(x)\equiv\frac{xf'(x)}{f(x)},$$

we obtain

$$\varepsilon(x)\big[(1-p)+p\alpha(\beta)\big]=p\,\alpha'(\beta)(1-\beta).$$

Finally, substituting back $x=(1-\beta)S$ yields

$$\varepsilon\big((1-\beta)S\big)\big[(1-p)+p\alpha(\beta)\big]=p\,\alpha'(\beta)(1-\beta). \tag{D.2}$$

This shows that farm size enters the smooth first-order condition only through the term $\varepsilon((1-\beta)S)$. If $\varepsilon(x)$ increases with cultivated area, then larger farms face a higher value of $\varepsilon((1-\beta)S)$ at any given β. In that case, the left-hand side of (D.2) rises relative to the right-hand side when S increases. Intuitively, the marginal opportunity cost of reallocating land away from production is steeper at larger operating scales. A single-crossing condition ensures that the equality in (D.2) is restored at a lower value of β, so that the privately optimal protection share becomes decreasing in farm size.

D.3.1 Homogeneity

If the elasticity is constant, $\varepsilon(x)=\lambda$, as under the homogeneous production function $f(x)=Ax^{\lambda}$, then (D.2) reduces to

$$\lambda\big[(1-p)+p\alpha(\beta)\big]=p\,\alpha'(\beta)(1-\beta), \tag{D.3}$$

which does not depend on S. The privately optimal interior protection share β^* would therefore be independent of farm size. In this formulation, all otherwise identical farms choose the same privately optimal protected share. Any interior β^* is therefore defined implicitly by the equation above. Under this homogeneous benchmark, the shape of $\alpha(\beta)$ affects the level of β^*, but not its dependence on farm size. As discussed in Section 7, the size-dependent protection result, that large farms optimally choose lower or zero protection, therefore does not follow from the smooth marginal analysis under standard IRS alone. Rather, it arises from the discrete adoption calculus we model.

D.4 Land inequality, and aggregate provision of semi-natural areas

Now we show how land inequality converts the individual responses presented above into aggregate provision of semi-natural areas. The central insight is simply that, when protection decreases with farm size, aggregate provision depends only on how much land mass lies in the part of the farm-size distribution receiving low protection. Inequality thus matters insofar as it reallocates land toward the upper tail of the farm-size distribution. We first explore the implications for a general farm-scale distribution, and then illustrate it with a Pareto parameterization that delivers closed-form comparative statics in standard inequality metrics, including the Gini coefficient.

Let S denote farm size, viewed as a random variable with cumulative distribution function F. Let N be the number of farms. Total arable land in the region is fixed and given by

$$L \equiv N\,\mathbb{E}[S].$$

Increasing marginal opportunity costs of foregone land imply that, under a flat per-hectare payment, there exists a cutoff S^* such that farms with $S \leq S^*$ maintain a protected share $\bar{\beta} > 0$ while farms with $S > S^*$ do not:

$$\beta(S) = \begin{cases} \bar{\beta} & \text{if } S \leq S^*, \\ 0 & \text{if } S > S^*. \end{cases}$$

Equivalently, $\beta(S) = \bar{\beta}\,\mathbf{1}\{S \leq S^*\}$. Aggregate semi-natural land is therefore

$$B \equiv N\,\mathbb{E}[S\,\beta(S)] = N\,\bar{\beta}\,\mathbb{E}[S\,\mathbf{1}\{S \leq S^*\}]\,.$$

Define the land share above the cutoff as

$$T_F(S^*) \equiv \frac{\mathbb{E}[S\,\mathbf{1}\{S > S^*\}]}{\mathbb{E}[S]}.$$

Since $S = S\mathbf{1}\{S \leq S^*\} + S\mathbf{1}\{S > S^*\}$, we have $\mathbb{E}[S\,\mathbf{1}\{S \leq S^*\}] = \mathbb{E}[S](1 - T_F(S^*))$, so

$$B = \bar{\beta}\,L\big(1 - T_F(S^*)\big).$$

Thus, conditional on $(L, \bar{\beta}, S^*)$, aggregate provision depends only on $T_F(S^*)$. Any distributional change that increases $T_F(S^*)$ reduces aggregate semi-natural land B.

To make this mechanism transparent and to connect it to standard inequality measures, suppose farm sizes follow a Pareto distribution on $[1, \infty)$[52]:

$$\Pr(S > s) = s^{-\gamma}, \qquad \gamma > 1.$$

In this family, a decline in γ corresponds to a fatter upper tail and greater land concentration. The associated Gini coefficient is

$$G(\gamma) = \frac{1}{2\gamma - 1},$$

which is strictly decreasing in γ. Holding L fixed, the upper-tail land share is

$$T_F(S^*) = (S^*)^{-(\gamma-1)},$$

so aggregate semi-natural land is

$$B(\gamma) = \bar{\beta}\,L\left(1 - (S^*)^{-(\gamma-1)}\right).$$

Since $S^* > 1$, $B(\gamma)$ is strictly increasing in γ and therefore strictly decreasing in inequality as indexed by the Gini.

Proposition. *Holding total arable land L, the cutoff S^*, and the protected share $\bar{\beta}$ fixed, ag-*

---

[52] Pareto distributions are defined with a shape parameter and a scale parameter, the first one regulates the fatness of the tail, while the second one shifts the starting point of the distribution. We set the starting point to 1, adjustable to any scale, to normalize and simplify the notation.

gregate semi-natural land B is strictly decreasing in land inequality within the Pareto family, whether inequality is indexed by a fatter tail (lower γ) or by a higher Gini coefficient G.

From the expression above,

$$\frac{dB(\gamma)}{d\gamma} = \bar{\beta}\, L\, (S^*)^{-(\gamma-1)} \ln(S^*) > 0 \quad \text{for } S^* > 1.$$

Since $G(\gamma) = 1/(2\gamma - 1)$ is strictly decreasing in γ, it follows that B is strictly decreasing in G.

The general decomposition $B = \bar{\beta}L(1 - T_F(S^*))$ clarifies that land inequality matters through the land mass located above the cutoff: consolidation or other distributional shifts that increase $T_F(S^*)$ mechanically reduce aggregate provision under a scale-neutral payment. The Pareto case illustrates this channel sharply because tail thickness is governed by a single parameter that is monotone in common inequality metrics. More generally, the same qualitative prediction holds for any distributional change that increases the upper-tail land share above S^*.[53]

Our modeled mechanism is triggered by distributional changes that reallocate land into the upper tail above S^*, as in consolidation-driven increases in land concentration. Empirically, however, inequality can rise through other channels as well, and not every upward shift in land inequality produces our effect. For instance, if inequality increases through fragmentation, e.g., an initially homogeneous set of large farms splits into smaller units below the cutoff, aggregate protected area could remain unchanged or even increase. At the same time, the classic structural-transformation view that links growth to consolidation largely reflects the historical trajectory of today's high-income economies and is most evident in cross-country comparisons; within-country time series exhibit more heterogeneous farm-size trajectories once land endowments and demographic pressure are taken into account (Lowder et al., 2016, 2025). In practice, consolidation at the top can also coexist with fragmentation below, or even with farmland expansion, so observed inequality movements may bundle multiple processes. Given the extreme skewness of operated land in levels across and within countries (e.g., the top percentile of farms operates close to 70% of global farmland, Lowder et al., 2021), we consider the empirically relevant notion of land inequality to be one in which upper-tail concentration is the dominant margin.

D.4.1 Spillovers and free-riding

When seminatural ecosystem services travel across plot boundaries there is underinvestment and the level of land inequality moderates spillover effects.

Let the expected pay-off for farm i be

$$\Pi_i \;=\; g(\beta_i)\, f\big((1-\beta_i)S_i\big) \;+\; h\big(B_i\big), \quad g(\beta) = (1-p) + p\,\alpha(\beta),$$

where $B_i = \sum_j w_{ij}\,\beta_j S_j$ is the local services provision around i with spatial weights $w_{ij} \geq 0$ which capture the protection-distance gradients, up to $\sum_j w_{ij} = 1$. The farmer's first-order condition is

$$\underbrace{-\,g(\beta_i)\, f'\big((1-\beta_i)S_i\big)\, S_i}_{\text{marginal opportunity cost}} + \underbrace{p\,\alpha'(\beta_i)\, f\big((1-\beta_i)S_i\big)}_{\text{private mitigation gain}} + \underbrace{h'\big(B_i\big)\, w_{ii}\, S_i}_{\text{self-generated spillover}} \;=\; 0.$$

A social planner internalizes all spillovers, replacing $h'(\cdot)\, w_{ii}$ by $h'(\cdot)\sum_j w_{ij}$. The wedge is

---

[53] For instance, the same qualitative prediction holds under a lognormal parameterization: if $\ln S \sim \mathcal{N}(\mu, \sigma^2)$ and inequality rises through a mean-preserving increase in dispersion, then aggregate protected area is decreasing in σ and the Gini coefficient, provided the cutoff lies above the mean, $S^* > \mathbb{E}[S]$.

therefore

$$[\sum_j w_{ij} - w_{ii}]\; h'(B_i)\, S_i \;>\; 0.$$

For example, for a small change in one of the parameters (e.g. a per-protected-hectare payment τ), the policy margin is

$$\frac{dB}{d\tau} \;=\; \sum_i S_i \,\frac{\partial \beta_i}{\partial \tau}.$$

This implies nonnegative weights

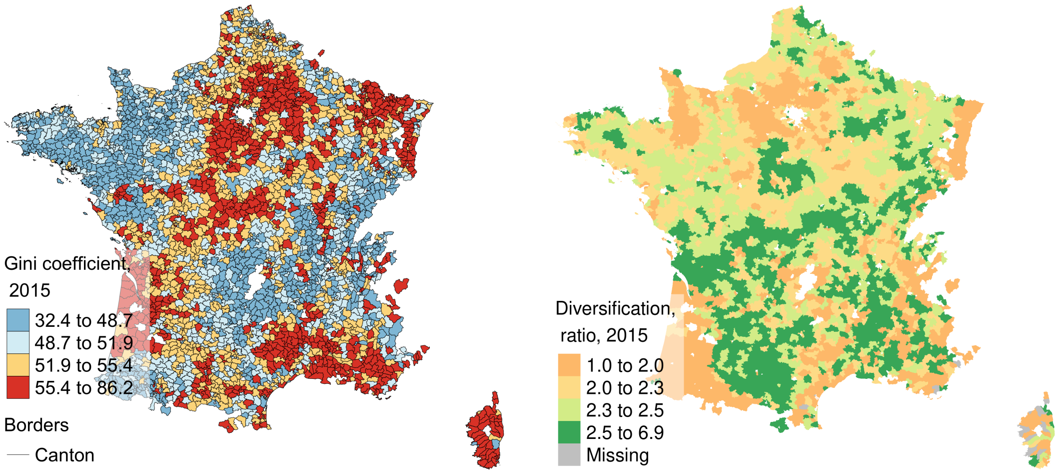
$$\mu_i \;\equiv\; \frac{S_i\,\left|\partial\beta_i/\partial\tau\right|}{\sum_j S_j\,\left|\partial\beta_j/\partial\tau\right|}, \qquad \sum_i \mu_i = 1.$$

Imagine that τ drops slightly. We can think of μ_i as the share of the "next" hectare of seminatural area controlled by farm i that will be removed. If μ_i is an increasing function of S_i under IRS this concentrates the margin in the upper tail of S_i.

E Additional tables and figures

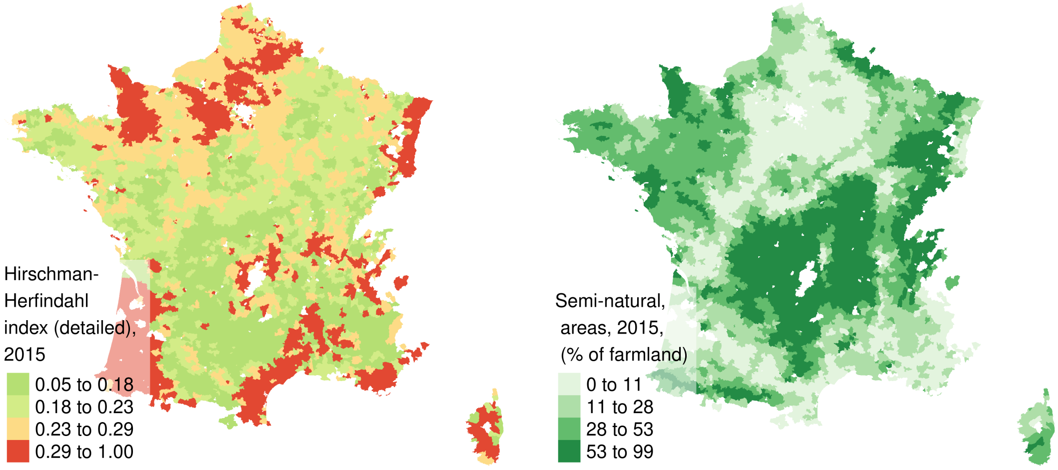
E.1 The spatial distribution of heterogeneity variables

Figure E.1: Different dimensions of diversity at the canton level



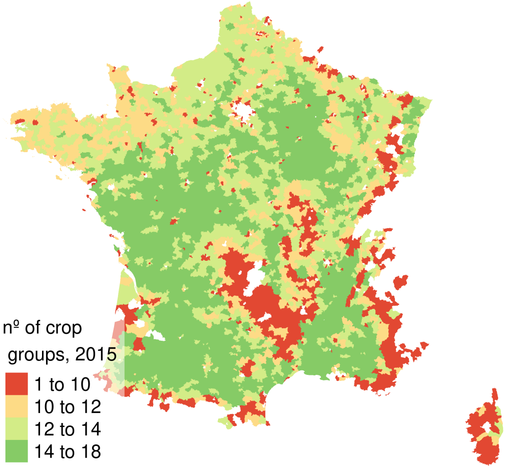
(a) Gini

(b) The diversification ratio



(c) The crop concentration index

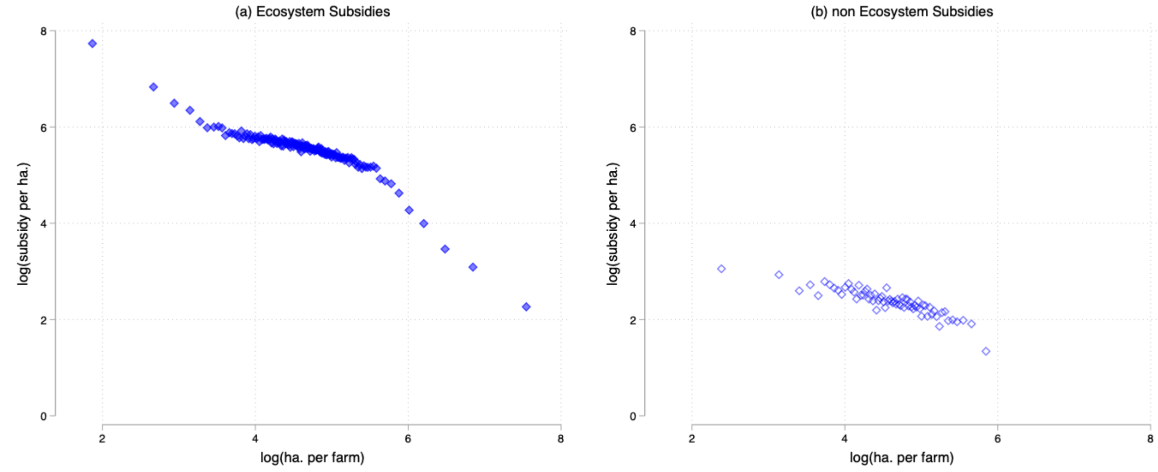
(d) The presence of semi-natural areas



(e) Counting crop groups

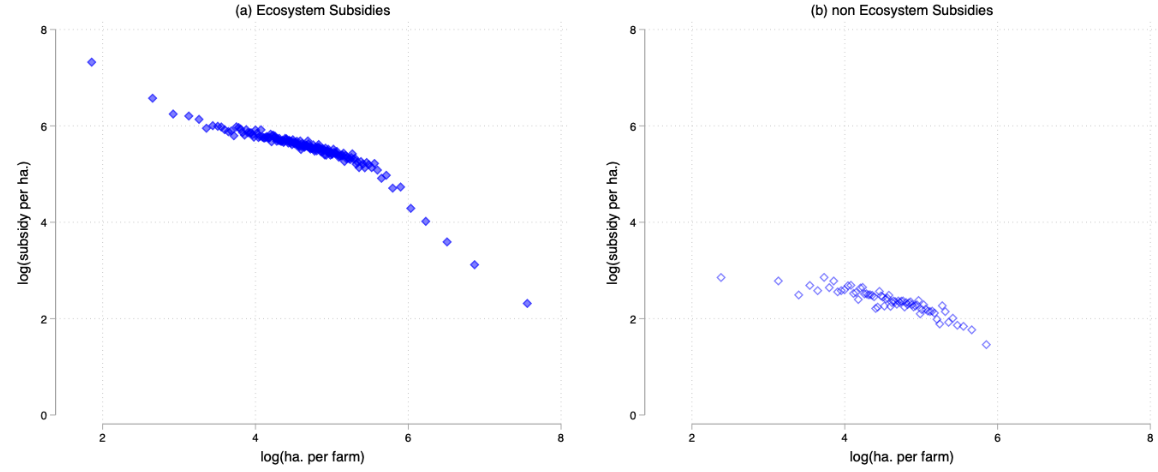
Note. Authors' elaboration based on the *Registre Parcellaire Graphique* (RPG) of 2015.

Figure E.2: PAC subsidies disbursed to municipalities with controls.



N bins: 160 and 71

(a) Conditional on departement fixed effects



N bins: 165 and 72

(b) Conditional on postal code fixed effects

Notes. N=16,220. Log subsidy amount received by communes over log mean farm size (SAU) of the commune in 2019. Optimal bin selection and partitioning follows the integrated mean squared error minimization algorithm of Cattaneo et al. (2024).