An Empirical Model of Agricultural Subsidies with Environmental Externalities*

Tristan Du Puy†

Job Market Paper

(Updated regularly, please click here for the latest version)

November 25, 2024

Abstract: *Agricultural subsidies are ubiquitous around the world, representing at least \$749 billion per year in direct transfers. In the European Union (EU), they account for 45% of all federal expenditures since 1980. At the same time, the agricultural industry is a major source of environmental disruptions, in part because of its use of chemicals, which translate into air, soil and water pollution. Subsidies are large, but their influence on the equilibrium use of agricultural chemicals is ambiguous. In this project I leverage farm-level administrative data to study their impact on chemical pollution and economic surplus. A shift-share design based on the 1992-1995 MacSharry reform of EU subsidies, the largest reform to date, shows how combined decreases in subsidy levels and changes in their design reduced farm profit, and led to exit and reallocations. The reform also lowered both farm-level chemical use and water pollution as measured via remote sensing. In an empirical model of dynamic land use, where producers differ in efficiency and propensity to pollute, I find that more efficient farms pollute on average more. Counterfactual analysis shows that subsidies which reallocate production towards low pollution producers have a small impact on aggregate pollution. Budget-equivalent subsidies that shift the incentives for the use of polluting inputs have larger effects and can reproduce part of the welfare gains of Pigouvian taxation.*

*I thank my advisors Suresh Naidu, Pietro Tebaldi, Wolfram Schlenker, Gautam Gowrisankaran and Jeffrey Shrader for their guidance during this project. I am also indebted to Matthew Backus, Mathias Reynaert, Bernard Salanié, Sylvain Chabé-Ferret, Eric Verhoogen, Jack Willis, Allan Hsiao, Robert Metcalfe, Rodrigo Soares, Karam Kang, Sebástian Otero, Diego Restuccia, Marguerite Obolensky and Nicolas Longuet-Marx for their generous feedback, as well as Yulia Kim for her support. I am grateful for comments from numerous participants in seminars and conferences at Columbia University, the Toulouse School of Economics, the CUNY Stone Center on Socio-Economic Inequality, NAREA and WEAI. All remaining errors are my own. I would like to acknowledge funding from the Columbia Center for Political Economy and the Columbia Center for Environmental Economics and Policy, and thank INSEE, the CASD and the French Ministry of Agriculture for data access.

†Du Puy: School of International and Public Affairs, Columbia University (td2631@columbia.edu)

1 Introduction

Agricultural subsidies are ubiquitous around the world. Recent estimates place net government agricultural transfers at $749 billion per year (Carisma and Boero (2024)). The outsize role of agricultural subsidies within public spending is particularly salient in the European Union (EU) where the Common Agricultural Policy (CAP) has accounted for 45% of all federal spending since 1980. A long literature has detailed the negative effects of agricultural subsidies for economic surplus, highlighting how they distort the farm-size distribution, and inflate the share of output allocated to small and economically inefficient establishments (Adamopoulos and Restuccia (2014)). The more elusive aspect of these subsidies is their interaction with the environmental footprint of agriculture, and its emission of local and global pollutants.[1] Agricultural fertilization accounts for 23% of the global annual fixation of nitrogen in marine and terrestrial ecosystems, with consequences for biodiversity and climate change (Fowler et al. (2013)). Recent research has also identified agricultural chemicals as the main force behind the 57% decline in the EU farmland bird population over 1980-2016 (Rigal et al. (2023)).[2] Like other government subsidies to polluting or nature-depleting industries (Davis (2017), Shapiro (2021)), the environmental consequences of agricultural subsidies is a central concern for environmental policy.

This paper studies the impact of the EU CAP on the environmental pollution stemming from the use of chemicals in agriculture, and compares it to its impact on economic surplus. I focus on chemical use, as it is both a fundamental input in modern agriculture, and a significant source of externalities. In France and for wheat, joint fertilizer and pesticide expenses account for 24% of all production costs (Guillermet (2015)). Farm chemical pollution intensity is also easy to measure using production data, and some of its consequences on water pollution can be tracked via remote sensing. To identify the role of subsidies on pollution, one needs to understand the drivers of aggregate chemical use. A key element for this is the relation between farms' reliance on chemicals and their production costs, and through this how the allocation of production across producers affects total chemical use. If more efficient producers pollute less, the efficient allocation of production across producers and the environmentally sustainable one should match (Qi et al. (2021), Ryan and Sudarshan (2021)). Distortions away from this allocation are not desirable, and externalities can be directly addressed by shifting within-farm incentives of production and input choice. In this world, the environmental concerns add to the usual efficiency argument for consolidation. If there are complementarities between low production costs and high pollution intensity, moving towards a more efficient equilibrium potentially entails moving away from a more

[1] Agricultural environmental externalities are diverse. They range from a reliance on slowly renewing aquifers for irrigation (e.g. Hornbeck and Keskin (2014), Ryan and Sudarshan (2021), Carleton et al. (2023), Taylor (2024)), air, soil and water pollution (Dias et al. (2023), Missirian (2020), Chabé-Ferret et al. (2021), Taylor (2022), Frank (2021)), significant greenhouse gas emissions in the form of N_2O and CH_4, to local ecosystem disruptions directly stemming from the reshaping of rural landscapes.

[2] Agriculture accounts for 21% of all emissions if I also include its impact on land use (Poore and Nemecek (2018)). Within the EU, 10% of emissions are associated with agricultural production, and nitrous oxide emissions linked to fertilizer use account for a third of them.

sustainable one. In this case, the welfare consequences of selecting farms on efficiency are hinged upon a trade-off between gains in economics surplus and increases in pollution. Because subsidies distort the allocation of production across producers, their consequences for pollution are theoretically ambiguous and depend on this empirical relation.

The environmental implications of agricultural subsidies will be context-specific. I study the EU CAP within the French cereal and oil crop market, which is the largest market of the EU and one of the most important world exporter of grains.[3,4] I make several contributions to the analysis of agricultural subsidies. I propose the first causal estimates for the 1992-1995 MacSharry reform, the largest CAP reform to date, and further develop a welfare analysis of the contemporary CAP subsidies encompassing economic surplus and environmental pollution. As a point of comparison, yearly CAP spending was on average 30% larger than the US Farm Bill around the reform (Johnson et al. (2010)). I rely on a dynamic model of land use, where I pay attention to cross-farm heterogeneity in efficiency and pollution intensity. This allows me to understand both the direct effect of subsidies on farm behavior, and the one on cross-farm reallocations. Importantly, farms are modeled as multi-product establishments. Within a farm, chemical-use per euro produced will consequently depend on its crop-mix composition, as I allow crops to have different chemical needs. It will also depend on the farm's idiosyncratic efficiency at using chemicals. Both crop-specific production technologies, and the distribution of chemical efficiency are recovered using French administrative micro-data. My functional form assumptions do not impose a sign for the relation between farm efficiency and pollution intensity. The recovered parameters entail it is a positive one. These relations are nested in a model where farms make dynamic decisions which endogenize the equilibrium allocation of production across producers.

I start by studying the impact of the MacSharry reform of the CAP on the French market. This reform was implemented in the EU to allow for the ratification of the Uruguay Round of World Trade Organization negotiations (1986-1994). The analysis highlights key channels through which subsidies impact farm behavior and aggregate pollution. The reform removed or reduced commodity price floors for cereal and oil crops, leading European prices to converge to international market levels. It further introduced a partial compensation in the form of a land subsidy. While the fall in prices was heterogeneous across crops, land subsidies were fairly homogeneous. As such, farms' ex-ante exposure to the reform varied with their pre-reform crop mix. I rely on a shift-share design which uses the heterogeneity across farms in pre-reform crop mix for identification, and control for the endogenous determinants of crop choice. I argue for causality following the framework of Goldsmith-Pinkham et al. (2020), where the shift-share approach relates to a difference-in-difference analysis. At the farm level, a one standard deviation increase in reform exposure leads to decreases in sales (25%) and profits (38%), as well as in total chemical use (19%). Within French municipalities,

[3]For practical reasons, I will consistently address the European Community or European Union as the European Union throughout the paper, regardless of whether I am talking of pre-1992 or post-1992 years. The EU was created by the 1992 Maastricht Treaty, and the pre-1992 corresponding entity was the European Community.

[4]See https://ourworldindata.org/grapher/cereals-imports-vs-exports?tab=table&time=2018.

a one standard deviation increase in the median farm exposure leads to a .5 decrease in farm population. For 36,000 municipalities in France, this represents a significant impact of additional exposure on exit. Further, increased exposure raises the minimum farm size within municipality by 1ha. I finally use remote sensing data from Landsat 5 to build an index for algal blooms, following the method of Taylor and Heal (2023). These blooms are regularly caused by over-fertilization and spikes in nitrogen which perturb an ecosystem's nutrient balance. They have detrimental consequences for ecosystem health, wildlife, human health and for climate change through methane emissions (Rossi et al. (2023)). The decrease in chemical use at the farm level is matched by a 7% decrease in algal blooms at the county level by 1999, with respect to their 1991 level.

These results illustrate how subsidies impact profitability and survival, and how they can also shift the incentives for relative input use. However, this analysis only recovers the effect of marginal additional reform exposure, and does not distinguish between the role of changes in subsidy level and design. My structural approach accounts for the equilibrium effects of subsidies, and allows me to compare the effect of different designs through counterfactual analyses. I develop a competitive partial equilibrium model of farm dynamics, paying particular attention to entry, exit and capital accumulation. The model follows the structure of Hopenhayn (1992), but allows for multi-product firms and capital accumulation, respectively in the spirit of Mayer et al. (2014) and Clementi and Palazzo (2016). I also specify entry costs which adjust with market size, using intuitions from Klenow and Li (2024) for their estimation. The model generates two key equilibrium relations: one between farm efficiency and pollution intensity, and another between aggregate production and total chemical use.

The structural analysis is organized around a demand system composed of constant elasticity demand curves, and the static and dynamic elements of supply: crop-specific production functions, and a set of parameters for the farms' dynamic decisions.[5] There are two important features to the estimation of production functions. The first one is the presence of chemical-biased productivity shocks on top of input-neutral ones. The second is that I model multi-product farms but observe most inputs at the farm-level. To recover chemical-biased productivity shocks, I use the first order conditions of the parametrized farm problem, following Doraszelski and Jomandreu (2018).[6] To account for multi-product producers, I follow De Loecker et al. (2016) and rely on single product farms for estimation. For these, the observed farm-level inputs are all allocated towards the same crop. This sample selection introduces an estimation bias which I address, adapting the selection correction of Olley and Pakes (1996) and De Loecker et al. (2016) to my context. I then use the estimated parameters to recover input allocations and unobserved shocks for the multi-product farms. I find a positive relation between the chemical intensity of production and farm profit. In

[5] The main specification does not model a possibility for cross-crop substitution in demand – but my extension using monopolistic competition gives me within and across crop categories substitution patterns and serves as a point of comparison.

[6] Their approach differs from the semi-parametric one of Ackerberg et al. (2015), by deriving estimating equations from the firms' profit maximization problem.

my competitive model, this also entails a positive relation between pollution intensity and efficiency. Because of the substitutability between chemicals and land, a farm with a relatively higher ability at using chemicals will switch its input shares towards chemicals and away from land. One might intuitively expect land and chemicals to be substitutes. Organic farming tends to have lower yields, and to compensate for these by increasing land use. I finally use indirect inference to estimate the costs of capital adjustment, a fixed cost paid by farms every period, and the parameters governing the adjustment of the entry cost to market size. Farm decisions are dynamic in at least two ways. Entry and exit rely on expectations about future profit. Capital accumulation is similarly forward-looking. It is also dynamic in the sense that accumulation might take time, especially if farms face convex costs of adjustment. In my context, capital is composed of buildings and machines. Larger investments might lead to larger disruptions in the production process, for example in the form of learning costs for new machinery. The relative volatility of farm-level input-neutral and chemical-biased shocks, and their interaction with capital accumulation determines how production gets allocated across producers through both within-period production decisions, and across-periods capital accumulation and exit decisions (Collard-Wexler (2013)).

Policy Implications: The counterfactual analysis is composed of three exercises. I first revisit the MacSharry reform. The simulations show it led to a 20% decrease in the average costs of production on the French market, and a 38% decline in its chemicals-to-output ratio. Both the reform-induced exit of low efficiency producers and decreases in production scales lowered costs. This same exit of low efficiency producers raised pollution, as these farms pollute less. However, this effect was dominated by the introduction of a land subsidy which significantly increased the relative price of chemicals.

I then study two different designs, which highlight two channels through which subsidies interact with pollution. First a land subsidy, and then a lump-sum payment for low pollution-intensity, or organic, producers. The subsidy leads to within-farm within-crop changes in the relative use of chemicals. It impacts all producers and has a large effect on aggregate chemical use. The lump sum payment allows low pollution producers to survive on the market. By doing so, it reallocates production away from high pollution farms. However this lump-sum only impacts aggregate outcomes through changes at the farm survival margin, and hence focuses on the lowest market share producers. Consequently, the lump-sum has a small impact on total chemical use. I recover the welfare impacts of these policies at different levels of intervention, as well as their marginal value of public funds (MVPF). In so doing, I follow the approach of Hahn et al. (2024) who study the effectiveness of climate change policies in the United States. For larger valuations of chemical pollution, the land subsidy has an MVPF above 1, comparable to the effects of the US electric vehicle policies and appliance rebates studied by Hahn et al. (2024).

I finally compare the land subsidy to a chemical tax. For every subsidy level, the tax leading to a similar change in the relative chemical-to-land price has a larger effect on total chemical use. Indeed, the tax shifts the price of chemicals relative to all inputs, and not only land. I characterize the extent to which a land subsidy can replicate the welfare gains of taxation. I compare the producers' willingness to support the two policies by tracking the evolution of their surplus. For a range of intervention levels, the subsidy potentially trades lower welfare gains with larger political support from producers.

Related Literature: This paper makes a series of contributions. I first estimate the elasticity of substitution between land and chemicals in French cereal and oil crop farming. I combine it with the distribution of chemical-efficiency across farms to obtain a relation between farm efficiency and pollution. Discussions of input-bias technological change in agriculture go back at least to Binswanger (1974) and the comparative study of Hayami and Ruttan (1971), and has been developed extensively to study the causes and consequences of agricultural productivity growth (recent examples include Bustos et al. (2016) and Clemens et al. (2018)). My production function is closed to the one outlined by Hayami and Ruttan (1971), which was estimated using country-level observations due to the unavailability of farm-level data. I can rely on micro-data to recover land-to-chemicals and fertilizers-to-pesticides elasticities, as well as the distribution of input-biased production efficiency, while addressing issues of selection and endogeneity in the estimation of production functions. By nesting this within-farm relation between profitability and pollution intensity in a model of farm dynamics, my analysis relates to the work of Oberfield and Raval (2021) who study how cross-firm heterogeneity in labor-biased productivity drives the fall of the aggregate labor share in U.S. manufacturing. Like them I recover an aggregate elasticity, here between land and chemicals, from a micro-level elasticity and the allocation of production across heterogeneous farms. In my implementation, I account for endogeneity in the estimation of elasticities, by following the approach of Doraszelski and Jomandreu (2018). Substitution between chemicals and land can be driven by both variation in relative input prices, or by changes in within-farm chemical efficiency which is unobserved. I rely on a series of timing and exogeneity assumptions to control for this second driver of the chemical-to-land ratio. I further use the method of De Loecker et al. (2016) to recover crop-specific production functions and account for the fact that farms are multi-product producers. Maue et al. (2020) was a recent implementation of production function estimation to agriculture. I build on their work by addressing these important features of agricultural production.

Farm-level heterogeneity and production functions are nested within a model of farm dynamics. Farms make a series of dynamic decisions related to entry, exit, crop mix and capital accumulation. Modeling multi-product farms with capital stocks adds to the growing industrial organization literature looking at dynamic land use, where farms often had to be modeled as atomistic fields because of data limitations (Scott (2013), Hsiao (2022), Burlig et al. (2024)). This paper also relates to the literature studying the environmental externalities associated with agricultural production: from chemical pollution (Frank (2021), Dias et al. (2023), Missirian (2020), Taylor (2022), Taylor and Heal (2023), Chabé-Ferret et al. (2021), Rossi et al. (2023)), deforestation (Balboni et al. (2023) for a review), impacts on local climate (Braun and Schlenker (2023) and Grosset-Touba et al. (2024)) to irrigation and groundwater depletion (Hornbeck and Keskin (2014), Burlig et al. (2024), Carleton et al. (2023), Ryan and Sudarshan (2021)). I propose a model for the equilibrium relation between efficiency and pollution intensity, which allows me to study the effects of different policy designs on pollution. In so doing, the paper also relates to the literature studying the environmental consequences of production subsidies. This literature has analyzed both the effects of input-biased subsidies favoring the use of polluting inputs such as fuel (Davis (2017), Englander et al. (2023)), and the indirect subsidization via trade regulation of relatively more polluting industries (Shapiro (2021)). Here I provide a reduced-form causal analysis of the EU MacSharry reform. I can recover the equilibrium effect of the reform thanks to my model, as the exogenous variation available for the reduced form only captures the consequences of marginal additional reform exposure. The model also yields the equilibrium environmental gains from EU agricultural subsidies, accounting for both marginal and infra-marginal effects. It allows me to compute the marginal value of public funds for different types of subsidy designs, following the comparative analysis of Hahn et al. (2024) studying U.S. climate change policies. Focusing on equilibrium outcomes, my analysis relates to the industrial organization literature looking at the equilibrium modeling of environmentally relevant markets, and paying particular care to agent heterogeneity and dynamics (Gowrisankaran et al. (2016), Souza-Rodrigues (2018), Blundell et al. (2020) Ryan and Sudarshan (2021), Rafey (2023), Aronoff and Rafey (2024), Aspelund and Russo (2024)). EU policy makers have been relying on the CAPRI general equilibrium model to study the effects of the CAP on production and the environment (Britz et al. (2007)). Compared to CAPRI, I propose a model of dynamic heterogeneous farms within the same crop market. My model is estimated using micro-data, addressing endogeneity, and links subsidies to both within farm decisions, and across farms allocations.

Finally, I study the impact of industrial policy intervention on market efficiency and social welfare, drawing from a rich literature on firm dynamics and specifically papers following Hopenhayn (1992) and Hopenhayn and Rogerson (1993). I highlight the comparative effect of policy on within-firm effects, and cross-firm reallocations, in that being close to Backus (2020). Here, policy can change the farms' behavior via changes in crop-level production, and in crop mix composition. It can also lead to reallocation effects through changes in market shares across incumbents, and through entry and exit movements. There is a large literature studying misallocations within agriculture (Adamopoulos and Restuccia (2014), Foster and Rosenzweig (2022), Lagakos and Waugh (2013), Gollin et al. (2021)). The French administrative data has the advantage to provide farm-crop level output, price and land use data. This allows me to recover precise measures of farm production efficiency which is important to recover the consequences of reallocations on aggregate efficiency.

This paper proceeds as follows: Section 2 provides background on the French market, the CAP, the MacSharry reform and describes the data. Section 3 presents reduced form evidence on the consequences of the reform, and the channels at play regarding selection and within-firm treatment. Section 4 outlines the structural model and the estimation strategy, Section 5 presents counterfactual simulations. Section 6 concludes

2 Context and Data

I discuss in turn the French and European agricultural markets, and the goals of the European Union Common Agricultural Policy within these markets. I then address the context in which the MacSharry reform was introduced, and its likely impact on French agriculture. I end by describing the data used for the analysis.

2.1 French and European Agriculture

The Common Agricultural Policy was first designed to foster the reconstruction of post-World War II European agriculture, encourage production within the union and support agricultural income (Petrick (2008)). It aimed at doing so by creating a common agricultural market within the European Union, and by implementing a common support scheme directly financed by the community's budget.[7] The free market for cereals was introduced in 1967. It was accompanied by a system of price support—floor prices implemented through public purchasing guarantees, export subsidies and import levies—jointly aimed at creating a buffer between world prices and European prices in order to support farm revenue. Purchasing guarantees were paired with minimum sales volumes, and were as such effective at the level of wholesalers – indirectly influencing farm-gate prices.[8] Price floors were seen as both stabilizing and supporting agricultural income, and as a way to boost productivity by favoring investments. The common market expanded to most of the other agricultural commodities over the following decades, with the introduction of intervention prices for oil crops in the late 1960s, and varying forms of price control for protein and textile crops, rice, sugar, meat, milk and milk products, and some vegetables and fruits. In the mid to late 1980s, however, the CAP faced a very different industry, and declining world prices following the boom of the 1970s (Gardner (2002)). European agriculture was now producing large surpluses, and internal pressure was building to reform the CAP.

2.2 The MacSharry Reform

The MacSharry reform was implemented between 1992 and 1995, after years of unsuccessful reform attempts, and can be considered as the largest reform of the EU CAP to date. The

[7]The 1950s in the US were dominated by debates surrounding the farm problem (Gardner (2002)), broadly defined as the combination of low earnings in agriculture, high income uncertainty, and an over-allocation of labor to agriculture (Schultz (1945)). While Western Europe shared the concerns regarding low and stagnating agricultural income, the focus on increasing productivity and decreasing the labor share of agriculture were amplified by the need to rebuild its industry, and by the desire to have local supply meet demand.

[8]Intervention prices were first regionalized, and later harmonized across the Community in the 1970s. They were first set in Unit of Accounts, a fictional common currency introduced for the purpose of the Common Agricultural Policy, later in ECUs, and then euros. For that purpose, specific agricultural exchange rates, called the "green rates", were implemented to prevent within community arbitrage following any re-evaluation of the members' currencies.

reform happened at a pivotal moment for the EU, and at the conjunction of the voting of the 1992 Maastricht Treaty creating the European Union, and of ongoing trade negotiations under the umbrella of the GATT Uruguay round. These external forces justify taking the timing and design of the reform as exogenous to local French agricultural conditions.

The Maastricht Treaty was signed in 1992, and prepared for the introduction of the European currency union and the deepening of the European single market. While the euro itself was only introduced between 1999 and 2002, the treaty meant that the EU budget needed sufficient reserves to provide financial transfers for the poorer member states in prevision of their entry in the currency union (Moyer (1993)). These requirements compounded with the growing European budgetary crisis. While originally manageable, the agricultural budget grew exponentially with European agricultural capacity. Over 1980-1986, EU agricultural spending doubled. In early 1991, expenses were expected to increase by an additional 25% in a single year. Figure A1 shows intervention stocks, spending and over-production projections from the EU Commission in 1990. While the European Commission strongly favored subsidy cuts, many member states opposed them. Among them was France, which was one of the primary beneficiary of the policy. The emerging risk of a trade war between the EU and the U.S., hinged upon questions of agricultural subsidy exports and internal price management, added additional pressure for the reform of European agricultural policy. Uruguay round negotiations had started in 1986, but had stalled. The US-led Cairns group insisted on the inclusion of agriculture in the negotiations, and more specifically for the treaty to regulate both direct trade policies, and internal support mechanisms likely to distort trade conditions across countries. This meant that European agricultural policy would need to be modified.

The reform consisted of the following points.[9] First, cereal intervention prices were moved to international price levels, while they were fully removed for oil and protein crops.[10] Second, these changes were partially compensated by the introduction of land use subsidies, in the form of payment per ha. Third, the reform paired these payments with compulsory set-aside requirements, effectively decreasing the rate of subsidization of land. In Figure 1 I illustrate the removal or decrease in intervention prices. I show the evolution of farm gate prices for a series of agricultural commodities. All variables are expressed in 2020 euros. The intervention prices are in red,[11] in grey the distribution of farm gate prices in France, as observed in the FADN. The solid grey line shows the median, and the ribbon the spread between the 25^{th} and 75^{th} percentiles. There is a clear co-variation in French farm gate prices and intervention prices.[12] Finally, the dashed and dot-dashed blue lines track US and

[9] A first reform plan was outlined by the Commission in 1991 – it proposed to cut intervention prices for cereals, oil and protein crops to world price levels, and to introduce a progressive compensatory subsidy scheme in the form of marginally decreasing payments per ha. This progressive structure was scrapped with pressure from the member states, to be replaced by the reform described here.

[10] Council Regulation 1766/1992 which introduces the reform for the cereal market states *"the objective can be achieved by lowering the target price to a level representing an anticipated rate on a stabilized world market"*, making very clear the goal of convergence to international prices.

[11] To the best of my knowledge, this corresponds to the first complete series of EU intervention prices, which I recovered from commodity and year specific regulations.

[12] The intervention prices did not translate in a one-to-one fashion into minimum prices at the farm gate, and that farm gate

Canadian farm-gate prices recovered from Faostat. Around the time of the reform, the U.S., Canada and France were by far the largest three exporters of cereals in the world – and North American prices are hence useful points of comparison.[13]

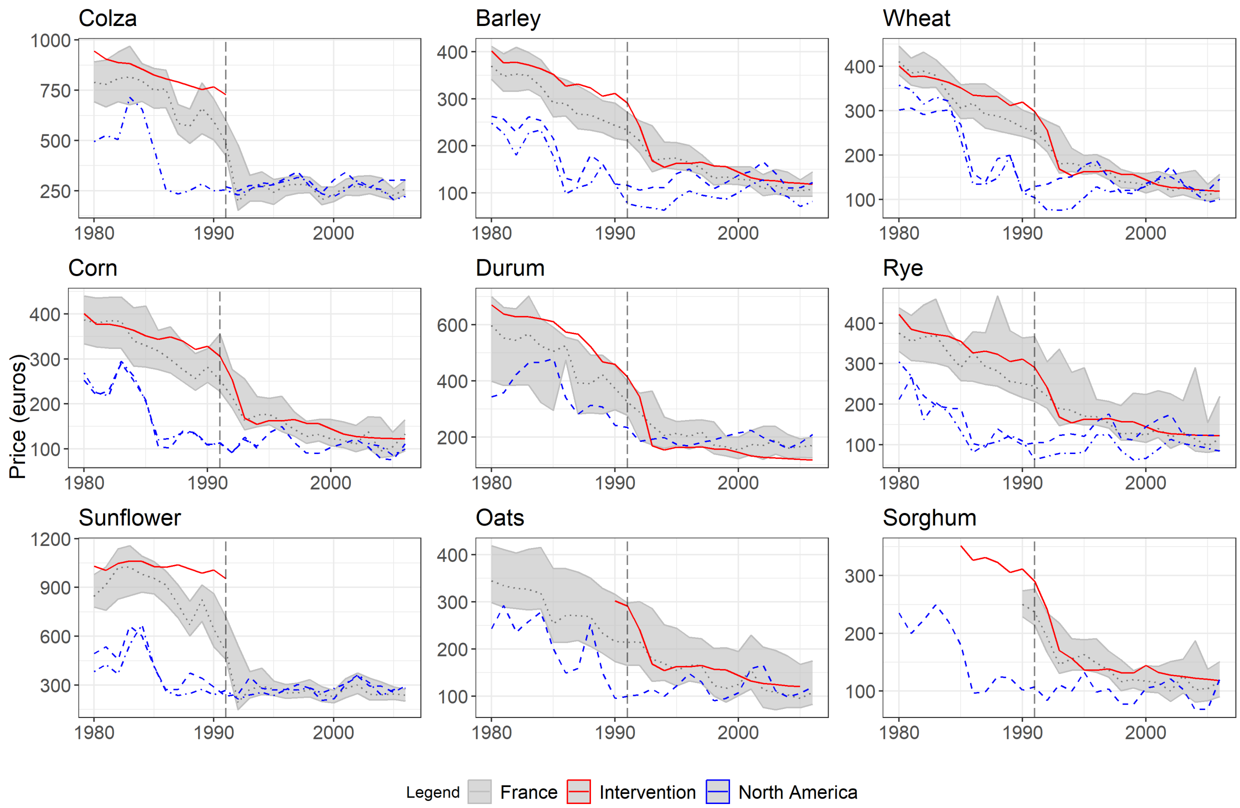


Figure 1: Convergence of French and North American Prices Following the MacSharry Reform

Notes: Evolution of French and North American farm gate prices, as well as EU intervention prices for most cereal and oil crops. French farm gate prices are shown with the grey ribbon, and taken from the FADN. The center dotted line corresponds to the weighted mean, and the edges to the bottom and top 25th and 75th centiles of the distribution of prices. Intervention prices are shown with the solid red line. The data is digitized, taken from the relevant EU directives published over time. US and Canadian farm-gate prices are taken from Faostat, and shown with the dashed and dot-dashed blue lines. All prices are converted into 2020 euros using the relevant exchange rates and correcting for inflation using the Insee's consumer price index.

This figure highlights two important points. The first is that intervention prices were effective at maintaining a large wedge between French and North American prices, and that this wedge disappeared with the reform. Prior to 1992, fluctuations in the gap between French and North American are dominated by movements in North American prices – and especially the 1984-86 commodity crisis (Morrison and Wattleworth (1986)). Around the reform, however, the closure of the price gap is driven by movement on the European and French side. Second, the size of these wedges significantly varied across crops. This implies

prices remain mostly lower than intervention prices. This is likely caused by the presence of minimum purchasing thresholds at which intervention happened, shifting intervention to the secondary market and leaving room for intermediary to only partially pass-through these price levels to farmers.

[13] See https://ourworldindata.org/grapher/cereals-imports-vs-exports?tab=table&time=1995

that farm-level EU-led protections varied according to the farms' crop mixes, and that the losses incurred by the reform also likely varied with their mix. I build an average output price across cereal and oil crops within farms, using land shares for the aggregation, and show the associated one-number summary in Figure A2. While prices have a similar slow decreasing trend pre and post-reform, there is a large jump happening between 1991 and 1993 corresponding to an average 33% decrease in prices, for which there is no recovery.

The second step of the reform was the introduction of land-use subsidies in the form of payments per ha. These acted as a partial compensation for the removal of price intervention. Land subsidies were considered to be sufficiently de-correlated from production because they would not scale with yields, but only land-use thought to be more sticky. I show the evolution of average land subsidies as measured in the FADN in Figure A3, and subsidies per unit of output across crops in Figure A4. The subsidies were computed as the product of an EU-wide price, multiplied by a regionalized historical yield. Because 1991-1995 output-price movement varied across crops, but land subsidies were more homogeneous, the reform can be seen as a relative homogenization of subsidies which were previously diversified across crops.[14]

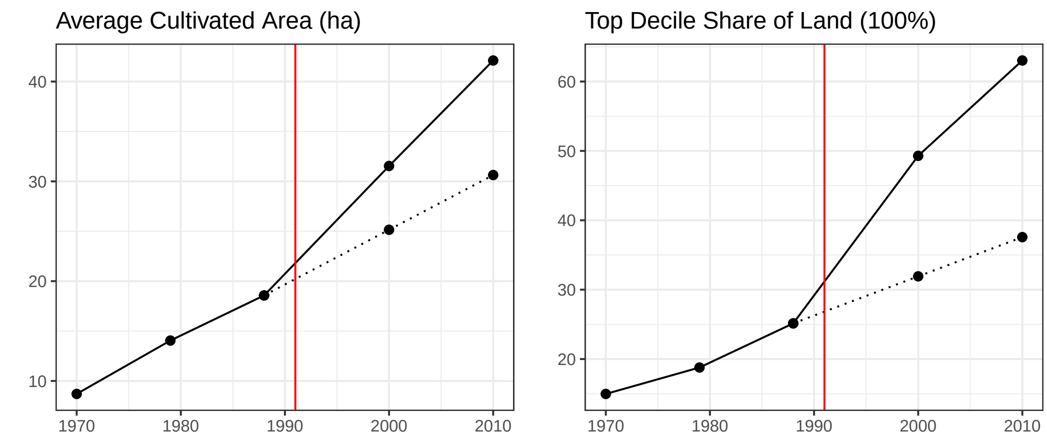
Finally, the reform implemented a set-aside requirement. Farms were required to set-aside a fixed proportion of their land in order to receive land subsidies. Specifically, 10% of their total land use had to be left fallow for them to receive subsidies. This fallow land was also compensated for by the land subsidy.[15] Small farms were not subject to this set-aside policy. I account for set-aside requirements in our measure of the reform, and control for the farms' 1991 land use in order to capture their exposure to the set-aside requirement.

I end this section by showing suggestive evidence for large reallocations in production happening in the decade surrounding the MacSharry reform. I show trends specific to the cereal and oil crop market, which is the one I will focus on—but give a comparison with the trends of the general agricultural market in Table A3. In Figure 2, I use the Census of Agriculture to plot the evolution of the average cultivated area across farms active on this market and the evolution of the land share allocated to the top decile of the farm-size distribution. These two variables show perennial trends across decades, but a kink at the time of the reform. The growth in average farm size accelerates between 1988 and 2000, while the share of land allocated to the top decile of farms jumps from 25% to 49%. These two trends are homogeneous across crops in the market, and I hence plot aggregate statistics across crops. On the contrary, exit dynamics are clearly differentiated across crop categories. There is no sharp change in exit trends around the reform for cereal crops, while the oil market moves between 1988 and 2000 from a net decade-to-decade increase in its number of producers, to a net decrease. Combining this with the differences in wedge decrease observed in Figure 1, it seems the cereal and oil crop markets were differently impacted

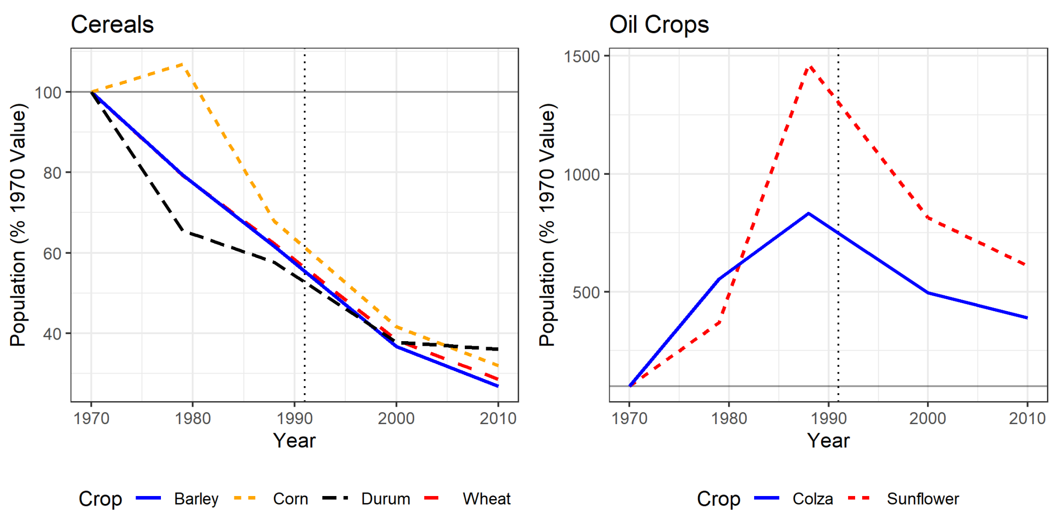
[14]This yield computation differed across crops in its exact implementation, but kept the same general structure.For wheat, it corresponded to a weighted average with two-thirds of the department-specific 1986-1990 yield, and one-third of the national yield for the same period.

[15]Regulation 1765/92 introduced the land payment scheme, and the set-aside requirement was set by Commission Regulation 2293/92.

by the MacSharry reform. Oil crops suffered a much larger decrease in profitability, and this difference will drive identification in our reduced form section. Figure A6 adds to this picture, and shows evidence for farm-level crop specialization accelerating around the time of the reform. Table A4 presents a more detailed accounting of the farms' crop mix composition, which is discussed further in the structural part of the paper.



(a) Farm Growth and Reallocations



(b) Differentiated Farm Exit across Markets

Figure 2: Farm Growth, Reallocations and Exit around the MacSharry Reform

Notes: The data is computed using the full count agricultural census for years 1970, 1979, 1988, 2000, 2010 and 2020. Statistics are only computed for farms active in the oil and cereal crop markets. Average farm size is measured in hectares, the share of land allocated to the top decile of farms (in size of land) is expressed in percent. The evolution of the respective farm populations is expressed wrt. to the 1970 baseline value, in order for the rates to be comparable across crops. The vertical line is for 1991, the last year prior to the reform. The dotted line for the first two figures show the average trend over 1970-1988.

2.2.1 Suggestive Evidence for Chemical-Biased Productivity Growth

I end the descriptive section by showing suggestive evidence for the mechanism through which I introduce cross-farm heterogeneity in pollution intensity in the model. I do so following the intuitions discussed by Doraszelski and Jomandreu (2018) in the context of labor-biased productivity growth in Spanish manufacturing. I use farm-level data to compute fertilizer-to-land and phytosanitary-to-land input ratios as observed in the FADN. They are normalized to their 1980 level. Price ratios are computed at the country level, using national price indices, and an average land price adjusted for land subsidies.

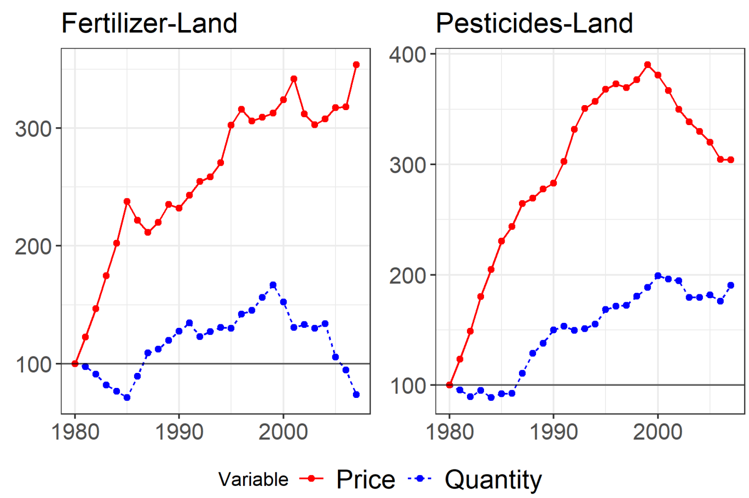


Figure 3: Suggestive Evidence for Chemical-Biased Productivity Changes

Notes: Evolution of relative input quantity ratios versus relative input price ratios. The binscatters are computed using combined data from the FADN (input quantities), the INSEE Ipampa price index series (chemical prices), and the Valeur Venale des Terres land price series (land prices). Land prices are corrected for the EU land subsidies in effect. I compute input-to-input price and quantity ratios. Inputs are measured as deflated bills for pesticides and fertilizers, and in hectares for land. Prices for chemicals are Laspeyres price indices, while land prices correspond to an average land price in France. All indices are normalized to 100 in 1980, in order to focus on their relative evolution over time.

The two relative input price ratios increase significantly over 1980-2000, while volume ratios increase more slowly but still significantly over the same period. This joint increase is likely indicative of input-biased technological change. This change is chemical-enhancing if land and chemicals are substitutes, or land-enhancing if they are complement. Because soil and climatic conditions can reasonably be considered constant over such a short period of time —meaning that land productivity should remain constant—these joint increases are likely suggestive of chemical-biased productivity change. The European Union Nitrate Directive was introduced in France between 1997 and 2000, limiting the quantity of nitrogen which could be used by farms in at-risk area. This could explain part of the decrease observed from 1999 onwards.[16] I provide additional descriptive statistics regarding the role of agricultural production in environmental pollution, and heterogeneity in chemical use

[16]See Chabé-Ferret et al. (2021) for an analysis of the causal effect of the Directive in France.

across farm size in appendix B.2.

2.3 Data

I describe the two main datasets here, and leave the description of the others to appendix B.1.

2.3.1 Farm Accountancy Data Network (FADN)

Most of the analysis relies on a repeated panel of agricultural production developed to be representative of European commercial farming. I will extensively use the French subset of this dataset. The FADN was established in 1968, and for France currently samples around 8,000 farms a year.[17] [18] The FADN is extraordinarily detailed for an agricultural dataset, and presents the clear advantage of being a repeated panel. It is built using accounting data of the farms' activities, and contains total and crop-specific sales, as well as sales volumes allowing me to backtrack farm-level crop prices. It also contains farm-level input usage, crop-specific land allocations and output production, and finally capital stock data. Finally, the data is geocoded at the department level. I will then be able to match farms to their department-level input prices as well as weather data.[19]

2.3.2 French Agricultural Census

The French agricultural census was first fielded extensively at the establishment level in 1955, following the 1951 law on statistical secrecy guaranteeing that none of the collected data would be used for fiscal purposes. The census is fielded approximately every ten years—and for this paper's purposes in 1970, 1979, 1988, 2000 and 2010. The French Census is notable in that it can be accessed at the farm level. As a census, it covers virtually every agricultural establishment in the country. Farms are defined as a unit of agricultural production with independent management, and reaching at least one of the following thresholds: a total agricultural area of .2ha, or a sufficient production level specific to the output type considered. The original list of establishments filtered with these requirements is built out of a combination of land registries and data from the farmers' family allowances mutual savings societies. The Census, however, does not provide actual output or input data beyond crop-specific acreage at the farm level. Furthermore, prior to the extension of stable fiscal numbers to farms in 1998, these cannot be tracked across the wages of the Census.[20] I will use the

[17]Many of the datasets I pair with the FADN are specific to France, preventing me to extend the analysis to the EU-wide FADN.

[18]The FADN is not representative of all agriculture, but only of commercial farming. The definition of a commercial farm changed in 2010, but this only led to the changing of the rules for choosing replacements for the farms leaving the sample, and not to an overhaul of the sample population itself. A commercial farm used to be defined as a farm with a unique manager, which sells more than half of their production, and whose manager's working hours corresponds to at least 75% of their total annual work hours. Finally, farms with less than 5ha of land were removed from the targeted population if they were not specialized. In 2000, there were 380k such farms recorded in the Agricultural Census out of 663,800, but together they accounted for 95% of the country's total agricultural production. From 2010 onward, the working hours requirement was removed, and the 5ha threshold was replaced by a requirement that farms have a production capacity (PBS or Production Brute Standard) of at least € 25,000.

[19]Farms are geocoded at the village level from 2000 onwards, too late to cover the MacSharry reform.

[20]These fiscal numbers, or Siren numbers, are the equivalent of the US EIN number.

Census to track net farm exit, and the distribution of farm size and crop mix composition at the municipality level.

3 Reduced-Form Analysis: Consequences of the MacSharry Reform

This section highlights some key channels through which agricultural subsidies can impact aggregate efficiency and pollution. At the farm level, increased exposure to the reform translated into decreases in output prices, sales and profit. The reform also caused decreases in total chemical use at the farm level. Post reform, subsidization happens through a land subsidy which increases the relative price of chemicals. This subsidy then reduces incentives to use chemicals in production. The reform also impacted the equilibrium allocation of production. Increased median exposure across farms within a municipality also comes with farm exit, likely at the bottom of the farm size distribution. Using a yearly remote-sensing based index of water pollution, I finally show increased exposure to the reform at the county level decreased pollution.

3.1 Empirical Strategy

The identification strategy relies on the variation in ex-ante exposure to the reform across farms. Pre-reform, price intervention implied that the revenue protection farms were getting depended on the wedges between EU intervention prices and international market prices. Figure 1 highlights these were very different across crops. Post-reform, French prices converged to international prices, and revenue protection happened through a land subsidy conditioned on land set-asides. This subsidy still varied across crops, but less systematically, and so did the cost of set-asides. I develop a design in order to isolate the exogenous part of this cross-farm variation.

Approximating European Policy: EU policy shields European agricultural revenues from world market competition. I write these crop-year revenue cushions (per unit of output) introduced by EU policy as:

$$cushion_{ct} = 1\{Intervention_{ct}\}\left(P_{ct}^{inter} - \overline{P}_{ct}^{I}\right) + \left(\frac{\overline{Sub}_{ct}}{\overline{Yield}_{ct}} - setaside_{ct}\right).$$

P_{ct}^{inter} is the intervention price for crop c in year t in France, and $\overline{P}_{ct}^{I}$ is the corresponding farm-gate price in North America.[21] The first element is simply the difference between these two, when intervention pricing is happening, which is denoted by the indicator. The second element corresponds to the post-1992 land subsidy, conditioned on land set-asides. $\frac{\overline{Sub}_{ct}}{\overline{Yield}_{ct}}$ translates the land subsidy into a price by dividing it by an average yield. I add to it the

[21] The US and Canada accounted for most of the export market over the 1980s to 2000s, making these prices a relevant reference point for our analysis. I use USA prices but for colza, for which I use Canandian prices – given that the USA was not a significant producer of colza at that time, but Canada was. All these prices are average farm gates prices from Faostat.

cost of compulsory set-asides. The cost of the set aside is the cost of holding 10% additional more land than the one used for the production of crop c (per unit of output). Its cost is that of the price of land, minus the land subsidy which is also given for land set-aside. I use the median subsidy per ha as observed in the FADN, and the average yield observed in the same dataset.[22]

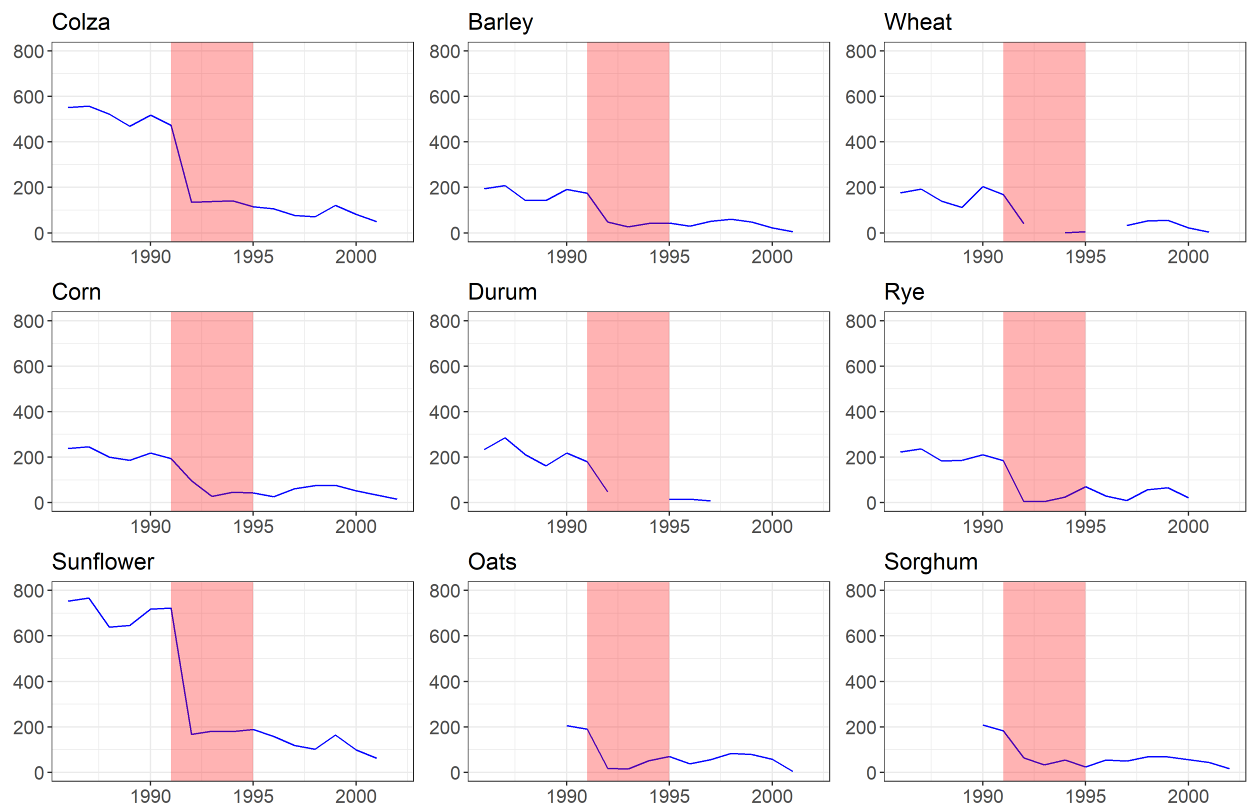


Figure 4: EU Subsidization per Unit of Output – Heterogeneity in Losses around the Reform

Notes: The figure plots crop-specific EU subsidies per unit of output over time, using data from the FADN, and digitized time series of EU intervention prices over time, as well as US and Canadian farm gate prices from Faostat. The shaded area is for 1991-1995, going from the last year pre-reform to the last year of the reform. Data on oats and sorghum is only introduced in 1980 in the FADN.

The evolution of crop-specific cushions is shown in Figure 4. The shaded area is for 1991-1995, the interval over which the cushion changes are evaluated, going from the last year pre-reform to the last year of the reform. All crops have a decrease in their cushion at the moment of the reform. For oil crops, the reform happens in 1992. The decrease in cushion size is of a much smaller magnitude for cereal than oil crops.[23]

Building the Instrument: I can use a weighted sum of the revenue cushions to

[22] I rely on the median and not average subsidies, as the subsidy data recorded in the FADN shows significant mis-measurement. The average subsidy per hectare decreases over time in the FADN, in a pattern that does not match EU policy, while the median value does.

[23] In 1992, land subsidies are still zero in the data, and I use the same land subsidy value as in 1993. Because I do not use the cushion values in 1992, this has no bearing on the following analysis.

obtain an average cushion at the farm level.[24] Setting the initial period to $t_0 = 1991$, I use the following aggregation formula:

$$Cushion_{jt} = \sum_{c \in \mathbb{C}_{j,t_0}} \frac{s^l_{j,c,t_0}}{\sum_{c'} s^l_{c',t_0}} cushion_{ct}.$$

1991 crop decisions were made in the fall of 1990, before any outline of the reform was set forward. Here s^l_{j,c,t_0} gives the land area allocated to crop c by j at t_0. Ex-ante farm-level exposure to the reform can then be captured by:

$$Exp_j = Cushion_{j,t=91} - Cushion_{j,t=95}.$$

In appendix B.3, I show additional results where the instrument is constructed using land shares to aggregate across crops for the land subsidy and set-asides, and output shares to aggregate across intervention prices. Results are similar to the ones presented in this section. Using the Rotemberg weights developed by Goldsmith-Pinkham et al. (2020), I can decompose the source of the variation in Exp_j across crops. I do so in Table A14. These weights show that variation in 1991 land shares across farms resp. sunflower, corn, wheat, durum and colza account for most of the cross-farm variation in exposure. As such, I cannot rely on the approach of Borusyak et al. (2022) to shift-share designs, which allows for endogenous shares but requires a shock-level law of large numbers. The MacSharry reform was not a set of many uncorrelated as-good-as-random shocks, but rather a few shocks with varying intensities that interacted with varying cropping patterns at the farm level. Hence, my design will follow the identification strategy outlined in Goldsmith-Pinkham et al. (2020), recently for example implemented by Ager et al. (2024)—and will also relate as such to the identification strategies of Topalova (2010), Ogeda et al. (2021) and Kovak (2013). It is similar to a difference-in-difference framework where agents are assigned a treatment intensity within a continuous interval of possible values, where the intensity (the average shock value) is based on their type (their 1991 crop mix). I discuss the validity of this identifying framework in turn at the farm and municipality levels.

3.2 Farm-Level Analysis

3.2.1 Design

The central design sets the analysis at the farm level and relies on FADN data. For this analysis, I will be using the following estimating equation:

$$\Delta^{1991} Y_{jt} = Exp_j \theta_t + X_j \delta_t + \eta_{d(j)t} + \varepsilon_{jt}. \tag{1}$$

[24] I include the following crops to compute the farm-level cushion: wheat, barley (winter), corn, rye, oats, sunflower, rapeseed, barley (spring), durum, sorghum.

$\Delta^{1991}Y_{jt}$ is the outcome of interest differenced with its 1991 level. With this formulation, every θ_t will give the effect of exposure in a given year relative to 1991. Exp_j is the farm-constant treatment, which is allowed to have a year-specific effect. The $\{\theta_t\}_{t<1992}$ will provide us with one test for the plausibility of the exogeneity of exposure. Figure A10 and Figure A9 resp. show the distribution of the standardized farm-level exposure, and its geographic department-level distribution within the estimation sample. The distribution of exposure varies quite systematically across space, likely because local climate and soil conditions create spatial auto-correlation in cropping patterns. It is then important to control for local trends, which could both influence cropping patterns and the farm-level outcomes I will be studying. I do so by including department-by-year fixed effects in the form of $\eta_{d(j)t}$. Finally, X_j is a vector of farm-level controls set to pre-reform values to avoid any confounding effect.[25] Standard errors are clustered at the department-by-year level.

In order for the $\{\theta_t\}_t$ to have a causal interpretation, I need both exogeneity in the shocks and shares which enter the computation of Exp_j. Regarding shocks, Figure A8 shows how the main driver in cross-crop heterogeneity in output price decrease was the level of the 1991 price wedge—itself driven by an accumulation of thirty years of price intervention decisions and world-market price variations. Figure 1 confirms there was relatively little variation across intervention price levels in the ten years before the reform. Land subsidies were composed of homogeneous prices within cereal and oil categories, and average yields which varied based on historical yields across departments. I use a country-level approximation of the land subsidies in order to obtain the average subsidy based on average historical yields. Once I control for department-by-year trends, these are unlikely correlated with finer scale auto-correlated shocks. Finally, the reform was European in scope and happened following external pressure from the US in the context of the WTO Uruguay round, which makes the magnitude of these shocks even less likely to correlate with within-region trends in the agricultural market.

Shares also need to be exogenous. The drivers of crop choice in fall 1990 need to be conditionally uncorrelated to the factors impacting the growth rates of the outcome variables studied. Crop choice reacts to supply and demand factors. In a multi-product setting such as Mayer et al. (2014), a higher capital stock or farm productivity should imply a more diverse and even crop mix. This would both directly change exposure and directly impact the farm's dynamic outcomes. It is then important to capture the elements of a farm's state which would relate to 1990 crop choice and outcome growth. Another source of supply-side endogeneity is local variation in input prices, likely correlated over time. If crops have different input needs, these would also lead to biases in the estimation results. On the contrary, the stickiness in crop choice decisions—for example because farms face switching costs related to cropping cycle, as discussed by Livingston et al. (2008) and Scott (2013)—could lead past transitory

[25] X_j contains the following elements: their 1991 capital stock, total labor used, total land use, their fertilizer and pesticide use, the number of crops they grew, their fertilizer-to-land and pesticides-to-land ratios, and the share of their production which corresponds to oil crops (colza and sunflower). I add this element in order to compare farms with relatively similar takes in oil crop production. I expect farms who only grow cereal crops to differ quite significantly from those that also grow oil crops. I also control for the farms' 1983-1984 adoption trends in chemicals measured as their evolution of their chemical use.

shocks to generate exogenous variation in the farms' crop mixes. On the demand side, auto-correlated preference shocks could also lead to biases in estimation. Figure A11 shows balance tests for farms with an above or below median exposure for a series of farm characteristics. While the impact of the reform is clearly different across groups, trends in labor and land use also differ across the groups prior to the reform—making it important to control for them in the analysis.

The first set of controls focuses on the farms' states in 1991: land use, capital stock, total output, the number of crops in their crop mix composition, their fertilizer-to-land and pesticides-to-land ratios. Static elements of the state with constant effects should be taken care of by the differencing of the outcome. I add all observed flexible input decisions that year (labor, fertilizer, pesticides) in order to target unobserved farm productivity. Following Olley and Pakes (1996), the production function literature argues that I can use flexible input decisions to control for a firm's unobserved productivity shocks. If the number of controls is higher than the dimensionality of the state, and under monotonicity and invertability assumptions, flexible input decisions could be inverted to give a control function for unobserved state shocks. While there is a lot of variation in the scale of exposure across oil and cereal crops, one can also expect farms to differ most significantly depending on whether they grow cereals or oil crops, as most cereal crops have relatively close production technologies, soil requirements and growing seasons. Following the recommendations of Goldsmith-Pinkham et al. (2020) to control for more aggregated shares pre-exposure, I add the relative share of output coming from oil crops in our set of controls. Because I also look at outcomes related to chemical use, I also want to make finer comparisons across farms with similar production technologies pre-reform. I thus control for the farms' 1983-1984 adoption trends in chemicals measured as their evolution of their farm-level chemical use. All of these variables are interacted with year-specific coefficients. I finally address auto-correlated shocks in demand with department-by-year fixed effects, under the assumption that agricultural markets in France are sufficiently integrated for finer geographic shock to have little bearing on farm decisions.

With these controls, the remaining identifying threat would take the form of auto-correlated shocks either specific to a farm and not captured by X_j, or local shocks taking place at a smaller scale than the department. Following Goldsmith-Pinkham et al. (2020), I compute the Rotemberg weights after the inclusion of all these controls. They are shown in table A13. The design will make comparisons with farms having relatively large or small shares of corn, wheat, sunflower and barley – which are the major crops on the French row crop market. I give balance tests comparing farms with an above and median crop share, for all crops in appendix B.3.2. I also provide regressions with crop-specific effects in the appendix. These are useful to test the heterogeneity of decreases in profitability across crops. In our case, effect are generally homogeneous across crops. I perform all the analysis in this section on a balanced panel of farms, present from 1985 to 2002, in order for changes in outcomes to be purely driven by within-establishment trends. I use FADN sampling weights in the regressions.

3.2.2 Farm-Level Results

Farm-level event studies results are shown in Figure 5.

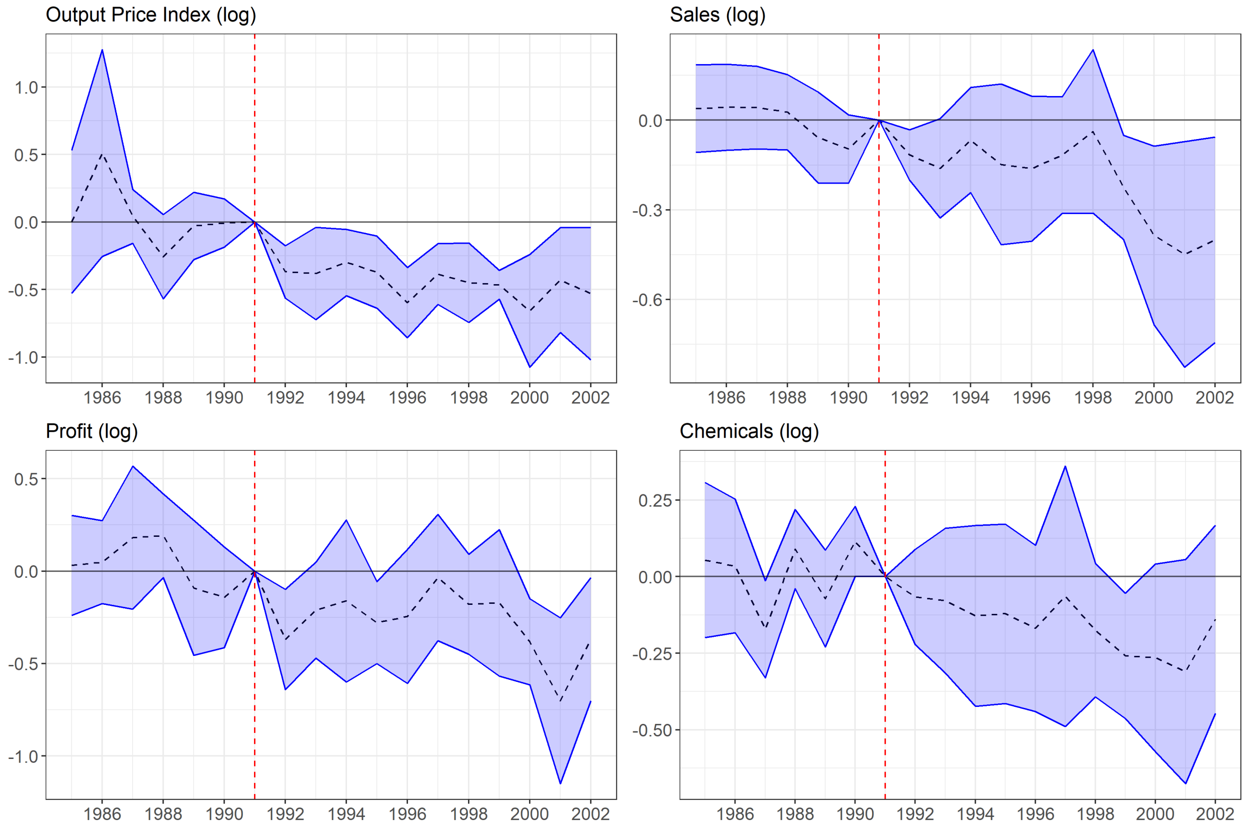


Figure 5: Exposure to the Reform – Farm-Level Event Study

Notes: I plot the result of the event studies for the following outcomes – all measured as differences in logs with respect to the farm's baseline 1991 level: a farm-level output price index (average price across row crops, using relative areas as weights), total subsidies received, profit measured as value added plus subsidies, and the farms' deflated chemical bill. All coefficients correspond to the year-specific coefficient associated to farm exposure, and give the effect of exposure in that given year relative to the effect in 1991. I control for observed characteristics of the farms pre-reform, and allow for a time-varying intercept for each of these controls. I add department-year fixed effects, and cluster the standard errors one-way at the department-year level. The table for these results in Table A5.

The first of these plots serves as a validation of the design. It shows that farms' average output price growth was unresponsive to their 1991 crop-mix prior to 1992, but significantly declines from 1992 onwards, to reach an about 40% decrease in the late 1990s.[26] Exposure is standardized here, hence a one standard deviation increase in exposure leads to this additional 40% decrease in average output price. The reform itself led to a full removal of intervention prices in oil crops, and an about 30% decrease in cereal crops. Figure A2 identifies a 33% decrease in average farm-level crop prices between 1991 and 1993, which

[26] I measure the farms' output price as the weighted average of its crop-specific prices, using relative land allocations as shares. I use the following crops to compute the average (those for which I observe a balance time series): wheat, winter barley, corn, rye, oats, sunflower, colza, spring barley, durum.

serves as a good benchmark for the potential total effect of the reform on prices. This larger decrease in output prices can come from additional reliance on oil crops in 1991, or crop switching away from oil crops. There are decreases in sales and profits as well, resp. of 25% for sales, and 38% in profits. Table A15 gives the difference-in-difference results matching the event-studies regressions. Lastly, there is an impact on farms' chemical use, here their total use of chemicals in production. A one standard deviation increase in exposure leads to a 19% decrease in chemical use. Additional event-study results in Figure A21 show the decrease in chemical use is driven by a decrease in fertilizer use specifically.

In appendix B.3, I discuss the robustness of the results, and show the sales, profit and chemical use regressions with crop-specific instruments to test the homogeneity of the effects across crops. I also show results with a difference-in-difference framework, and results using our alternative instrument for exposure.

3.3 Municipality-Level Analysis

I move to the Census of Agriculture to study the consequences of the reform on farm exit and the local distribution of land. The Census covers the entire population of French farms. Effects on the farm population will hence be best identified in this dataset. Prior results were based on the FADN as the Census does not contain any output, sales, profit or chemical use data. The Census does not allow me to track farms across waves. I then aggregate our measure of exposure at the municipality level, specifically I rely on the median exposure across farms and within the municipality, denoted $\widetilde{Exposure}_k$ for municipality k.[27] This gives a measure of exposure which is less sensitive to tail exposure values. I discuss in the appendix results using different centiles of the distribution, as well as mean exposure. Results are similar across these different measures. I use the following design: for municipality k in year t located in department d, $\Delta^{1988} Y_{kt}$ denoting outcomes differenced with municipality-specific 1988 levels:

$$\Delta^{1988} Y_{kt} = \widetilde{Exposure}_k \theta_t + X_k \delta_t + \eta_{d(k)t} + \varepsilon_{kt}. \tag{2}$$

X_k accounts for the following 1988-valued municipality-level controls: average, minimum, maximum and standard deviation in farm size, total municipality-level agricultural area, total crop count at the municipality level and average crop count across farms, municipality-level crop evenness and average crop evenness across farms, average area cultivated for oil and cereal crops, as well as minimum and maximum cultivated area, and the fraction of farms in the municipality affected by the reform. I cluster our standard errors at the department-by-year level.

[27] Note there are about 36,000 municipalities in France, and that they correspond to the smallest geographic unit at which I can observe data above the farm. I only use data from metropolitan France for this analysis, and results are robust to the inclusion of Corsica.

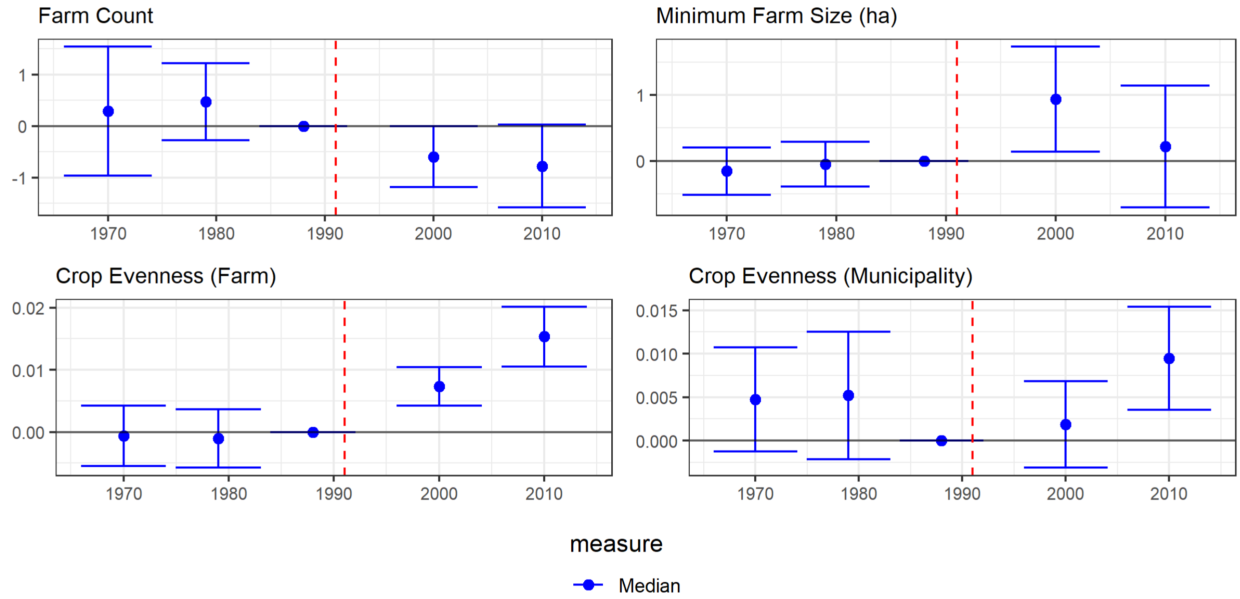


Figure 6: Exposure to the Reform – Municipality-Level Event Study

Notes: This figure gives the results for our municipality-level event study. Outcomes are differenced-out with respect to 1988 values. Farm count corresponds to the number of farms operating in the row crop market within the municipality, and minimum size gives the evolution of the minimum farm size within the municipality. Crop evenness is an index for the evenness of the distribution of land across crops, measured both as the average value across farms, and as the municipality-level evenness. The regression includes a series of controls set to their level in 1988 within the municipality, and interacted with a time-varying coefficients, as well as county-by-year fixed effects. I weight observations by their relative share of total French agricultural area in 1988. Standard errors are clustered at the department-by-year level. The tables associated to these results are Table A6 and Table A7.

I show balance tests for this measure of exposure, and its distribution in appendix B.3.5. Municipalities with a median exposure to the reform below and above the French median value have similar trends in farm count, average farm size and total agricultural area prior to the reform, and slightly different trends in the homogeneity of their cropland. I control for the 1988 farm and municipality-level evenness of cropland, in order to account for this.

Figure 6 plots results. The red line is for 1991, the last year pre-reform. Outcomes are first the evolution of the farm count at the municipality level, then the evolution of the minimum farm size in log terms, finally the evenness of land allocations across crops (the average across farms within the municipality, and the evenness directly computed at the municipality level). All are differenced with respect to the municipality's 1988 value. θ_t gives us here the relative effect of increasing median exposure by one standard deviation. The estimated effects in 2000 of increasing median exposure by one standard deviation are: a decrease in the number of farms by .5 within the municipality (with 36,000 municipalities on average), an increase in minimum farm size of 1ha, and an increase in average crop evenness of .01 (when the average value is .3, for a potential range between 0 and 1). I show results using different methods of aggregation from farm to municipality-level exposure in appendix B.3.5, results are comparable across measures. Figure A6 shows that larger farms are more diverse. The fact that the average evenness goes up, but not the municipality-level one is suggestive that farm exit happens at the bottom of the farm-size distribution.

3.4 County-Level Pollution

I use the same design as the one for municipality level, but aggregate it further to the county level to look at the impact of the reform on algal blooms. I focus on algal bloom as a visible form of agricultural soil and water pollution, in part linked with the excessive use of fertilizers, and specifically nitrogen, in row crop production (Rossi et al. (2023)). Algal blooms can be measured using remote sensing data, by looking at the evolution of the spectral signature of inland water bodies over time, blooms being characterized by a greening of water. I follow the methodology of Taylor and Heal (2023), and use Landsat 5 data to compute a county-year measure of blooms. The methodology is further described in appendix B.1.4. I aggregate our data to the county level, as municipalities tend to be very small, and many will not overlap with significant water bodies on which I can measure an index value. France has about 2,000 counties. I use the same set of controls from the Census this time measured at the county level in 1988. I also add the average county-level precipitation, squared precipitation and growing degree days. For county c, located in department d in year t, the regression follows:

$$\Delta^{1988} Y_{ct} = \widetilde{Exposure}_c \theta_t + X_c \delta_t + \eta_{d(k)t} + \varepsilon_{ct}. \tag{3}$$

The outcome is the evolution in log terms of the algal bloom index over time. Starting in 1995, higher exposure to the reform leads to a decrease in algal blooms within the county of around 7% by 1999 (compared to 1991 levels). I show in Figure A26, the same results but using different aggregations of exposure at the county level. Results are similar in levels and signs.[28]

Overall, the results presented in this section indicate that the reform had important effects on market dynamics: it decreased profits, and generated exit at the bottom of the farm size distribution. It also reduced the use of chemicals, in so doing also reducing agricultural pollution across France. If the farms that stayed were on average more productive, and either more or less polluting, these reallocations likely had additional consequences for aggregate efficiency and pollution. I investigate this equilibrium effect, as well as that of different subsidy designs using the empirical model I present in the next section.

[28]Chabé-Ferret et al. (2021) analyze the effects of the EU-wide regulation of fertilizer use, within France, and from 2001 onwards. I stop the analysis in 2000 in order for the effects not to be confounded with this policy, as potential treatment areas could correlate with counties more exposed to the 1992 reform.

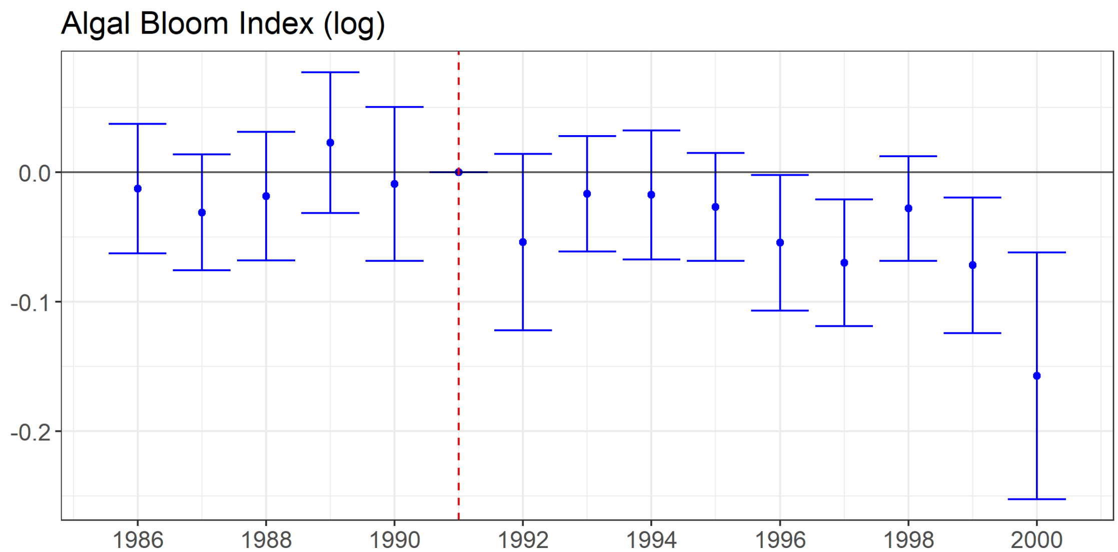


Figure 7: Exposure to the Reform – County-Level Event Study

Notes: This figure gives the results for our county-level event study. The outcome is differenced-out (in log terms) with respect to the 1991 value, and is a Landsat-5 based index of algal bloom intensity on the within-county water bodies. Algal blooms are generally caused by the over-fertilization of agricultural land. The regression includes a series of controls set to their level in 1988 within the county (last year of the Census prior to the reform), and interacted with a time-varying coefficients, as well as department-by-year fixed effects. Standard errors are clustered at the department-by-year level. The table associated to these results are Table A8.

4 Empirical Model

The structural section is organized as follows. I start by outlining the model. First demand for agricultural commodities, which I model as static. Then supply, which is a combination of static input choices for production, and dynamic entry, exit, capital accumulation and crop choice. I then discuss the estimation strategy, the recovered parameters, and their implications.

4.1 Model Outline

The aim of the model is to deliver an equilibrium use of chemicals and economic surplus. Both are then shifted by policy intervention. To introduce heterogeneity in pollution intensity across farms, I will model variation in their efficiency at using chemicals. I distinguish this inherent variation in ability from other factors which generate variation across farms in pollution intensity—specifically crop choices, capital stocks, total factor productivity and scale.

4.1.1 Static Demand

I consider demand for agricultural commodities. The set-up of our production function estimation requires aggregating crops into categories. I then specify a demand function for each of these aggregated crop category c. For a crop c with price P_{ct} in year t, demand is isoelastic and e_{ct} is an idiosyncratic mean zero demand shocks:

$$Q_{ct}^{D} = exp(\alpha_c + e_{ct}) P_{ct}^{\beta_c}.$$

Appendix B.9 discusses an alternative modeling path using a linear quadratic demand for horizontally differentiated agricultural commodities themselves located within crop markets (eg. two producers will produce two varieties of wheat sold in a wheat market, which is a subset of the general row crop market), following the work of Mayer et al. (2014).

4.1.2 Supply

Supply contains two key elements: the production function, and parameters related to capital investment, entry and exit decisions.

Agents & Timing: Farms act as single agents in a competitive row crop market with $\mathbb{C}$ differentiated crops. In a given period, a farm's state is defined by the set: $\Upsilon_{jt} = \{\omega_{jt}^h, \omega_{jt}^{ch}, S_{jt}^o, K_{jt}, \sigma_j, \mathbb{C}_{jt}\}$. This set first contains productivity shocks: ω_{jt}^h for a farm-level TFP shock, and ω_{jt}^{ch} for a farm-level chemical-biased shock. Both ω_{jt}^h and ω_{jt}^{ch} are exogenous to the farm's decisions, I later impose some structure on their process for identification purposes. The set also characterizes the farm's input stocks, the land it owns S_{jt}^o, as well as its capital K_{jt}. Finally, $\mathbb{C}_{jt}$ is the farm's crop mix for period t, and σ_j is its competence ladder endowed at entry. The ladder is used to differentiate TFP shocks across crops within the farm. The TFP has a geometric rate of decrease along the ladder, using the functional form of Mayer et al. (2014). σ_j is a permutation of $\mathbb{C}$ which ranks a farm's productivity at growing crops. Given a crop's rank, its TFP is set using the formula: $\omega_{jct}^h = \omega_{jt}^h \lambda^{\sum_{n=0}^{dim(\mathbb{C})-1} n 1\{\sigma_j(n)=c\}}$. For example, if wheat is the first crop (rank $n = 0$) on farm j's ladder, the farm will be most efficient at producing wheat, and its TFP for wheat will be $\omega_{jt}^h \lambda^0$. The TFP for its second crop will be $\omega_{jt}^h \lambda$, which will be strictly lower as I assume $\lambda \in]0, 1[$.

Farms also face an aggregate market state Ω_{jt}, which is defined as follows: $\Omega_{jt} = \{\mu_t, N_t^e, f_t^e, \{P_{jct}\}_{\mathbb{C}}, \{P_{jt}^x\}_x, P_{jt}^K, Policy_t\}$, μ_t is the measure describing the distribution of incumbents over the market space, N_t^e the mass of entrants, f_t^e the fixed cost of entry, and then vectors of output and input prices, as well as the state of the policy $Policy_t$.

I assume decisions are made in the following order: incumbents first observe the evolution of their productivity and of the state Ω_{jt}, they then make their production decisions for all the crops in their mix. Finally, they decide whether or not to remain on the market, a decision I denote with ξ_{jt}^x. If they remain on the market, they pay a fixed cost of incumbency f_k, and update their capital and land stocks, as well as decide on their crop mix for the next period. The timeline is represented as follows:

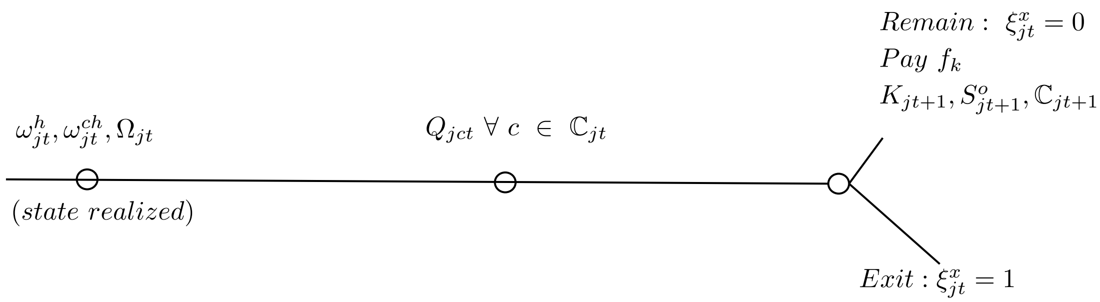


Figure 8: Timing for Incumbent Farms

Entrants make decisions in the following order: they make an entry decision after observing Ω_{jt}, but prior to knowing their initial productivity and their competence ladder. This entry decision is denoted ξ^e_{jt}, and associated with the fixed cost of entry f^e_t. If they enter, they observe their initial state, and end their entry period by deciding whether or not to stay on the market—if they do stay they make their initial investment decisions and set their crop mix for the next period.

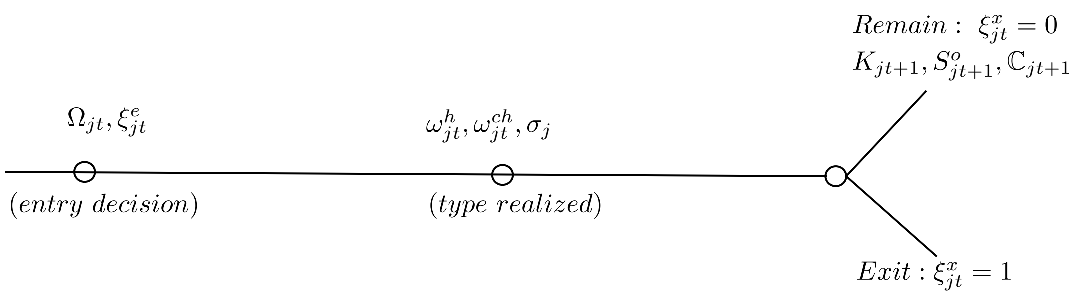


Figure 9: Timing for Entrants

I provide the Bellman equations linked with the incumbent and entrant problems, as well as the law of motion of the measure μ_t and the equilibrium definition for the model in appendix B.4.

Production: I consider that farms maximize profits independently across crops. Production decisions are made independently across production lines, with flexible inputs fully allocated across these lines.[29] This means both that all flexible inputs are fully private when it comes to their use, that input prices do not depend on the quantity purchased, and that

[29]The assumption that flexible inputs are fully attributable to production lines is frequent in the literature – implied in single product settings, and formally made in multi-product ones by De Loecker et al. (2016), Orr (2022) or Valmari (2023). I discuss and estimate an alternative joint production framework where inputs are shared publicly across products – with a penalization – in Annex appendix B.6. I later discuss how this assumption impacts potential economies of scope in the spirit of Panzar and Willig (1981) and Baumol (1977). The assumption that crop-specific profits can be maximized independently from each other captures our assumption on flexible input allocations, on exogenous output prices – unaffected for example by cannibalization concerns of Eckel and Neary (2010) and Nocke and Schutz (2018) – and of limited economies of scope. This also for example assumes away the overhead costs modelled by Foster and Rosenzweig (2022) for agricultural labor.

productivity shocks are exogenous to the size and composition of the crop mix. On the contrary, capital is taken to be fully shared across crops without penalization.[30] The production function for each crop c has the following outward Cobb-Douglas shape:

$$Q_{jct} \;=\; e^{\omega^h_{jct}}\, K_{jt}^{\alpha_k^c}\, L_{jct}^{\alpha_l^c}\, \tilde{S}_{jct}^{\alpha_s^c}\, e^{\varepsilon_{jct}}. \tag{4}$$

L_{jct} is the amount of labor used in production, and $\tilde{S}_{jct}$ a composite input I use to relate land, fertilizers and pesticides – and that I describe further below. ε_{jct} is a mean-zero random shock uncorrelated over time, across firms. ε_{jct} is not known to firms prior to making production decisions. The composite input takes the form:

$$\tilde{S}_{jct} = \left\{\delta_s^c S_{jct}^{\rho} + (1-\delta_s^c)\left(e^{\omega_{jt}^{ch}}\left\{\delta_p^c Pest_{jct}^{\rho_2} + (1-\delta_p^c) Fert_{jct}^{\rho_2}\right\}^{\frac{1}{\rho_2}}\right)^{\rho}\right\}^{\frac{1}{\rho}}.$$

$\tilde{S}_{jct}$ is an embedding of two CES nests. As discussed by Doraszelski and Jomandreu (2018), CES are a flexible way to include input-biased productivity shocks within a production function, where the impact of the shocks on equilibrium input ratios and input shares are set by the elasticities of substitution of the CES. Here, there is a nest which relates land and our two chemicals, for which substitution is governed by ρ. The chemical-biased productivity shocks shifts the effective amount of chemicals used in production. The second nest relates pesticides and fertilizers, following ρ_2. I further note that all input shares are crop-specific, and assume that $\sum_x \alpha_x^c = 1$. An increase in ω_{jt}^{ch} will raise a farm's chemical-to-land and chemical-to-output ratios at the optimal level of production if $\rho \in]0,1[$. I discuss this in further details in appendix B.5.

Dynamics: Entry, Exit and Capital Investment: The final model makes some additional simplifying assumptions—which are not made prior in order to keep the production function as general as possible. I remove the possibility for farms to own land S^o_{jt}, and consider all land used is rented. I assume that all farms share the same competence ladder $\sigma_j = \sigma, \forall j$. I finally aggregate fertilizers and pesticides in a unique input which I call $Chemicals_{jct}$. As a consequence, σ drops out from a farm's state. Additionally, input and output prices will now be unique and homogeneous across all farms within a period. This makes the exogenous market state Ω_t identical across farms.

I allow for convex costs of capital adjustment, and for the cost of entry to adjust with the mass of farms on the market, which is denoted by $\mathbb{M}_t$. Costs of capital adjustments follow:

$$C(K_{jt}, K_{jt+1}) = P_t^K i_{jt} + C_k^Q 1\{K_{jt} > 0\} K_{jt}\left(\frac{i_{jt}}{K_{jt}}\right)^2.$$

[30] Its law of motion is $K_{jt+1} = (1-\delta_k)K_{jt} + i_{jt+1}$, where δ_k is the rate of capital depreciation, and i_{jt} is investment. Because land rental markets are perfect, I consider that S^o_{jt} is fully rented out, and that crop-specific land uses are flexible and unconstrained. Land and labor are in reality ambiguous inputs, with an intermediary dynamic status. Labor is a composite of farm owner and permanent labor, and of seasonal work sometimes paid hourly. Land is also a composite of owned and rented land, with a likely wedge in price between the two. When addressing the choice of instruments, I will be careful in only using lagged labor and land decisions to avoid any issues related to these dynamics from impacting estimation.

The cost of entry f_t^e follows:

$$f_t^e = \frac{1}{\alpha_e A_e} \left(\frac{\mathbb{M}_t}{A_e} \right)^{\frac{1-\alpha}{\alpha}}.$$

The idea behind this modeling choice is that as the number of farms increase in France, the initial investment necessary to open another one goes up. This captures the scarcity of the current stock of cleared agricultural land and built farm buildings.[31] Practically, this choice of modelization allows for the model to close when prices are set through intervention. Indeed, under price intervention, the government provides an infinite purchasing guarantee at a fixed price, and there is no market clearing equation. The adjustment of the cost of the entry good provides a channel to pin down market size (mass of producers and total production volume) under intervention.

Role of Dynamics: Farms make different dynamic decisions—capital adjustment, crop choice, and entry and exit decisions. They also face an exogenous varying state in the form of their TFP and chemical efficiency shocks, over which they form expectations when it comes to making dynamic decisions. Crop choice and capital stock matter for production. A farm's chemical use and chemical-to-land ratio respond to its crop mix composition, while a farm's chemical-to-output ratio will decrease with its capital stock. Hence, these are important elements to account for, if I want to evaluate the impact of subsidy design on production externalities. Capital and its relation to productivity shocks also matters to look at selection. Asker et al. (2014) discuss the relation between volatility and investment patterns, and how too little and too much volatility de-correlate investment behavior from shocks. Finally, all of these dynamic considerations are better accounted for in a model with many periods, as under convex costs of capital adjustment, capital accumulation takes time.[32] As Collard-Wexler (2013) shows, recovering both costs of capital adjustment and the volatility of productivity are key elements to correctly account for selection on productivity within a market. Large capital stocks protect farms from low productivity shocks, and lead to differential exit behavior – for a given productivity shock – across the capital stock distribution. Dynamics in the form of varying productivity types, capital adjustment, crop choice, entry and exit, are then all important to recover how different subsidy design will interact with selection on productivity.

[31]Anecdotally, Table A2 shows that at the aggregate level and over 1970-2000, there is a small albeit monotone decrease in the average ruggedness of farm land in France which accompanies the decline in the number of farms – this decrease in ruggedness likely tracks with decreasing costs of entry.

[32]See Figure A14 showing the relation between farm size (in land use measured in ha) and their age in years of tenure. I show this relation in the 2010 census. The relationship in prior years is heavily impacted by a trend in entry size growth, which biases the overall size-tenure relation. I measure size in land use here, as this is the only available relevant variable observed in the Census.

4.2 Estimation & Results

The estimation of the model is done for a market with two crop categories: $\mathbb{C} = \{wheat, other\}$.[33] This coarse aggregation is required for the production function, as estimation is done using only data from single crop farms. It however matches well the French row crop market structure where wheat is the most common row crop grown, and is also the crop the most frequently grown on its own. In the market, two-thirds of the farms grow at least two crops, and about 90% of multi-crop farms grow wheat. See Table A4 for proportions taken from the Census of Agriculture.

I present all the parameters in a joint table at the end of the section.

4.2.1 Demand

I estimate one demand equation per crop. They are recovered using data from the FADN which I aggregate at the national level using sampling weights, I then rely on weather data to construct supply side shifters and instrument for prices—weather shocks will impact the production of row crops, but should not affect demand. Importantly, I build measures of yearly local deviations from climate averages, that I interact with weather realizations and use as instruments. These deviations are more likely to be orthogonal to demand—as they are more likely to be orthogonal to EU-wide weather events. I also run a second design using only post-reform observations, and additionally using seasonal weather forecasts as instruments on top of realizations. I include a linear time trend as a control in all designs, and use the Newey-West correction for standard errors in order to account for auto-correlation in the error terms. The regressions with their different designs are shown in Table A9.

4.2.2 Production Function

The estimation of the production function is done in two steps, and estimating equations are derived from the first order conditions for the farm's present-discounted value maximization. Because static input choice has no dynamic consequences, these conditions correspond to the first order conditions for the maximization of static profit Π_{jt}. Independence of crop profits, and the fact that all flexible inputs are fully allocated to product lines finally means that I will be looking at the first order conditions for crop-specific static profit maximization. Throughout this section, I will assume interior solutions to input choice decisions for profit maximization, and will hence remove FADN observations with zero flexible input use.

I will outline specific issues related to the FADN data structure and how they call for additional care when deriving estimating equations. I then proceed to the derivations giving me the two estimating equations, as well as the estimation procedure itself. The aim of decomposing the estimation in two steps is that I can first recover an equation where the non-Hicksian shock is the only endogenous unobservable—once recovered—I can then

[33] List of crops in *other* category: sunflower, colza, oats, barley, rye, triticale, corn, sorghum and durum. Splitting the crop groups into finer categories, or along other lines dramatically reduces the amount of available observations. Adding more categories also increases the time taken for the model to converge in a significant way.

recover TFP shocks and remaining parameters in a second step with yet again a unique endogenous unobservable.[34]

Input Observability: My analysis runs into the usual difficulty of the level of input measurement in multiproduct production panels. Besides land, all the inputs are measured at the farm level.[35] My approach to this problem follows the method proposed by De Loecker et al. (2016), and I estimate the parameters using only data from single product farms. This bypasses the input allocation question, and I later use simple algebra and the problem's first order conditions—both described in appendix B.7—to recover the unobserved shocks for the multi-product farms. Contrary to De Loecker et al. (2016), and because I observe crop-specific land allocations, I can back-out these shocks while allowing for TFP shocks to be heterogeneous within a farm across its crops. Using single product farms for estimation is likely to induce a selection issue—farms with more products are likely to be more productive.[36] Both Levinsohn and Petrin (2003) and De Loecker et al. (2016) argue that using an unbalanced panel is helpful in reducing this selection issue. De Loecker et al. (2016) argue one should keep observations related to both always single-product firms, and sometimes single product firms when they happen to be single product. I follow their dataset construction guidelines for all my estimation. I also extend their modification of the Olley and Pakes (1996) selection correction to this specific selection issue. The procedure is discussed in appendix B.8. As Olley and Pakes (1996), De Loecker et al. (2016) and Backus (2020) argue, the selection-induced omitted variable bias will affect the coefficients of variables correlated with the threshold rule used for resp. exit or the introduction of a second product. Such variables are only present in the second step in the form of the farms' capital stocks and chemical productivity, justifying using this selection correction in our second step only.

Additional Considerations: A frequent issue posed by firm surveys is that they only contain output data measured in sales rather than volume. The FADN is useful in that it contains crop-level output in both volume and value, allowing us to recover precise output data and firm-specific prices. Given the context, I also assume that cross-farm differences

[34] This follows the methodology of Doraszelski and Jomandreu (2018).

[35] This implies that the analysis has either to be run at the farm-level—as Doraszelski and Jomandreu (2018) do—adopt a more general transformation function approach where input allocations do not have to be recovered—as Dhyne et al. (2022) and as I do in appendix B.6—or use a method that gives us these input allocations. Traditionally, the production function literature has relied on revenue shares (e.g. Foster et al. (2008) and De Loecker and Collard-Wexler (2015)), or on the number of products observed at the establishment level (De Loecker (2011)) to recover these allocations. More recent work has used information from demand to jointly recover input allocations and production function parameters (Orr (2022) or Valmari (2023)). My case differs from these advances in that I want to recover non-Hicksian productivity shocks, while accounting for the farms' varying crop mixes and the related heterogeneity in the crops' reliance on chemicals for production.

[36] The production function literature has for a long time discussed how market exit introduces a selection bias in observed firm data. In their seminal paper, Olley and Pakes (1996) develop a selection correction which accounts for the fact that larger firms are relatively more likely to withstand negative productivity shocks, and hence that they are more likely to be observed in the data with a low productivity shock. The literature has since advanced that this correction does not significantly shift estimated parameters, as long as researchers do not attempt to balance their panel. As argued by Levinsohn and Petrin (2003) and Orr (2022). However, De Loecker et al. (2016) note that by selecting single product firms for estimation, a second type of selection issue might be introduced, which is that if TFP shocks are shared or correlated across product lines, I might focus on firms that repeatedly draw low productivity TFPs. One model making this selection issue transparent is the one of Mayer et al. (2014) where a firm's scope is a decreasing step function in one's constant marginal cost of production.

in output quality will not be too large. Organic farming was less developed in the years surrounding the reform, and the general market is itself quite narrowly defined.

I observe both labor and land in volume, but only observe input bills for pesticides and fertilizers.[37] I deflate these bills into volumes using input-specific Laspeyres price indices at the national level.[38] As I observe aggregate input bills for both the chemical categories, I cannot address issues of varying consumption basket composition—across farms or across time. Homogeneous input quality updates at the country level might be reflected in prices. Any upgrade in quality that comes with an increase in price not reflected in the national price index—for example a farm starting to buy better and more expensive products—will be manifested as a higher volume of chemicals used rather than an increase in quality. As shown above, a higher chemical-to-land ratio unexplained by varying prices or a change in crop mix will be classified as an increase in chemical productivity. In that sense, the non-Hicksian productivity will not discriminate between for example input upgrading, learning or changes in farm management that lead to increased chemical use.

First Stage: The first stage relies on the first order condition for the use of land and chemicals, and follows the methodology of Doraszelski and Jomandreu (2018). Profit maximization yields:

$$\begin{aligned}\rho\omega_{jt}^{ch} =& log\Big(\frac{P_{jt}^{p}}{P_{jt}^{s}}\Big) + log\Big(\frac{\delta_s^c}{\delta_p^c[1-\delta_s^c]}\Big) + (1-\rho_2)pest_{jct} - (1-\rho)s_{jct} \\ &+ (1-\frac{\rho}{\rho_2})log\Big[\delta_p^c Pest_{jct}^{\rho_2} + (1-\delta_p^c)Fert_{jct}^{\rho_2}\Big].\end{aligned} \tag{5}$$

This expression gives me a relation for the unobserved non-Hicksian shock depending only on parameters and observables. Assuming that the non-Hicksian shock follows a first-order Markov process, I can then create moment conditions on which can rely for estimation. Specifically: $\omega_{jt+1}^{ch} = \mathbb{E}[\omega_{jt+1}^{ch}|\omega_{jt}^{ch}] + \zeta_{jt+1}^{ch} = g_{ch}(\omega_{jt}^{ch}) + \zeta_{jt+1}^{ch}$. This expression is combined with Equation 5. I can then approximate g_{ch} with a high-order polynomial. I build moments—with $A_{jt}^{z,ch}$ standing for our instrument matrix—and use them in a GMM estimation to recover the following parameters $\{\rho, \rho_2, \{\delta_s^c, \delta_p^c\}_{\mathbb{C}}\}$:

$$\mathbb{E}\Big[\zeta_{jt+1}^{ch} A_{jt}^{z,ch}\Big] = 0. \tag{6}$$

Second Stage: The estimating equation for the second stage relies on the first order condition for labor. It gives me a parametric function for the TFP shock that depends on

[37] I observe respectively labor in hours, and land in hectares.

[38] These Laspeyres indices are themselves built using an agricultural price survey to which I have access. However, an initial inquiry ran in 1995 by Insee established that the geographic dispersion in agricultural input prices in France was too small to warrant a stratification of the survey which would allow to build representative sub-national price indices. As such, the survey has been run since in a way that only achieves representativeness at the country-level, and I similarly do not differentiate fertilizer and pesticide prices across space. Labor and land prices will be specified at the department-by-year level using other data sources.

observed variables and parameters that I will recover. It is written below. This expression is useful to build an estimate of the predicted part of the TFP process.

$$\omega^h_{jct} = p^l_{jt} - p_{jct} - \alpha^c_k k_{jt} - log(\alpha^c_l) + (1-\alpha^c_l) l_{jct} - \alpha^c_s \tilde{s}_{jct}. \tag{7}$$

I would ideally like to follow an estimation procedure similar to the one used for the first step. That is, specify a (uncorrected) Markov process of the form: $\omega^h_{jct+1} = \mathbb{E}[\omega^h_{jct+1}|\omega^h_{jct}] + \zeta^h_{jct+1}$, and then use the following expression (a combination of the log of the production function and this Markov process) to form moments, where Equation 7 serves as a baseline to construct the parametric one-period ahead prediction of the TFP:

$$\zeta^h_{jct+1} + \varepsilon_{jct+1} = q_{jct+1} - \alpha^c_k k_{jt+1} - \alpha^c_l l_{jct+1} - \alpha^c_s \tilde{s}_{jct+1} - \mathbb{E}[\omega^h_{jct+1}|\omega^h_{jct}].$$

The issue, however, is that I introduced a selection bias in focusing on single product farms—and that this bias is likely to interact with the coefficients of capital α^c_k and of aggregated land α^c_s—as both the farm's capital stock and non-Hicksian productivity will enter the threshold used by farms to decide whether to produce a second crop. If I use Ξ_{jct} to denote whether farm j only produces c in t, then the actual process observed for the TFP in the unique-crop farm sample is: $\omega^h_{jct+1} = \mathbb{E}[\omega^h_{jct+1}|\omega^h_{jct}, \Xi_{jct+1} = 1] + \zeta^h_{jct+1}$. The bias introduced by using the uncorrected process rather than this one will be correlated to both k_{jt+1} and $\tilde{s}_{jct+1}$, and would bias the estimates of $\{\alpha^c_k, \alpha^c_l\}$. The first stage had no variable correlated with this bias—which, conditional on using an unbalanced sample—allowed me to use an uncorrected specification of the non-Hicksian process. Here, I need a correction, and rely on a procedure which follows the lines of Olley and Pakes (1996) and De Loecker et al. (2016). I describe it in appendix B.8. With the additional control for selection, I denote $A^{z,h}_{jct}$ the matrix of instruments—I use the following moments to recover $\{\alpha^c_k, \alpha^c_l, \alpha^c_s\}$:

$$\mathbb{E}\left[\left(\zeta^h_{jct+1} + \varepsilon_{jct+1}\right) A^{z,h}_{jct}\right] = 0. \tag{8}$$

Estimation: For both steps, I recover the parameters using the generalized method of moments (GMM). I absorb the parameters related to the approximated functions $g_{ch}(.)$ and $g_h(.)$, following the approach of Wooldridge (2010) and Doraszelski and Jomandreu (2018). The moments rely on the timing assumption related to input choice combined with the assumed structure of the productivity processes, as well as on the exogeneity of the policy environment specific to the CAP to these farm-level shocks. Instruments for the first stage are: a constant, the lag pesticide-to-fertilizer input ratio, lag land use, the current capital stock, hourly wage and land prices, and a measure of current-farm exposure to EU agricultural policy using lagged crop land shares as weights. For the second stage, I use: current and lag log capital stocks, lag log labor use, log wage and log land prices, the log of the chemicals price ratio (pesticide to fertilizer relative prices), a lag log output price index at the farm level, the current land subsidy, the lag log land use, and the log of our measure of farm exposure to EU agricultural policy.

I use Amemiya (1974) GMM weights and a Nelder-Mead minimization algorithm—searching over the space of starting parameters to ensure homogeneous convergence to these values.

Sensitivity of Results to Moments: I implement the procedure of Andrews et al. (2017) in appendix B.11 to shed some light on the respective role of the moments in driving the estimation of the production function parameters. I recover the sensitivity matrix associated with each of the estimation steps, which allows to understand the extent to which any moment-specific failure of exogeneity would impact the estimates. For the first step, the analysis shows the estimates are particularly sensitive to the assumption regarding the exogeneity of the non-Hicksian innovation shock, as well as assumptions regarding the exogeneity of contemporaneous input prices faced by farms. Interestingly, their procedure shows the crucial role played by the relative fertilizer-to-pesticides price ratio in the identification of the fertilizer-to-pesticides parameter of substitution. The same analysis for the second estimation step also shows the role of input prices in driving estimation. The role of relative input prices in the estimation of substitution parameters is thoroughly discussed by Doraszelski and Jomandreu (2018) as central to their identification strategy. More generally, the production function literature has largely discussed the potential for input prices as powerful instruments for the identification of production parameters, under the condition that they can be assumed to be exogenous to establishment unobserved characteristics (Griliches and Mairesse (1998)).

4.2.3 Dynamic Parameters

There are four remaining parameters to recover: the cost of incumbency f_k, the convex cost of capital adjustment C_k^Q, and the parameters for the production of the entry good: $\{A_e, \alpha_e\}$. Together, these parameters will give the survival threshold in the farms' type space, the farms' behavior when it comes to capital accumulation, and the speed of adjustment of entry costs to market size. I rely on indirect inference for estimation. This means that I perform a search over the parameter space, targeting moments which vary one-to-one with the parameters and hence provide useful information for their estimation.

It is important to note that there are two categories of parameters, and hence that I need two categories of moments. Given the parameters already estimated for the model, $\{C_k^Q, f_k, f_e\}$ jointly determine a stationary equilibrium under a specific policy design. Different pairs of $\{A_e, \alpha_e\}$ can then relate f_e to equilibrium market size. As such, moments that relate to firm behavior and distribution within a stationary equilibrium can help recover the effective cost of entry for that equilibrium, but not $\{A_e, \alpha_e\}$. To recover them, I need moments that relate to transitions across equilibria, and can exploit observed changes in the French market as EU policy switched from price intervention to the land subsidy.

I use the following moments related to within-equilibrium behavior: two moments related specifically to investment—the auto-correlation in investment rates within establishment, and the correlation between investment rates and profit shocks, and two moments more specifically targeting our fixed costs—the coefficient of variation in profits across establishments, and the coefficient of variation in tenure. I use the transition in market size observed in the Census of Agriculture between 1988 and 2000 to get to the parameters that govern the speed of adjustment of the cost of entry with market size.

The indirect inference procedure could directly recover these four parameters, targeting the necessary moments for it. I however use a two-step procedures to reduces the dimension of the search space and, given the model is slow to converge, significantly increases the precision that can be reached with the estimation. I rely on the intuition of Klenow and Li (2024) to split the estimation in two stages. The authors remark that under the assumption that the zero-profit condition holds, one can backtrack the fixed cost of entry from the expected value of entry. Under the price intervention regime, the expected value of entry only depends on $\{C_k^Q, f_k\}$. I can thus focus the indirect inference procedure on recovering $\{C_k^Q, f_k, f_e^{land\ subsidy}\}$, with $f_e^{land\ subsidy}$ the effective cost of entry under the land subsidy regime. Knowing $f_e^{land\ subsidy}$ and $f_e^{price\ intervention}$, the market size under the land subsidy regime and the target I need to match for the change in market size going from price intervention to the land subsidy, I can then recover $\{A_e, \alpha_e\}$ by matching simulated and observed change in market size—where the market size under price intervention is an analytical expression which depends on these two parameters.

I use the following routine for the first step: for each candidate vector of parameters, I solve the model to its stationary equilibrium under the land subsidy policy regime, recover the policy functions and simulate a dataset akin to the actual farm-level data. I then recover key moments in this simulated data and compare them to the observed data (for the period when the land subsidy was effective). The estimates of the true parameters are be the one minimizing the distance between the two. The use of indirect inference to recover dynamic parameters is common in the literature. Cooper and Haltiwanger (2006) recover costs of capital adjustment in US manufacturing using a similar estimation routine, and so do Asker et al. (2014) for country-specific adjustment costs, and Johnston (2020) for the US paper industry. Collard-Wexler (2013) and Gowrisankaran et al. (2024) similarly use it in dynamic oligopolistic frameworks.

I discuss in turn the algorithmic routine used, the identification argument, and the relevance of these chosen moments.

Algorithm: For a potential vector of parameters Θ, I solve the model to its stationary equilibrium, and generate associated policy functions regarding capital investment, crop choice and market incumbency. I then create a sample of 1,000 farms, drawing their initial state at random from the equilibrium distribution of incumbents. I simulate their behavior for $T = 50$ periods, and compensate for exit by continuously drawing in new entrants to keep the sample size constant.[39] In order to address the impact of those random draws on the composition of the data, I bootstrap this exercise $B = 100$ times, and then compute average simulated moments over these farms, periods and bootstrap runs. I denote these moments estimate $\Phi^s(\Theta)$, and the observed moments in the FADN and the Census Φ^d. I then use a grid search to find the vector of parameters that solves:

$$\underset{\Theta}{min}\left[\Phi^d - \Phi^s(\Theta)\right]' W \left[\Phi^d - \Phi^s(\Theta)\right]. \tag{9}$$

W is the weighting matrix. Because the moments are similarly scaled, I use the identity matrix—following Asker et al. (2014).

Table 1: Moments for the Indirect Inference

Moment	Observed Value	Simulated Value
Auto-Correlation Investment	-0.085	-0.091
Correlation Investment-Profit	0.175	0.142
Coef. Variation in Tenure	0.756	0.607
Coef. Variation in Profits	0.774	0.661
Change in Mass of Producers 1988-2000	1.711	1.770
S^2 (scaled sum of squared errors)		.990

Notes: The coefficient of variation in tenure is computed in the 2000 Census of Agriculture, which is the first census post-reform for which land subsidies are effective. All other moments are computed using the FADN survey over the 1995-2007 period. The auto-correlation in investment tracks the correlation in investment rates within farms over time, the correlation between investment and profit the correlation between farm investment and profit levels. The coefficients of variation in tenure and profits capture the spread in farm tenure and profits within the market at equilibrium.

Identification: Following Gouriéroux et al. (1993), identification relies on two elements. The first one is that the binding function $\Phi^s(.)$ is one-to-one between the parameter vector and the moments – this one-to-one property defines a strong relevance requirement for the chosen moments, conditional on the targeted parameters. The second one is that the simulation routine should yield a consistent functional estimator of the binding function. What is important for this part, is that the impact of the initial draws for the states of the simulated farms wash away and do not impact the convergence of the binding function to its asymptotic value at Θ.[40]

[39] New entrants are drawn from the distribution of entrants. Cooper and Haltiwanger (2006) and Asker et al. (2014) also rely on indirect inference to recover their model's costs of capital adjustment. They generate a sample and simulate it for a larger number of periods, to only keep the last of them for the estimation of their simulated moments. This is aimed at addressing the biases introducing by the random sampling of initial states. Because the firms can exit, I have to compensate for this by drawing new firms – I cannot rely on the removal of early periods to address this bias. I note that Clementi and Palazzo (2016) and Cao (2007) also face this issue of exit, and balance their simulated panel before computing moments. Depending on the proposed value of fixed costs, balancing the panel is not a feasible option in this case, as firms potentially always exit the market at some point. I address this by bootstrapping the simulation, rather than balancing the simulated dataset.

[40] Collard-Wexler (2013) notes that while maximum-likelihood estimation will be biased in the presence of simulation error

Relevance of moments: While the parameters are recovered jointly from the matching of all moments, it is useful to discuss the parameter-specific relevance of each moment. I show in Figure A16 the distribution of the three variables from which I build moments (investment rate, profit and tenure). Two of these moments relate more specifically to costs of capital adjustment: the within-farm auto-correlation in investment, and the correlation between investment rates and profit. Without any cost of capital adjustment, investment would perfectly correlate with variations in profit—leading to both positive and negative investment shocks, little inaction, and little auto-correlation over time. On the contrary, larger convex costs call for a smoothing of investment—this implies more frequent and smaller investments, and a larger auto-correlation over time.

The last two moments are cross-establishment coefficients of variation in resp. realized profits and incumbency. Conditional on entry, tenure will depend on the value of f_k, the volatility of productivity shocks which I recover from production function estimation (Collard-Wexler (2013), and the ease with which farms can accumulate capital The less it is easy to accumulate it, the larger the option value of holding capital and the more farms can sustain a large negative productivity shock relative to the value of f_k. The higher f_k, the larger the productivity exit threshold and the shorter the tenure of farms—hence the shorter the dispersion in tenure in the data.

This variation in tenure across farm is itself a good measure of turnover within the market. In that sense, it is also related to the value of $f_e^{land\ subsidy}$—as a higher f_e relative to f_k will create a larger option value of incumbency. The spread in realized profits relates in a similar way to the model's fixed costs. A larger spread in profits indicate that farms can sustain smaller profits and remain on the market, it also means that smaller profit levels enter a farm's expected value of entry—to which the fixed cost of entry is compared. As discussed previously, in an economy where one expects the zero profit condition to bind, one can also track entry costs by looking at expected profits of entry. In this context, these expectations are driven by two key elements—the expected duration of firm tenure, and the spread and level of realized profits. Recovering the spread for both of these should then be extremely relevant in the effort to recover the cost of entry.

4.2.4 Results

I obtain the following set of results, and give in Table A10 the values I use for the additional set of model parameters that I directly calibrate.

(Pakes et al. (2007)), indirect inference will be consistent as simulation error will wash out in expectation when computing conditional means and moments.

Table 2: Estimated Model Parameters

Coefficient	Parameter	Estimate	Std. Error
Demand Parameters			
Demand constant (wheat)	α_{wheat}	20.67	(0.83)
Demand constant (other)	α_{other}	20.22	(1.03)
Demand elasticity (wheat)	β_{wheat}	-.29	(0.16)
Demand elasticity (other)	β_{other}	-.24	(0.20)
Production Function			
Substitution Land-Chemicals	ρ	0.36	(0.22)
Substitution Fertilizers-Pesticides	ρ_2	-2.35	(1.09)
Land Share (Wheat)	δ_s^{wheat}	0.83	(0.30)
Land Share (Others)	δ_s^{other}	0.16	(0.28)
Pesticide Share (Wheat)	δ_p^{wheat}	0.35	(0.29)
Pesticide Share (Others)	δ_p^{other}	0.64	(0.35)
Labor Share (Wheat)	α_L^{wheat}	0.24	(0.09)
Labor Share (Others)	α_L^{other}	0.18	(0.084)
Land-Nest Share (Wheat)	α_S^{wheat}	0.39	(0.099)
Land-Nest Share (Others)	α_S^{other}	0.52	(0.083)
Capital & Fixed Costs			
Convex Adjustment Cost	C_k^Q	0.00	/
Fixed Cost of Incumbency	f_k	.75	/
Returns to Scale (entry good)	α_e	.12	/
Scaling (entry good)	A_e	11.4e5	/

Notes: For demand, the parameters come from IV regressions of aggregate quantity sold on prices, using local weather deviations as supply shocks. Standard errors are corrected for auto-correlation using the Newey-West procedure. For the production function, the parameters are obtained from a two-step estimation, each step performed by GMM. The estimation is run on the FADN French sample restricted to 1980-2007. Observations prior to 1980 are removed, as they do not contain output price data. Observations post-2007 are removed, to focus on a period with significant variation in EU agricultural subsidies, both across time and across crops. I keep farms observed for at least three periods in a row[41], and that produce either only crops in the wheat group, or in the other crops group. I remove farms that are not observed with positive input values for our set of considered inputs (land, labor, capital, fertilizers, pesticides). standard errors are obtained using a block bootstrap procedure, where I draw all the observations related to a farm at a time, using $B = 1000$*. For capital and fixed costs, parameters are obtained using indirect inference, and a random grid search for which the standard errors will be computed numerically.*

The parameters estimated jointly imply that more profitable firms are more pollution-intensive, as shown in Figure A28. In the competitive market, profitability is also a direct measure of efficiency, hence there is also a positive relation between production efficiency and pollution intensity.

Recovered Input Substitution: The elasticities of substitution across resp. land and chemicals, and fertilizers and pesticides are the key parameters recovered. The first elasticity is $\sigma = 1.57$, indicating that land and chemicals are substitutes, and the second

[41] This is done to smooth potential measurement issues, and follows from De Loecker et al. (2016) and Doraszelski and Jomandreu (2018).

is $\sigma_2 = .29$, meaning that fertilizers and pesticides are complements. These elasticities can be compared to those found in the literature.[42] Carpentier and Rainelli (1997) for France, McGuirk and Mundlak (1991) for India, and Hayami and Ruttan (1971) generally, all discuss the role of technical change in agriculture post-WW2 as a bundle of higher yield seed varieties coming with either higher water, fertilizer or pesticides needs—this calls for technological change to be chemical-using or land and chemicals to be substitutes. Carpentier and Rainelli (1997) also discuss how new French wheat varieties could absorb more fertilizers at the cost of more pest exposure, and hence higher pesticides needs. It does seem necessary to allow for this complementarity in the production function, and both the specification and the recovered value of σ_2 go in this direction. In terms of the substitution between land and chemicals, I obtain a relatively large value, which can be related to the studies of input-biased technical change in agriculture discussed by Hayami and Ruttan (1971), and for example the elasticity of 1.71 between labor and capital recovered for US agriculture by Kislev and Peterson (1982). For Hayami and Ruttan (1971) this represents a move along the meta-production function (a change in the production function structure) rather than an elasticity of substitution across inputs holding production technology fixed. As such, larger elasticities of substitution for them was the mark of input-biased technological change, which matches very much my interpretation. σ here represents the substitution between land and effective chemicals within the farm, and an increase in effective chemicals can come from a rise in chemical use, or an increase in chemical productivity (arguably a change in production practices and technology within the farm).[43]

5 Policy Analysis

I start this section by revisiting the effects of the MacSharry reform, then study in more details the effects of different policy designs holding budget expenses constant. Finally, I compare the welfare consequences of a land subsidy to a chemical tax, and do so for different levels of subsidization.

[42] Note that there are debates surrounding how to specify production functions that include pesticides as inputs. Lichtenberg and Zilberman (1986) for example argue for modeling pesticides as aiding for pest abatement rather than entering production functions as a direct input. My approach aims at allowing for fertilizers and pesticides to potentially be complements, and for chemicals intensity to be associated with increases in productivity.

[43] Discussions around the extension of the original Hicksian elasticity concept Hicks (1932) to more than two inputs originally led to different elasticity concepts. Allen and Hicks (1934) introduced two extensions, one that Blackorby and Russell (1989) call the Hick's elasticity of substitution—applying the original two-input formula to each pair of inputs, holding all other input quantities constant, as well as output. In the case of a CES function—additive and homothetic—this extension yields constant elasticities across all inputs. This is the relation I recover here. The second extension is the Allen Uzawa partial elasticity of substitution. As Blackorby and Russell (1989) discuss, this elasticity does not easily compare to the estimates I recover here, and in a more than two inputs context, it does not map clearly with the original Hicksian concept. Specifically, the Allen elasticity is not a measure of curvature of the isoquant, provides no information about relative factor shares, and cannot be interpreted as a logarithmic derivative of a quantity ratio with respect to a price ratio. Most of the agricultural economics literature which focused on input-biased technical change was relying on the Allen partial elasticities, making larger comparisons more difficult – and which is why I focus on Kislev and Peterson (1982) here. The central Binswanger (1974) study of input-biased technical change in agriculture for example relies on Allen elasticities. I also note that other results, such as those of Hayami and Ruttan (1971) rely on national account data to estimate production function, recovering very different objects from the establishment-level elasticities obtained here.

5.1 Evaluating the MacSharry Reform

I first show the impact of the reform on the average costs of production, and chemical intensity of production. Net changes are expressed with respect to pre-reform values.

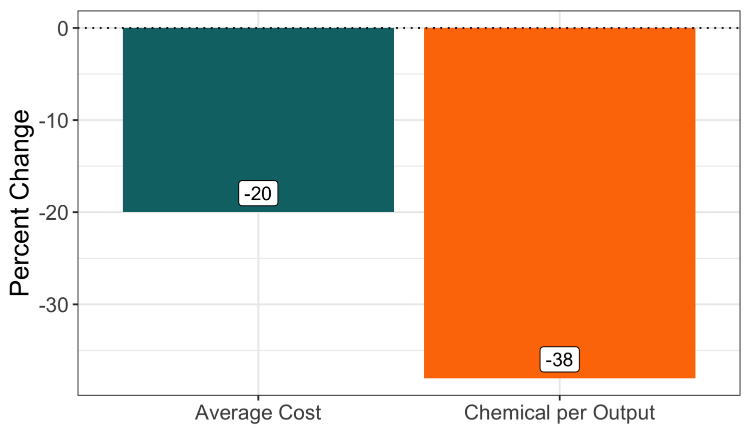


Figure 10: Costs and Pollution Intensity Effects of the MacSharry Reform (Net Changes)

Notes: The figure describes the evolution of the average cost of production, and the average chemical-to-output ratio, comparing the French market in its pre- and post-MacSharry stationary equilibrium.

The reform generated both decreases in the costs of production, and in chemical intensity. On the cost side, there are have two mechanisms going in the same direction. First, the market became less profitable. Inefficient producers exited the market which reduced the average costs of production. Second, production scale was also reduced, which further decreased costs. The impact of the reform on chemical intensity is more surprising, and is the combination of two countervailing effects. First, because inefficient producers are on average less pollution intensive, their exit raised the average chemical-output ratio. But second, the land subsidy had a large impact on the relative chemical-to-land price ratio, incentivizing farmers to use less chemicals in production. This second effect dominated, and led to a total decrease in pollution intensity.

Welfare: I can study the effect of the reform on welfare. For this, I define welfare as economic surplus plus environmental damages – with τ^c the marginal cost of chemical pollution:

$$Welfare = \underbrace{Revenue - Costs}_{Producer\ Surplus} + \underbrace{\int_{p^\star}^{\infty} Q^D(p)dp}_{Consumer\ Surplus} - Subsidies - \tau^c Chemicals.$$

It is difficult to obtain a valuation for chemical pollution, as chemicals have varying environmental impacts—even within fertilizer and pesticide categories—and because effects vary across space as discussed by Rossi et al. (2023). I benchmark τ^c with the marginal tax for the use of glyphosate under the French tax for pesticides (1.36€/kg), and for the same product under the Danish tax for pesticides (9.79€/kg). While the French tax makes sense given this context, the Danish one is largely seen as the most comprehensive within the EU and was recently proposed as a blueprint for a potential EU-wide tax. I get the following changes from the MacSharry, where I express changes in percent of the total value of production under price intervention.

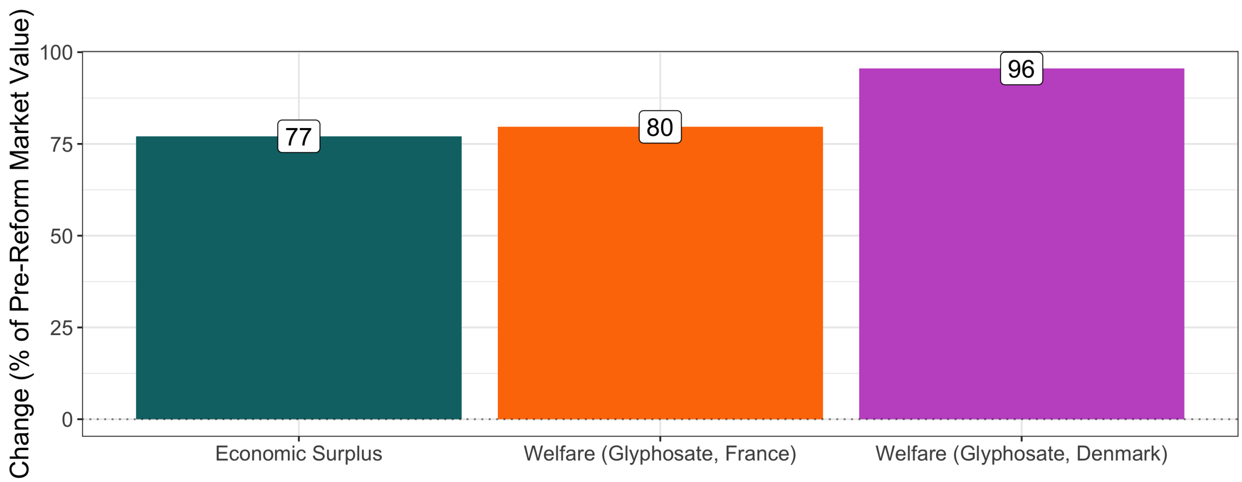


Figure 11: Welfare Effects of the MacSharry Reform (Net Changes)

Notes: The figure describes the net evolution of economic surplus and welfare following the MacSharry reform. For example, 77% indicates a 77% gain in economic surplus. Economic surplus is defined as producer surplus, consumer surplus and the cost of the policy. Welfare corresponds to economic surplus minus the cost from pollution. All values are expressed in percent of the total market value pre-reform.

The MacSharry reform led to large economic gains, and even larger welfare gains depending on the chosen valuation for chemical pollution. Importantly, I do not give any weight here for its impact on farm exit, potential associated impacts on inequalities and labor markets, nor other environmental externalities also impacted by the reform.

5.2 Relative Efficiency of Subsidies Targeting Chemical Use

I then investigate two types of agricultural subsidies which highlight mechanisms through which policy can reduce environmental externalities. The first one is a land subsidy, which is still the primary form of subsidization of agriculture in the EU, and will shift the relative price of chemicals and land for producers. The second is a lump sum payment to organic or low pollution producers. The CAP has introduced a "Greening Payment" in 2014, which supplements the land subsidy for farmers respecting a series of pro-environmental measures. Here, I model two subsidies to highlight two different channels of intervention: subsidies that distort production choices within the farm, and lump-sum payments that distort the continuation decision of farms by favoring lower pollution producers and hence reallocating production across producers. Lump-sum payments are here payments to all the farms with a lower chemical efficiency than the lowest surviving farm under no-intervention. I express the value of the lump-sum in terms of this farm's equilibrium profit under no-intervention. This section discusses the relative importance of these "within" and "between" channels. Table A11 and Table A12 describe outcomes for different levels of intervention, and I plot below the main conclusions drawn from this analysis.

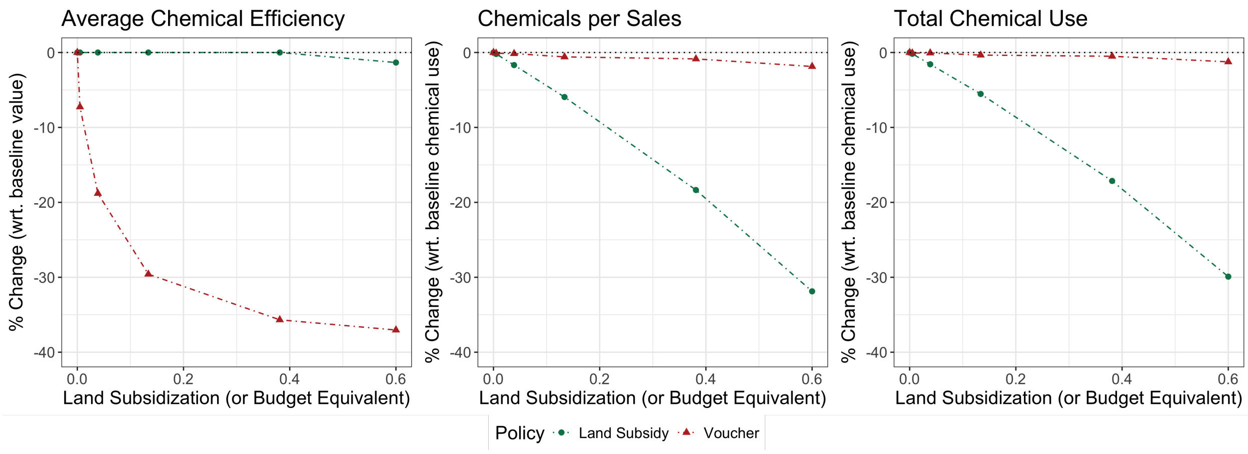


Figure 12: Consequences on Chemical Use (Net Changes)

Notes: I first plot net changes in the average chemical efficiency across farms in the market, as well as in the amount of chemicals per euro of output produced. I finally show the net change in total chemical use. All changes are expressed as net changes relative to no-intervention levels. Lump-Sums are computed at a level which is ex-post budget equivalent to the different land subsidies.

The first panel of Figure 12 illustrates the impact of subsidies on selection along chemical efficiency, which is the externality-inducing dimension of farm productivity. The lump-sum heavily reduces selection along that dimension, which is precisely its goal, while the land subsidy has almost no impact on selection. However, these effects on selection do not really carry through to the chemical intensity of production. The lump-sum only leads to a small decrease in chemical intensity, and consequently to a small decrease in total chemical use. On the contrary, the land subsidy leads to large changes in chemical intensity and total chemical use. Our conclusion is that for the French agricultural market, subsidies that target selective forces and lead to reallocations only have a small impact on aggregate externalities. On the contrary, subsidies that shift production behavior within the farm, for all farms on the market, have a large impact on these aggregate outcomes. This pattern is similar to the one highlighted by Backus (2020) when looking at the link between market competition and productivity—there within-firm changes in production induced by changes in the market's competitiveness have a large impact on aggregate productivity, while reallocations across firms through changes in competition only have a small aggregate impact. In this case, the lump sum changes the type of farm which can remain on the market at the margin. These farms are small and only account for a small fraction of the total market. Hence, subsidies that target them have marginally small effects. On the contrary, subsidies which target all producers on the market like the land subsidy matter for aggregate outcomes. I conclude this section by looking at the welfare consequences of these policies. I also give the marginal value of public funds (MVPF) associated to these policies. The MVPF is a useful tool to compare the effectiveness of policies across different contexts, and has been used by Hahn et al. (2024) to compare a large set of policies targeting climate change in the U.S.. It corresponds to the net benefits of a policy relative to its net government cost. In our context, the MVPF is given by the change in the sum of consumer surplus, producer surplus and chemical pollution, relative to the cost of the policy. As for welfare, I give an MVPF for both the valuations of chemical use implied by the French and Danish taxes for glyphosate.

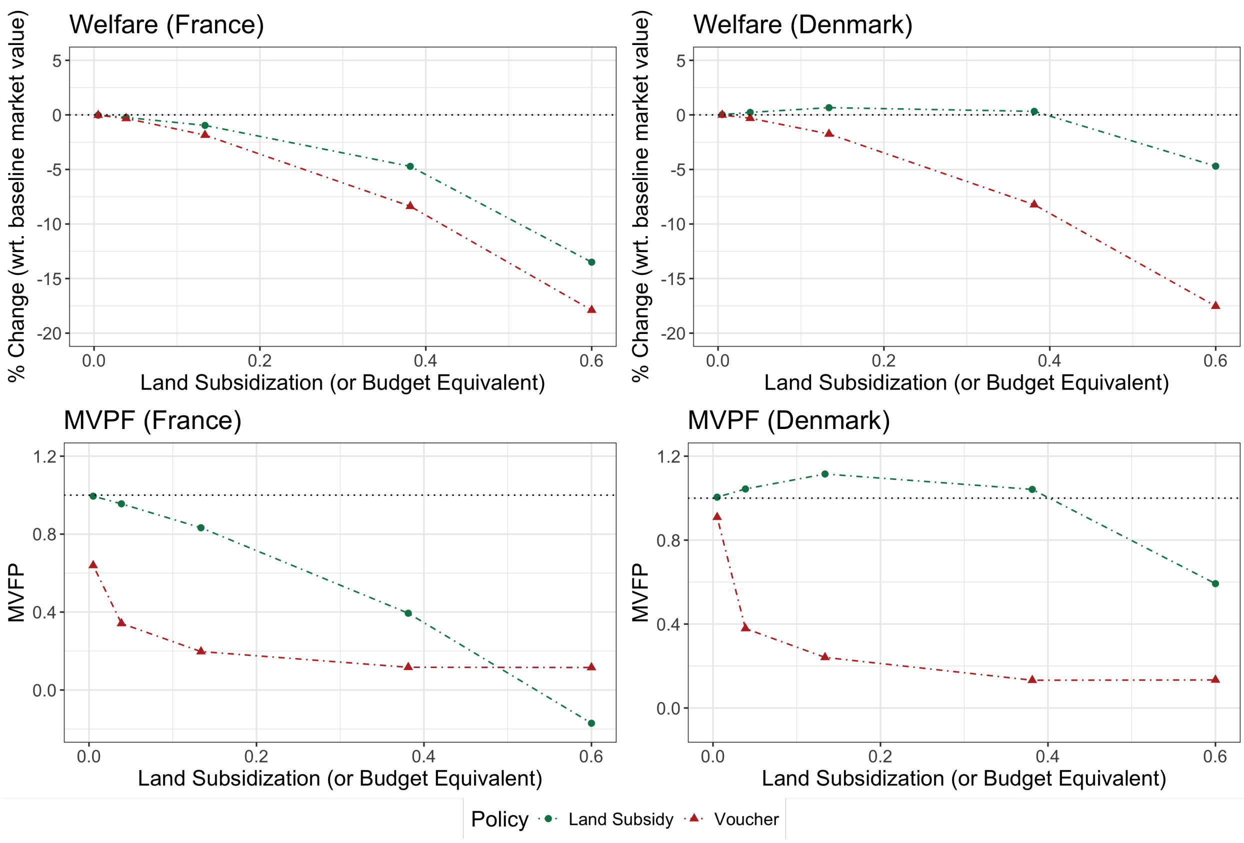


Figure 13: Consequences on Welfare (Net Changes) & Policy MVPF

Notes: I plot the welfare and MVPF implied by the different policies, resp. for the valuation of chemical pollution implied by the French and Danish taxes for glyphosate use. Lump-Sums are computed at a level which is ex-post budget equivalent to the different land subsidies.

For the low valuations of chemical pollution implied by the French tax, all the policies have negative effects on welfare, and consequently have MVPF values under 1 (their costs is larger than their benefits). For the higher valuation implied by the Danish tax, the land subsidy can lead to net welfare increases, up to a value of about 40% of the price of land. Higher valuations for pollution would imply larger gains. Under the Danish tax, the MVPF of 1.11 for a 13% land subsidy is comparable to the range of MVPF found by Hahn et al. (2024) for US electric vehicle policies, as well as appliance rebates.

5.3 Taxes versus Subsidies

In the last section, I compare the effect of a land subsidy to the one of a chemical tax. Specifically, I want to know how close a land subsidy can get us to a chemical tax. The EU recently debated the introduction of a tax for chemical use in agriculture, before withdrawing this proposal under heavy pressure from large farm lobbies. A land subsidy is both easier to implement politically, and potentially entails lower information requirements if the tax is to be implemented with a lump-sum repayment to farmers.

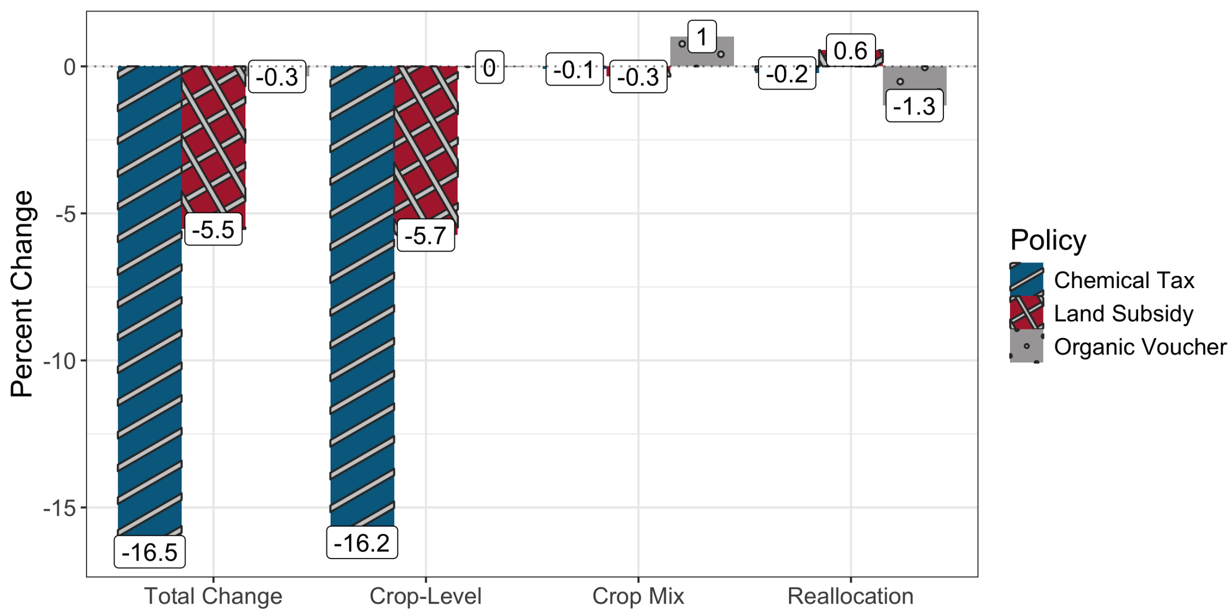


Figure 14: Decomposing Effects on Total Chemical Use (Net Changes)

Notes: I decompose the effects of a 13.4% land subsidy, the budget equivalent lump-sum payment to organic farms (5.5% of the marginal survivor's profit under no-intervention), and a 15.5% chemical tax inducing a similar shift in the chemical-land price ratio as the land subsidy. Effects are decomposed in changes in chemical-per-output intensity within-farm within-crop, changes in chemical-per-output intensity within-farm across-crop, and changes in chemical use across farms. I give the formula for the accounting decomposition in subsection B.10.

For each land subsidy level, I also study the impact of a chemical tax inducing a similar change in the chemical-to-land price ratio. Prior to showing these results, I first describe the channels through which the tax, the subsidy and the organic voucher impact total chemical use. I modify for this the decomposition of De Loecker and Collard-Wexler (2015), for which I give the formula in subsection B.10. There are three channels through which a policy can impact total chemical use: changes within-farm within-crop in production practices (scale and input choice), within-farm and across crops (reallocations within the farm), and across farms (entry, exit and reallocations across producers).

As discussed above, the lump-sum payment to organic farms has a small effect on total chemical use through reallocations across farms. Lower pollution producers can remain on the market and produce, leading to a small decrease in total chemical use. Most of the effect is however counter-acted by equilibrium effects changing relative crop prices and farms' crop mixes in favor of the chemical-intensive crop. The land subsidy impacts total chemical use through the same channel as the tax, which is within-farm within-crop changes in production choices. Specifically, farms shift away from using the now relatively more expensive chemicals. However, the tax changes the price of chemicals relative to all other inputs, while the land subsidy only impacts the chemical-land price ratio. This explains that the tax has a larger effect on total chemical use. The second difference of course, is that this tax is budget neutral, while the subsidy has a cost. I conclude, by comparing the relative performance of the land subsidy with respect to the tax.

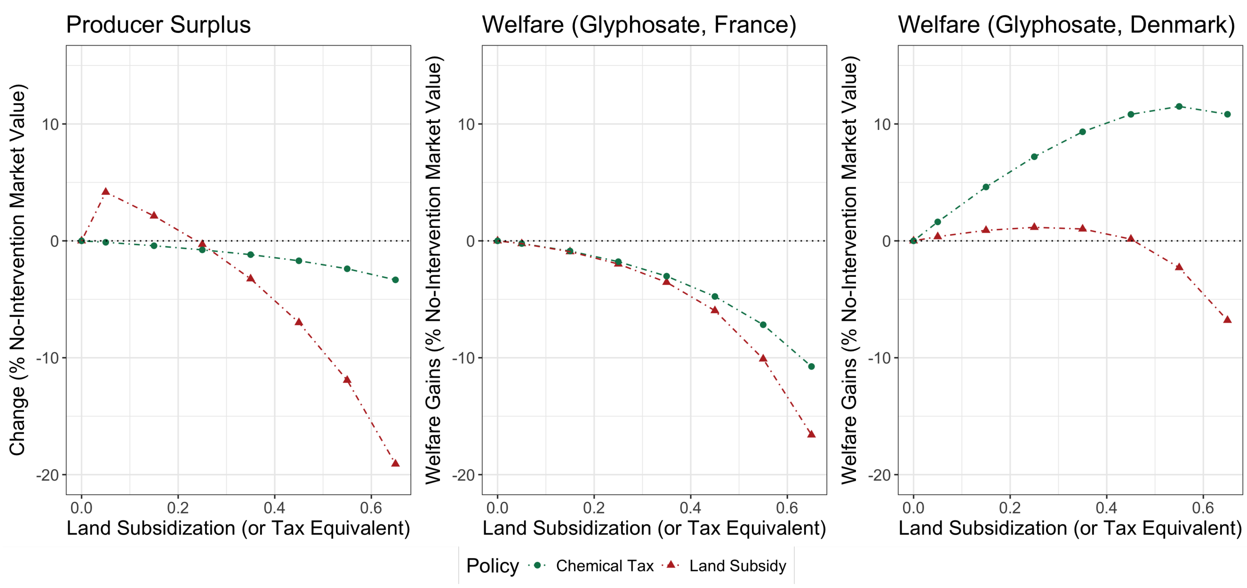


Figure 15: Tax and Subsidy Welfare Effects (Net Changes)

Notes: The figure describes the effect of different taxes and subsidies on producer surplus and welfare. I plot the net changes compared to the no-intervention equilibrium, and express net changes in percent of the no-intervention market value. Welfare values for two different valuations of chemical pollution are shown, using resp. the pesticide taxes for glyphosate in France and Denmark.

Welfare Considerations: Producer surplus can be taken as a measure of producer willingness for the adoption of the policy (without considering the fiscal pressure coming from the cost of the subsidies). For low levels of subsidies, and if the tax is paired with a lump-sum payment, subsidies will be favored by producers over a tax. For a low valuation of pollution reduction, the land subsidy and tax have comparable negative welfare effects.

For the marginal valuation of chemicals implied by the Danish tax for glyphosate, both the subsidy and the tax can generate positive welfare gains. This is always the case for the tax for the values I study, but there is an inflection point in the curve, suggesting very high levels of chemical taxation, or a complete ban on the use of chemicals might have a negative welfare impact under this valuation for pollution. The subsidy starts having a negative impact of welfare at 45%, or slightly above the current level of subsidization.

6 Conclusion

In this paper, I study the effect of EU agricultural subsidies on environmental pollution within the French cereal and oil crop markets, and compare it to the impact of the CAP on its economic surplus. This environmental effect is ambiguous, and depends both on how subsidies impact production behavior at the level of the farm, and the allocation of production across producers. Externalities come from the use of chemicals in production, and the heterogeneity across producers in their propensity to pollute depends on their relative ability at using chemicals. I first use a shift-share design to study the economic and environmental effects of the 1992-1995 MacSharry reform of the CAP. I show it decreased farm profits, led to exit and decreases in pollution. I then develop a model of the French market to propose both an equilibrium study of the reform, and to run counterfactual analyses which focus on the role of subsidy design on chemical use. Estimating my model, I show that in this market there is a positive relation between farm efficiency and pollution intensity. I then investigate how subsidies shift the allocation of production across producers, and incentivize producers to use more or less chemicals in production. Land subsidies lead to decreases in economic surplus, but raise the relative price of chemicals and decrease pollution. Their effects are large, as they affect all producers on the market. On the contrary, lump-sum payments to low pollution producers only have a small effect on aggregate pollution, through a marginal reallocation of production away from high pollution producers towards these smaller farms. This is the case, because these low pollution producers only account for a small total market share. A natural direction along which to expand this research would be to study the spatial reallocation of agriculture across EU member states, for example focusing on the effects of the 2004 and 2007 enlargements to Eastern Europe. This phenomenon likely led to land abandonment and potential forest growth in some areas of the EU, and in the intensification of agriculture and pollution in others.

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015): "Identification Properties of Recent Production Function Estimators," *Econometrica*.

Adamopoulos, T. and D. Restuccia (2014): "The Size Distribution of Farms and International Productivity Differences," *American Economic Review*.

Ager, P., J. J. Feigenbaum, C. W. Hansen, and H. R. Tan (2024): "How the Other Half Died: Immigration and Mortality in U.S. Cities," *Review of Economic Studies*.

Allen, R. and J. Hicks (1934): "A Resonsideration of the Theory of Value Pt. II," *Economica*.

Amemiya, T. (1974): "The Nonlinear Two-Stage Least-Squares Estimator," *Journal of Econometrics*.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *The Quarterly Journal of Economics*.

Aronoff, D. and W. Rafey (2024): "Conservation Priorities and Environmental Offsets: Markets for Florida Wetlands," *NBER Working Paper*.

Asker, J., A. Collard-Wexler, and J. De Loecker (2014): "Dynamic Inputs and Resource (Mis)Allocation," *Review of Economic Studies*.

Aspelund, K. M. and A. Russo (2024): "Additionality and Asymmetric Information in Environmental Markets: Evidence from Conservation Auctions," *Working Paper*.

Backus, M. (2020): "Why is Productivity Correlated with Competition?" *Econometrica*.

Balboni, C. A., A. Berman, R. Burgess, and B. A. Olken (2023): "The Economics of Tropical Deforestation," *Annual Review of Economics*.

Baumol, W. J. (1977): "On the Proper Cost Tests for Natural Monopoly in a Multiproduct Industry," *American Economic Review*.

Binswanger, H. (1974): "The Measuremeng of Technical Change Biases with Many Factors of Production," *American Review of Economics*.

Blackorby, C. and R. R. Russell (1989): "Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities)," *American Economic Review*.

Blundell, W., G. Gowrisankaran, and A. Langer (2020): "Escalation of Scrutiny: The Gains from Dynamic Enforcement of Environmental Regulations," *American Economic Review*.

Borusyak, K., P. Hull, and X. Jaravel (2022): "Quasi-Experimental Shift-Share Research Designs," *Review of Economic Studies*.

Braun, T. and W. Schlenker (2023): "Cooling Externality of Large-Scale Irrigation," *NBER Working Paper*.

Britz, W., I. P. Dominguez, A. Zimmermann, and T. Heckelei (2007): "Definitions of the CAPRI Core Modelling System and Interfaces with Other Components of SEAMLESS-IF," *Seamless report no. 26*.

Burlig, F., L. Preonas, and M. Woerman (2024): "Groundwater and Crop Choice in the Short and Long Run," *Working Paper*.

Bustos, P., B. Caprettini, and J. Ponticelli (2016): "Agricultural Productivity and Structural Transformation: Evidence from Brazil," *American Economic Review*.

Cao, S. (2007): "Three Essays on Capital Adjustment, Reallocation and Aggregate Productivity," *PhD Dissertation, University of Texas at Austin*.

Carisma, B. and V. Boero (2024): "Government expenditures in agriculture 2001–2022 – Global and regional trends," *FAOSTAT Analytical Briefs*.

Carleton, T., L. Crews, and I. Nath (2023): "Agriculture, Trade, and the Spatial Efficiency of Global Water Use," *Working Paper*.

Carpentier, A. and P. Rainelli (1997): "Agriculture, Pesticides et Environnement: quelles politiques? Etude de cas de la France," *INRA Research Paper*.

Carson, R. (1962): "Silent Spring," *Houghton Mifflin*.

Chabé-Ferret, S., A. Reynaud, and E. Tene (2021): "Water Quality, Policy Diffusion Effects and Farmers' Behavior," *Working Paper*.

Clemens, M. A., E. G. Lewis, , and H. M. Postel (2018): "Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion," *American Economic Review*.

Clementi, G. L. and B. Palazzo (2016): "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations," *American Economic Journal: Macroeconomics*.

Collard-Wexler, A. (2013): "Demand Fluctuations in the Ready-Mix Concrete Industry," *Econometrica*.

Cooper, R. W. and J. C. Haltiwanger (2006): "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*.

Crossley, M., K. Burke, S. Schoville, and V. Radeloff (2021): "Recent collapse of crop belts and declining diversity of US agriculture since 1840," *Global Change Biology*.

Davis, L. W. (2017): “The Cost of Global Fuel Subsidies,” *The Energy Journal*.

De Loecker, J. (2011): “Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization,” *Econometrica*.

De Loecker, J. and A. Collard-Wexler (2015): “Reallocation and Technology: Evidence from the US Steel Industry,” *American Economic Review*.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016): “Prices, Markups, and Trade Reform,” *Econometrica*.

Dhyne, E., A. Petrin, V. Smeets, and F. Warzynski (2022): “Theory for Extending Single-Product Production Function Estimation to Multi-Product Settings,” *NBER Working Paper*.

Dias, M., R. Rocha, and R. R. Soares (2023): “Down the River: Glyphosate Use in Agriculture and Birth Outcomes of Surrounding Populations,” *Review of Economic Studies*.

Doraszelski, U. and J. Jomandreu (2018): “Measuring the Bias of Technological Change,” *Journal of Political Economy*.

Eckel, C. and J. P. Neary (2010): “Multi-Product Firms and Flexible Manufacturing in the Global Economy,” *Review of Economic Studies*.

Englander, G., J. Zhang, J. C. Villasenor-Derbez, Q. Jiang, M. Hu, O. Deschenes, and C. Costello (2023): “Input Subsidies and the Destruction of Natural Capital: Chinese Distant Water Fishing,” *NBER Working Paper*.

Foster, A. D. and M. R. Rosenzweig (2022): “Are there too many Farms in the World? labor Market Transaction Costs, Machine Capacities, and Optimal Farm Size,” *Journal of Political Economy*.

Foster, L., J. Haltiwanger, and C. Syverson (2008): “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*.

Fowler, D., M. Coyle, U. Skiba, M. A. Sutton, J. N. Cape, S. Reis, L. J. Sheppard, A. Jenkins, B. Grizzetti, J. N. Galloway, P. Vitousek, A. Leach, A. F. Bouwman, K. Butterbach-Bahl, F. Dentener, D. Stevenson, M. Amann, and M. Voss (2013): “The Global Nitrogen Cycle in the Twenty-First Century,” *Philosophical Transactions of the Royal Society: Biological Sciences*.

Frank, E. (2021): “The Economic Impacts of Ecosystem Disruptions: Costs From Substituting Biological Pest Control,” *Working Paper*.

Gardner, B. (2002): “Changing Economic Perspectices on the Farm Problem,” *Journal of Economic Literature*.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020): "Bartik Instruments: What, When, Why and How," *American Economic Review*.

Gollin, D., C. W. Hansen, and A. M. Wingender (2021): "Two Blades of Grass: The Impact of the Green Revolution," *Journal of Political Economy*.

Gouriéroux, C., A. Monfort, and E. Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*.

Gowrisankaran, G., A. Langer, and M. Reguant (2024): "Energy Transitions in Regulated Markets," *Working Paper*.

Gowrisankaran, G., S. S. Reynolds, and M. Samano (2016): "Intermittency and hte Value of Renewable Energy," *Journal of Political Economy*.

Griliches, Z. and J. Mairesse (1998): "Production Functions: The Search for Identification," *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*.

Grosset-Touba, F., A. Papp, and C. Taylor (2024): "Rain Follows the Forest: Land Use Policy, Climate Change, and Adaptation," *Working Paper*.

Guillermet, J. (2015): "Couts de Production en Grandes Cultures," *Technical Report*.

Hahn, R. W., N. Handren, R. Metcalfe, and B. Sprung-Keyser (2024): "A Welfare Analysis of Policies Impacting Climate Change," *Working Paper*.

Hayami, Y. and V. W. Ruttan (1971): *Agricultural Development: An International Perspective*, John Hopkins University Press.

Hicks, J. (1932): "Theory of Wages," *Macmillan*.

Hopenhayn, H. (1992): "Entry and Exit in Long Run Equilibria," *Econometrica*.

Hopenhayn, H. and R. Rogerson (1993): "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*.

Hornbeck, R. and P. Keskin (2014): "The Historically Evolving Impact of the Ogallala Aquifer: Agricultural Adaptation to Groundwater and Drought," *American Economic Journal: Applied Economics*.

Hsiao, A. (2022): "Coordination and Commitment in International Climate Action: Evidence from Palm Oil," *Working Paper*.

Johnson, R., C. E. Hanrahan, and R. Schnepf (2010): "Comparing U.S. and EU Program Support for Farm Commodities and Conservation," *CRS Report for Congress*.

Johnston, S. (2020): "Dynamic Inputs, Misallocation, and Policy: Evidence from Electricity-Intensive Manufacturing," *Working Paper*.

KISLEV, Y. AND W. PETERSON (1982): "Prices, Technology, and Farm Size," *Journal of Political Economy.*

KLENOW, P. J. AND H. LI (2024): "Entry Costs Rise with Growth," *Working Paper.*

KLETTE, T. J. AND S. KORTUM (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy.*

KOVAK, B. (2013): "Regional Effects of Trade Reform: What is the Correct Measure of Liberalization?" *American Economic Review.*

LAGAKOS, D. AND M. E. WAUGH (2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review.*

LEVINSOHN, J. AND A. PETRIN (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies.*

LICHTENBERG, E. AND D. ZILBERMAN (1986): "The Econometrics of Damage Control: Why Specification Matters," *American Journal of Agricultural Economics.*

LIVINGSTON, M., M. ROBERTS, AND J. RUST (2008): "Optimal Corn and Soybean Rotations," *Working Paper.*

MAUE, C. C., M. BURKE, AND K. J. EMERICK (2020): "Productivity Dispersion and Persistence Among the World's Most Numerous Firms," *NBER Working Paper.*

MAYER, T., M. J. MELITZ, AND G. I. OTTAVIANO (2014): "Market Size, Competition, and the Product Mix of Exporters," *American Economic Review.*

MCGUIRK, A. AND Y. MUNDLAK (1991): "Incentives and Constraints in the Transformation of Punjab Agriculture," *IFPRI Research Report.*

MELITZ, M. J. AND G. I. OTTAVIANO (2008): "Market size, trade, and productivity," *Review of Economic Studies.*

MEYNARD, J.-M., T. DORÉ, AND P. LUCAS (2003): "Agronomic Approach: cropping systems and plant diseases," *Comptes Rendus Biologies.*

MISSIRIAN, A. (2020): "Yes, in your backyard: Forced technological adoption and spatial externalities," *Working Paper.*

MORRISON, T. AND M. WATTLEWORTH (1986): "Causes of the 1984-86 Commodity Price Decline: What were the relative roles of supply and demand factors?" *International Monetary Fund, Finance and Development.*

MOYER, H. W. (1993): "EC Decisionmaking, the MacSharry Reform of the CAP, Mastricht, and the GATT Uruguay Round," *Third Biennial International Conference of the European Community Studies Association.*

NOCKE, V. (2006): "Gap for Me: Entrepreneurs and Entry," *Journal of the European Economic Association.*

NOCKE, V. AND N. SCHUTZ (2018): "Multiproduct Firm Oligopoly: An Aggregative Games Approach," *Econometrica.*

OBERFIELD, E. AND D. RAVAL (2021): "Micro Data and Macro Technology," *Econometrica.*

OGEDA, P., E. ORNELAS, AND R. SOARES (2021): "Labor Unions and the Electoral Consequences of Trade Liberalization," *CEP Working Paper.*

OLLEY, G. S. AND A. PAKES (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica.*

ORR, S. (2022): "Within-Firm Productivity Dispersion: Estimates and Implications," *Journal of Political Economy.*

PAKES, A., S. BERRY, AND M. OSTROVSKY (2007): "Simple Estimators for the Parameters of Discrete Dynamic Games, with Entry/Exit Examples," *RAND Journal of Economics.*

PANZAR, J. C. AND R. D. WILLIG (1981): "Economies of Scope," *American Economic Review.*

PETRICK, M. (2008): "The Co-Evolution of Semantics and Policy Paradigms: 50 Years of Europe's Common Agricultural Policy," *Intereconomics.*

PIERRE, G. (2004): "Agriculture Dépendante et Agriculture Durable," *Publications de la Sorbonne.*

POORE, J. AND T. NEMECEK (2018): "Reducing food's environmental impacts through producers and consumers," *Science.*

QI, J., X. TANG, AND X. XI (2021): "The Size Distribution of Firms and Industrial Water Pollution: A Quantitative Analysis of China," *American Economic Journal: Macroeconomics.*

RAFEY, W. (2023): "Droughts, Deluges, and (River) Diversions: Valuing Market-Based Water Reallocation," *American Economic Review.*

RIGAL, S., V. DAKOS, H. ALONSO, Z. BENKŐ, L. BROTONS, T. CHODKIEWICZ, P. CHYLARECKI, E. DE CARLI, J. C. DEL MORAL, C. DOMŞA, V. ESCANDELL, B. FONTAINE, R. FOPPEN, R. GREGORY, S. HARRIS, S. HERRANDO, M. HUSBY, C. IERONYMIDOU, F. JIGUET, J. KENNEDY, A. KLVAŇOVÁ, L. K. PRIMOŽ KMECL AND, P. KURLAVIČIUS, J. A. KÅLÅS, A. LEHIKOINEN, ÅKE LINDSTRÖM, R. LORRILLIÈRE, C. MOSHØJ, R. NELLIS, D. NOBLE, D. P. ESKILDSEN, J.-Y. PAQUET, M. PÉLISSIÉ, C. PLADEVALL, D. PORTOLOU, J. REIF, H. SCHMID, B. SEAMAN, Z. D. SZABO, T. SZÉP, G. T. FLORENZANO, N. TEUFELBAUER, S. TRAUTMANN, C. VAN TURNHOUT, Z. VERMOUZEK,

T. Vikstrøm, P. Voříšek, A. Weiserbs, and V. Devictor (2023): "Farmland Practices are Driving Bird Population Decline across Europe," *PNAS*.

Rossi, D., G. Hoque, Y. Ji, and C. Kling (2023): "The Economics of Nutrient Pollution from Agriculture," *Annual Review of Resource Economics*.

Ryan, N. and A. Sudarshan (2021): "Rationing the Commons," *Journal of Political Economy*.

Schultz, T. (1945): "Agriculture in an Unstable Economy," *McGraw-Hill*.

Scott, P. (2013): "Dynamic Discrete Choice Estimation of Agricultural Land Use," *Working Paper*.

Shapiro, J. S. (2021): "The Environmental Bias of Trade Policy," *The Quarterly Journal of Economics*.

Souza-Rodrigues, E. (2018): "Deforestation in the Amazon: A Unified Framework for Estimation and Policy Analysis," *The Review of Economic Studies*.

Taylor, C. and G. Heal (2023): "Fertilizer and Algal Blooms: A Satellite Approach to Assessing Water Quality." *NBER Risks in Agricultural Supply*.

Taylor, C. A. (2022): "Cicadian Rhythm: Insecticides, Infant Health, and Long-term Outcomes," *CEEP Working Paper*.

——— (2024): "Irrigation and Climate Change: Long-Run Adaptation and its Externalities," *Working Paper*.

Topalova, P. (2010): "Factor Immobility and Regional Effects of Trade Liberalization: Evidence on Poverty and Inequality from Indian Districts," *American Economic Journal: Applied Economics*.

Valmari, N. (2023): "Estimating Production Functions of Multiproduct Firms," *Review of Economic Studies*.

Watkins, B. and A. van Niekerk (2019): "A comparison of object-based image analysis approaches for field boundary delineation using multi-temporal Sentinel-2 imagery," *Computers and Eletronics in Agriculture*.

Wooldridge, J. M. (2010): "Econometric Analysis of Cross Section and Panel Data," *MIT Press, Cambridge*.

Appendix for online publication:
An Empirical Model of Agricultural Subsidies with Environmental Externalities

A Additional Figures and Tables

A.1 Figures

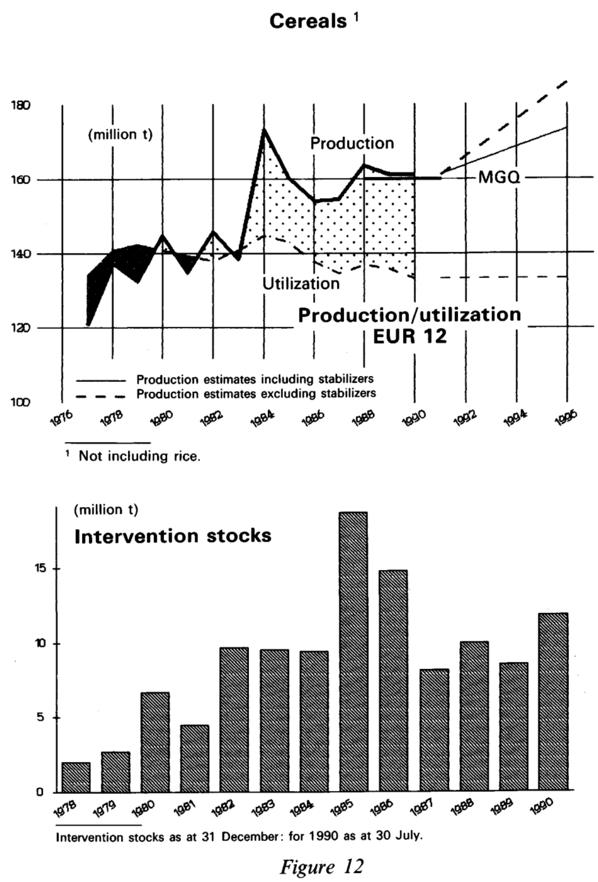


Figure A1: Intervention on the Cereal Market

Notes: This figure taken from a 1990 report of the European Commission describes the growing gap between supply and demand for cereals in the European Union, as well as the evolution of the stock of cereals purchased by the governmental agencies following the purchasing guarantee set on the market. The figure comes from the 1990 The Agricultural Situation in the Community report. The report can be accessed via the University of Pittsburgh at https://aei.pitt.edu/31386/1/CM5890934ENC_002.pdf.

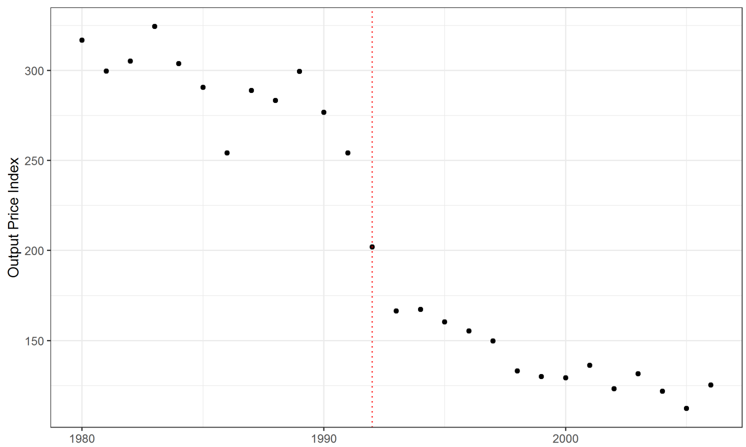


Figure A2: Average Farm Price

Notes: Binscatter for a farm-level average output price, computed over oil and cereal crops, using relative land shares as weights.

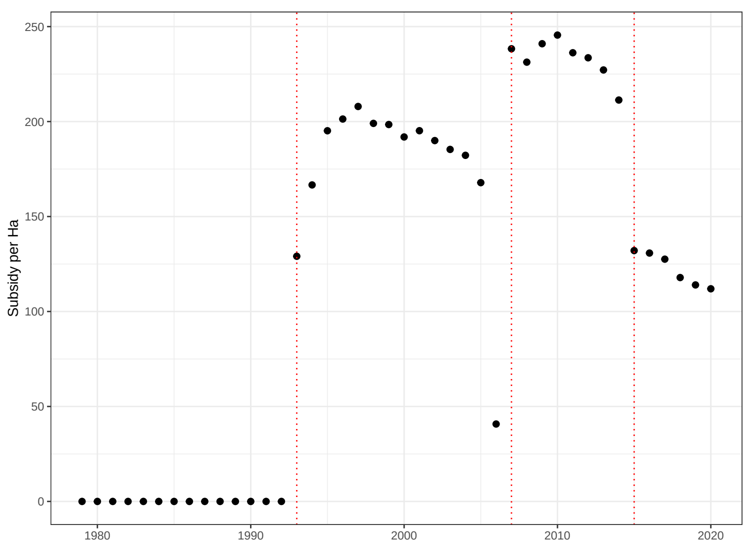


Figure A3: EU Land Subsidies under the CAP

Notes: I show a binscatter for the land subsidies given under the CAP, as observed in the FADN survey of farms. Each vertical red line indicates a reform of the subsidization scheme, the first one corresponding to the MacSharry reform.

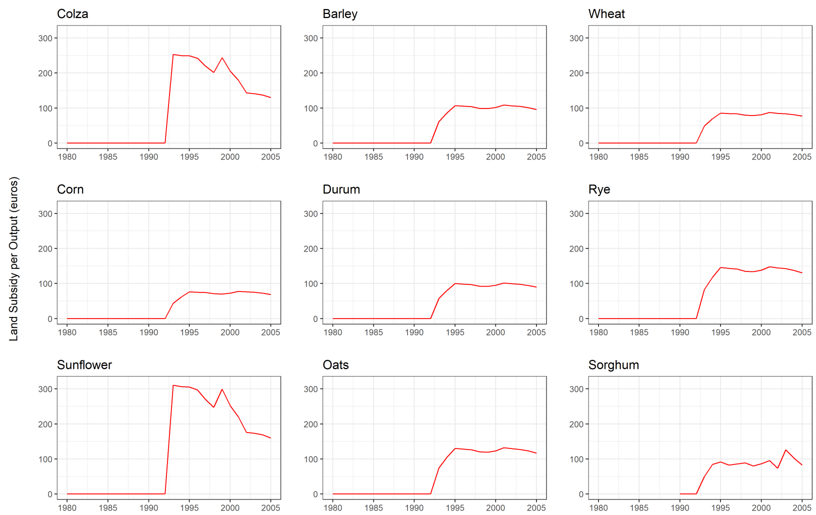


Figure A4: EU Land Subsidies per Crop under the CAP

Notes: The figure shows the evolution of the average land subsidy per unit of output over time, as measured in the FADN. I measure the average subsidy per hectare for cereals and oil crops, dividing the farms' total subsidy for each crop category, by their land allocated to each category. I then divide this measure by the average crop-specific yield observed in France in that year, in order to recover a subsidy per unit of output. Subsidies per land are relatively stable across years, while yields increase, making it so that subsidies slightly decrease over time absent any policy adjustment.

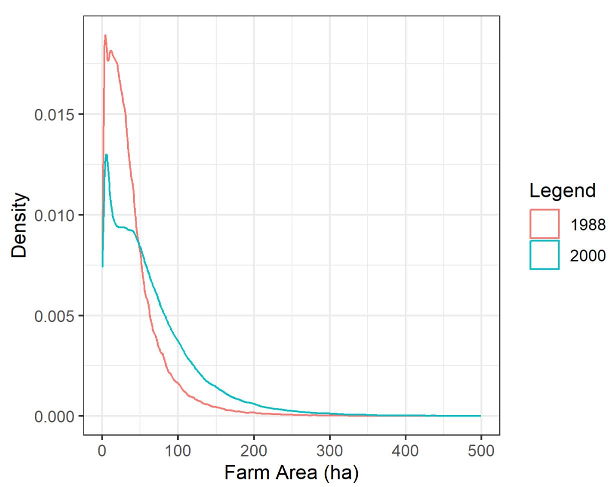


Figure A5: Farm Size Distribution in 1988 and 2000

Notes: The two distributions of the farm size distribution are recovered from the French Census of Agriculture, specifically the 1988 and 2000 waves. Farm area is measured using their total used agricultural area measured in ha. The distributions are recovered using an Epanechnikov kernel. The distribution shifts significantly over time, and that density is reallocated in the right tail of the distribution. This period corresponds to significant farm exit and average farm size growth, and this figure confirms most of the exit is probably located among small farms.

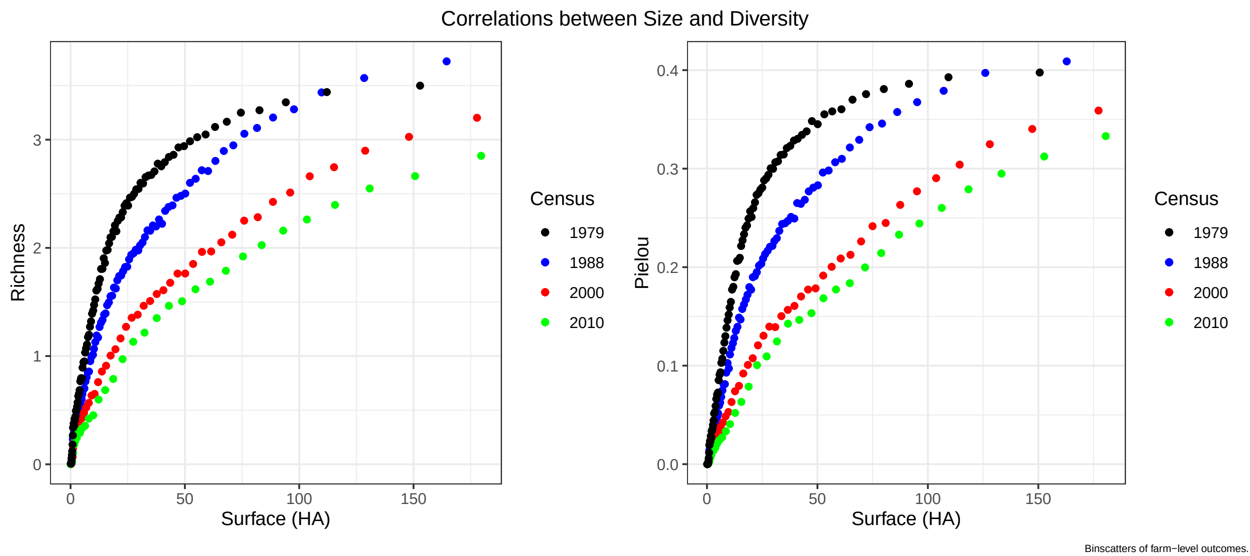


Figure A6: Evolution of Farm Crop Diversity

Notes: These binscatters are computed using data from the Agricultural Census. I take into account the maximum number of crops which can be tracked over the 1970-2010 waves of the Census, which are: wheat, durum, barley, oats, rye, corn (grain), corn (seeds), sorghum, rice, beetroot, rapeseed, sunflower, soy. Crop richness is a simple count measure of the number of crops being grown in a given farm. The Pielou index in a measure of how evenly land is allocated across the crops grown within a farm[44].

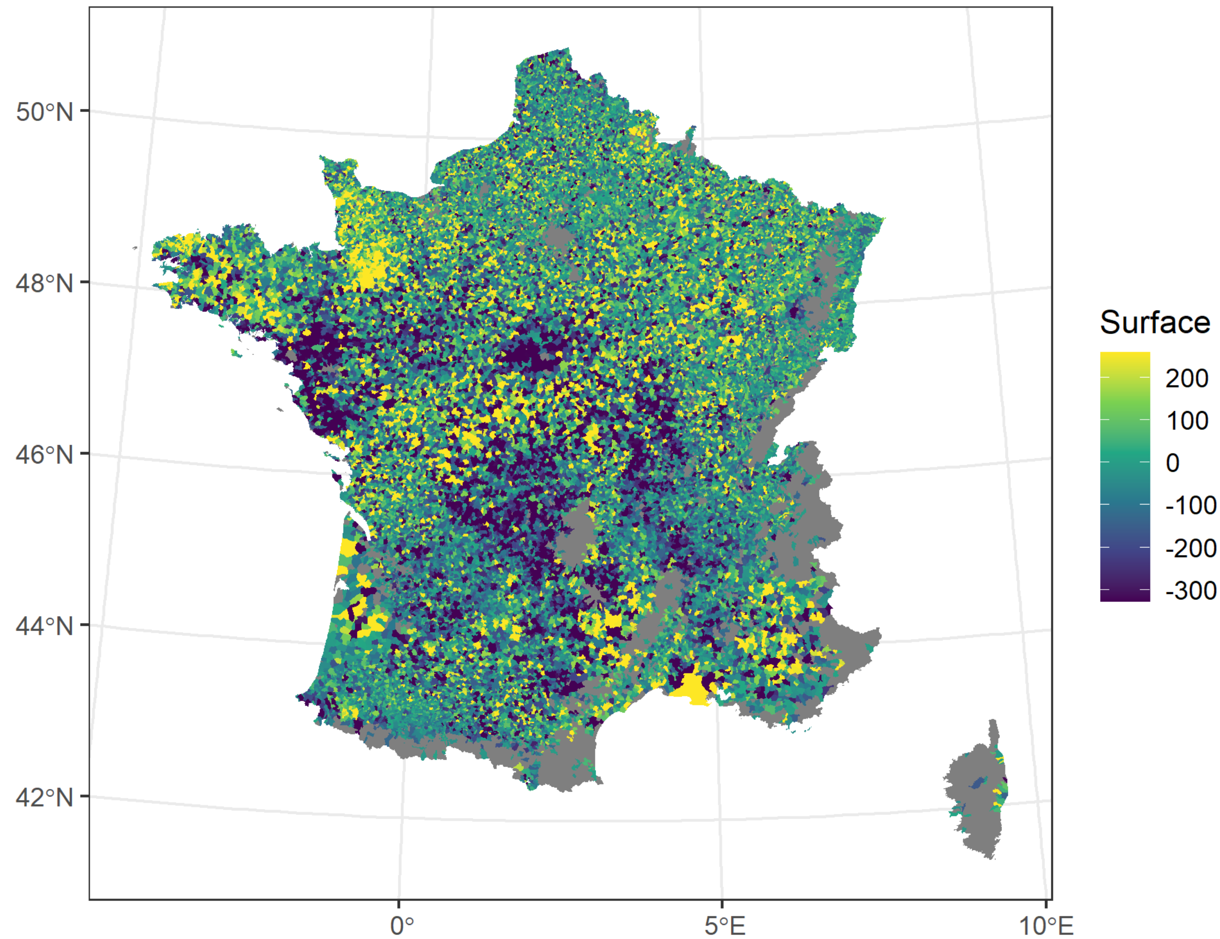


Figure A7: Evolution of Village-level Total Agricultural Area

Notes: This map shows the evolution of within-village total agricultural area between 1988 and 2000. I use the data from the relevant waves of the French Agricultural Census. Decreases in agricultural area mainly happen in the center of France around the Massif Central, as well as in Vendée and possibly lower Brittany.

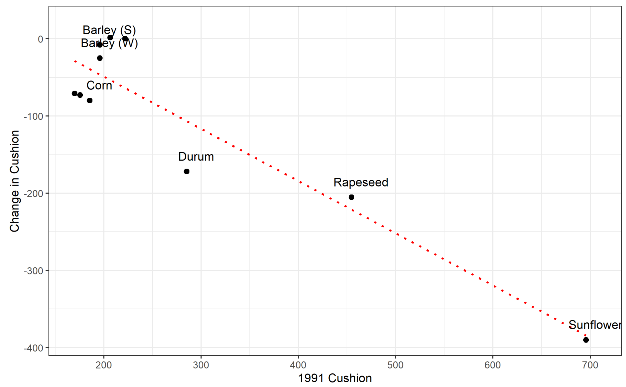


Figure A8: Convergence of Crop Cushions Post Reform

Notes: The figure illustrates the convergence of crop cushions post reform. Here the cushion is the French average of the crop-specific cushions in the relevant year, computed according to the formula outlined in **??**. *The change in cushion is the change observed between 1991 and 1995, corresponding to the period of the MacSharry reform.*

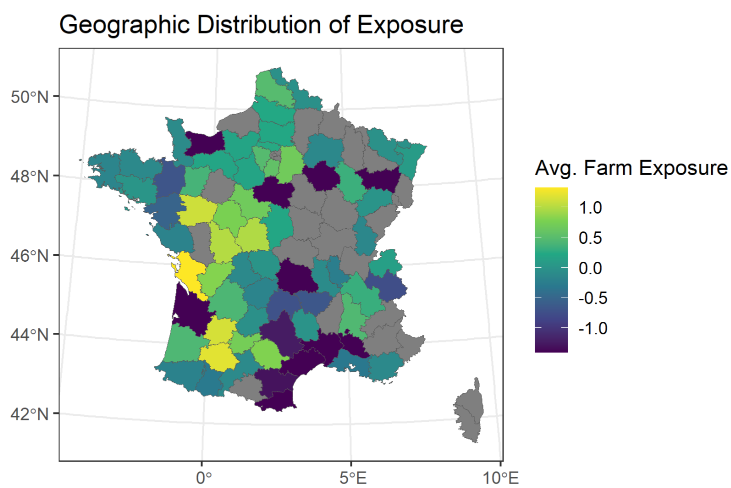


Figure A9: Department Weighted Averages of Farm-Level Exposure

Notes: The figure shows the geographic distribution of our standardized measure of farm-level exposure. The exposure is computed using data from the FADN, and department-level averages are computed using the extrapolation weights provided there. I plot the geographic variation of the standardized exposure here.

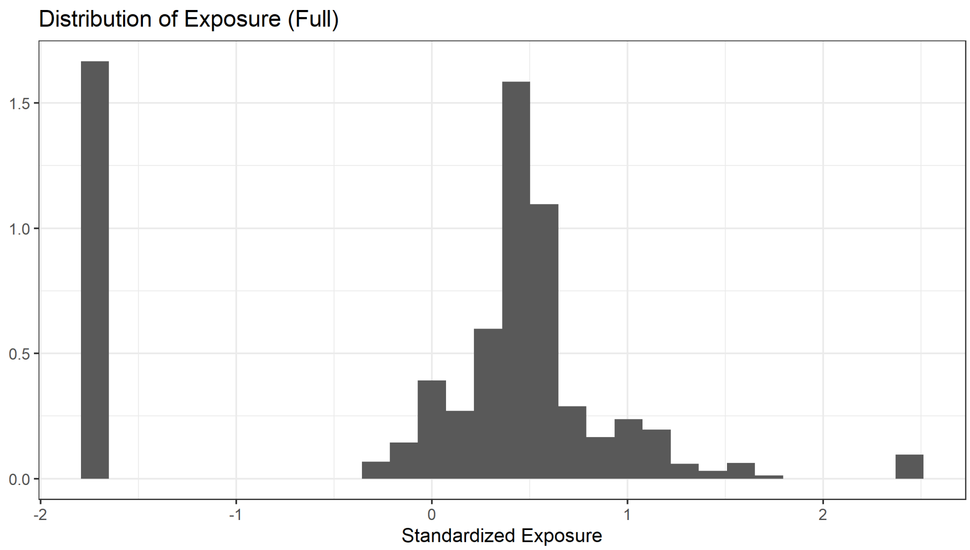


Figure A10: Distribution of Farm-Level Exposure

Notes: The figure shows the distribution of our standardized measure of farm-level exposure (our preferred instrument). The exposure is computed using data from the FADN.

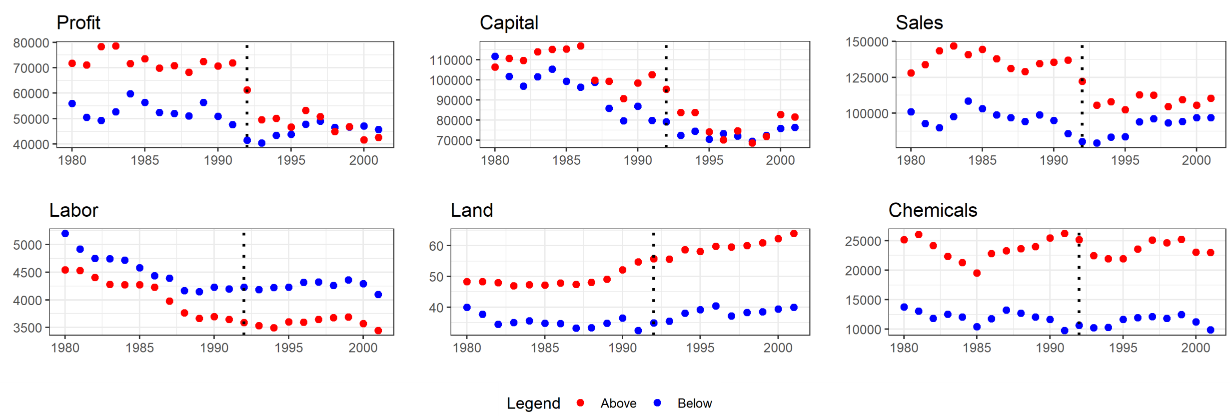
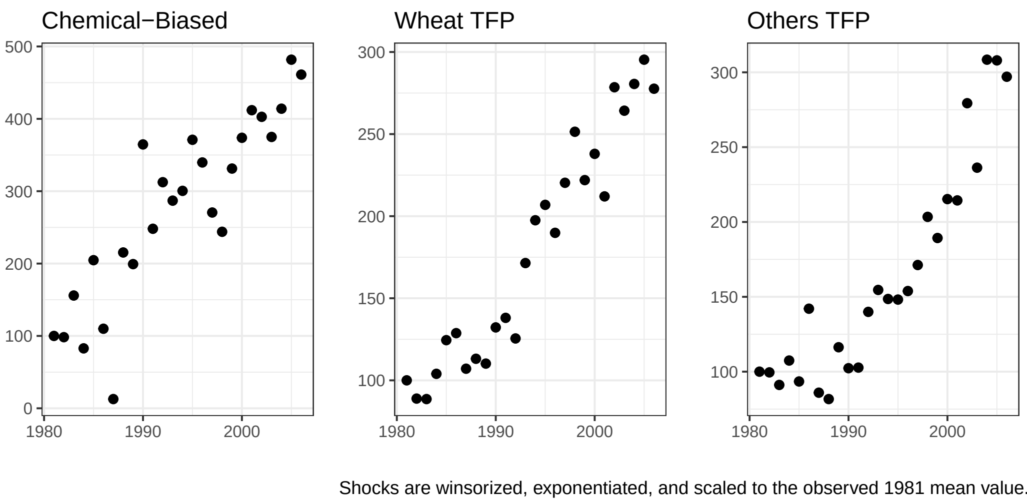
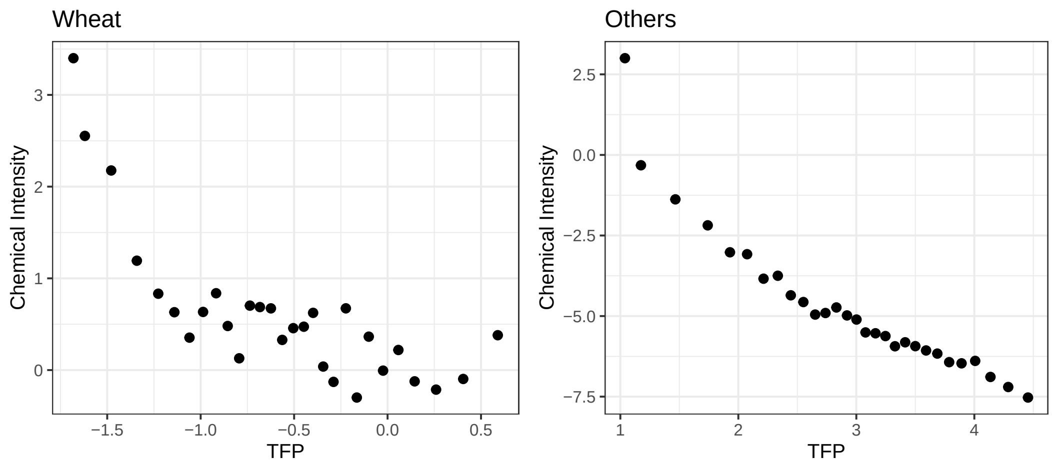


Figure A11: Balance in Levels and Trends

Notes: These binscatters are computed using data from the FADN, and use sampling weights to compute bin-specific averages within each category. Profit corresponds to total farm profit, capital is the sum of total farm value in buildings and machinery, sales is total sales. All of these values are computed in 2020 euros. Labor is total labor in hours per year, and land total utilized agricultural area in hectares.



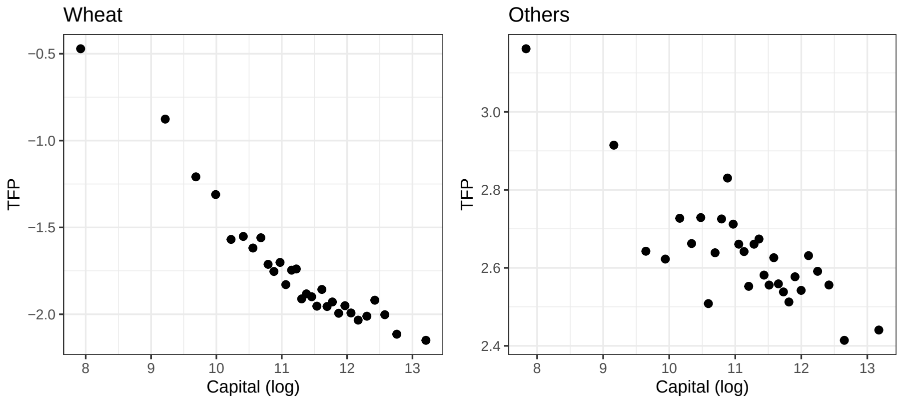
(a)



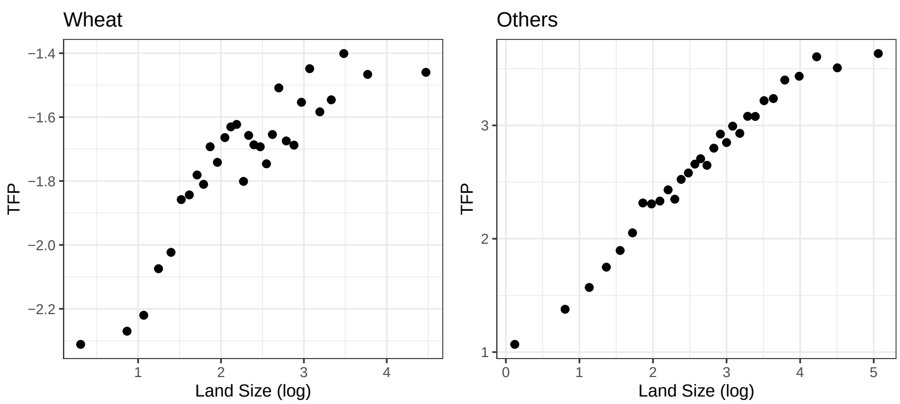
(b)

Figure A12: Recovered Productivity Shocks

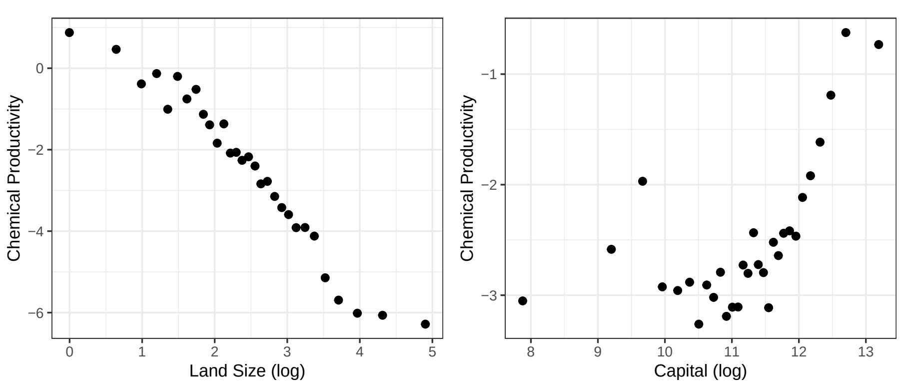
Notes: The first row of figures shows the evolution of productivity shocks as measured in the FADN, resp. the non-Hicksian shock which I introduce as farm-specific and shared across crops, and the two TFP shocks which are differentiated across crops within farms. Shocks for the first row of figures are standardized to their 1980 values, in order to make their evolution over time easier to read. The second row of figures shows the correlations within farms across crop-specific TFP shocks and the non-Hicksian shock. I use binscatters for that purpose, which give us the average values within bins of equal population size along the two-dimensional distribution of our variables. Shocks here are directly measured as ω_{jt}^{ch} *and* $\{\omega_{jt}^{h,wheat}, \omega_{jt}^{h,other}\}$*, explaining that they sometimes take negative values.*



(a)



(b)



(c)

Figure A13: Recovered Productivity Shocks - Relations with Farm Stocks

Notes: The figures show correlations of within farms across crop-specific TFP shocks, the non-Hicksian shock and resp. the farms' capital stock and land use. I use binscatters for that purpose, which give the average values within bins of equal population size along the two-dimensional distribution of our variables. Shocks here are directly measured as ω_{jt}^{ch} *and* $\{\omega_{jt}^{h,wheat}, \omega_{jt}^{h,other}\}$*, explaining that they sometimes take negative values.*

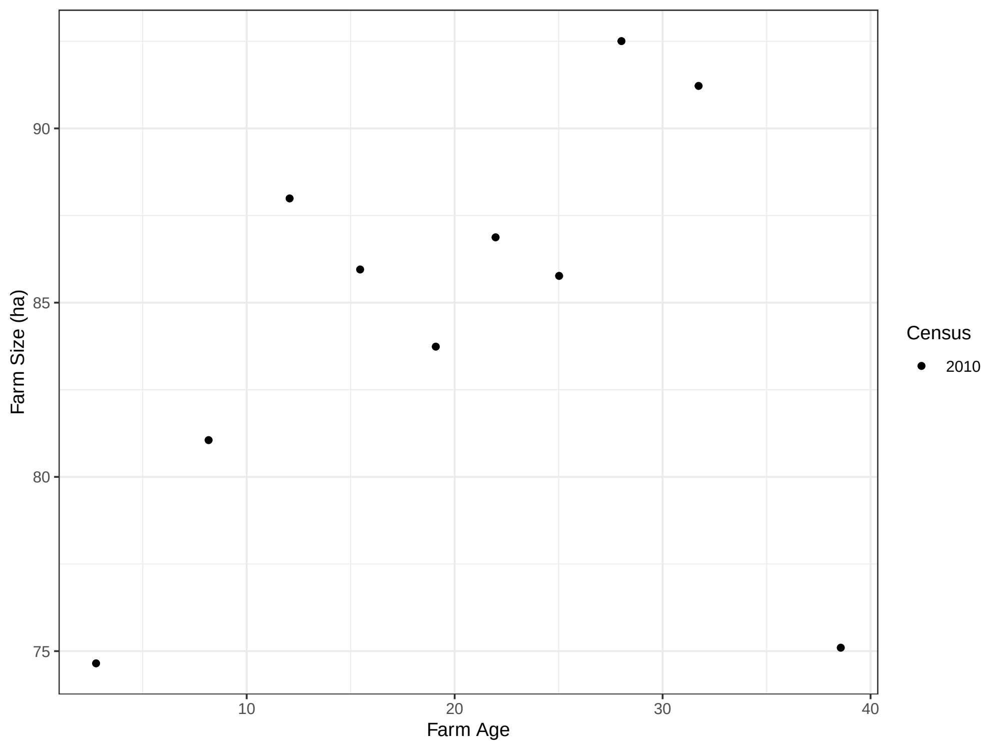
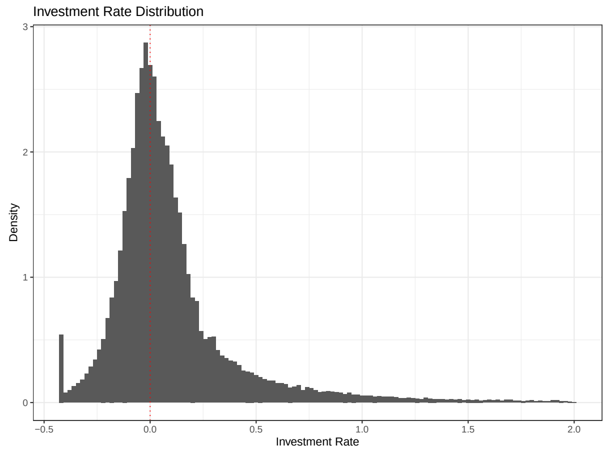
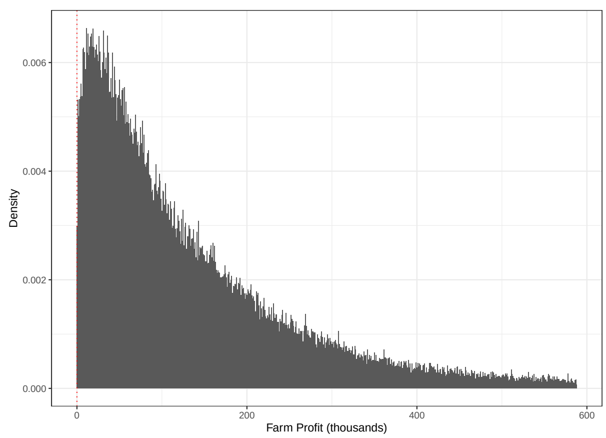


Figure A14: Relation between Farm Size and Farm Age in the 2010 Census

Notes: The figure shows a binscatter showing the relation between farm size and farm tenure, as observed in the 2010 French agricultural census. I use the 2010 Census to capture this relation, as the total number of farms ceases to decrease as quickly between 2000 and 2010, and one can interpret the market as relatively more stable than in the previous decades. In previous waves of the Census, older farms had likely entered the market under very different conditions, and were more likely to be small non-commercial farms with different management styles.

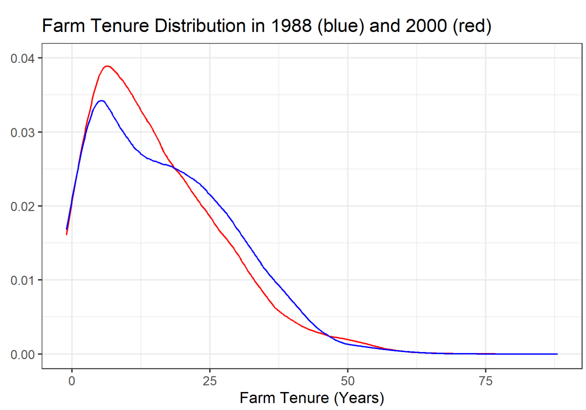


(a) Distribution of Investment Rates (FADN)



(b) Distribution of Profits (FADN)

Figure A15



(a) Distribution of Tenure (Census)

Figure A16

Notes: I plot the distributions from which I obtain our moments used for the indirect inference estimation routine. The investment rate is recovered as the difference between lagged depreciated capital using an 8ppt depreciation rate, and the current capital level. Profits are directly observed in the data, as well as farm tenure. The first two graphs are obtained from FADN data, while the last one is built with 1988 and 2000 Census data.

A.2 Tables

Table A1: Descriptive Statistics - Farm-Level Dataset

Statistic	N	Mean	St. Dev.
Output Volume (sum)	308,984	1,791.687	3,031.357
Output Volume (wheat)	191,357	1,498.631	2,044.714
Output Volume (corn)	95,925	1,162.133	2,046.903
Output Volume (sunflower)	36,263	309.987	338.692
Output Price Index	169,219	213.341	94.838
Farm Surface	308,982	71.021	67.640
Capital	308,984	128,135.600	167,984.500
Total Labor Hours	308,982	4,406.731	3,738.925
Profit	308,984	161,576.600	264,341.700
Phytosanitary (bill)	268,893	8,071.386	11,066.410
Fertilizer (bill)	268,893	16,040.360	17,643.240

Table A2: Evolution of Agricultural Land Ruggedness

Year	Average	$10th$	$25th$	$75th$	$90th$	$99th$
1970	33.2264	13.3583	19.3155	41.8313	58.5489	107.5599
1979	33.1215	13.3772	19.3356	41.6604	58.3028	105.9168
1988	33.0303	13.3583	19.2822	41.5127	58.0048	105.8695
2000	32.9716	13.3321	19.2807	41.4354	57.8952	105.8146

Notes. The data is computed using the full count agricultural census for years 1970, 1979, 1988 and 2000. Ruggedness data is extracted from the European Environment Agency's "Elevation map of Europe based on GTOPO30" on a 1km grid. I compute the average ruggedness of land within each French village, and then compute area weights for the villages using their year-specific share of total agricultural land within France.

Table A3: Agricultural Market Trends

Panel A	All Farms (0/1)								
	Year	N	Total Farm Surface	Total Cultivated Land	Mean Farm Surface	Mean Cultivated Land	Gini (all land)	Gini (cultivated)	Land Share of Top 10^{th} Decile
	1970	1583279	29904735.3300		18.8878		0.5813		.2323
	1979	1257168	29496571.6300		23.4627		0.5792		.2952
	1988	1006120	28595798.5400		28.4219		0.5816		.3744
	2000	653090	27856312.8900		42.6531		0.6130		.6252
	2010	518925	27832907.9000		53.6357		0.6263		.7390
Panel B	Row Crops Farms (1/1)								
	Year	N	Total Farm Surface	Total Cultivated Land	Mean Farm Surface	Mean Cultivated Land	Gini (all land)	Gini (cultivated)	Land Share of Top 10^{th} Decile
	1970	1034687	25984231.5400	9006112.4200	25.1131	8.7042	0.4916	0.6571	.1497
	1979	823432	25882409.8300	11568370.0200	31.4324	14.0490	0.4836	0.6578	.1878
	1988	602476	23827571.5600	11189216.8300	39.5494	18.5721	0.4771	0.6534	.2514
	2000	352098	22884643.2800	11105923.7700	64.9951	31.5421	0.4754	0.6295	.4929
	2010	260731	21827942.3500	10979980.2300	83.7182	42.1123	0.4608	0.6008	.6303

Notes. The data is computed using the full count agricultural census for years 1970, 1979, 1988, 2000 and 2010. The average farm size and total agricultural land are both computed in hectares, and the total agricultural area for France corresponds to the sum of all total used agricultural areas at the farm level. Cultivated land for row crop farms is computed for the largest set of row crops that I can track across the years of the Census: wheat, durum, barley, oats, rye, corn, sorghum, rice, beetroot, rapeseed, sunflower and soy. The Gini-Coefficient is computed at the country level, for the distribution of farm size.

Table A4: Crop Mixes within French Farms

Year	Only Wheat	Non-Wheat Only	Wheat +	Other +	Only Barley	Only Corn	Only Oats	Only Rye	Only Sunflower	Only Colza
1979	0.0414	0.1856	0.6380	0.1354	0.0476	0.1148	0.0106	0.0071	0.0002	0.0004
1988	0.0938	0.1788	0.6351	0.0925	0.0542	0.0850	0.0136	0.0107	0.0058	0.0023
2000	0.1294	0.1855	0.6197	0.0659	0.0454	0.0960	0.0126	0.0051	0.0090	0.0022
2010	0.1425	0.1690	0.6259	0.0629	0.0462	0.0787	0.0078	0.0057	0.0078	0.0042

Notes. The data is computed using the full count agricultural census for years 1970, 1979, 1988, 2000 and 2010.

Table A5: Farm-Level Shift Share Results

Dependent Variables:	ΔPrices (log)	ΔSales (log)	ΔProfit (log)	ΔChemicals (log)
Model:	(1)	(2)	(3)	(4)
Variables				
$Exposure_j \times 1985$	0.0015	0.0384	0.0314	0.0541
	(0.2703)	(0.0747)	(0.1376)	(0.1292)
$Exposure_j \times 1986$	0.5111	0.0432	0.0488	0.0347
	(0.3910)	(0.0735)	(0.1143)	(0.1112)
$Exposure_j \times 1987$	0.0423	0.0419	0.1816	-0.1719^{**}
	(0.1017)	(0.0705)	(0.1975)	(0.0808)
$Exposure_j \times 1988$	-0.2572	0.0271	0.1921^{*}	0.0900
	(0.1591)	(0.0643)	(0.1156)	(0.0661)
$Exposure_j \times 1989$	-0.0289	-0.0579	-0.0903	-0.0718
	(0.1269)	(0.0777)	(0.1867)	(0.0807)
$Exposure_j \times 1990$	-0.0075	-0.0964^{*}	-0.1410	0.1147^{**}
	(0.0909)	(0.0584)	(0.1390)	(0.0583)
$Exposure_j \times 1992$	-0.3697^{***}	-0.1161^{***}	-0.3695^{***}	-0.0668
	(0.0994)	(0.0428)	(0.1382)	(0.0792)
$Exposure_j \times 1993$	-0.3816^{**}	-0.1614^{*}	-0.2110	-0.0783
	(0.1748)	(0.0847)	(0.1323)	(0.1203)
$Exposure_j \times 1994$	-0.2999^{**}	-0.0666	-0.1612	-0.1278
	(0.1250)	(0.0899)	(0.2240)	(0.1504)
$Exposure_j \times 1995$	-0.3718^{***}	-0.1477	-0.2792^{**}	-0.1212
	(0.1366)	(0.1370)	(0.1131)	(0.1495)
$Exposure_j \times 1996$	-0.5969^{***}	-0.1625	-0.2450	-0.1689
	(0.1328)	(0.1239)	(0.1845)	(0.1381)
$Exposure_j \times 1997$	-0.3853^{***}	-0.1167	-0.0347	-0.0637
	(0.1148)	(0.0995)	(0.1746)	(0.2166)
$Exposure_j \times 1998$	-0.4490^{***}	-0.0373	-0.1793	-0.1748
	(0.1506)	(0.1398)	(0.1378)	(0.1111)
$Exposure_j \times 1999$	-0.4658^{***}	-0.2249^{**}	-0.1714	-0.2583^{**}
	(0.0546)	(0.0889)	(0.2021)	(0.1044)
$Exposure_j \times 2000$	-0.6579^{***}	-0.3857^{**}	-0.3818^{***}	-0.2648^{*}
	(0.2135)	(0.1526)	(0.1189)	(0.1564)
$Exposure_j \times 2001$	-0.4307^{**}	-0.4491^{**}	-0.7017^{***}	-0.3099^{*}
	(0.1980)	(0.1928)	(0.2291)	(0.1868)
$Exposure_j \times 2002$	-0.5315^{**}	-0.4002^{**}	-0.3678^{**}	-0.1387
	(0.2504)	(0.1757)	(0.1699)	(0.1566)
Fixed-effects				
Department-Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	1,626	2,685	2,577	2,685
R^2	0.93118	0.72107	0.73446	0.66998

Clustered (Department-Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes. I show the results of our shift-share event-study design at the farm level, for our three outcomes of interest. These coefficients correspond to the year-specific coefficient associated to farm exposure. Outcomes are differenced with respect to their 1991 value. I control for the farms' 1991 total sales, capital stock, agricultural area and labor, allowing for a time-varying intercept for each of these controls. I add department-year fixed effects, and cluster the standard errors at the department-year level.

Table A6: Village-Level Shift Share Results

Dependent Variable:	ΔFarm Count			
Model:	(1)	(2)	(3)	(4)
Variables				
$Exposure_k$ (66th) × 1970	0.5932			
	(0.5231)			
$Exposure_k$ (66th) × 1979	0.7326**			
	(0.3269)			
$Exposure_k$ (66th) × 2000	-0.5148**			
	(0.2359)			
$Exposure_k$ (66th) × 2010	-0.6191**			
	(0.3085)			
$Exposure_k$ (mean) × 1970		-0.4966		
		(0.5695)		
$Exposure_k$ (mean) × 979		-0.1244		
		(0.3341)		
$Exposure_k$ (mean) × 2000		-0.4552*		
		(0.2711)		
$Exposure_k$ (mean) × 2010		-0.6845*		
		(0.3539)		
$Exposure_k$ (median) × 1970			0.2874	
			(0.6402)	
$Exposure_k$ (median) × 1979			0.4705	
			(0.3816)	
$Exposure_k$ (median) × 2000			-0.5998**	
			(0.3015)	
$Exposure_k$ (median) × 2010			-0.7791*	
			(0.4101)	
$Exposure_k$ (33rd) × 1970				-0.4275
				(0.6767)
$Exposure_k$ (33rd) × 1979				-0.1279
				(0.3836)
$Exposure_k$ (33rd) × 2000				-0.2878
				(0.3252)
$Exposure_k$ (33rd) × 2010				-0.4485
				(0.4204)
Fixed-effects				
Department-Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	147,157	147,157	147,157	147,157
R^2	0.80519	0.80505	0.80513	0.80501

Clustered (Department-Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes. I show the results of our shift-share event-study design at the municipality level. These coefficients correspond to the year-specific coefficient associated to farm exposure. Outcomes are first-differenced with respect to the level value in 1988, standard errors are clustered at the year-department level.

Table A7: Village-Level Shift Share Results

Dependent Variable:	ΔMin. Farm Size (ha)			
Model:	(1)	(2)	(3)	(4)
Variables				
$Exposure_k$ (max) × 1970	-0.2187			
	(0.1855)			
$Exposure_k$ (max) × 1979	-0.0180			
	(0.1788)			
$Exposure_k$ (max) × 2000	1.250^{***}			
	(0.4316)			
$Exposure_k$ (max) × 2010	0.5040			
	(0.4700)			
$Exposure_k$ (mean) × 1970		-0.1916		
		(0.1910)		
$Exposure_k$ (mean) × 1979		0.0050		
		(0.2077)		
$Exposure_k$ (mean) × 2000		0.9959^{*}		
		(0.5217)		
$Exposure_k$ (mean) × 2010		0.0440		
		(0.5762)		
$Exposure_k$ (median) × 1970			-0.1551	
			(0.1829)	
$Exposure_k$ (median) × 1979			-0.0479	
			(0.1739)	
$Exposure_k$ (median) × 2000			0.9383^{**}	
			(0.4051)	
$Exposure_k$ (median) × 2010			0.2179	
			(0.4707)	
$Exposure_k$ (min) × 1970				-0.1131
				(0.1566)
$Exposure_k$ (min) × 1979				0.0383
				(0.1646)
$Exposure_k$ (min) × 2000				0.6867
				(0.4663)
$Exposure_k$ (min) × 2010				-0.1898
				(0.5100)
Fixed-effects				
Department-Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	147,157	147,157	147,157	147,157
R^2	0.40543	0.40528	0.40530	0.40525

Clustered (Department-Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes. I show the results of our shift-share event-study design at the municipality level. These coefficients correspond to the year-specific coefficient associated to farm exposure. Outcomes are first-differenced with respect to the level value in 1988, standard errors are clustered at the year-department level.

Table A8: County-Level Shift Share Results

Dependent Variable:	ΔBloom (log)			
Model:	(1)	(2)	(3)	(4)
Variables				
$Exposure_c \times 1986$	-0.0125	-0.0027	-0.0222	-0.0112
	(0.0256)	(0.0289)	(0.0243)	(0.0310)
$Exposure_c \times 1987$	-0.0310	-0.0179	-0.0370*	-0.0306
	(0.0228)	(0.0251)	(0.0219)	(0.0280)
$Exposure_c \times 988$	-0.0182	-0.0216	-0.0179	-0.0199
	(0.0253)	(0.0273)	(0.0247)	(0.0294)
$Exposure_c \times 1989$	0.0231	0.0088	0.0340	0.0212
	(0.0278)	(0.0289)	(0.0269)	(0.0299)
$Exposure_c \times 1990$	-0.0089	-0.0265	-0.0003	-0.0221
	(0.0304)	(0.0316)	(0.0303)	(0.0346)
$Exposure_c \times 1992$	-0.0539	-0.0368	-0.0578	-0.0507
	(0.0347)	(0.0343)	(0.0354)	(0.0398)
$Exposure_c \times 1993$	-0.0164	-0.0100	-0.0246	-0.0211
	(0.0227)	(0.0258)	(0.0224)	(0.0285)
$Exposure_c \times 1994$	-0.0173	-0.0100	-0.0274	-0.0194
	(0.0254)	(0.0275)	(0.0240)	(0.0274)
$Exposure_c \times 1995$	-0.0266	-0.0324	-0.0213	-0.0308
	(0.0213)	(0.0236)	(0.0239)	(0.0243)
$Exposure_c \times 1996$	-0.0544**	-0.0519*	-0.0676**	-0.0716**
	(0.0267)	(0.0274)	(0.0270)	(0.0324)
$Exposure_c \times 1997$	-0.0698***	-0.0622**	-0.0724***	-0.0773***
	(0.0249)	(0.0277)	(0.0231)	(0.0286)
$Exposure_c \times 1998$	-0.0278	-0.0325	-0.0186	-0.0312
	(0.0206)	(0.0216)	(0.0217)	(0.0258)
$Exposure_c \times 1999$	-0.0717***	-0.0621**	-0.0722***	-0.0802**
	(0.0267)	(0.0293)	(0.0254)	(0.0312)
$Exposure_c \times 2000$	-0.1571***	-0.1451***	-0.1593***	-0.1663***
	(0.0486)	(0.0500)	(0.0445)	(0.0533)
Measure	Median	66th	33rd	Mean
Fixed-effects				
Department-Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	23,693	23,693	23,693	23,708
R^2	0.54261	0.54220	0.54282	0.54257

Clustered (Department-Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes. This table gives the results for our county-level event study. The outcome is differenced-out (in log terms) with respect to the 1991 value, and is a Landsat-5 based index of algal bloom intensity on the within-county water bodies. The regression includes a series of controls set to their level in 1988 within the county (last year of the Census prior to the reform), and interacted with a time-varying coefficients, as well as department-by-year fixed effects. Standard errors are clustered at the department-by-year level.

Table A9: Demand Results

Dependent Variables:	Wheat (log)	Others (log)	Wheat (log)	Others (log)
Model:	(1)	(2)	(3)	(4)
Variables				
Constant	21.03***	21.14***	20.67***	20.22***
	(0.4976)	(0.5401)	(0.8259)	(1.031)
Wheat Price (log)	-0.3507***		-0.2934*	
	(0.0868)		(0.1577)	
Other Price (log)		-0.3836***		-0.2391
		(0.0906)		(0.2039)
Time Trend	0.0087***	0.0101***	0.0107***	0.0154***
	(0.0023)	(0.0028)	(0.0029)	(0.0036)
Fit statistics				
Observations	42	42	25	25
R^2	0.81418	0.72615	0.27735	0.47193

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes. I obtain estimates of total quantity sold and average prices for France for our two crop categories using the FADN and its sampling weights. I then run regressions of sales on prices, instrumenting the prices with weather shocks in order which should shift supply but not demand. I include a linear time trend in all regressions, and use the Newey-West correction for serial auto-correlation. Regressions in column 3 and 4 use weather forecasts as additional instruments, and only use post 1993 years for estimation – matching both the post-reform period, and historical forecast availability.

Table A10: Model Parameters

Coefficient	Parameter	Value
Other Estimated Parameters		
Competence Ladder	λ	.80
Baseline TFP	μ	.58
Auto-Correlation TFP	ρ_h	.71
Shape TFP Pareto Distribution	σ_h	.12
Location TFP Pareto Distribution	δ_h	0
Auto-Correlation Non-Hicksian shock	ρ_{ch}	.83
Shape Non-Hicksian Pareto Distribution	σ_h	.16
Location Non-Hicksian Pareto Distribution	δ_h	0
Calibrated Parameters		
Time Preference	$\frac{1}{1-r}$	.9
Capital Depreciation	δ_k	.08
Price Capital	p_k	1
Price Land	p_s	1.77
Price Chemicals	p_c	4.15
Price Labor	p_l	.5

Notes: I use the recovered productivity shocks to estimate these other estimated parameters. Specifically, I demean TFP shocks, and use the average as an estimate of μ. Demeaned TFP shocks are used to compute the following parameters. The competence ladder corresponds to the average ratio of the highest crop TFP to lowest crop TFP within a farm, across farms and years. The Pareto distribution parameters are recovered across farms and years, and the auto-correlation in both shocks are computed using within-farm variation. For the additional calibrated parameters, the rate of time preference follows from the modelling choice of Scott (2013), the rate of capital depreciation is set to 8%, the price of capital is normalized to 1 – and such that one unit of capital accounts for a stock of 1,000€ in capital. The relative price of land, to chemicals to labor follows from estimates of per-hectare expenses in each category for wheat in France in Guillermet (2015), and their ratio to the cost of capital is normalized so that the average type farm in our model has the average observed labor to capital ratio in the FADN.

Table A11: Budget Equivalent Policy: Aggregate Outcomes

		Aggregate Statistics				
Land Subsidy	Lump-Sum	Mass of Farms	Average Profit	Average Size	Average. TFP	Average Chemical Productivity
0.005%	0	99.899	100.857	100.549	100.000	100.000
0.039%	0	99.132	101.273	104.502	100.000	100.000
0.134%	0	96.876	101.672	117.400	100.000	100.000
0.381%	0	89.525	101.971	171.736	100.000	100.000
0.600%	0	86.038	94.965	264.693	99.930	98.663
0	2.7%	126.708	80.149	78.910	98.274	92.744
0	4%	192.313	54.995	51.901	91.685	81.191
0	5.5%	419.749	29.059	23.694	82.696	70.416
0	6.8%	1123.900	16.408	8.774	75.457	64.323
0	8.2%	1942.900	14.261	4.999	73.555	62.962

Notes. The table details the aggregate consequences of the different policies. First different levels of land subsidies, and then different levels of lump sum payments to organic farms, each one computed to be budget equivalent to a land subsidy. Land subsidies are expressed in % of the price of land, and the lump sum as a percent of the marginal farm's profit under no-intervention. I express all the statistics wrt. to their no-intervention value, 100% implying no change.

Table A12: Budget Equivalent Policies: Crop-Level Outcomes

Panel A	Wheat Market								
Land Subsidy	Lump-Sum	Price	Supply	Cost	Land	Chemicals	Chemicals/Unit	Chemicals/Land	Cost/Unit
0.005%	0	99.893	100.031	87.514	100.438	99.652	99.625	99.217	100.078
0.039%	0	99.199	100.236	87.926	103.535	97.313	97.080	93.990	100.671
0.134%	0	97.135	100.888	89.338	113.491	90.607	89.821	79.836	102.658
0.381%	0	90.870	102.883	95.590	152.582	71.865	69.879	47.100	111.509
0.600%	0	83.698	105.360	108.900	224.191	53.224	50.521	23.741	130.044
0	2.7%	99.911	100.056	99.967	99.979	99.795	99.750	99.816	99.911
0	4%	99.662	100.099	99.761	99.787	99.367	99.291	99.579	99.662
0	5.5%	98.950	100.310	99.257	99.361	97.658	97.372	98.286	98.950
0	6.8%	97.918	100.650	98.555	98.581	98.147	97.539	99.559	97.918
0	8.2%	95.764	101.278	96.988	97.049	96.060	94.869	98.981	95.764
Panel B	Aggregated Crop Market								
0.005%	0	99.980	100.030	99.269	100.642	99.854	99.824	99.217	100.055
0.039%	0	99.788	100.022	99.551	104.892	98.588	98.567	93.990	100.379
0.134%	0	99.250	100.132	100.629	119.021	95.022	94.897	79.836	101.460
0.381%	0	97.686	100.529	104.893	179.257	84.429	83.986	47.100	105.786
0.600%	0	96.040	101.059	112.302	305.340	72.492	71.735	23.741	113.575
0	2.7%	99.989	99.973	99.962	100.119	99.947	99.974	99.816	99.990
0	4%	99.986	100.067	100.053	100.389	100.019	99.955	99.579	99.986
0	5.5%	100.121	99.959	100.079	101.472	99.941	99.981	98.286	100.121
0	6.8%	99.563	100.100	99.662	99.367	99.691	99.595	99.559	99.563
0	8,2%	98.838	100.275	99.110	98.812	99.140	98.871	98.981	98.839

Notes. The table details the crop-level consequences of the different policies. First different levels of land subsidies, and then different levels of lump sum payments to organic farms, each one computed to be budget equivalent to a land subsidy. Land subsidies are expressed in % of the price of land, and the lump sum as a percent of the marginal farm's profit under no-intervention. I express all the crop-level statistics, from the output price to the average cost per unit in percent terms of the no-intervention levels, 100% implying no change.

B Appendix

B.1 Data Description

B.1.1 Intervention Prices

For the purpose of this paper, and to the best of our knowledge, I created the first database of agricultural commodities' EU intervention prices covering the end of the 1960s to the early 2000s when intervention was completely removed. Intervention prices denominated in ECUs, euros or francs were gathered from the yearly and sometimes commodity-specific directives published by the European Union over that period of time. When prices were denominated in ECUs, I used the CAP-specific exchange rate which was then used to translate ECUs to each member state's currency. I similarly gathered data on commodity-specific land subsidies, which I later combine with farm-specific land subsidies as recorded in the FADN.

B.1.2 Input Prices: Indices and Price Surveys

Input prices are drawn from the IPAMPA agricultural input price indices, and from the Enquête sur l'observation des Prix des Consommations Intermédiaires nécessaires aux Agriculteurs (EPCIA), the price survey used to develop the IPAMPA. I use the IPAMPA indices for pesticides and fertilizers interpolated by INSEE back to 1980.

The EPCIA has been run since 1996, and provides price data for pesticides and fertilizers. Its sample is based on a 1995 survey, which helped identify a sample of products representing 50% of the total sales within their product category. The sampling of establishments selling these products was done based on the establishments' respective market share within that product category. Finally, within a product and establishment, the series chosen corresponded to the product's most frequent sales conditions. As such, series are good, company and mode of sale specific. When series disappear, they are replaced with their closest equivalent. The EPCIA is made of 4165 series, sold by 250 companies. The EPCIA is aggregated into national and regional price indices using 1990 plutocratic weights, and the following index of category-specific month-to-month price evolution:

$$i_t = i_{t-1}\left(\prod_{i}^{n} \frac{P_{i,t}}{P_{i,t-1}}\right)^{\frac{1}{n}} \tag{10}$$

I also recover agricultural hourly wages from the continuous labor survey "Enquête Emploi", which is a survey fielded for a sampled set of households, each household being drawn once and then followed for the next six trimesters. I consider the hourly wage data provided by workers whose occupation is classified as that of an agricultural worker or farmer.

Finally, I use land prices taken from the Land Market Value survey (Valeur Venale des Terres), a yearly and department-level survey, which is fielded every year by the statistical services of the French departmental administration for agriculture and forestry. These are based on data provided by the public company in charge of land management (SAFER), which authorizes agricultural land purchases and consolidations when transactions surpass a given threshold. This data is then complemented by data provided by local notaries, and several local administrations. I use this data for 1994-2015. The data was digitized from scanned data catalogues for the first years of the series.

B.1.3 Weather Data: Realizations and Forecasts

Our realized climate data comes from the European Centre for Medium-Range Weather Forecasts (ECMWF)'s ERA5 reanalysis product. ERA5 gives hourly estimates of climate data, out of which I use precipitation and temperature (temperature 2m above the surface of the Earth). ERA5 combines observational weather data with model-based data into a 0.25*0.25 gridded dataset.

I extract that data at the French department level, cropping the grid with department shapes, and averaging the data using simple area weights. Using the time separability assumption common in the literature, I aggregate the hourly data into growing-season observations [45]: growing degree days (GDD) and heat degree days (HDD) for temperature[46], and total precipitation for rainfall. In order to match the forecast dataset, I only use 4 daily measurements of temperature to compute the GDDs and HDDs, specifically at midnight, 6h, 12h and 18h.

The forecast data is taken from ECMWF's SEAS5 seasonal forecasting system. Forecast are produced on the first of each month for the following 5,160 hours [47]. For temperature, they are produced at a 6h interval, and give an instantaneous prediction of temperature, while for rainfall, they give the accumulation of rainfall every 24h. As such, the rainfall forecasts for the second day of the month will be the following: a 24h ahead forecast, a 30 to 32 days ahead forecast produced on the first day of the previous month, and so on until the lead value exceeds 5,160 hours. Temperature forecasts work in a similar way, but are just produced with more granular time steps.

The main issue for our purpose is that given that forecasts are produced every first of the month, different days within a month will not be provided a forecast with the same lead. I would ideally like to build the forecast-equivalents of our growing-season aggregates for realized weather, for different lead times. For example, the rainfall forecast for the growing season, with a constant one-month lead throughout the season (or the equivalent of the farmer's knowledge about rainfall one month in advance throughout the growing season). I approximate this by bundling together forecasts produced one calendar month ago (our one month lead for the rest of the paper), produced two calendar months ago, up to five

[45] I use an extensive definition of the agricultural growing season for France, running from October of the previous year, to July of the current one.

[46] Growing degree days are computed over the $[4°, 30°]$ degree interval, and heat degree days sum the realized temperature above 30°C.

[47] ECMWF provides an ensemble of 25 forecasts, which I average.

months ago. As such, the forecasts that I aggregate into growing season observations are not homogeneous in terms of lead value, but are the closest equivalent of it that I can get.

Similarly to the weather realization data, I extract the gridded data into department-level observations, using area weights.

B.1.4 Landsat 5 Remote Sensing Data

Both our remote sensing indices are based on Landsat 5 imagery. Landsat 5 was a low orbit satellite jointly managed by the US Geological Survey and NASA, and ran between 1984 and 2013. Its unusual longevity making it a good source of data to build long time series. It had a repeat cycle of about 16 days, and was equipped among else with a Thematic Mapper (TM) and a Multi-Spectral Scanner.

Our algal bloom index follows the methodology of Taylor and Heal (2023). For every year between 1985 and 2001, I filter Landsat data for images taken between June and August, which I crop for surface water and treat for clouds and cloud shadows. I then run the following function on each pixel:

$$Bloom = NIR - 1.03 * SWIR \tag{11}$$

NIR corresponds to the near infrared band of Landsat 5 (.77 to .90μm), and SWIR to the shortwave infrared band (1.55 to 1.75μm). In both cases, these come from atmospherically corrected surface reflectance data produced by the Landsat TM series. The index is then averaged across the selected months of the year, and averaged at the county level.[48]

Our measure of edge density follows recent remote sensing work developed for agricultural field delineation, and specifically the general procedure outlined in Watkins and van Niekerk (2019). I use the following workflow. I first create three composite images per year, respectively for the periods January-April, May-August and September-December. Clouds and cloud shadows are masked, and images are averaged within each period. I then use first a Gaussian filter, and then a Canny edge detection filter using four bands: red, blue, green and near infrared. I then aggregate the resulting edge image across bands, and across seasons within the year. Filtering per season prior to running the filter allows to account for varying landscape patterns within the year, and for example to better differentiate fields that might have different seasonal patterns but might be similarly green during some time of the year. I finally filter the images to remove human settlements, and compute the resulting average edge density per county. Very roughly, this procedure should give us an index of landscape fragmentation, and should correlate with the homogenisation of landscapes driven by increasing field size and an increased encroaching of agriculture on previously non-agricultural land.

[48] There are 2,054 counties in France, making them much smaller geographic entities than US counties, and the smallest geographic entity above villages (communes).

B.2 Rationale for our Chosen Environmental Focus

The environmental impacts of agricultural production are many and well documented – from the consequences of pesticide use on human health (Carson (1962),Dias et al. (2023),Taylor (2022),Missirian (2020),Frank (2021)), those of fertilizer use (Rossi et al. (2023) for a review), on groundwater depletion (Burlig et al. (2024),Ryan and Sudarshan (2021),Carleton et al. (2023)), as well as the likely consequences of agricultural intensification and local crop homogenization (Crossley et al. (2021)). There are also many ways in which changes in the allocation of production across heterogeneous farms towards the more efficient ones could impact the prevalence of environmental externalities on the overall agricultural market. When productivity differences are input-neutral, more efficient farms will have a higher input-efficiency for all inputs. If one thinks for example of water needs, a farm with a higher TFP will produce more from a set level of water use than others, reducing the water footprint of its products. Following models of multi-product firms such as Mayer et al. (2014), more efficient firms will also be more diverse, and have a higher number of crops per hectare than smaller ones – potentially benefiting local ecosystems. These relations change, however, when the source of productivity differences across farms stops being input-neutral. Studying intensive agriculture, one might actually think that some farms are more productive than others because of their higher reliance on irrigation, or because of their specialization in certain crops which could interact with scale effects, or facilitate the accumulation of crop-specific capital.

I focus here on a specific environmental externality – the chemical intensity of agricultural pollution. This choice is motivated by several reasons. The chemical intensity of production is readily observable in the production data, relevant to the French agricultural market, and has well documented consequences both for the local and global environment, and for public health. While irrigation is also easily observable, it is not relevant for France – at least in the 1990s – where the vast majority of row crop farms relied on rainfall for irrigation. Other externalities, such as the impact of field size, hedge density, or the clustering of fields within space cannot be measured in our data, and likely require additional research when it comes to their environmental impact. Finally, there is a natural way to model the heterogeneity in chemical intensity across farms which fits the general agro-science research focused on French agriculture in the 1990s. I discuss it below.

The 1992 reform found French agriculture in a state of high productivity growth, and amid a strong trend of intensification (Carpentier and Rainelli (1997),Meynard et al. (2003)). The late 1980s saw France as the second world consumer of pesticides, with an unusually high use of fungicides justified in part by the role of wine production, in part by its large culture model for the production of cereals and oil crops. Intensification then corresponded to the implementation of a series of practices among which: earlier sowing dates (specifically for wheat), higher sowing density, shorter crop rotations allowing for a more frequent cropping of wheat, the use of higher productivity cultivars, and a more intensive use of fertilizers and irrigation. This technological nexus was characterized by the joint use of different practices, and of a more intensive reliance on both pesticides and fertilizers. This complementarity between productivity gains and higher chemical intensity can readily be modeled as non-Hicksian productivity growth. Along with these changes, the late 1980s and the 1990 saw a rise of no-till agriculture with reduced or no ploughing. All of these techniques increased yields (or decreased costs for the case of no-till), but relied on a heavier use of pesticides, and more specifically fungicides, as they all increased the exposure of crops to diseases and pests. In addition, raising potential yields also increased the risks posed by damaging pest shocks, and so further justified a strict protection of fields (Meynard et al. (2003)). While the 1990s also saw the development of sustainability concerns for agriculture, the 1992 cutting of intervention prices is widely seen as having further pressured farmers to reduce their marginal costs of production (Pierre (2004)). In this specific context and for cereals and oil crops, this mainly entailed a further intensification of production, and a heavier reliance on pesticides and fertilizers. The pivot towards a conception of agriculture as multi-functional[49], which started in the 1990s, hence coincided with the most significant pro-competitive reform of European agricultural policy, which at the time presented little to no features to orient agriculture towards more sustainable means of production.

I discuss next some suggestive evidence surrounding the role of agriculture in driving chemical-led pollution in France, heterogeneity across farms in pollution, and some evidence backing the use of non-Hicksian shocks to model this heterogeneity.

B.2.1 Correlates of Agricultural Intensification

I start by providing suggestive evidence for the role of agricultural expansion between 1988-200 and algal blooms. Over the past years, the specific role of the CAP in the increase in eutrophication and algal blooms in Europe has been heavily discussed.[50]

I rely on a remote-sensing based index for algal blooms on French waterways and inland water bodies. The index is computed at the county-year level, using Landsat 5 imagery.[51] See subsection B.1 for a description of how this index is constructed. Algal blooms are generally driven by an over-fertilization of soil – and through leaching – of waterways. They are disruptive for local ecosystems and can produce toxins toxic to animals or even humans.[52] They can also deplete oxygen levels in water-bodies, and lead to so-called dead zones where no other organism can survive. Current work is also studying their potential impact in terms of increased N_2O emissions (Rossi et al. (2023)).

I look at the 1988-2000 evolution of county-level total agricultural area, and how it correlates with yearly percentage changes in our remote sensing index.[53] I do so controlling

[49]By this I mean the understanding of agriculture both as an industry aimed at producing quality food at low prices, but also as a central institution for the management of land-use, for the protection of ecosystems, and for the development of rural spaces.

[50]See https://www.nytimes.com/interactive/2019/12/25/world/europe/farms-environment.html

[51]France has over 2,000 counties, making them relatively small geographic entities, and the smallest ones above the village (commune) one. French counties are much smaller than their average US counter-part.

[52]Nitrate pollution can be harmful to humans, notably through the blue baby syndrome. In France, animals and three persons are also believed to have died over the last decades due to direct exposure to decomposing algae, and hydrogen sulphide, produced by algal blooms.

[53]I show in the annex Figure A7 a map for the evolution at the village level of total agricultural area. There are clear

for region-by-year fixed effects to further isolate the effect of agricultural growth. I see that there is a strong positive relation locally between the expansion of agriculture and increases in algal blooms. This pattern matches the one observed in the US by Rossi et al. (2023) between agricultural land expansion and fertilizer pollution.

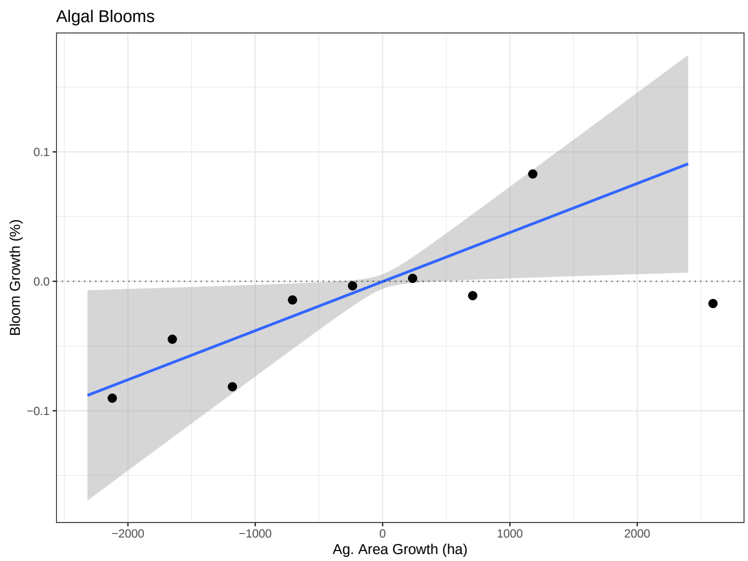


Figure A17: Correlates of Agricultural Area Expansion

Notes: The two figures show the correlation between the annual growth in remote sensing based indices and the 1988-200 growth of total agricultural land use (in ha) at the county level and as seen in the French agricultural Census, after controlling for region-by-year fixed effects. The blue line corresponds to a GAM model, with the associated standard error in grey. The dots represent mean values for each bin of the joint distribution of the variables.

I take this result as indicative of the potential disruptive impact of local agricultural growth, and hence for the potential impact that large reallocations in agricultural production can have on ecosystems.

B.2.2 Heterogeneity in Pollution: Larger Farms are More Chemical-Intensive

Knowing that agricultural activity comes with soil and water pollution brings the important question of whether establishments show some heterogeneity in their propensity to pollute. Whether large farms pollute more or less than smaller farms is theoretically ambiguous – and crucially depends on the structure of production.

Below I show binscatters obtained using FADN data, for resp. the fertilizer-to-land and pesticides-to-land ratios of production across the distribution of farm size measured as total sales value. I observe a clear positive and linear relation between farm size and these two ratios. I focus on the chemical-to-land ratio here, and not the chemical-to-output ratio, as the second one will respond both to potential heterogeneity in non-hicksian productivity,

geographic patterns of local increases or decreases in total agricultural area – the center of France and the lower part of Brittany and Vendée losing agricultural area.

but also scale choices, capital accumulation and heterogeneity in TFP. Under fairly general choices of modelization for production, the chemical-to-land ratio will be orthogonal to these three additional factors and respond only to non-hicksian productivity differences.

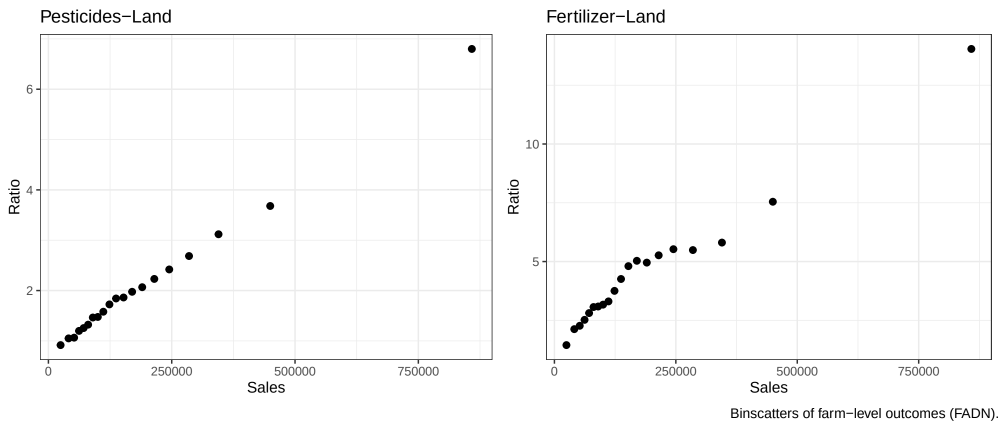


Figure A18: Chemical-to-Land Ratios in relation with Farm Size

Notes: I show in these two figures the relation between resp. the fertilizer-to-land and pesticide-to-land ratios and farm size, measured by farm-level sales in 2020 euros. Both fertilizers and pesticides are measured as deflated input bills using the INSEE input-specific price indices, while land is measured as total land used for agriculture (here in ares for easier-to-read ratios). The data is taken from the FADN. The binscatters show a clear linear positive relation between the chemical intensity of farming and farm size.

The chemical-intensity to size relation is of course a complex one – and this correlation could be driven by different mechanisms. I next argue for the presence of non-Hicksian productivity growth in the French row crop market. This productivity – under substitutability between land and chemicals – can explain that establishments which are more efficient at using chemicals have both higher chemical-to-land ratios than others, and end up being larger.

B.3 Robustness and Additional Results for the Reduced Form Analysis

I start this robustness section by showing the Rotemberg weights for our shift-share instrument, following the formula provided by Goldsmith-Pinkham et al. (2020), and additionally show the correlation between shock size, the Rotemberg weights, and the variance of the shares across farms. I compare them to the weights obtained when including all our controls – and specifically how the inclusion of the farms' 1991 relative share of land allocated to oil crops changes the variation used for identification. I then provide a more in-depth analysis of our conditional parallel-trends assumption driving our difference-in-difference design. I show trends in the same set of variables as shown in Figure A11, but for farms with a higher or lower crop shares – and this for the main crops that drive the variation in our shift-share instrument – as given by the value of the Rotemberg weights. Finally, I show the event studies for the farms' value added – compare it to the effects on prices and on total profit, provide the complete set of results in a difference-in-difference format, and give the shock-by-shock regression for value added and profits.

B.3.1 Rotemberg Weights

I give the Rotemberg weights below – with and without the inclusion of the controls used in our farm-level regressions. This has the advantage of both showing what cross-crop variation is used for identification, and how this variation is influenced by the controls.

Table A13: Rotemberg Weights: With Controls

Panel A	Rotemberg Weights		
Crop	Weight	Shock	Share Variance
Sunflower	0.5571	533.7	0.1171
Corn	0.2271	150.2	0.0307
Wheat	0.1641	163.7	0.0829
Barley - Spring	0.0526	127.2	0.0150
Oats	0.0134	120.3	0.0236
Durum	0.0098	165.4	0.0057
Sorghum	0.0025	158.7	0.0104
Colza	-0.0058	357.3	0.0000
Barley - Winter	-0.0063	131.4	0.0336
Rye	-0.0147	114.3	0.0105
Panel B	Correlation Matrix		
Weights	1	0.7338	0.8765
Shocks	0.7338	1	0.5841
Variance Shares	0.8765	0.5841	1

Notes. The weights are computed after residualizing the instrument on the fixed effects used in our farm-level regressions. These Rotemberg weights are obtained when using only land shares to measure farm exposure. The shock column indicates the value of the crop-specific shock, and the share variance corresponds to the variance observed in the Fadn sample of the crop-specific share. The correlation matrix gives the correlation between the value of the shocks, the weights implicitly used in the shift-share summation, and the variance in crop share across farms. Higher correlations means the instrument better captures the variation of the reform, and the cross-sectional variation in exposure across farms.

As I can see, the weights vary in a quite striking way as I include the 1991 relative oil share. Most importantly I move from an instrument which is essentially a difference-in-difference with a continuous treatment for the share of land allocated to sunflower in 1991, to one comparing farms growing relatively more sunflower, durum, wheat and corn. Wheat and corn being the most planted crops in France, this seems like a source of variation yielding more representative results. I note however that some negative weights also appear in the process, specifically for colza, which could raise issues in the case of heterogeneous treatment effects. I will investigate this later with a crop-by-crop iv strategy.

Table A14: Rotemberg Weights: Without Controls

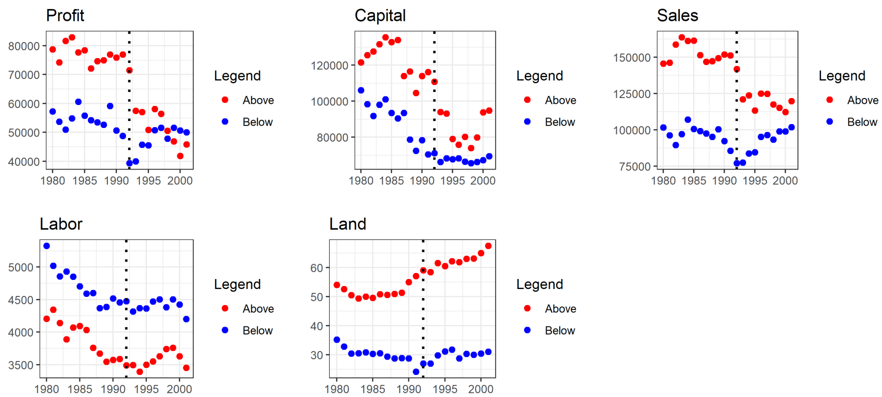
Panel A	Rotemberg Weights		
Crop	Weight	Shock	Share Variance
Sunflower	0.6277	533.7	1.171e-01
Corn	0.1355	150.2	3.073e-02
Wheat	0.0946	163.7	8.299e-02
Colza	0.0616	357.3	1.508e-02
Durum	0.0405	165.4	2.367e-02
Rye	0.0152	114.3	5.732e-03
Barley - Winter	0.0102	131.4	1.049e-02
Barley - Spring	0.0096	127.2	9.583e-05
Oats	0.0019	120.3	3.365e-02
Sorghum	0.0028	158.7	1.050e-02
Panel B	Correlation Matrix		
Weights	1	0.8669	0.8446
Shocks	0.8669	1	0.6713
Variance Shares	0.8446	0.6713	1

Notes. The weights are computed after residualizing the instrument on the fixed effects used in our farm-level regressions. These Rotemberg weights are obtained when using only land shares to measure farm exposure. The shock column indicates the value of the crop-specific shock, and the share variance corresponds to the variance observed in the Fadn sample of the crop-specific share. The correlation matrix gives the correlation between the value of the shocks, the weights implicitly used in the shift-share summation, and the variance in crop share across farms. Higher correlations means the instrument better captures the variation of the reform, and the cross-sectional variation in exposure across farms.

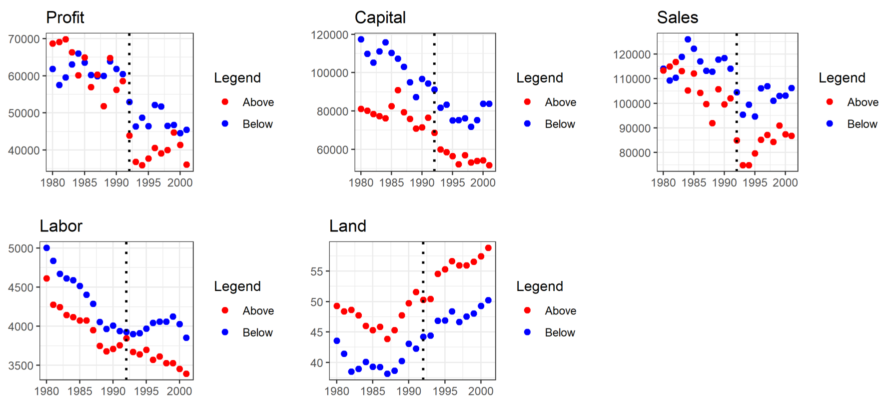
B.3.2 Balance Tests: Across Crop Shares

Next, I show our balance tests across farms with an above and below median share of wheat, corn, winter barley, rye, sunflower and colza. when looking at these comparison, one should keep two things in mind. The first thing is that everything else equal, I expect larger farms to have lower crop shares than smaller ones, as they tend to be more diverse – and this to be especially the case for the most common crops like wheat or corn. The second, is that this heterogeneity in crop shares should otherwise correlate with local production structures

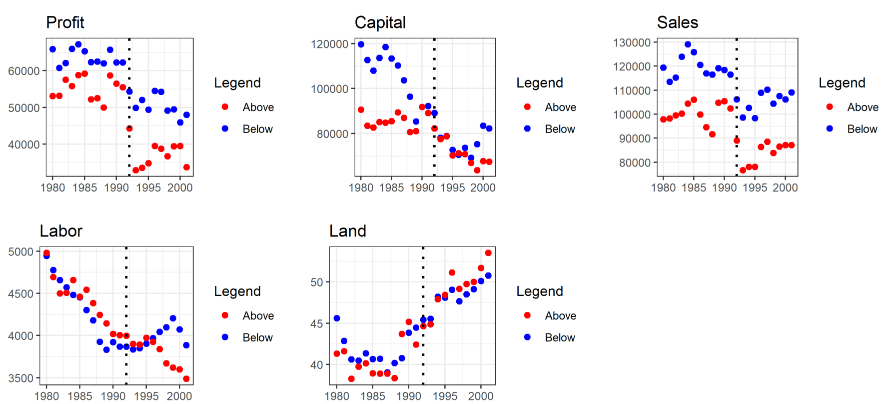
– Pierre (2004) describes the difficulties of starting to grow a given crops when the local distribution system is not already in place – and it should also correlate with local climatic and soil conditions. The graphs are then useful in order to understand the extent to which these factors can change the levels and dynamics across farms. I note that because I look at impacts on outcome growth rates, rather than outcome levels – differences in levels are not worrisome in our context. The main threat to identification lies with the presence of unobservables that are linked with heterogeneity in exposure, and which affect farm-level growth. The following elements stand out: across all crops, capital is the one variable that differs the most in trends across groups prior to the reform – making it a necessary control in our design. Profit, sales and labor show relatively comparable trends prior. Trends in land use also show some notable differences in the cases of rye and colza. While the differences for rye are quite significant, I also know from the previous Rotemberg weights that rye plays an almost absent role in the variation of the shift-share instrument, and hence should not drive results too much. Differences in colza are potentially more important, again justifying the control for land use pre-reform at the farm-level.



(a) Balance Tests - Wheat



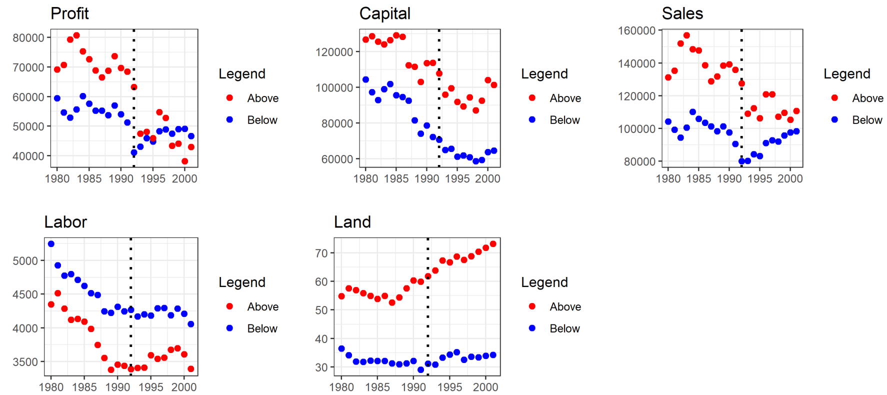
(b) Balance Tests - Sunflower



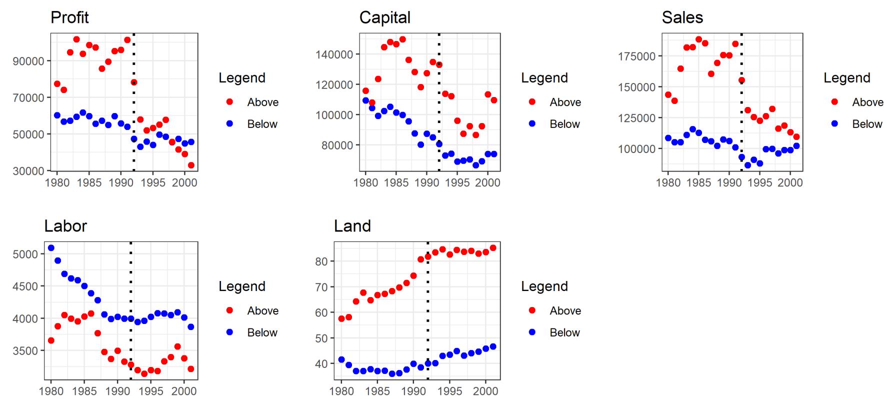
(c) Balance Tests - Corn

Figure A19: Trends per Category

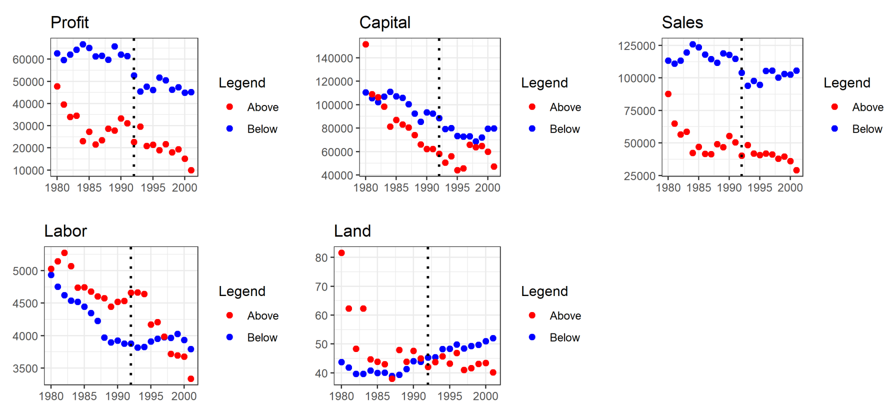
Notes: I show the trends of relevant farm characteristics for the bottom and top third of farms in terms of their share of land allocated to the relevant crop in 1991. I use FADN data for this exercise.



(a) Balance Tests - Barley



(b) Balance Tests - Colza



(c) Balance Tests - Rye

Figure A20: Trends per Category (continued)

Notes: I show the trends of relevant farm characteristics for the bottom and top third of farms in terms of their share of land allocated to the relevant crop in 1991. I use FADN data for this exercise.

B.3.3 Additional Farm-Level Results

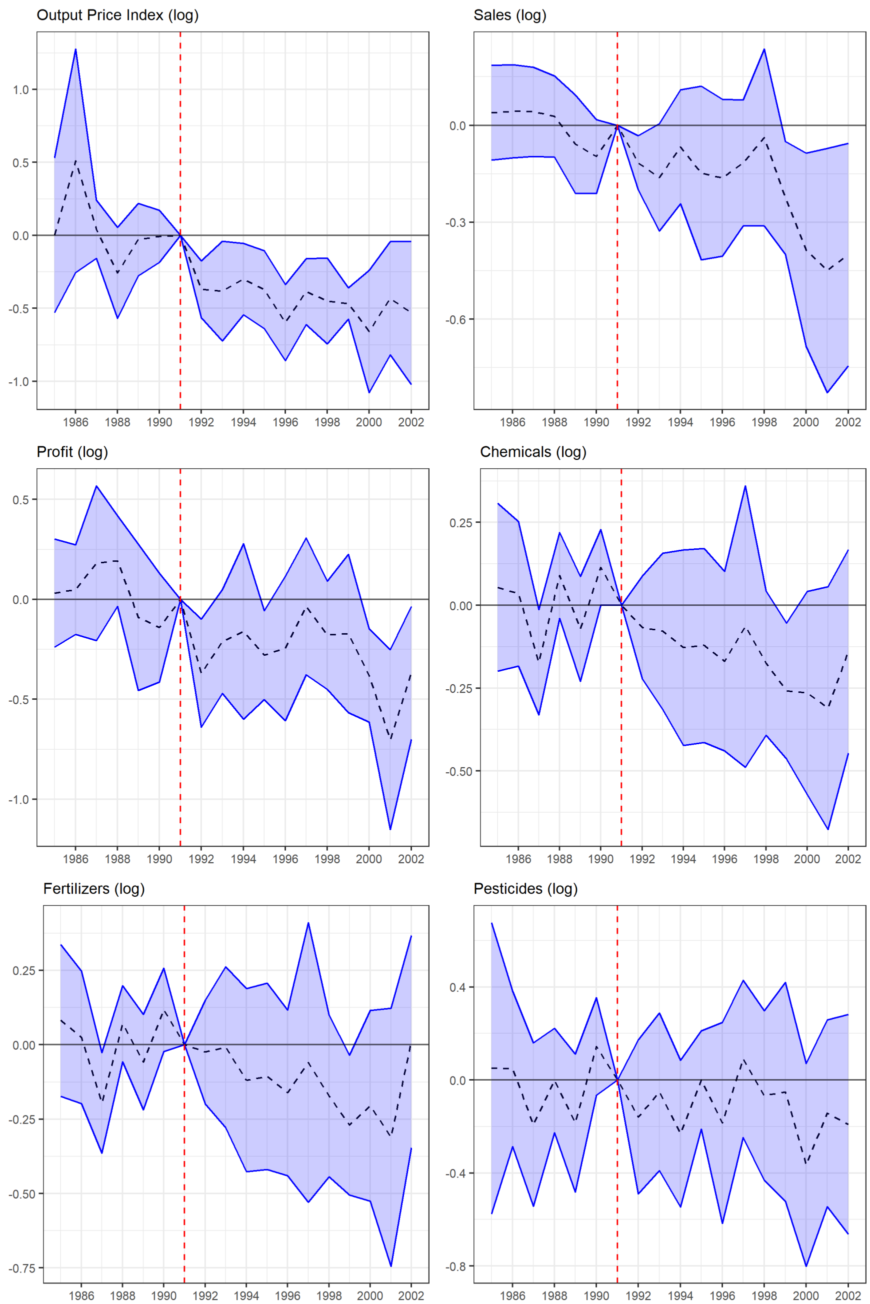


Figure A21: Farm-Level Event Study: Full Set of results

Notes:I plot the result of the event studies for the following outcomes – all measured as differences in logs with respect to the farm's baseline 1991 level: a farm-level output price index (average price across row crops, using relative areas as weights), total subsidies received, profit measured as value added plus subsidies, the farms' deflated chemical bill, the deflated bills for fertilizers and for pesticides. All coefficients correspond to the year-specific coefficient associated to farm exposure, and give the effect of exposure in that given year relative to the effect in 1991. I control for observed characteristics of the farms pre-reform, and allow for a time-varying intercept for each of these controls. I add department-year fixed effects, and cluster the standard errors one-way at the department-year level.

It is useful to recast our design in a pure difference-in-difference design, in order to get a unique coefficient summarizing the effect of the reform. I can do so using a modified estimating regression of the following form:

$$\Delta^{1991} Y_{jt} = Exp_j \beta^0 + Exp_j Post_t \beta^1 + X_j \Gamma^0 + X_j Post_t \Gamma^1 + \eta_{r(j)t} + \varepsilon_{jt} \tag{12}$$

The implied difference is that Exp_j now follows a binary form of heterogeneity, a unique effect pre-reform until 1991, and a unique effect post 1992. I also run this with an alternative construction of our exposure instrument, where price intervention shocks are averaged across crops using output-based weights, and land subsidy ones using land weights. I obtain the following results which match the ones obtained using our event study design. Results across the two forms of instruments are also comparable.

Table A15: Difference-in-Difference Results

Dependent Variables:	ΔPrice (log)	ΔSales (log)	ΔProfit (log)	ΔChemicals (log)	ΔFertilizers (log)	ΔPesticides (log)
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
$Exposure_j$	-0.0512	0.0273	0.0789	0.0142	0.0136	-0.0108
	(0.0783)	(0.0295)	(0.0627)	(0.0520)	(0.0548)	(0.0942)
$Exposure_j \times Post_t$	-0.3971***	-0.2477***	-0.3818***	-0.1927***	-0.1552**	-0.1129
	(0.0926)	(0.0575)	(0.0872)	(0.0713)	(0.0779)	(0.1138)
Mean Level	223.7	11.29	49,982.3	17,031.6	12,876.1	4,148.5
Fixed-effects						
Department-Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	1,626	2,685	2,577	2,685	2,685	2,685
R^2	0.91390	0.67309	0.69452	0.61699	0.60736	0.63383

Clustered (Department-Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes. I show the difference-in-difference results on farm-level: prices, sales, profit, chemicals, fertilizers and pesticides. The output price index is computed as a weighted average of the farm's output prices, using the crops' relative output as weights. These coefficients correspond to the year-specific coefficient associated to farm exposure. All variables are differenced with respect to their 1991 values, in log term, and hence represent a relative growth rate. I control for the farm's 1991 stock of capital, total labor, total land, total fertilizer and total pesticide use, their fertilizer-to-land and pesticides-to-land ratio, the number of crops they grow, as well as the share of oil crops in their output. I also control for their 1983-1984 adoption trends in chemicals measured as the evolution of their total chemical use. I allow for a time-varying intercept for each of these controls. I add department-year fixed effects, and cluster the standard errors at the department-year level.

Table A16: Difference-in-Difference Results (Alternative Instrument)

Dependent Variables:	ΔPrice (log)	ΔSales (log)	ΔProfit (log)	ΔChemicals (log)	ΔFertilizers (log)	ΔPesticides (log)
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
$Exposure_j$ (alt.)	-0.0435	0.0324	0.0886	0.0199	0.0237	0.0018
	(0.0590)	(0.0299)	(0.0618)	(0.0496)	(0.0529)	(0.0914)
$Exposure_j$ (alt.) $\times$ $Post_t$	-0.2921***	-0.2488***	-0.3825***	-0.1832***	-0.1538**	-0.1245
	(0.0698)	(0.0571)	(0.0867)	(0.0691)	(0.0761)	(0.1116)
Fixed-effects						
Department-Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	1,626	2,685	2,577	2,685	2,685	2,685
R^2	0.91411	0.67266	0.69418	0.61583	0.60690	0.63380

Clustered (Department-Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes. I show the difference-in-difference results on farm-level: prices, sales, profit, chemicals, fertilizers and pesticides. The output price index is computed as a weighted average of the farm's output prices, using the crops' relative output as weights. These coefficients correspond to the year-specific coefficient associated to farm exposure. All variables are differenced with respect to their 1991 values, in log term, and hence represent a relative growth rate. I control for the farm's 1991 stock of capital, total labor, total land, total fertilizer and total pesticide use, their fertilizer-to-land and pesticides-to-land ratio, the number of crops they grow, as well as the share of oil crops in their output. I also control for their 1983-1984 adoption trends in chemicals measured as the evolution of their total chemical use. I allow for a time-varying intercept for each of these controls. I add department-year fixed effects, and cluster the standard errors at the department-year level.

B.3.4 Independent Instruments

Finally, I decompose the shift share instrument into crop-specific instruments. Each crop-specific instrument corresponds to the crop-specific change in subsidization between 1991 and 1995, interacted by the land share allocated to that crop by the farm in 1991. The regression is otherwise similar to the difference-in-difference regressions presented previously. I only show the coefficients for the instruments interacted with the post-1991 dummy. These indicate the effect of crop-specific exposure to the reform after the reform happened. I see that all crop-specific instruments but sorghum have negative effect on sales, and almost all of these are statistically significant. Sorghum is similarly the one coefficient with a positive effect for profit, and for chemical use. This means that the shift-share instrument hides some degree of heterogeneity in terms of the effects of exposure. I note however, that sorghum production in France corresponds to about 1.5% of the production of wheat, and hence does not correspond to a very significant share of the market.

Table A17: Heterogeneity in Crop-Specific Effects

Dependent Variables:	ΔSales (log)	ΔProfit (log)	ΔChemicals (log)
Model:	(1)	(2)	(3)
Variables			
$Wheat_j \times Post_t$	-0.0026***	-0.0032***	-0.0021**
	(0.0008)	(0.0010)	(0.0008)
$BarleyWinter_j \times Post_t$	-0.0003	-0.0013	-0.0021
	(0.0012)	(0.0020)	(0.0013)
$Corn_j \times Post_t$	-0.0030***	-0.0045***	-0.0026***
	(0.0008)	(0.0013)	(0.0010)
$Rye_j\text{st} \times Post_t$	-0.0050***	0.0003	-0.0022
	(0.0015)	(0.0022)	(0.0021)
$Sunflower_j \times Post_t$	-0.0033***	-0.0041***	-0.0004
	(0.0010)	(0.0015)	(0.0011)
$Colza_j \times Post_t$	-0.0046***	-0.0048**	-0.0038***
	(0.0012)	(0.0021)	(0.0013)
$Durum_j \times Post_t$	-0.0117***	-0.0300***	-0.0130***
	(0.0031)	(0.0102)	(0.0029)
$Sorghum_j \times Post_t$	0.0486***	0.0675***	-0.0051
	(0.0100)	(0.0216)	(0.0106)
$BarleySpring_j \times Post_t$	-0.0008	-0.0032**	-0.0018
	(0.0008)	(0.0013)	(0.0014)
$Oats_j \times Post_t$	-0.0019*	-0.0049**	0.0020
	(0.0011)	(0.0023)	(0.0014)
Fixed-effects			
Department-Year	Yes	Yes	Yes
Fit statistics			
Observations	2,685	2,577	2,685
R^2	0.70244	0.71366	0.62818

Clustered (Department-Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes. I show the heterogeneity of our difference-in-difference results across crop-specific shocks. The instrument is decomposed in multiple crop-specific instruments. I include the crop shares in levels, but do not show the coefficients in the table. I control for the farm's 1991 stock of capital, total labor, total land, total fertilizer and total pesticide use, as well as their relative share of land allocated to oil crops. I also control for their 1983-1984 adoption trends in chemicals measured as their evolution of chemical-to-land and chemical-to-output ratio growth. I allow for a time-varying intercept for each of these controls. I add department-year fixed effects, and cluster one-way the standard errors at the region-year level.

B.3.5 Municipality-Level Results

In this section, I first present balance tests comparing municipality with relatively more and relatively less exposure to the reform, and then the distribution of the measure of exposure used for our municipality-level design. I end by showing our shift-share results using different aggregate measures to go form farm to municipality.

Starting with balance tests, I show below the trends (averages for each wave of the census) for four municipality characteristics, splitting them nationally between municipalities with a median exposure above or below the French median in 1988. Trends are overall similar, apart from the evenness variable. In our design, I control for both evenness and crop count measured at the municipality, and their average value across the farms in the municipality in 1988.

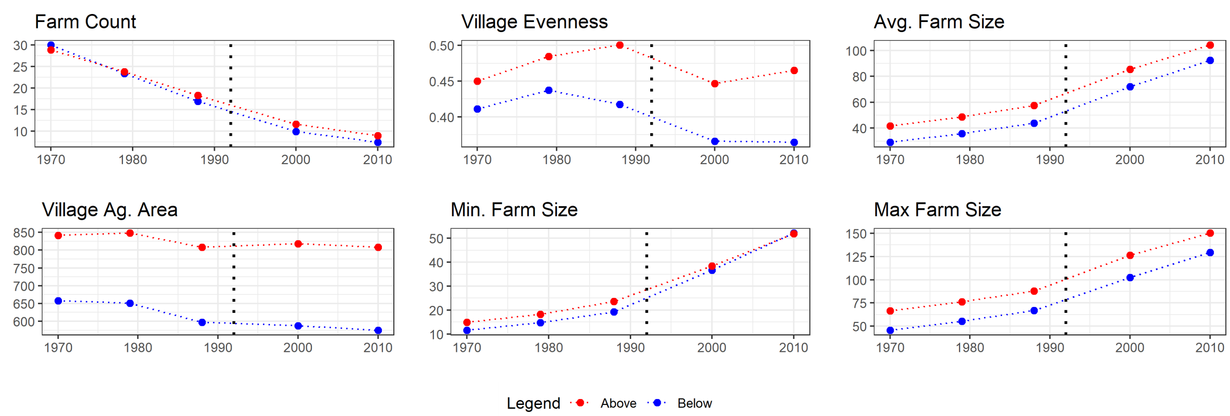


Figure A22: Municipality-Level Trends

Notes: This figure gives the evolution of municipality level outcomes over time, for the group of municipalities for which median exposure is above or below the median exposure in France as a whole.

Next, I show the distribution of median municipality exposure. As mentioned previously, all our measures of municipality-level are computed after winsorizing the farm-level measure of exposure in the Census for the bottom and top 1ppt values. Exposure is mostly distributed between 50 and 100€ per output unit, with a large right tail which likely corresponds to municipalities which have some of their land allocated either to colza, durum and sunflower, which are the crops with a shock value higher than 100 € per unit. These are mostly located in two areas of France, as shown in Figure A24, along the Mediterranean coast, as well as the main grain region of France around the Beauce region below Paris.

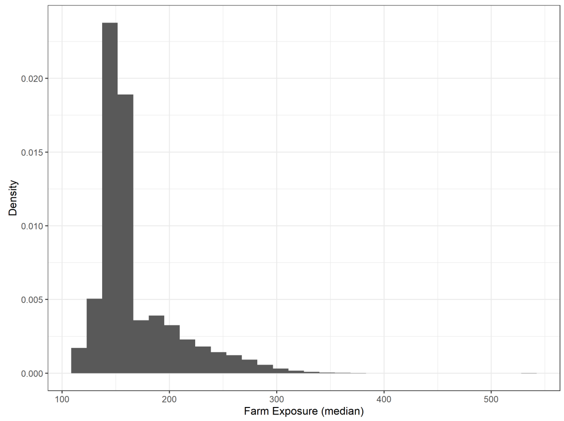


Figure A23: Distribution of Median Exposure

Notes: This figure gives the distribution of the municipality-level exposure to the MacSharry reform.

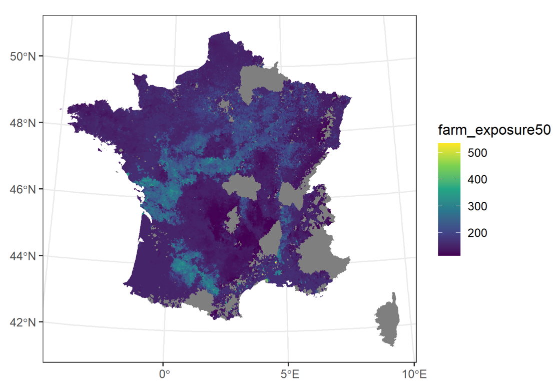


Figure A24: Geographic distribution of Median Exposure

Notes: This figure gives the distribution of the municipality-level exposure to the MacSharry reform.

Finally, I show our shift-share results using different aggregations of the farm-level exposure, resp. the median, 33rd centile, 66th centile, and the average value. I note that apart from the average, and as should be expected, the centile values are not sensitive to winsorizing the farm-level exposures. The different centiles show similar responses for the evolution of the minimum farm size, while farm exit really only responds to the median and 33rd centile. Because all our measures of exposure are standardized, this implies that a one standard deviation increase in exposure of the 66th centile does not have the same impact as a one standard deviation increase in the median or 33rd centiles. I further note that farm exit is a discrete decision which likely responds to a threshold value of exposure (given a farm's type). If these standard deviation increases do not map into similar shifts in exposure that lead to one farm within the municipality to cross its exit threshold, they will not have the same effect in our regression, which is likely what is happening here.

Finally, I note that average exposure has no effect on farm count, being flat around zero for all our periods, while the results on minimum farm size are of the same size but a much smaller magnitude. Standard deviations in average values are likely larger than standard deviations in centile values for centiles not located in the extremes of the tails, which could explain these differences in the levels of the coefficients.

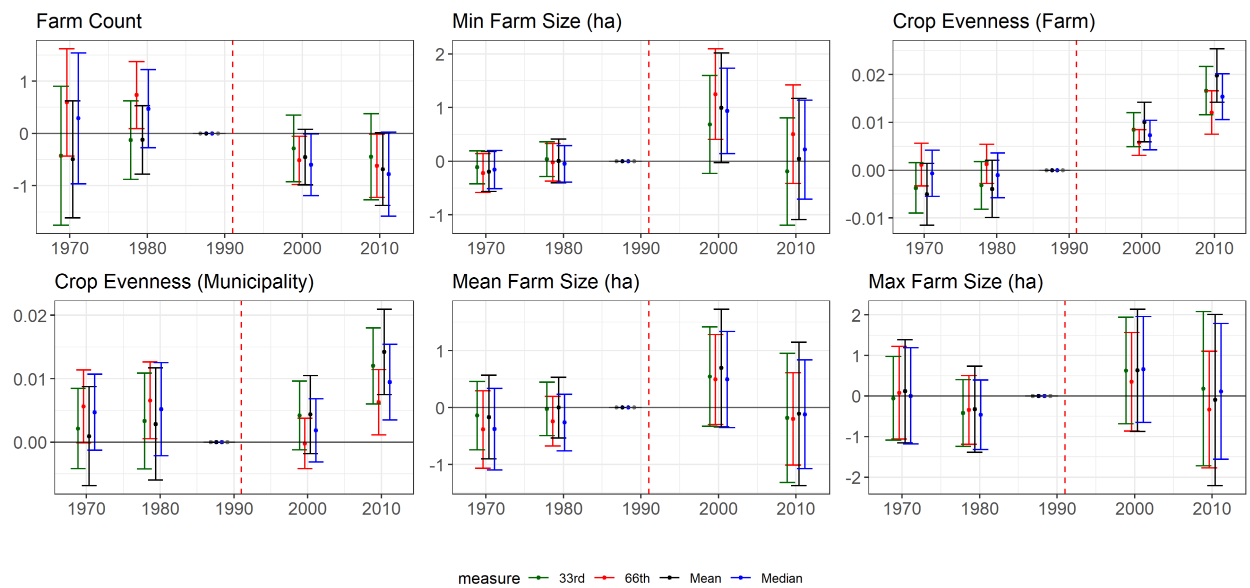


Figure A25: Municipality-Level Outcomes

Notes: This figure gives the results for our municipality-level event study. Outcomes are first-differenced. Farm count corresponds to the number of farms operating in the row crop market within the municipality, and min. size gives the evolution of the minimum farm size within the municipality. The regression includes a series of controls set to their level in 1988 within the municipality, and interacted with a time-varying coefficients, as well as department-by-year fixed effects. Standard errors are clustered at the department-by-year level. The table associated to these results is **??**.

B.3.6 County-Level Results

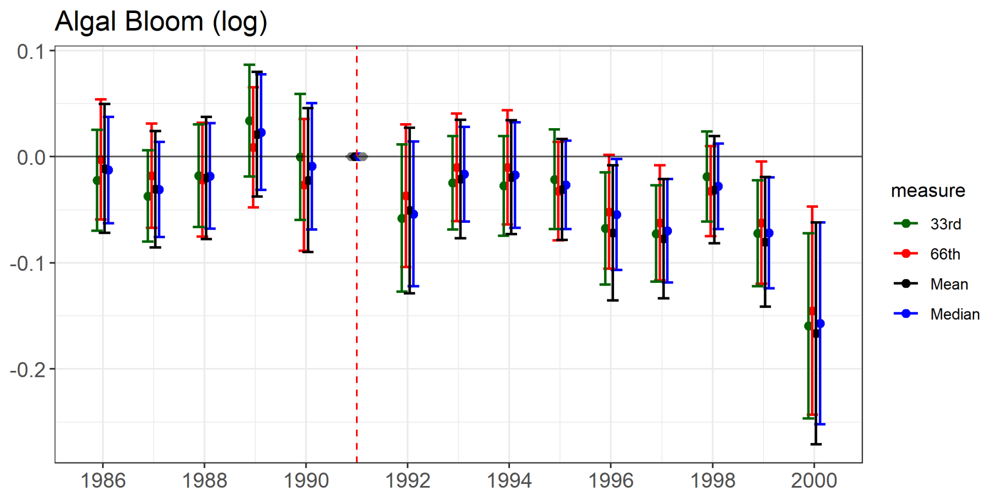


Figure A26: County-Level Algal Blooms

Notes: This figure gives the results for our county-level event study. The outcome is differenced-out (in log terms) with respect to the 1991 value, and is a Landsat-5 based index of algal bloom intensity on the within-county water bodies. The regression includes a series of controls set to their level in 1988 within the county (last year of the Census prior to the reform), and interacted with a time-varying coefficients, as well as department-by-year fixed effects. Standard errors are clustered at the department-by-year level. The table associated to these results are Table A8.

B.4 Model Details

I start with the Bellman decisions describing the farms' decisions – both incumbent and entrant farms – which correspond to the general outline of the model from section 4.1.2.

The incumbent's decision can be characterized by:

$$V(\Upsilon_{jt};\Omega_{jt}) = \max_{\{X_{jct}\}_{\mathbb{C}_{jt}},\xi^x_{jt},K_{jt+1},S^o_{jt+1},\mathbb{C}_{jt+1}} \underbrace{\Pi_{jt}}_{\text{Static Profit}} - \underbrace{C(K_{jt},K_{jt+1})}_{\text{Capital Adjustment}} - \underbrace{\left(S^o_{jt+1} - S^o_{jt}\right)P^s_t}_{\text{Land Purchase}}$$
$$+\, \xi^x_{jt}\underbrace{\left[(1-\delta_k)K_{jt}P^K_{jt} + S^o_{jt}P^s_t\right]}_{\text{Scrap Value if Exit}} + (1-\xi^x_{jt})\underbrace{\left(\beta\mathbb{E}\left[V(\Upsilon_{jt+1};\Omega_{jt+1})|\Upsilon_{jt},\Omega_{jt}\right] - f_k\right)}_{\text{Expected Value of Continuation}}$$

And I can further write static profit as:

$$\Pi_{jt} = \underbrace{S^O_{jt}P^S_{jt}}_{\text{Land Rental}} + \sum_{c\in\mathbb{C}_{jt}} \underbrace{\left[P_{jct}Q_{jct} - P^L_{jt}L_{jct} - (1-\tau^c_s)P^S_{jt}S_{jct} - P^F_{jt}F_{jct} - P^P_{jt}Pest_{jct}\right]}_{\text{Crop Specific Profit: } \pi_{jct}}$$

The entrant's problem takes the form:

$$V_e(\Omega_{jt}, f_t^e) = \max_{\{\xi_{jt}^e, \xi_{jt}^x, K_{jt+1}, S_{jt+1}^o, \mathbb{C}_{jt+1}} \mathbb{E}\Bigg[\xi_{jt}^e \Bigg\{ - f_t^e - \underbrace{C(0, K_{jt+1})}_{\text{Capital Adjustment}} - \underbrace{S_{jt+1}^o P_t^s}_{\text{Land Purchase}}$$

$$+ (1 - \xi_{jt}^x) \underbrace{\beta \mathbb{E}\Bigg[V(\Upsilon_{jt+1}; \Omega_{jt+1}) | \Upsilon_{jt}, \Omega_{jt}\Bigg]}_{\text{Expected Value of Continuation}} \Bigg\} | \Omega_{jt}, f_t^e \Bigg]$$

After I make our simplifying assumptions regarding the homogeneity of σ_j, the ownership of land and on input and output prices, I can further write the the transition operators which describe the evolution of entrant and incumbents farms as: $\phi_t(\omega', \omega^{ch'}, K', \mathbb{C}' | \omega, \omega^{ch}, K, \mathbb{C})$ and $\phi_{e,t}(\omega', \omega^{ch'}, K', \mathbb{C}' | \omega, \omega^{ch})$. They require the additional introduction of the transition operator $J^h(.|.)$ and $J^{ch}(.|.)$ for both shocks, and of the policy functions κ^k for incumbent capital investment, $\kappa^{\mathbb{C}}$ for crop choice, and κ_e^k for entrant capital investment. These transition operators now write as:

$$\begin{aligned} \phi_t(\omega', \omega^{ch'}, K', \mathbb{C}' | \omega, \omega^{ch}, K, \mathbb{C}) =& J^h(\omega' | \omega) J^{ch}(\omega^{ch'} | \omega^{ch}) 1\{K' = \kappa^k(\omega, \omega^{ch}, K | \Omega_t)\} * \\ & 1\{\mathbb{C}' = \kappa^{\mathbb{C}}(\omega, \omega^{ch}, K | \Omega_t)\} 1\{\xi^x(\omega, \omega^{ch}, K | \Omega_t) = 0\} \end{aligned}$$

I note that crop choice is a dynamic decision simply because of our timing assumption – which allows us to address the selection bias induced by focusing on single crop farms for estimation – but crop choice has no impact on a farm's transition. By this I mean that the current state of a farm's crop mix in a given period has no bearing on the dynamic decisions made by the farm concerning the next period, including its crop mix decision. The analogue transition operator for entrants writes $\phi_{e,t}(\omega', \omega^{ch'}, K' | \omega, \omega^{ch})$:

$$\begin{aligned} \phi_{e,t}(\omega', \omega^{ch'}, K', \mathbb{C}' | \omega, \omega^{ch}) =& J(\omega' | \omega) J^{ch}(\omega^{ch'} | \omega^{ch}) 1\{K' = \kappa_e^k(\omega, \omega^{ch}, 0 | \Omega_t)\} * \\ & 1\{\mathbb{C}' = \kappa^{\mathbb{C}}(\omega, \omega^{ch}, 0 | \Omega_t)\} 1\{\xi^x(\omega, \omega^{ch}, 0 | \Omega_t) = 0\} \end{aligned}$$

Once I have defined these transition operations, I can then define the law of motion for the measure μ_t. $G_1(.)$ and $G_2(.)$ are the distributions from which entrants draw resp. their original TFP and chemical productivity. With N_t^e the mass of entrants, the allocation of farms over the market space is then fully described by the measure μ_t, whose transition operation $T(.)$ can be written as:

$$\begin{aligned} \mu_{t+1}(\omega', \omega^{ch'}, K', \mathbb{C}') =& \int \phi_t(\omega', \omega^{ch'}, K', \mathbb{C}' | \omega, \omega^{ch}, K, \mathbb{C}) d\mu_t(\omega, \omega^{ch}, K, \mathbb{C}) + \\ & N_t^e \int \int \phi_{e,t}(\omega', \omega^{ch'}, K', \mathbb{C}' | \omega, \omega^{ch}) dG_1(\omega) dG_2(\omega^{ch}) \end{aligned}$$

With all these elements, I can finally define the stationary equilibria on which I will focus as follows. I look at stationary competitive equilibria composed of the tuple of $\Omega^\star = \{\mu^\star, N^{e,\star}, f^{e,\star}, \{P_c^\star\}_{\mathbb{C}}, \{P^x\}_x, P^K, Policy\}$ such that – for a given policy state and given input prices:

- $V^e(\Omega^\star) \leq 0$, and $V^e(\Omega^\star) = 0$ if $N^{e,\star} > 0$. (**Zero Ex-Ante Profit**)
- $\mu^\star = T(\Omega^\star)$. (**Stationary Market**)
- $P^\star$ s.t. $Q_c^S(\Omega^\star) = Q_c^D(P_c^\star)\ \ \forall c \in \mathbb{C}$. (**Output Market Clearing**)
- $f^{e,\star}$ s.t. $f^{e,\star} = Q^{e,\star}(\mathbb{M}(\mu^\star))^{-1}$. (**Entry Goods Market Clearing**)

B.5 Role of Chemical-Biased Productivity Shocks

When I combine the first order conditions related to optimal pesticide and land use in the maximization of each crop's flexible profit, I obtain the following relation – where small letters denote log terms. This relation highlights the role of the production structure in informing the prevalence of chemicals in production, and the drivers of this prevalence:

$$\begin{aligned}\rho\omega_{jt}^{ch} =& log\Big(\frac{P_{jt}^p}{P_{jt}^s}\Big) + log\Big(\frac{\delta_s^c}{\delta_p^c[1-\delta_s^c]}\Big) + (1-\rho_2)pest_{jct} - (1-\rho)s_{jct} \\ &+ (1-\frac{\rho}{\rho_2})log\Big[\delta_p^c Pest_{jct}^{\rho_2} + (1-\delta_p^c)Fert_{jct}^{\rho_2}\Big]\end{aligned}$$

It is useful to then consider the case where chemicals form a single input $Chemicals_{jct}$ in order to gain some intuition. In that case, this first equation can be re-written as:

$$log\Big(\frac{Chemicals_{jct}}{S_{jct}}\Big) = \frac{1}{1-\rho}\Big(\frac{P_{jt}^s}{P_{jt}^x}\Big) + \frac{1}{1-\rho}log\Big(\frac{1-\delta_s^c}{\delta_s^c}\Big) + \frac{\rho}{1-\rho}\omega_{jt}^{ch}$$

This equation makes the drivers of the chemical-to-land ratio very clear: a decreasing relative price of chemicals, the production of a more chemical-intensive crop, and – if land and chemicals are substitutes – a higher productivity at using chemicals: ω_{jt}^{ch}. While the sign of the first two relations is independent from the value of ρ, the influence of chemical productivity crucially depends on whether $\rho \in]0,1[$, or $\rho < 0$, ie. whether chemicals and land are substitutes or complements in the production process.[54] The MacSharry reform is likely to play a role along these three dimensions.

The relation between the chemical-to-output ratio (at the production optimum) and the chemical-biased productivity shock ω_{jt}^{ch} is also signed by the value of ρ. These simulations illustrate how the chemical-to-output ratio increases if $\rho \in]0,1[$.

[54] For a CES function to be appropriately defined, I indeed need $\rho<1$. This implies that $\frac{1}{1-\rho}$ is always positive.

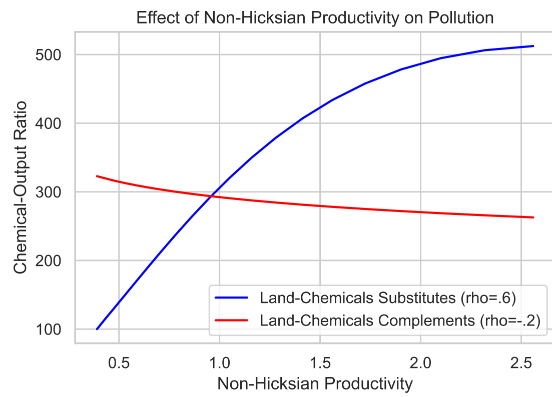


Figure A27: Relation Between Land-Chemicals Elasticity and non-Hicksian Productivity

Notes: This figure shows the effect of non-Hicksian productivity shocks on firms' chemical-to-output ratios at their optimal production level. I use simulations to shows the effect of changing the parameter of substitution between land and chemicals from land and chemicals being substitutes, to them being complements.

I finally get the following relation between farms' profitability and pollution intensity. In our competitive model, profitability is also a direct measure of efficiency, and hence I have a positive relation between farm efficiency and pollution intensity. To facilitate the aggregate at the farm-level, I show pollution intensity as the ratio of the amount of chemicals used over the value of production.

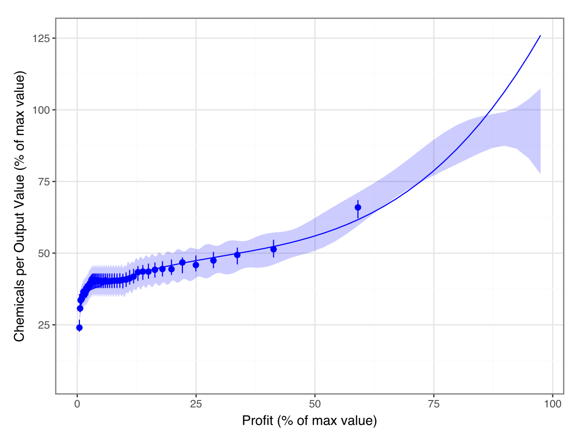


Figure A28: Relation Between Pollution Intensity and Profitability

Notes: This figure shows the relation between farm profitability and pollution intensity (measured as total chemicals used over the value of production) implied by our model parameters. I express both values in percents of their maximum values.

B.6 Joint Production Estimation Framework

Here I describe an alternative joint production framework, for which I also provide estimates. This approach combines the work of Dhyne et al. (2022) and of Doraszelski and Jomandreu (2018). From the first, I use the idea that inputs can be shared across product lines, and that the parametrization of a firm-level transformation function will define the degree of penalization that accompanies this sharing of inputs. I then draw from Doraszelski and Jomandreu (2018) to model and estimate non-hicksian shocks, within that multi-product production framework. Because the recovery of non-hicksian shocks relies heavily on the first-order conditions taken from the firm-level optimization problem, I move away slightly from the exact specification used by Dhyne et al. (2022).

In effect, consider a firm j which produces the set of crops $\mathbb{C}_{jt}$ within a production period t. The firm is endowed with product-specific input-neutral shocks $\tilde{\omega}_{jct}$. I write the transformation function as:

$$f_{jct} = exp(\tilde{\omega}_{jct}) g_{jct}^{\gamma_{cc}} \prod_{c' \in \mathbb{C}_{jt}, c' \neq c} g_{jc't}^{\gamma_{cc'}}$$

Here f_{jct} is the quantity of crop c produced. That quantity depends on a specific TFP shock, and on product-specific functions g_{jct} which are evaluated at the level of inputs used for each crop grown in the season. The parameters γ_{cc} parametrize the rivalry or complementarity in production from jointly producing c and c', and potentially sharing some inputs across these two product lines. I keep the functions g_{jct} fairly general for the moment, and parametrize them further later. One additional set of assumptions I make is that:

$$g_{jct} = g_c(K_{jt}, L_{jt}, Fert_{jt}, Pest_{jt}, S_{jct}, \omega_{jt}^p, \omega_{jt}^f)$$

This means that each function g_{jct} depends on the firm-level chosen amounts of capital K_{jt}, labor L_{jt}, fertilizer $Fert_{jt}$, pesticides $Pest_{jt}$, on a crop-specific chosen amount of land S_{jct}, and on firm-level pesticide and fertilizer productivity shocks that I denote ω_{jt}^p and ω_{jt}^f. This implies that all inputs but land are selected at the firm-level, and shared across product-lines according to the γ matrix. I also consider that the g_c function only varies across crops, but not periods or farms, and that these two non-hicksian shocks fully account for the heterogeneity in these functions across farms. This assumption of land being a crop-specific input, while others are shared with the same penalty γ allows us to derive an estimating equation to recover the non-hicksian shocks.

With this specification, our firm-level profit is as follows:

$$\Pi_{jt} = \left(\sum_{c \in \mathbb{C}_{jt}} P_{jct}(f_{jct}, D_{jct}) f_{jct} - P_{jt}^S S_{jct} \right) - P_{jt}^L L_{jt} - P_{jt}^F Fert_{jt} - P_{jt}^p Pest_{jt}$$

Taking $\{L_{jt}, Fert_{jt}, Pest_{jt}, S_{jct}\}$ as flexible inputs, their setting will only impact within period profit. I write $P_{jct}(f_{jct}, D_{jct})$ for the inverse demand function, known to the farmers, which depends on the volume produced and sold f_{jct}, and a demand shock D_{jct}. Writing X_{jt} for any of the public inputs $\{L_{jt}, Fert_{jt}, Pest_{jt}\}$, I get the following expression:

$$\frac{\partial \Pi_{jt}}{\partial X_{jt}} = 0 \Leftrightarrow P^X_{jt} = \sum_{c\in\mathbb{C}_{jt}} \frac{\partial g_{jct}}{\partial X_{jt}} g^{-1}_{jct} \sum_{c'\in\mathbb{C}_{jt}} \gamma_{cc'} f_{jc't} P_{jc't}\Big[1 - \frac{1}{\eta_{jct}}\Big]$$

I denote by η_{jct} the absolute value of the elasticity of demand. I then look at land, the only private input in this set-up:

$$\frac{\partial \Pi_{jt}}{\partial S_{jct}} = 0 \Leftrightarrow P^S_{jt} = \sum_{c'\in\mathbb{C}_{jt}} P_{jc't}\Big[1 - \frac{1}{\eta_{jct}}\Big]\gamma_{cc'} \frac{\partial g_{jct}}{\partial S_{jct}} g_{jct} f_{jc't}$$

Combining these two equations, I get:

$$\frac{P^X_{jt}}{P^S_{jt}} = \sum_{c\in\mathbb{C}_{jt}} \frac{\frac{\partial g_{jct}}{\partial X_{jt}}}{\frac{\partial g_{jct}}{\partial S_{jct}}}$$

This relation is quite intuitive, the relative allocation of inputs has to be such that the ratio of their marginal impacts on production across production lines equals the ratio of their prices.

To go further, I assume a specific shape for the g_{jct} function[55]. I take:

$$g_{jct} = K^{\alpha_k}_{jt} L^{\alpha_L}_{jt} \Big\{ \delta^c_s S^{\rho}_{jct} + \delta^c_p \big[e^{\omega^p_{jt}} Pest_{jt}\big]^{\rho} + \delta^c_f \big[e^{\omega^f_{jt}} Fert_{jt}\big]^{\rho} \Big\}^{\frac{\alpha_s}{\rho}}$$

Without relying on a translog specification, this is the simplest specification which allows us to model non-hicksian productivity shocks. I note that the heterogeneity in the g_{jct} across crops only relies on the $\{\delta^c_s, \delta^c_f, \delta^c_p\}_c$ parameters. Because I are mostly worried about variations in crop-composition impacting the chemicals-to-land ratios at the farm-level, this is the main form of heterogeneity I are interested. A more expansive heterogeneity could specify crop-specific elasticities ρ_c, which I assume away here.

With this parametrization, I obtain the following estimating equation, where small cap letters represent logs, with $x \in \{pest, fert\}$:

$$p^x_{jt} - p^s_{jt} = \rho\omega^x_{jt} + (\rho - 1)x_{jt} + log\big(\sum_{c\in\mathbb{C}_{jt}} \frac{\delta^c_x}{\delta^c_s} S^{1-\rho}_{jct}\big)$$

This equation mirrors the first stage estimating equation from Doraszelski and Jomandreu (2018), adapted to our parametrized multi-product setting. Here, the input ratio can either be affected by variations in the input price ratio, by changes in the establishment's production mix, or by changes in the etablishment's profficiency in using input x in the production process. Similarly to them, when our inputs of interest are substitutes, and holding everything else constant, an increase in ω^x_{jt} will increase the ratio of x_{jt} to land.

[55] Note that this framework can also accomodate a full CES specification. Given the difficulty of estimating the parameters within a CES nest, I keep the CES structure to a minimum - our land nest being itself estimated within an equation which is linear in part of the parameters.

I propose to recover the $\{\delta_s^c, \delta_f^c, \delta_p^c\}_c$, and ρ in a first estimation step, and to recover the remaining parameters in a second step. To simplify the estimation process, I assume that $\gamma_{cc} = 1$, and $\gamma_{cc'} = \gamma$, and verify post-estimation that $\gamma_{cc'} \in]-\frac{1}{dim(\mathbb{C})}, 0[$ to guarantee that the second order conditions from the optimization problem are met.

First Step: I specify a structure for the two input-biased productivity shocks. For each of them, I assume an AR(1) structure such that, with ζ_{jt+1}^x an exogenous i.i.d. innovation shock:

$$\omega_{jt+1}^x = \mathbb{E}[\omega_{jt+1}^x | \omega_{jt}^x] + \zeta_{jt+1}^x = g^x(\omega_{jt}^x) + \zeta_{jt+1}^x$$

I take $g^x(.)$ to be a third-order polynomial. I use the following moments, with A_{jt}^1 our matrix of instruments:

$$\mathbb{E}[(\zeta_{jt}^p + \zeta_{jt}^f)A_{jt}^1] = 0$$

Our instruments match the production function literature, and correspond to lagged firm-level observables presumably uncorrelated to the productivity innovations, as well as the current values of farm-level land prices and hourly agricultural wages.

Second Step: For our second step, I take the log of our production function, and re-order the elements to get:

$$\tilde{\omega}_{jct} = log(f_{jct}) - \left(\sum_{c' \in \mathbb{C}_{jt}} \gamma_{cc'}\right)\left[\alpha_K k_{jt} + \alpha_L l_{jt}\right] - \sum_{c' \in \mathbb{C}_{jt}} \gamma_{cc'} \tilde{s}_{jc't}$$

Where $\tilde{s}_{jc't}$ is the log of the land-fertilizer-pesticide nest, which I can compute using the estimates from the first step. Now, I use crop-specific policy cushions, as well as land prices and agricultural wages as instruments denoted A_{jt}^2 in a GMM estimation, with moments:

$$\mathbb{E}\left[\tilde{\omega}_{jct} A_{jt}^2\right] = 0$$

I obtain the following results:

Table A18: Joint Production - Parameters

	Coefficient	se
α_k	0.241595	0.000602
α_l	0.481252	0.001580
α_s	0.093129	0.000022
γ	-0.090429	0.000126
ρ	0.388665	0.000730
δ_f - Cereals	0.242870	0.002716
δ_f - Oil/Protein	0.291691	0.003467
δ_f - Industrial	0.028215	0.000214
δ_p - Cereals	0.178954	0.002593
δ_p - Oil/Protein	0.292520	0.004082
δ_p - Industrial	0.102460	0.002069

Standard errors from the second step are corrected for the two-step procedure, following Doraszelski and Jomandreu (2018), and the coefficients associated to the input-biased productivity processes are concentrated out, and not estimated. I note a negative value for γ indicating that the more you share the public inputs across production lines, the less a given crop will benefit from that input (either because of an increase in scope, or an increase in the scale of the other product's production levels). I also note that $\rho \in]0,1[$, indicating that land, pesticides and fertilizers are substitutes in production, which was expected.

I show below the time series obtained for the pesticide and fertilizer specific productivity shocks. I note that the trends are not significantly different from the ones obtained in the disjoint production framework.

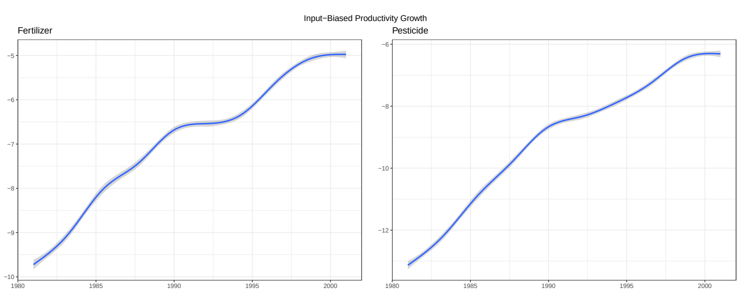


Figure A29: Input-Biased Productivities - Joint Production Framework

Notes: I show GAM models that interpolate the farm-level trends of resp. fertilizer and pesticide specific productivity shocks in France over time.

B.7 Recovering Input Allocations

As discussed in subsection 4.2, I estimate our production parameters on single crop-group firms, following De Loecker et al. (2016). Because I observe farm-crop level land allocations, knowing the parameters of our production specification, I can then recover both input allocations and hicksian and non-hicksian productivity shocks for multi-product firms. Our production function has the following shape:

$$Q_{jct} = e^{\tilde{\omega}^h_{jct}} K_{jt}^{\alpha^c_K} L_{jct}^{\alpha^c_L} \left\{ \delta^c_s S^\rho_{jct} + (1-\delta^c_s) \left(e^{\omega^{ch}_{jt}} \left[(1-\delta^c_p) Fert^{\rho_2}_{jct} + \delta^c_p Pest^{\rho_2}_{jct} \right]^{\frac{1}{\rho_2}} \right)^\rho \right\}^{\frac{\alpha^c_S}{\rho}}$$

With: $\tilde{\omega}^h_{jct} = \omega^h_{jct} + \varepsilon_{jct}$

Drawing from the first-order conditions, I can write the following relation between pesticides and fertilizers:

$$Fert^{1-\rho_2}_{jct} = \frac{P^P_{jt}}{P^F_{jt}} \frac{(1-\delta^c_p)}{\delta^c_p} Pest^{1-\rho_2}_{jct} \tag{13}$$

I can use the relation to recover a relation between land and pesticide allocations:

$$\frac{P^s_{jt} S^{1-\rho}_{jct}}{\delta^c_S} = \frac{P^P_{jt}}{\delta^c_p(1-\delta^c_s)} \left(e^{-\omega^{ch}_{jt}} \right)^\rho Pest^{1-\rho}_{jct} \left[(1-\delta^c_p) \left(\frac{P^P_{jt}(1-\delta^c_p)}{P^F_{jt}\delta^c_p} \right)^{\frac{\rho_2}{1-\rho_2}} + \delta^c_p \right]^{1-\frac{\rho}{\rho_2}}$$

For farms producing the two crop groups, I can then take the ratio of this expression for their two crop categories c and c':

$$\frac{Pest^{1-\rho}_{jct}}{Pest^{1-\rho}_{jc't}} = \frac{S^{1-\rho}_{jct}}{S^{1-\rho}_{jc't}} \frac{\delta^c_p(1-\delta^c_s)\delta^{c'}_s}{\delta^{c'}_p(1-\delta^{c'}_s)\delta^c_s} \frac{\left[(1-\delta^c_p) \left(\frac{P^P_{jt}(1-\delta^c_p)}{P^F_{jt}\delta^c_p} \right)^{\frac{\rho_2}{1-\rho_2}} + \delta^c_p \right]^{\frac{\rho}{\rho_2}-1}}{\left[(1-\delta^{c'}_p) \left(\frac{P^P_{jt}(1-\delta^{c'}_p)}{P^F_{jt}\delta^{c'}_p} \right)^{\frac{\rho_2}{1-\rho_2}} + \delta^{c'}_p \right]^{\frac{\rho}{\rho_2}-1}}$$

Once I have estimated the parameters of our production function, the right hand side of this expression is fully known. As I know total farm pesticide use, I can then recover the volumes of pesticide used for each of the crops that they grow. I can then use Equation 13 to recover the fertilizer allocations as well. With these, I can construct the crop-specific land nest values, and use the outer Cobb-Douglas shape of our production function to recover crop-specific labor allocations, from these crop-specific land nests.

B.8 Selection Correction for the Production Function Estimation

I need to account for selection while building moments for the second step of our estimation procedure. Using $\Xi_{jct} = 1$ a dummy for the fact that j only produces crop c in period t, I want to build moments based on the following corrected process:

$$\omega^h_{jct+1} = \mathbb{E}[\omega^h_{jct+1} | \omega^h_{jct}, \Xi_{jct+1} = 1] + \xi^h_{jct+1} \tag{14}$$

With I_{jt} the farm's information set at the end of period t, I want to base our moments on the following derivation – with $\overline{\omega}^h_{jcc't+1}$ the threshold used at the end of period t to decide on the inclusion of crop c' on top of c. Here c is the ex-ante most profitable crop to grow in period t, and c' the second (out of two crops). I use a competence ladder, which means that there will be a strict relation between the farm's TFP between c and c'. I also note that because demand and input prices play a role in crop choice – the core competence of a farm might not always match the identity of the unique crop it grows.

$$
\begin{aligned}
\xi^h_{jct+1} + \varepsilon_{jct+1} &= q_{jct+1} - \alpha^c_k k_{jt+1} - \alpha^c_l l_{jct+1} - \alpha^c_s \tilde{s}_{jct+1} - \mathbb{E}[\omega^h_{jct+1} | I_{jt}, \Xi_{jct+1} = 1] \\
&= q_{jct+1} - \alpha^c_k k_{jt+1} - \alpha^c_l l_{jct+1} - \alpha^c_s \tilde{s}_{jct+1} - \mathbb{E}[\omega^h_{jct+1} | \omega^h_{jct}, \overline{\omega}^h_{jcc't+1}] \\
&= q_{jct+1} - \alpha^c_k k_{jt+1} - \alpha^c_l l_{jct+1} - \alpha^c_s \tilde{s}_{jct+1} - g_h(\omega^h_{jct}, \overline{\omega}^h_{jcc't+1})
\end{aligned}
$$

Where moving from line 1 to 2 relies on the first-order Markov process assumption, and the definition of the threshold $\overline{\omega}^h_{jcc't+1}$.[56] I can then specify $g_h(.)$ as a polynomial of its two terms. While Equation 7 provides an expression for ω^h_{jct}, I also need one for the threshold $\overline{\omega}^h_{jcc't+1}$. For this I derive an expression for the conditional probability of producing only crop c as a function of farm productivity and its state, focusing on the crop-addition threshold.[57] Our competence ladder structure means that the threshold for the production of the second crop c' can be expressed using the farm's observed TFP shock for crop c combined with its competence ladder σ_j – specifically by its rank on the farm's ladder $\sigma_j^{-1}(c)$, as well as additional variables that influence the value of the threshold – input and output prices, capital stock and chemical productivity:

$$
\begin{aligned}
\mathbb{P}\Big(\Xi_{jct+1} = 1\Big) &= \mathbb{P}\Big(\omega^h_{jct} \leq \overline{\omega}^h_{jcc't+1}(\omega^{ch}_{jct}, K_{jt+1}, S^o_{jt+1}, \sigma_j^{-1}(c), \Omega_{jt}) | \overline{\omega}^h_{jcc't+1}(.), \omega^h_{jct}\Big) \\
&= h(\overline{\omega}^h_{jcc't+1}(.), \omega^h_{jct}) \\
&= h(\omega^h_{jct}, \omega^{ch}_{jct}, K_{jt+1}, S^o_{jt+1}, \sigma_j^{-1}(c), \Omega_{jt}) \\
&= h(\{X_{jct}\}, \omega^{ch}_{jct}, K_{jt+1}, S^o_{jt+1}, \sigma_j^{-1}(c), \Omega_{jt}) \\
&= SP_{jt}
\end{aligned}
$$

Equation 7 is used to move from the third to the fourth line. While I do not observe $\sigma_j^{-1}(c)$, I approximate it by the interaction between the farm's location and a dummy for the crop it produces – assuming competence ladders are determined by local soil and climatic factors. As in Olley and Pakes (1996) and De Loecker et al. (2016), under some regularity conditions on the density of ω^h_{jct}, I can invert SP_{jt}, which is first recovered non-parametrically, to obtain a proxy for the threshold that I use in our polynomial approximation of $g_h(.)$.

B.9 Monopolistic Competition

An alternative to perfect competition is to adopt a monopolistic competition framework – which preserves the single agent setting – but allows for price dispersion and some horizontal differentiation across goods. While agricultural products are commodities usually thought of as homogeneous, exogenous local conditions can lead to similar crops having different

[56] Because of our assumed competence ladder structure, firms face a unique TFP process that propagates across their crops. The process g_h is similar across crops as a direct consequence of this.

[57] The threshold for not including crop c as well corresponds to the usual selection issue outlined by Olley and Pakes (1996) when no production is equated with exit, and is addressed in the usual way by keeping an unbalanced panel of farms.

moisture or nutrient content, making them more or less suitable for different uses. In that sense, a model with differentiated products might be preferable. Monopolistic competition can also allow for mark-ups that vary across producers, and along with changes in market structure.

I adopt for this section the quadratic demand of Melitz and Ottaviano (2008). The demand structure goes as follows: in a closed economy, a representative consumer has the following utility at period t:

$$U_t = q_{0,t} + \alpha \int_{i\in\Omega} q^c_{i,t} di - \frac{1}{2}\gamma \int_{i\in\Omega} (q^c_{i,t})^2 di - \frac{1}{2}\eta\Big(\int_{i\in\Omega} q^c_{i,t} di\Big)^2 \tag{15}$$

$q_{0,t}$ corresponds to the quantity of outside good consumed. Agricultural varieties $i \in \Omega$ are differentiated. α and η control the substitution between differentiated varieties and the outside good, while γ accounts for the degree of differentiation across varieties. Here, I note that two varieties can be two types of corn grown by two different farms, or wheat and corn grown by the same farm. I discuss next a clustering of the agricultural market, where each market correspond to a crop type, and firm produce different varieties within each crop market.

I take the outside good to be the numeraire. This preference structure implies the following demand for variety i - with $\overline{p}_t$ the average price of agricultural varieties on the market, and M_t the mass of consumed varieties:

$$q_{i,t} = \frac{\alpha}{\eta M_t + \gamma} - \frac{1}{\gamma} p_{i,t} + \frac{\eta M_t}{\eta M_t + \gamma}\frac{1}{\gamma}\overline{p}_t$$

I see that M_t and $\overline{p}_t$ – market-level statistics on which atomistic farms have no individual impact – determine the characteristics of the demand curve faced by farms. From this relation, I can also derive a choke price p^{max}_t above which farms will not face any positive demand.

Estimating Demand: I use the following expression to recover the parameters of demand. From now on I denote by m the crop grown, to match the competence ladder structure that I introduce next. Because crops are differentiated in the model, each of a farm's variety is uniquely produced by that farm.

$$q_{mt} = \frac{1}{\gamma}\alpha - \frac{1}{\gamma} p_{mt} - \frac{\eta}{\gamma} Q_t$$

I take for Q_t total agricultural production sold in France, approximated using the weighted sum of sold output in the FADN, for the set of crops that I consider. q_{jmt} and p_{jmt} are respectively the quantity sold, and associated price, for each variety produced by a farm in a given period. I estimate the following regression using a two-stage least squares strategy, with exogenous shocks to firm supply as instruments. Specifically, I use local realized and forecasted weather shocks as supply shocks.

$$q_{jmt} = \beta_0 + \beta_1 p_{jmt} + \beta_2 Q_t + \varepsilon_{jmt}$$

I recover the following parameters. I only use data post-2003 for the estimation, in order to obtain demand parameters that are not impacted by intervention pricing.

Table A19: Demand - Parameters

Coefficient	Parameter	Estimate	Std. Error
Substitution with Numeraire	η	1.32e-07	6.92e-08
Substitution within Varieties	γ	3.19e-02	1.01e-02
Substitution with Numeraire	α	203.5	71.42

Notes: I recover the parameters from a 2SLS regression, using FADN data fro 2003-2020, and using realized and forecasted weather data constructed at the department-year level as supply shocks. The realized weather data is constructed from the ECMWF ERA-5 hourly weather data series, and the forecasts come from ECMWF SEAS-5. Both are aggregated into growing season observations.

The standard errors for the transformed parameters are computed using the delta method. This gives us a decreasing choke price over time, which drops significantly around the time of the reform.

Extension to Differentiated Crop Markets: It is useful to extend the model to account for different agro-ecological zones, for which different crops are more or less suitable. Doing so can allow us to draw some conclusions on the impact of different policy designs on respectively the localisation of agriculture, and on the evolution of local crop diversity and pollution.

The first step in developing a model with differentiated locations is to extend our demand system allow for differentiated crop markets. In the previous version of the model, all the varieties produced within and across farms were considered as different agricultural varieties. Translated to the data, this meant that two crops growing wheat would grow different agricultural varieties, to the same extent that one farm growing corn and wheat would also grow two differentiated varieties. All the agricultural commodities were then facing the same choke price, and the same aggregate demand. I now differentiate the crop markets. Specifically, agricultural varieties c can now be located in different crop markets k. There are K crop markets, and their set is represented by $\mathbb{K}$, with $k \in \mathbb{K}$. Demand takes the following form:

$$U_t = q^c_{0,t} - \frac{1}{2}\gamma \int_{i\in\Omega} (q^c_{i,t})^2 di - \frac{1}{2}\eta\Big(\int_{i\in\Omega} q^c_{i,t} di\Big)^2 + \sum_k \alpha_k \int_{i\in\Omega_k} q^c_{i,t} di \tag{16}$$

This corresponds to the simplest way in which I can introduce some heterogeneity in demand across crop types. Extending our notation to M_k, the mass of produced and consumed varieties in market k, and $\overline{p}_k$ their average price, each crop market will have the following choke price:

$$p^k_{max} = \alpha_k - \eta \frac{\sum_k M_k(\alpha_k - \overline{p}_k)}{\left(\eta \sum_k M_k\right) - 1}$$

Quite simply, the choke prices will vary across crop markets in an additive way according to the variation in the $\{\alpha_k\}_k$.

Firms will now relate to these differentiated markets according to what I call their crop schedule, or a permutation of $\mathbb{K}$. Farm j will be endowed by a time-constant schedule σ_j, which relates the rank of their varieties grown to a specific crop market – and hence to a specific residual demand curve. The first crop grown by a farm is now $\sigma_j(0)$, and two farms with different schedules will then have differentiated core competences. There are $K!$ such possible schedules, and a farm draws its schedule at random upon entry, from a uniform distribution over the set of possible schedules. Entrants then do not decide which market to enter, but are endowed with a schedule randomly. While the firm dynamics literature allows for a wider range of assumptions regarding the market-directedness of entry, such as Nocke (2006) where entrepreneurs enter a unique targeted market to produce a unique product, or Klette and Kortum (2004) where innovation happens randomly in a new market across a continuum, our assumption allows for a very smooth integration of differentiated markets within the Mayer et al. (2014) multi-product setting. Here, firms keep producing an integer amount of products, specifically they produce at a maximum K goods, and given the continuum of producers, I preserve the atomicity of each of them, and the single agent framework.

Our model then follows the same structure as the one previously outlined, although the zero profit condition now sets a set $\{p^k_{max}\}_k$ of choke prices. Given the additive structure of the choke prices, a unique zero profit condition at entry is sufficient to define the K different choke prices.

The Bellman equation characterizing an incumbent's problem can now be written as follows:

$$\begin{aligned}
V(\delta_{jt}, K_{jt}, \sigma_j; \{p^{kt}_{max}\}_k) = & \max_{\mathbb{C}_{jt}, \{S_{jct}, X_{jct}, L_{jct}\}_{c \in \mathbb{C}_{jt}}, \zeta_{jt}, K_{jt+1}} \sum_{c \in \mathbb{C}_{jt}} \Pi_{jct} \\
& - C(K_{jt}, K_{jt+1}) + \zeta_{jt} P^K (1-\delta_k) K_{jt} \\
& + (1-\zeta_{jt}) \frac{1}{R} \left\{ \int\int V(\delta_{jt+1}, K_{jt+1}, \sigma_j; \{p^{kt+1}_{max}\}_k) dJ\big(\delta_{jt+1}|\delta_{jt}\big) dF\big(\{p^{kt+1}_{max}\}_k | \{p^{kt}_{max}\}_k\big) - f_c \right\}
\end{aligned}$$

Where I note $\mathbb{C}_{jt} = \{\sigma_j(0), ..., \sigma_j(M_{jt}-1)\}$, for M_{jt} the farm's scope in the given period. F is now extended to account for the evolution of the complete choke price vector, rather than the unique one. However, given the additive structure of the choke price heterogeneity,

F retains the same dimension as previously, and only models the evolution of a slightly different object.

The entrant's problem now has to account for the drawing of one's crop schedule as well. I write the entrant's Bellman as follows:

$$V^e(\{p_{max}^{kt}\}_k) =$$
$$\max_{\zeta_{jt}^e} \zeta_{jt}^e \Big\{ \max_{K_{jt+1}} \frac{1}{R} \int\int\int\int V\big(\delta_{jt+1}, K_{jt+1}, \sigma_j; \{p_{max}^{kt+1}\}_k\big) dJ\big(\delta_{jt+1}|\delta_{jt}\big) dF\big(\{p_{max}^{kt+1}\}_k|\{p_{max}^{kt}\}_k\big) dG(\delta_{jt}) dU(\sigma_j)$$
$$- P^K K_{jt+1} - f_e \Big\}$$

Here, I denote by $U(.)$ the uniform distribution over the different crop schedules, which corresponds to the set of K-cycles, without repetition.

Again, I can follow the notation of Hopenhayn and Rogerson (1993) to write the law of motion which describes the evolution of the market. I denote this measure by μ, a measure defined over the three dimensions that characterize the state of a farm – meaning productivity, capital and crop schedules. I denote by $\phi_\sigma(\delta', K'|\delta, K)$ the probability of transition from state (δ, K) to (δ', K'), for a farm with crop schedule σ. With $\kappa(\delta, K, \sigma; \{p_{max}^{kt}\}_k)$ the investment policy function, and $\zeta(\delta, K, \sigma; \{p_{max}^{kt}\}_k)$ the incumbency policy function, I have:

$$\phi_\sigma(\delta', K'|\delta, K) = J(\delta'|\delta) 1\{K' = \kappa(\delta, K, \sigma; \{p_{max}^{kt}\}_k)\} 1\{\zeta(\delta, K, \sigma; \{p_{max}^{kt}\}_k) = 0\}$$

And for an entrant, it is also useful to denote the transition probability:

$$\phi_\sigma^0(\delta', K'|\delta) = J(\delta'|\delta) 1\{K' = \kappa(\delta, K = 0, \sigma; \{p_{max}^{kt}\}_k)\} 1\{\zeta(\delta, K = 0, \sigma; \{p_{max}^{kt}\}_k) = 0\}$$

The stationary measure then solves, with N_e the mass of entrants, $\forall$ (δ', K', σ):

$$\mu(\delta', K', \sigma) = \int\int\int \phi_\sigma(\delta', K'|\delta, K) d\mu(\delta, K, \sigma) + N_e \int\int \phi_\sigma^0(\delta', K'|\delta) dG(\delta) dU(\sigma)$$

This relation defines an operator T for the evolution of the state measure μ.

I can then write the mass of produced varieties in each crop market, and their respective average price:

$$M_k = \int\int\int 1\{c_k \in \mathbb{C}(\delta, K, \sigma)\} \; d\mu(\delta, K, \sigma)$$
$$\overline{p}_k = \frac{1}{M_k} \int\int\int 1\{c_k \in \mathbb{C}(\delta, K, \sigma)\} p_k(\delta, K, \sigma) \; d\mu(\delta, K, \sigma)$$

With $\mathbb{C}(\delta, K, \sigma)$ the crop mix of a farm with this given type, derived from that farm's scope $M(\delta, K, \sigma)$ combined with its crop schedule σ. I also denote by $p_k(\delta, K, \sigma)$ the optimal price for crop k, set by a farm of type (δ, K, σ) actually producing this crop k.

The stationary equilibrium of this differentiated demands model is then a tuple formed by choke prices $\{p^k_{max,\star}\}_k$, a stationary measure $\mu^\star$, and an equilibrium mass of entrants $N_e^\star$ such that:

- $V^e(\{p^{k,\star}_{max}\}_k) \leq 0$, and $V^e(\{p^{k,\star}_{max}\}_k) = 0$ if $N_e^\star > 0$.
- $T\left(\mu^\star, N_e^\star, p^{k,\star}_{max}\}_k\right) = \mu^\star$.
- $p^{k,\star}_{max} = \alpha_k - \eta \frac{\sum_j M_j(\alpha_j - \overline{p}_j)}{\left(\sum_j \eta M_j\right) - 1}$.

B.10 Decomposition

I propose a slight modification of the dynamic decomposition of Olley and Pakes (1996) and De Loecker and Collard-Wexler (2015) to highlight the role of different channels in driving the evolution of total chemical use on the market. I write evolution of market-level total chemical use between t and $t+1$:

$$\begin{aligned}
\Delta Chemicals = &\sum_j Q_{jt}\left(s_{jc_1t}\Delta \frac{C}{Q}_{jc_1} + s_{jc_2t}\Delta \frac{C}{Q}_{jc_2}\right)\\
&+ Q_{jt}\left(\Delta s_{jc_1}\frac{C}{Q}_{jc_1t} + \Delta s_{jc_1}\Delta\frac{C}{Q}_{jc_1} + \Delta s_{jc_2}\frac{C}{Q}_{jc_2t} + \Delta s_{jc_2}\Delta\frac{C}{Q}_{jc_2}\right)\\
&+ \left(\Delta Q_j \frac{C}{Q}_{jt} + \Delta Q_j \Delta \frac{C}{Q}_{jt}\right)
\end{aligned}$$

I write $\frac{C}{Q}_{jtc_2}$ for the chemical-to-output ratio of farm j in period t for crop c_2, s_{jtc_1} denotes the output share of crop c_1 in farm j at t, and Q_{jt} is for total output. The first line of the decomposition is for within-farm within-crop changes in chemical intensity associated with changes in input price ratios. The second line is for changes in chemical-use intensity coming from reallocations of production within the farm. The last line is for cross-farm reallocations.

B.11 Sensitivity Analysis for the Production Function Estimation

I rely on the work of Andrews et al. (2017) to discuss the role of our identifying assumptions regarding instrument exogeneity in driving the estimation of our production function parameters. I first summarize the purpose and implementation of their method, and then describe our estimates and their implications for the sensitivity of our estimates to instrument exogeneity.

The purpose of the method developed by Andrews et al. (2017) is to recover the sensitivity of model parameters to the moments used in structural estimation, and in this extending the omitted variable bias formula to structural cases. They propose to recover a matrix they call the "sensitivity" matrix, which maps the relation between the parameters of the model, and the moments used for their estimation. When combined with alternative assumptions about instrument validity – recasted as the impact of alternative assumptions on moment values – this sensitivity matrix allows to predict the impact of alternative assumptions on the estimated parameters. In other words, this analysis aims at shedding light on the role of identifying assumptions in driving estimation.

I denote by $\hat{\theta}$ the parameter vector minimizing the criterion function:

$$\hat{g}(\theta)'\hat{W}\hat{g}(\theta)$$

In our case, $\hat{\theta}$ can the our estimate for the first stage and second stage of our production function estimation. For any a in the set A of alternative assumptions, they define a local perturbation of the model in the direction of a such that the estimate $\hat{\theta}$ has first-order asymptotic bias – where Λ is the sensitivity matrix:

$$\mathbb{E}[\tilde{\theta}(a)] = \Lambda \mathbb{E}[\hat{g}(a)]$$

An alternative assumption a should be interpreted as an alternative assumption regarding the relation between our chosen instruments and the structural errors for resp. our first and second step of estimation. The sensitivity matrix can be written as:

$$\Lambda = -\Big(G'WG\Big)^{-1} G'W$$

Where W is the probability limit of $\hat{W}$, the weight matrix used in our GMM criterion, and G is the Jacobian of the probability limit of $\hat{g}(\theta)$ evaluated at the true parameter vector θ_0. Λ serves as a local approximation to the mapping from moments to estimated parameters.

I provide an estimate of ΛK for each of our estimation step for the production function, where K is a weighting matrix, which weights the sensitivity matrix by the inverse of the standard deviation of the relevant instruments. The values for the constant are standardized by the standard deviation of the structural error, in order to have a similar interpretation. As Andrews et al. (2017) discuss, their units of Λ are contingent on the units of the different moments – which for us means it will vary across moments depending on the values of the instruments I interact with the GMM structural errors. As such, this weighting ensures that Λ elements can be read as the effect of a one standard deviation violation of the given moment condition on the asymptotic bias in $\hat{\theta}$. I present and discuss below our estimates of ΛK for each estimation step.

I see from the first table that our first stage estimates are particularly sensitive to assumptions regarding farmer expectations on the structural error (related to the constant), to the exogeneity of department-level agricultural wages in the current period, and to the exogeneity of the lag value of the chemical (fertilizer-to-pesticides) ratio. The production function literature has highlighted that input prices make for powerful instruments if their exogeneity to firm decisions is believable. Given the atomistic nature of farms, I expect these department wages to be exogenous to farm-level innovation in non-Hicksian productivity. As should be expected, the ratio of fertilizers to land is particularly useful to identify the parameter of substitution between fertilizers and pesticides ρ_2. The lag values of this ratio should be correlated to current production choices through auto-correlation in the relative prices of fertilizers and pesticides faced by farms – which I can also assume to be exogenous. Because I focus on single-product farms, the composition of the farms' crop mix should not play a role in driving this ratio.

The second table shows second-stage parameters are less sensitive to any specific identifying assumptions. The moments which play a largest role in the estimation are the current input prices - resp. the hourly wage, the current price of chemicals, and the lag of the output price.

Table A20: Sensitivity for First Step

	Constant	$log(Chem.Ratio_{t-1})$	$Capital_t$	$Wage_t$	$LandPrice_t$	$Land_{t-1}$	$FarmExposure_t$
$log(1-\rho)$	3.02	-0.2420	0.0000	-2.0303	-0.0000	0.0007	-0.0003
ρ_2	-68.86	9.6731	-0.0000	44.0966	0.0000	-0.0276	0.0046
δ_s (group 1)	-4.60	0.3820	-0.0000	3.0439	0.0000	-0.0009	0.0004
δ_s (group 2)	1.17	-0.1288	0.0000	-0.7720	-0.0000	0.0002	-0.0001
δ_p (group 1)	-7.00	0.8927	0.0000	5.0727	0.0000	-0.0017	-0.0002
δ_p (group 2)	23.88	-3.1918	0.0000	-15.6575	-0.0000	0.0072	-0.0017

Table A21: Sensitivity for Second Step

	Constant	$log(Capital_t)$	$log(Labor_{t-1})$	$log(Wage_t)$	$log(LandPrice_t)$	$log(Capital_{t-1})$	$log(Chem.Prices_t)$	$log(Price_{t-1})$	$LandSubsidy_t$	$log(Land_{t-1})$	$log(FarmExposure_t)$
α_l (group 1)	0.0000	0.0000	-0.0025	-0.0203	0.0009	-0.0000	0.0281	0.0039	0.0000	-0.0004	-0.0010
α_s (group 1)	0.0000	-0.0000	0.0010	0.0078	-0.0002	0.0000	-0.0097	-0.0015	-0.0000	0.0002	0.0004
α_l (group 2)	0.0000	-0.0002	0.0012	0.0067	-0.0004	0.0000	-0.0118	-0.0015	-0.0000	0.0001	0.0004
α_s (group 2)	0.0000	0.0000	-0.0003	-0.0017	0.0001	-0.0000	0.0026	0.0004	0.0000	-0.0000	-0.0001