Welfare Consequences of Production Subsidies in a Competitive Market: Efficiency and Pollution in European Agriculture*

Tristan Du Puy†

September 2, 2024

Abstract: *I study the economic incidence of agricultural subsidies in a dynamic industry model with pollution externalities, and do so in the context of the European Union's Common Agricultural Policy (CAP) and the French row crop market. Relying first on a shift-share design, I analyse the role of subsidy design and level on firm survival, profit and chemical use. I then develop and estimate an empirical partial equilibrium dynamic model of the market to recover the welfare impact of different subsidy designs common in agricultural markets – output price intervention and a land subsidy. I compare both to a chemicals tax and the no-intervention equilibrium. Subsidies cause misallocation, but the costs of misallocation are attenuated due to complementarities between farm productivity and pollution-intensive inputs. Additionally, the land subsidy shifts incentives away from using chemicals in production, mimicking some of the treatment effect of a chemicals tax. We use the marginal taxation rates for pesticide use effective in resp. Denmark and France to rank policy designs, and find that a chemicals tax delivers the highest reductions in pollution for the lowest effect on prices. A moderate land subsidy still compares favorably to a no-intervention case, while output price intervention yields the highest pollution intensity and prices of all equilibria.*

*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, Rodrigo Soares, Sebástian Otero and Diego Restuccia for their generous feedback, as well as Marguerite Obolensky, Vincent Bagilet, Nicolas Longuet-Marx and Yulia Kim for their support. I am grateful for comments from numerous participants in seminars and conferences at Columbia University, the Toulouse School of Economics, NAREA and WEAI. All remaining errors are my own. I would like to acknowledge funding support 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

Among OECD countries, agricultural producers are usually atomistic, goods fairly homogeneous, costs of entry low, and input markets well functioning. As such, market failures related to efficiency concerns are relatively scarce. Yet, agricultural subsidies still constitute one of the most widespread and largest form of industrial policy intervention within these countries.[1] Standing on the side of these efficiency arguments are growing concerns for the environment, concerns particularly relevant to the agricultural industry given the breadth of its environmental production externalities.[2] OECD agricultural markets can then be characterized as markets where policy intervention should be limited under efficiency considerations, but likely warranted under environmental ones. We investigate the extent to which the joint valuation of efficiency and environmental outcomes changes our understanding of these subsidies' welfare consequences.

Subsidies can affect market outcomes through the reallocation of production across heterogeneous producers, and by treating producers directly in their production choices. Compared to taxes, subsidies will likely reduce firm selection – and hence shift the role of reallocation in driving market outcomes. The nature of producer heterogeneity is key to understand the relevance of reallocations, and the relation between heterogeneity in efficiency and in externality intensity is particularly important. If these are negatively correlated, the efficient allocation of production and the environmentally sustainable ones should match. Distortions away from this allocation are not desirable, and externalities can be directly addressed by shifting within-firm incentives of production and input choice. In such a world, smaller farms are likely less efficient and more polluting, and environmental concerns add to the usual efficiency ones for land consolidation. If there are complementarities between pollution and productivity, moving towards a more efficient equilibrium potentially entails moving away from a more sustainable one. Specifically, the welfare consequences of selecting establishments on efficiency are less obvious.

We focus on European agricultural subsidies under the umbrella of the European Union's (EU) Common Agricultural Policy (CAP), and their impact on the French row crop market.[3] There are several reasons for this choice of context. The CAP is the EU's largest industrial policy – its subsidy regime one of the largest in the world. France is one of the largest exporter of cereals.[4] This context also benefits from agricultural data of unusual

[1] The most prevalent efficiency concerns surrounding agricultural markets relate to the volatility of agricultural income stemming from climate variability (Schultz (1945)), buyer monopsony power, or credit constraints (Gardner (2002)). All can be addressed with specific tools – in turn yield insurance programs, low output floor prices, or agricultural loan programs. These programs however only account for a small share of these governments' agricultural expenses.

[2] 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.

[3] For practical reasons, we will consistently address the European Community or European Union as the European Union throughout the paper, regardless of whether we are 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

precision in the form of producer panels with precise input and output data, as well as the farm-level French Agricultural Census. Finally, we can leverage the 1992-1995 MacSharry reform of the CAP for identification, which shifted subsidization from a decades-long output price intervention regime, to the current land subsidy one. The French row crop market itself can properly be modelled as competitive, and is a setting where economic efficiency concerns are relatively scarce, and environmental ones many.

We address our question in two steps. We recover the structure of production within the French row crop market, paying particular attention to the distribution of producer heterogeneity, the relation between efficiency and pollution intensity, and the role of crop mix composition in driving that heterogeneity. We then nest it within a partial equilibrium model of firm dynamics, which endogenizes the main farm decisions relevant to concerns of equilibrium production allocations: market entry and exit, as well as capital accumulation. We recover a positive correlation between the chemical intensity of production and farm efficiency in the form of non-Hicksian heterogeneity in chemical-use efficiency, combined with an estimated substitutability between land and chemicals. A farm with a relatively higher ability at using chemicals will switch its input shares towards chemicals and away from land. This land-chemicals substitutability matches the intuition that organic farming tends to be relatively more land extensive. The relative volatility of farm-level Hicksian and non-Hicksian shocks, and their interaction with capital accumulation then determines how production gets allocated across producers through both within period production decisions, and across period capital accumulation and exit decisions (Collard-Wexler (2013), Asker et al. (2014)). We use this framework to evaluate a series of existing and alternative CAP designs.

We preface the structural analysis with a reduced form study of the consequences of the MacSharry reform in France. We rely on a shift-share design for that section, casting EU policy as setting wedges between European and international farm gate prices. This identification setting is nested within the framework of Goldsmith-Pinkham et al. (2020), and relates to the work of Topalova (2010), Kovak (2013) or Ager et al. (2024). The MacSharry reform had heterogeneous implications for crop-level subsidization. Farms were then heterogeneously exposed depending on the structure of their pre-reform crop mix. The design builds an argument for the conditional exogenous variation in farm exposure by controlling for the determinants of crop choice, and in that sense relates to a difference-in-difference study in a quasi-experimental setting. We show that as agricultural subsidization is reduced through the reform, farms' profits decrease, leading to significant farm exit at the lower tail of the farm size distribution. Additionally, the reform switches from an input-neutral subsidy to one specifically targeting land. This impacts the relative input price ratio of chemicals and land, and decreases farms' reliance on chemicals in production. Two channels are hence highlighted: subsidies impact profitability and survival, and can shift how selective a market is, subsidies can also shift incentives for input choice. Our design however only highlights the effect of marginal additional reform exposure, and cannot capture equilibrium consequences

of subsidy design. Our structural approach is able to account for both this effect, the average treatment effect across farms, and the effect of cross-establishment reallocations.

We develop a competitive partial equilibrium model of firm dynamics, where we pay particular care to entry, exit and capital accumulation decisions. The model generally follows the structure of Hopenhayn (1992), but allows for multi-product firms and capital accumulation, resp. in the spirit of Mayer et al. (2014) and Clementi and Palazzo (2016). We also allow for entry costs to adjust with market size.

The structural analysis is organized around three key elements. The first and most straightforward one is a demand system composed of constant elasticity demand curves.[5] The shape of demand curves is recovered using local weather forecasts and realizations as exogenous shocks to supply orthogonal to demand. The second and third elements relate to supply. A production function, and a set of parameters relating to the farms' dynamic decisions. There are two important features to the production function. The first one is the presence of chemical-biased productivity shocks, which drive cross-farms heterogeneity in production externalities. The second is that we model crop-specific production functions, while observing all inputs but land at the farm-level. To recover chemical-biased productivity shocks, we use the two steps parametric estimation strategy of Doraszelski and Jomandreu (2018).[6] To account for multi-product producers, we need a method to recover crop-specific input allocations. Here the approach of De Loecker et al. (2016) is most suited to allow both for crop-specific input coefficients, and for non-Hicksian shocks. Our two-step estimation thus relies on single product firms, and we later use the estimated parameters to recover input allocations and unobserved shocks for multi-product firms. Finally, we recover the costs of capital adjustment, a fixed cost of incumbency and the parameters governing the speed of adjustment of the cost of entry to market size. We recover these parameters using indirect inference, and identification relies on the selection of key moments which shift one-to-one with the parameters we wish to recover. This strategy relates to the work of Cooper and Haltiwanger (2006), Asker et al. (2014) and Johnston (2020) to recover costs of capital adjustment.

Policy Implications: The model allows us to revisit the welfare consequences of prior and existing CAP subsidy designs, and compare them in turn to the no-intervention equilibrium, and to the equilibrium under a chemicals tax directly targeting production externalities. The land subsidy regime dominates price intervention under any welfare criterion. It has a large impact on relative input choices by reducing the land-to-chemicals price ratio. This design achieves reductions in the chemical-to-output ratio on both crop markets. Williamson (1968) noted that in merger cases, the productive efficiencies of competition are infra-marginal, and hence likely larger than the effect of allocative efficiencies. The relative

[5] Our main specification does not model a possibility for cross-crop substitution in demand – but our extension using monopolistic competition gives us 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.

stronger importance of infra-marginal treatment channels is exemplified by Backus (2020) in the context of the US concrete industry. Our results echoes his, in the sense that direct policy changes to input price ratios have a much stronger impact on aggregate pollution here, than selective forces and their interaction with cross-farm heterogeneity in propensity to pollute. Under the valuation of pesticide-driven pollution implied by both France's and Denmark's current pesticide taxes, the post-1995 level of land subsidization dominates no-intervention as well. The production inefficiencies it introduces are compensated by its impact on pollution. This targeted input subsidy acts as a partial substitute for a tax on the polluting inputs, and its efficiency effects are limited as output prices adjust to the decreases in production costs it generates. We finish by modelling a chemical tax which induces a similar shift in the land-to-chemical price ratio as the land subsidy, and add to it a lump sum compensation for farmers based on their expected level of taxation. This chemical tax outperforms the land subsidy. It presents two advantages compared to subsidies. First, it directly targets chemicals, and shifts the price ratio of chemicals with all other inputs. Second, the subsidization it induces is only distortive through its effect on selection and input choice, but it does not impact scale decisions. Inefficient farms remain on the market thanks to the lump-sum, but do not produce when it is not efficient to do so.

Related Literature: This paper makes a series of contributions. First, we propose and estimate a model for the equilibrium relation between farm size and environmental pollution intensity stemming from the use of fertilizers and pesticides in production. In our modelling approach, we bring to the longstanding literature on input-biased technical change which goes back at least to both 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)). Methodologically, we rely on advances in the parametric estimation of production function with unobserved non-Hicksian shocks made by Doraszelski and Jomandreu (2018), and also used to study environmental outcomes by Hawkins-Pierot and Wagner (2023). This also links our paper to the growing 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)). Our paper is one of the firsts proposing an equilibrium modelling of the relation between agricultural production and chemical-led pollution, which also allows to study the effects of different policy designs on pollution.

Second, we propose a model of farms making joint static production decisions regarding scale and input choices, and dynamic crop mix, capital accumulation and entry/exit decisions. We add to a literature modelling dynamic farming and planting decisions, most

often to recover long term crop choice and land use elasticities (Livingston et al. (2008), Scott (2013), Hsiao (2022), Burlig et al. (2024)). Our detailed panel of producers allows us to recover price crop-specific production functions, as do Maue et al. (2020), but also costs of capital adjustment. Our data environment also allows us to model farms as multi-product producers – matching the reality of the French row crop market. This also sets our analysis within a series of recent papers focusing on equilibrium modelling of environmentally relevant markets, and paying particular care to agent heterogeneity and dynamics (Aspelund and Russo (2024), Rafey (2023), Aronoff and Rafey (2024), Ryan and Sudarshan (2021)).

Finally, we 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) We also use the approach of Mayer et al. (2014) to model multi-product firms, and set it within a dynamic setting. We highlight the comparative effect of policy on within firm effects, and cross-firm reallocations, in that being close to Backus (2020). Here, policy can have within-farm treatment effects via within-farm within-crop changes in modes of production, and within farm across-crops reallocations, as well as across incumbents reallocation effects through changes in market shares and entry/exit movements. This finally links our paper to the large literature on misallocations within agriculture (Adamopoulos and Restuccia (2014), Foster and Rosenzweig (2022), Lagakos and Waugh (2013), Gollin et al. (2021)), although we model an economy without efficiency-related market failures, where heterogeneity across establishments in productivity and pollution both interact with a homogeneous policy. We combine reduced form and structural evidence to show the effects of subsidy-driven shifts in market profitability on farm selection, and the comparative effects of technology-neutral and input-specific subsidies on efficiency and pollution.

This paper proceeds as follows: section 2 provides background on the French row crop market, the EU CAP, the MacSharry reform, and describes our 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 our structural model, our estimation strategy and finally counterfactual simulations. section 5 concludes

2 Data and Context

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

2.1 French and European Agriculture

The Common Agricultural Policy was first designed to foster the reconstruction of post-WW2 European agriculture, reduce the output deficit and support agricultural income (Petrick

(2008)). It aimed to do so by creating a common agricultural market within the European Community, and by implementing a common support scheme for all of its member states, which would be directly financed by the community's budget.[7] The free market for cereals was introduced in 1967, effectively opening a significant market for the large French cereal surplus, and leading to cross-border price adjustments between France and Germany. It was accompanied by a system of price support - export subsidies, import levies and internal intervention prices - jointly aimed at creating a buffer between world prices and European prices in order to support farm income. Intervention prices themselves were introduced as a series of commodity-specific fixed price purchasing guarantees for wholesalers, which would then indirectly influence 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.[9]

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 crops, rice, textile crops, 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, and came after years of unsuccessful attempts at reforming EU agricultural policy. It 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 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 passed 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)). The passing of the treaty hence compounded with the growing European budgetary crisis. While originally manageable, 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.

[9] Their high level compared to world market prices highlight the relative importance of income and production support in the CAP framework, especially if we compare them to the concurrent US system of price floors - where the floors were used as protections against large negative price shocks, but not meant to inflate actual output prices. These were implemented via a loan rate ensuring prices against market price fluctuations (Gardner (2002)).

agricultural budget grew exponentially with European agricultural capacity. Over 1980-1986, EU agricultural spending doubled. Following the collapse of the world cereal market in early 1991, expenses were expected to increase by an additional 25% in a single year. Indeed, because an important feature of price intervention was the subsidization of exports relative to international prices, large drops in these prices had very negative budgetary consequences. Figure A1 shows intervention stocks, spending and over-production projections from the EU Commission in 1990. The European Commission, the supra-national decision making body of the union was then strongly in favor of reform and subsidy cuts. Many member states opposed them however, among them France which remained one of the primary beneficiary of the policy within the union.

The emerging risk of a trade war between the EU and the US, 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.

As a result, a first reform plan was outlined by the Commission in 1991 – which aimed for the cutting of intervention prices for cereals, oil and protein crops to world price level, and for the introduction of a progressive compensatory subsidy scheme in the form of marginally decreasing payments per hectare. This progressive structure was scrapped with pressure from the member states, and the reform finally consisted of the following points. First, cereal intervention prices were moved to international price levels, while they were fully removed for oil and protein crops.[10] In Figure 1, we illustrate this part of the reform by showing the evolution of farm gate prices for a series of agricultural commodities. All the data is translated in 2020 euros. In red, we show the intervention prices, which corresponded to the price level at which the Union was providing a purchasing guarantee – these guarantees were mostly relevant for the secondary market (intermediary level) given the presence of minimum purchasing thresholds.[11] In grey, we show 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} centiles. We observe a clear co-variation in French farm gate prices and intervention prices.[12] Finally, the blue and blue dotted lines indicate US and Canadian farm-gate prices recovered from Faostat.

[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 our knowledge, this corresponds to the first complete series of EU intervention prices, which we recovered from commodity and year specific regulations.

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

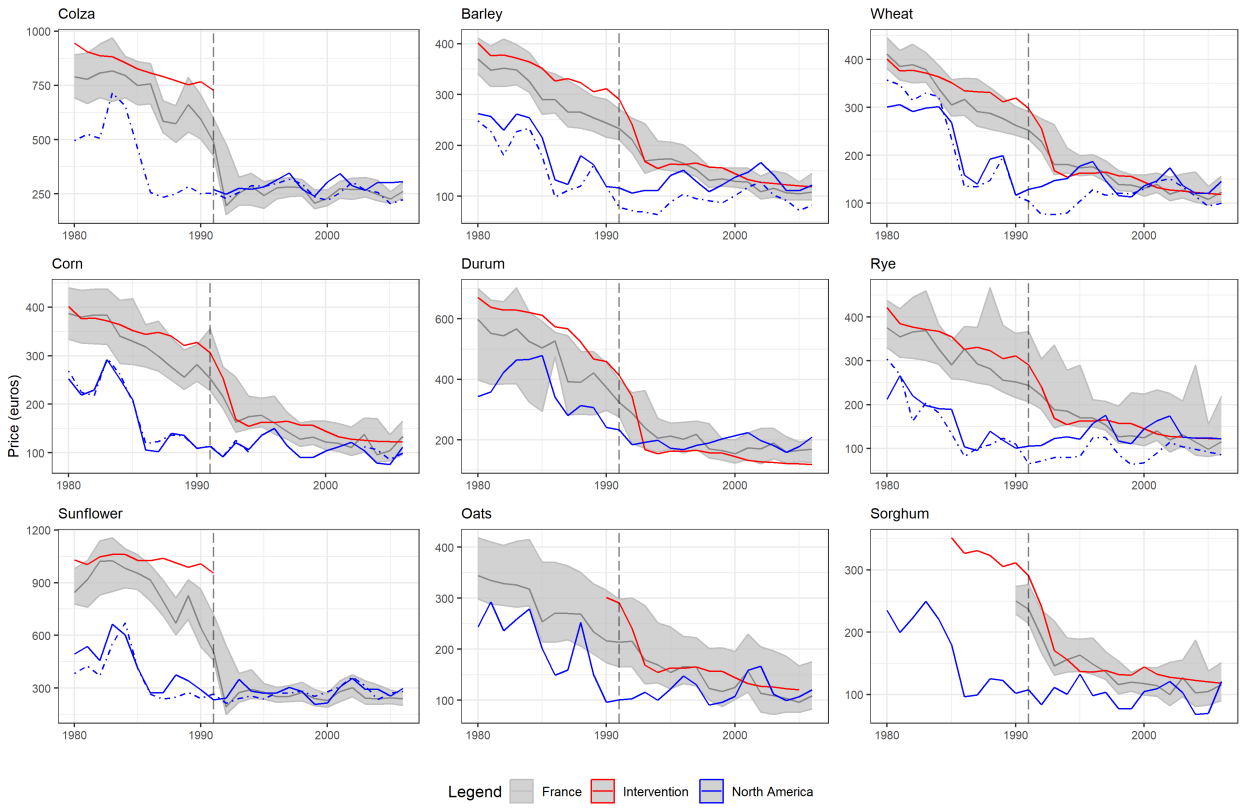


Figure 1: Convergence of French and North American Prices

Notes: French farm gate prices in grey are taken from the FADN and correspond to the weighted mean, bottom and top 25th and 75th centiles of the distribution of prices observed in the survey, where prices are recovered from the ratio of value to quantity sold. Intervention prices in red are taken from the relevant EU directives published over time. US and Canadian farm-gate prices are taken from Faostat. 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 seemed effective at maintaining a large wedge between French and international market prices, and that their removal led to that wedge to disappear. Prior to 1992, fluctuations in the gap between French and international prices are dominated by movements in international prices – and especially the 1984-86 drop in international prices (Morrison and Wattleworth (1986)). Around the reform, however, the closure of the gap between French and international prices is only driven by movement on the European and French side. This illustrates how losses in the international-French wedges surrounding 1992 are likely reform-driven. Second, we note that the size of these wedges significantly varied across crops. This implies 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. We build an average output price across row crops within farms, using land shares for the aggregation, and show the associated one-number summary in Figure A2. This is useful to see that while prices have a similar slow decreasing trend pre and post-reform, there is a large jump happening between 1991 and

1993, where prices are on average reduced by 33%.

The second step of the reform consisted in the introduction of land use subsidies in the form of payments per hectare, which acted as a partial compensation for the removal of price intervention. These subsidies were considered to be sufficiently de-correlated from production because they would not scale with yields but only land extent.[13] We show the evolution of average land subsidies as measured in the FADN in Figure A3. These were computed as the product of an EU-wide price, multiplied by a regionalized historical yield. 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.[14]

Finally, the reform implemented an indirect 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 should be left fallow for them to receive subsidies, this fallow land being also compensated for by the land subsidy.[15] Small farms were not subject to this set-aside policy. In our reduced form section, we will account for the heterogeneity in land subsidization induced by this set-aside requirement by measuring farm-level average land subsidies per hectare – where land is both land directly used in production and land set-aside. We also control for the farms' 1991 land use. In our structural section, we will model a land subsidy without set-aside requirements.

We end by showing evidence for the likely large scale reallocation in production happening in the decade surrounding the MacSharry reform. We show trends specific to the row crop market, which is the one we will focus on – but give a comparison with the trends of the general agricultural market in Table A3. We also provide level values there.

[13] In that sense, they related to post-1996 US deficiency payments also introduced after the passing of the Uruguay round.

[14] Reduced form result using this design to investigate the impact of the land subsidy on farm outcomes are available upon request.

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

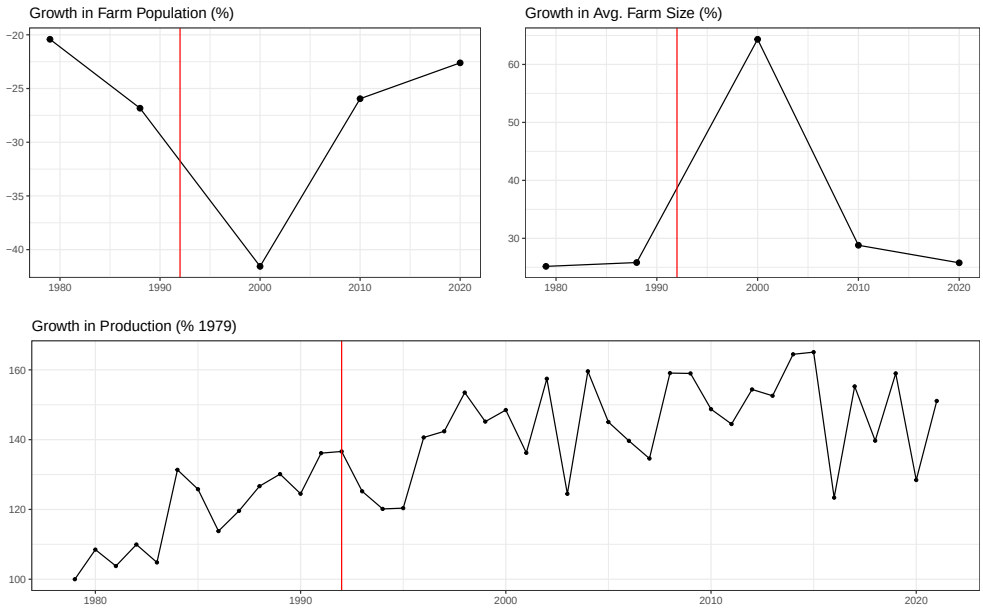


Figure 2: Agricultural Dynamics – Row Crop Market

Notes: The data is computed using the full count agricultural census for years 1970, 1979, 1988, 2000, 2010 and 2020. The data for the third facet comes from Our World in Data. The exit rate shown on the left is computed as the negative of the growth rate in the number of farms registered in France, with the $t-1$ value as baseline. The average farm-size growth rate corresponds to the unweighted average farm-level total used agricultural area growth rate. The third graph shows total production in cereals, measured in tons and expressed with respect to the 1979 baseline level.

The two first facets of Figure 2 show growth rates in between the Census waves, hence 1970-1979, 1979-1988, 1988-2000, 2000-2010 and 2010-2020, while the last one gives the evolution of production at yearly intervals.[16] The first graph shows that while the number of row-crop farms has been heavily decreasing over time in the last decades, there is a sharp jump around the period of the reform. Between 1988 and 2000, the number of producers goes from 602k to 352k – accounting for over two-thirds of the total decrease in the French farm population. This wage of exit comes with a sharp rise in farm size observed in the second facet (a truncation of the left tail, or a rightward shift of the entire farm size distribution), with an average farm size growing from 40ha to 65ha, and an average cultivated area from 19ha to 32ha. Table A3 shows these movements are larger in the row crop market, than in the overall agricultural market. Finally, we note that the gini coefficient for the distribution of farm size remains constant. This should imply that the growth in farm size is homogeneous over the distribution. It is interesting to note that the gini coefficient for farm land size over the entire distribution of farms – not only row crop ones – goes up, implying that this growth in the row crop market is not matched by a similar growth in size outside of it. Figure A5 adds

[16]The Census is the most precise source of data for land use for the late 20th century, while production data is more easily accessible at a yearly frequency.

to this picture, and shows evidence for farm-level specialization – farms showing decreasing crop-mix complexity over time, with both decreases in the number of crops they grow, and a relative concentration of their land allocation within the crops that they keep growing. We give in Table A4 a more detailed accounting of the farms' crop mix composition – which we discuss in more details when we estimate our model of the market. We conclude by noting that total production immediately collapses after the reform, but starts increasing again from 1995 onwards. Bringing all of these together, the timing of the reform seems associated with an exceptional amount of farm exit, increases in average farm size, but no decreases in total production capacity. All of these point to likely large-scale reallocations in production across establishments, as well as increases in production efficiency.

2.3 Data

We describe the two main datasets used for our analysis here, and leave the others to the annex in subsection B.1.

2.3.1 Farm Accountancy Data Network (FADN)

Most of our analysis will rely on a repeated panel of agricultural production developed to be representative of European commercial farming. We will extensively use the French subset of the dataset, but will also perform some analysis on its EU-wide counterpart, noting that other data sources are only available for France, preventing the extension of our entire analysis to the EU. The FADN was established in 1968, and for France currently samples around 7,000 farms a year.[17] 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 us to backtrack farm-level crop prices. It also contains farm-level input usage, crop-specific land allocations and capital stock data. Finally, while geocoded at the department level, its sampling is stratified at the region level. We will then be able to match the farm data to department-level input prices and weather data, but will rely on region-level fixed effects for estimation.[18]

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

[17] 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.

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

data would be used for fiscal purposes – as is thus understood to be significantly more precise since. The census is fielded approximately every ten years - and for our purposes in 1970, 1979, 1988, 2000 and 2010.

The French Census is notable in that it provides disaggregated data at the farm-level, and 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.

We note however that the Census 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.[19] We will then use the Census to track net farm exit, and the distribution of farm size and crop mix composition at the village level. The Census will then be useful to get a clear and externally valid picture of agricultural market dynamics, which will complete the data coming from the FADN.

3 Reduced Form Analysis: Consequences of Decreased Agricultural Subsidization

We describe in this section the impact of the 1992-1995 decrease in agricultural subsidization on farm-level and village-level dynamics in France. Using a farm-level panel, we show how increased exposure to the reform translated into decreases in output prices and sales, and in spite of targeted increases in subsidies, in decreases in overall profits as well. We also show that the reform is followed by a decrease in the within-farm intensity of chemical use, measured either by total chemical use, chemical-to-output and chemical-to-land ratios. This is likely caused by the post-reform subsidization being a land subsidy, which shifts the farms' input choice incentives. We then highlight how sales and profit effects are likely selective – by which we mean heterogeneous across important farm-level characteristics – as the largest price and profit effects are concentrated among smaller farms. Relying on the French agricultural census, we then go further by showing higher local exposure to the reform led to farm exit at the lower tail of the farm size distribution. This section sets the stage for the structural analysis, by highlighting the importance of two channels. First, when subsidies shift the profitability of farming, they will prevent or encourage exit and likely change the selectivity of the market. If productivity is in part input-using, this selection channel can impact the prevalence of externalities related to input use. Second, subsidies that distort the incentives for input choices will likely impact the prevalence of specific externalities related to input use.

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

3.1 Empirical Strategy

Model Outline: We start by outlining the general structure of our model, as it will allow us to clearly present our identification argument. A farm j in period t is characterized by a type $\Upsilon_{jt} = \{\omega_{jt}, K_{jt}, S^o_{jt}, \mathbb{C}_{jt}\}$: ω_{jt} is its productivity (potentially a vector), K_{jt} its capital stock, S^o_{jt} its land stock, and a crop mix $\mathbb{C}_{jt} \subseteq \mathbb{C}$, with $\mathbb{C}$ the entire set of crops grown in the row crop market. Every period, the farm will observe the evolution of its productivity, observe a series of demand shocks that we write as $\{D_{jct}\}_{\mathbb{C}}$, and then decide on its production levels for its given crop mix, as well as whether to stay on the market, and if it does, its next period capital, land and crop mix. Throughout, we will assume farms are price takers, although subsection B.9 in the annex extends the model to a monopolistic competition framework. Production levels as well as input mix composition decisions are set to maximize a flexible profit that we write as follows:

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

We also make the simplifying assumption that farms are not constrained in production by the land they own. Farms have access to perfectly functioning land rental markets, where the land they own and the land they rent have the same rental price.

European Policy: EU policy enters the problem in two ways. Price intervention happening on the secondary market will appear as demand shocks pushing up the crop prices faced by farms. We note these shocks $\{D^{eu}_{jct}\}_{\mathbb{C}}$. Land subsidies enter through a decrease in the effective price of used land, here the wedge τ^c_s, which also varies across crops. We assume this subsidy does not pass-through into the equilibrium price of land, which is a simplifying assumption.[20] We will control for the farm's 1991 land use and directly measure average land subsidies per hectare as seen in the data, in order to adequately impact for the interactions between the land subsidy and set-aside requirements.

Approximating European Policy: We can consider that the role of EU agricultural policy is to shield European farmers from world market prices. Specifically, we can write the crop-year cushions in France that are introduced by price intervention ($cushion^q_{ct}$) and by land subsidies ($cushion^l_{ct}$) as:

$$\begin{cases} cushion^q_{ct} = P^{inter}_{ct} - \overline{P}^I_{ct} \\ cushion^l_{ct} = \frac{\overline{Sub}_{ct}}{\overline{Yield}_{ct}} \end{cases}$$

Here P^{inter}_{ct} is the intervention price for crop c in France, and $\overline{P}^I_{ct}$ is a world market average

[20] See Femenia et al. (2010) for a careful accounting of the role of land subsidies in increasing land values.

farm gate price.[21] $cushion^q_{ct}$ is simply the difference between these two, and is an approximation of the true farm-level effect of intervention. If we could decompose demand shocks into the ones shared at the world level ($\{D^I_{jct}\}_{\mathbb{C}}$), the ones driven by EU policy ($\{D^{eu}_{jct}\}_{\mathbb{C}}$), and others ($\{D^o_{jct}\}_{\mathbb{C}}$), then:

$$P_{ct}(D_{jct}) = P_{ct}(D^I_{jct}\, D^{eu}_{jct}, D^o_{jct})$$

$P_{ct}(D^I_{jct}\, D^o_{jct})$ is then the correct counterfactual at the farm level, which we approximate with $\overline{P}^I_{ct}$. We note again that the main regulation introducing the MacSharry reform also explicitly stated its goal to have European prices converge to international levels post-1992.[22] $cushion^l_{ct}$ translates the land subsidy, here $\overline{Sub}_{ct}$ – an average subsidy per ha as observed in the FADN, into a price by dividing it by an average yield – also observed in the FADN. $\overline{Sub}_{ct}$ itself is an approximation of $\tau^c_s P^s_{jt}$ across French farms. These two cushions approximate the impact of the reform on crop-specific subsidization. We show the cross-temporal variation of these two cushions for two key crops – sunflower and wheat – in Figure A8. The heterogeneity across crop cushions pre-reform will justify the use of our shift-share design.

Building the Instrument: We can sum these cushions, and obtain an average one at the farm-level which will depend on the farm's crop mix composition.[23] Setting our initial period to $t_0 = 1991$, hence crop decisions made in the fall of 1990 before any outline of the reform was set forward, we use the following aggregation formula:

$$Cushion_{jt} = \sum_{c \in \mathbb{C}_{j,t_0}} \frac{s^l_{j,c,t_0}}{\sum_{c'} s^l_{c',t_0}} \left\{ cushion^q_{c,t} + cushion^l_{c,t} \right\}$$

Here s^l_{j,c,t_0} gives the land area allocated to crop c by j at t_0.[24] Robustness tests and alternative designs – including for the instrument design – are discussed in subsection B.3.

Ex-ante farm-level exposure to the reform can then finally be captured by $Exp_j = Cushion_{j,t=91} - Cushion_{j,t=95}$. Figure A9 shows how the main driver in cross-crop heterogeneity in subsidy loss was the level of the 1991 price wedge - itself arguably driven by an accumulation of thirty years of price intervention decisions and world-market price variations. This justifies using these shocks as exogenous ones. There is however a limited amount of shocks explaining most of the variation in Exp_j, preventing us from adopting the approach outlined in Borusyak et al. (2022). The MacSharry reform was not a set of many uncorrelated

[21] This world price is computed as a weighted average of farm gate prices across the main non-EU agricultural prices, with their relative share of world exports as weights. Farm gate prices come from Faostat, and we include the following countries: Canada, India, Indonesia, Japan, USA, China, the Russian Federation and Ukraine. The export shares are shown in **??** for the main four crops.

[22] 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"*.

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

[24] We perform a robustness check test using resp. land shares and output shares for resp. the land subsidy and price-intervention driven cushions. Results are similar, while this complicates the reading of our Rotemberg weights, which is why we use this simpler design for our main specification.

as-good-as-random shocks, but rather a few exogenous shocks with varying intensities that interacted with varying cropping patters at the farm-level. Hence, our shift-share 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 compares 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). We discuss the validity of this identifying framework in turn at the farm and village levels.

3.2 Farm-Level Analysis

3.2.1 Design

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

$$\Delta^{1991} Y_{jt} = Exp_j \theta_t + X_j \delta_t + \eta_{r(j)t} + \varepsilon_{jt} \tag{2}$$

$\Delta^{1991} Y_{jt}$ is the outcome of interest differenced with its 1991 level. With this formulation, every θ_t will give us the effect of exposure in a given year relative to 1991. Exp_j is our 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 A11 and Figure A10 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 – making it important to control for local trends. This is what we do by including region-by-year fixed effects in the form of $\eta_{r(j)t}$.[25] Finally, X_j is a vector of farm-level controls set to pre-reform values to avoid any confounding effect.[26] We note that we will cluster standard errors at the region-by-year level.

In order to causally interpret the $\{\theta_t\}$, we need for Exp_j to be conditionally exogenous. We have previously argued for the exogeneity of the shocks, and that their scale was directly linked to pre-1991 cushion sizes. We add that the exact timing of the MacSharry reform was driven by factors external to the French agricultural market, as outlined in subsection 2.2. We also need for the shares to be exogenous – specifically for the drivers of crop choice in fall 1990 to be conditionally uncorrelated to the factors impacting outcome growth. Going back to Equation 1, we see there are two sets of drivers of crop choice. The first set relates

[25] While we can locate farms within departments, the geographic unit below regions, the FADN survey is stratified at the region level. This combined with the balancing of the panel for the period of the event study makes for a very uneven balancing of farms across departments. Region-by-year fixed effects correspond to more realistic fixed effects which we can recover without too much error.

[26] X_j contains the following elements: their 1991 capital stock, total labor used, total land use, the local price of land and local hourly wage for agricultural workers, the evenness of their land allocation across crops, their fertilizer and pesticide use, and the share of their production which corresponds to oil crops (colza and sunflower). We add this element in order to compare farms with relatively similar takes in oil crop production. We expect farms who only grow cereal crops to differ quite significantly from those that also grow oil crops. We also control for the farms' 1983-1984 adoption trends in chemicals measured as their evolution of chemical-to-land and chemical-to-output ratio growth.

to the farm's state. In a multi-product setting like Mayer et al. (2014), a higher capital stock or productivity should imply a richer and more diverse crop mix. This would both directly change exposure and directly impact the farm's dynamic outcomes. The second set of drivers corresponds to varying input prices (combined with cross-crop heterogeneity in input shares) and demand shocks.

We propose to control for the endogenous and unobserved determinants of 1991 cropping decisions by controlling for the farm's state in 1991. We do so by controlling for all of its observed flexible input decisions that year (labor, fertilizer, pesticides), its land use, its capital stock and the evenness of its crop mix composition. Static elements of the state with constant effects should be taken care of by the differencing of the outcome. The fact that flexible input decisions in 1991 should be useful to control for the farm's state then relates to the identifying assumptions used in the production function literature – 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. We further control for the relative share of output coming from oil crops (sunflower and colza) – as we expect this characteristic of crop mix composition to be the most related to underlying unobserved heterogeneity across farms – while farms growing cereals but different ones within the category should be more similar. Because we also look at outcomes related to chemical use, we also want to make finer comparisons across farms with similar production technologies pre-reform. We thus control for the farms' 1983-1984 adoption trends in chemicals measured as their evolution of chemical-to-land and chemical-to-output ratio growth. All of these controls are interacted with yaer-specific coefficients. Auto-correlated shocks in demand and input prices are addressed with region-by-year fixed effects, and by controlling for the department 1991 land prices and agricultural hourly wages.[27] 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 region. We argue first that our set of farm controls should be sufficient to capture all first-order farm-level unobserved dynamic elements, and second that the French row crop market is sufficiently homogeneous and competitive for region-by-year fixed effects to be adequate controls. Formally, we want shares in 1991 to be orthogonal to our residual drivers of outcome growth, or $\mathbb{E}[\varepsilon_{jt}|\{s^l_{jct}\}_{\mathbb{C}_{jt}}] = 0$.

We perform all the analysis in this section on a balanced panel of farms, present from 1984 to 2002, in order for changes in outcomes to be purely driven by within-establishment trends. The structural analysis will encompass selection effects, and across-farms reallocations.

Figure A12 shows balance tests for farms with an above or below median exposure for a series of farm characteristics of interest. Importantly, we see that profit shows parallel pre-trends prior to the reform across the two groups, but that more-exposed farms have much larger profit levels – and that this wedge completely disappears post-reform. This is the

[27] These two input prices also have year-varying coefficients.

effect we are interested in capturing. The first-differencing of outcomes should remove these differences in levels, and allow us to capture sharp changes in slope if they are correlated with exposure. We further see in these figures that capital, land and labor trends diverge between the two groups prior to the reform, and that it is hence important to control for them in our regressions. We provide further balance tests – looking at variations in trends across farms with larger or smaller crop-specific land shares – and a series of alternative designs used as robustness tests in subsection B.3.

3.2.2 Farm-Level Results

We show below the event studies obtained using the design from Equation 2, and when looking at the farms': average output price, sales, profit and chemical use (all in log terms). Chemical use is measured as the sum of the resp. deflated bills for fertilizers and pesticides, each deflated with an input-specific input price index. Our measure of profits includes subsidies and is thus a correct measure of the overall effect of additional exposure to the reform.

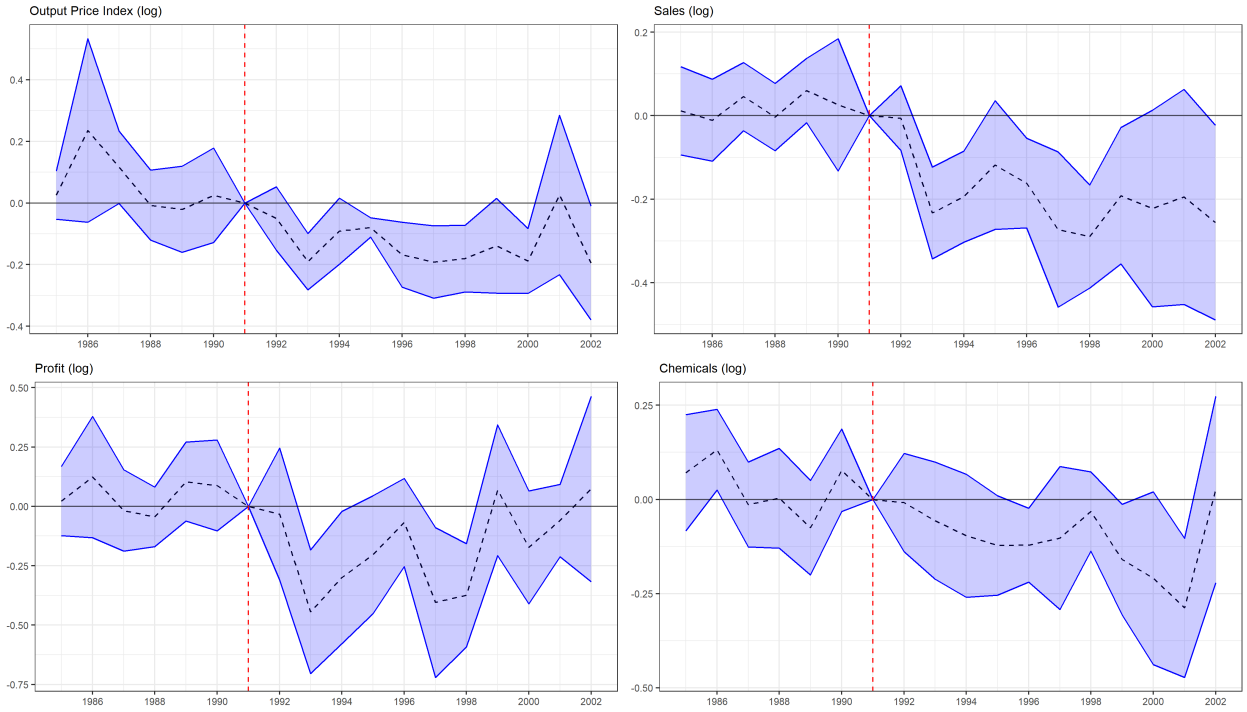


Figure 3

Notes: We plot graphically the result of our 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' chemical-to-land ratio where chemicals are measured as the deflated farm-level input bill, and land is measured in ha. 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. We control for: 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, and the evenness of their crop mix in 1991. We also control for their 1983-1984 adoption trends in chemicals measured as their evolution of chemical-to-land and chemical-to-output ratio growth. We allow for a time-varying intercept for each of these controls. We add region-year fixed effects, and cluster the standard errors one-way at the region-year level. The table for these results in Table A6.

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 16% decrease in the late 1990s.[28] Exposure is standardized here, hence a one standard deviation increase in exposure leads to this additional 16% decrease in average output price. The reform itself manifested as 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 row crop prices between 1991 and 1993, which serves as a good benchmark for the potential total effect of the reform. This implies a lot of heterogeneity in exposure, compared to the potential average effect. The second facet of the figure shows this decrease in prices is followed by a similar decrease in sales – a one standard deviation increase in exposure leads to a 17% decrease in sales. We

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

then look at farm profits – there a one-standard deviation increase leads to a 16% decrease in profits. Table A12 gives the difference-in-difference results matching the event-studies regressions. Lastly, we show the impact of exposure on farms' chemical use, here their total use of chemicals in production. A one standard deviation increase in exposure leads to a 16% decrease in chemical use. Additional event-study results in Figure A31 show this decrease does correspond to a change in the intensity of chemical use, and not just a decrease in the scale of production. Indeed, the farms' chemical-to-output ratio subsequently decreases by 25%, and their chemical-to-land ratio by 35%. The chemicals-to-land ratio is a useful target here, as it should both be orthogonal to changes in TFP, and in scale. This likely indicates that as land subsidy increases, farms switch away from using chemicals towards land.

In subsection B.3, we also compute the Rotemberg weights associated with our shift-share instrument, and use them to decompose the source of the variation used for identification. We show the sales, subsidies and profit regressions with crop-specific instruments to test the homogeneity of the effects across crops. Finally, we propose an alternative way to compute our exposure instrument, and replicate our findings with it.

3.3 Village-Level Analysis

3.3.1 Design

We move to the Census of Agriculture to study the consequences of the reform on farm exit, and the use and local distribution of farm land. The census is more appropriate for this study, as it covers the entire population of French farm and thus provides results with a greater external validity. Prior results were based on the FADN as the Census does not contain any accounting data.

We will aggregate our measure of exposure at the village level, to account for the fact that the Census does not allow us to track farms across waves.[29] This aggregation can be performed in different ways. We are interested in the effect of having more highly exposed farms within a village on village outcomes. As farms are quite heterogeneous within a village, we choose a measure of exposure that should be robust to having this heterogeneity drive most of the variation in our instrument, i.e. having one large farm with a very large exposure driving total village exposure upwards. When we look at exit specifically, having a unique farm drive the average village exposure level is unlikely to increase exit beyond one, while having many farms having a lesser but still significant exposure will. Our main aggregation approach should allow for a more monotone tracking of exposure-intensity across farms to farm-level outcomes.[30] Hence, our measure of exposure will rely on the 1988 within-village share of farms having an above-median exposure to the reform, which we write $Share_k$ for village k. $Share_k$ is the relative share within the set of farms having any exposure to

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

[30] In 1988, the mean across villages of the within village standard deviation in farm size was 33.5ha for about 27.7 farms. The coefficient of variation in exposure was .28 – but ranging from .08 ($10th$ centile) to .53 ($90th$ centile).

the reform – and we will control for the share of farms with no exposure (farms that do not operate in the row crop market). The median itself is computed for all France.[31] In the appendix Table A17 we compare results using this measure, the average exposure and maximum exposure within the village. While results on land use are quite homogeneous across aggregations, farm exit responds only to our share measure, and to the maximum exposure within the village. In the share case, the analysis corresponds to a difference-in-difference design with a continuous measure of treatment ranging from 0 to 1. We show in Figure A13 the distribution of average village exposure, and of the share of farms with an above median exposure, and in Figure A14 the truncated and non-truncated geographic distributions of exposure. These comparisons are useful for two things. The first one shows how our measure of exposure averaged at the village level shows a high upper tail, mostly concentrated along the Mediterranean coast of France.[32] Using our share of farms with an above median exposure keeps this geographic feature, but smooths the distribution. Second, we see that the patterns remain similar between the two measures – with the Massif Central, Brittany and the North of France receiving lower treatment and having a lower share of farms with an above median exposure.

We show balance tests for this measure of exposure in Figure A16. Farm count and average farm size show similar trends prior to the reform, while village evenness and total village agricultural area differ slightly. More exposed village have a stable level of crop evenness, while those below have a decreasing one over 1979-1988. Total village agricultural area also decreases more quickly pre-reform for less exposed villages. It is as such important to control for pre-reform village characteristics in our analysis.

With this in mind, we will use the following design – for village k in year t located in county c, $\Delta^{1988}Y_{kt}$ denoting outcomes differenced with village-specific 1988 levels:

$$\Delta^{1988}Y_{kt} = Share_k\theta_t + X_k\delta_t + \eta_{c(k)t} + \varepsilon_{kt} \tag{3}$$

X_k accounts for the following 1988-valued village-level controls: share of farms not operating on the row-crop market, the total agricultural area of the village, the average and maximum farm size, the total area allocated to the crops we use to build exposure, the standard deviation in farm size. We cluster our standard errors at the department level.[33] In each year of the census, we compute our outcomes only for the farms active on the row crop market.

[31] We normalize the shares, so that they are computed only among the set of farms with a positive exposure to the reform. Other farms are likely not active on the cereal-oil-protein market, and not a good point of comparison. We however include the share of farms with zero exposure (out of the total number of farms) as a control in all our regressions. We also compute all our village-level variables uniquely among the set of farms growing at least one oil, protein or cereal crops – out of the maximal set we can track over the waves of the Census.

[32] Histograms for the distribution of exposure are shown in Figure A15.

[33] There are 101 departments in metropolitan France, which means this level of clustering should avoid the issue of having too few clusters.

3.3.2 Village-Level Results

We obtain the following results – where the red vertical line indicates the omitted 1988 census which serves as a baseline as the closest pre-reform wave of data collection. Outcomes are in order: the evolution of the farm count at the village level, average farm area within the village measured in ha, and the $99th$ centile of the within village farm size measured in ha. All are first-differenced, and hence represent a trend rather than a level. θ_t gives us here the effect of moving $Share_k$ from 0 to 1, in 1980, 2000 and 2010 on the growth rate of our variables of interest relative to its 1988 value.

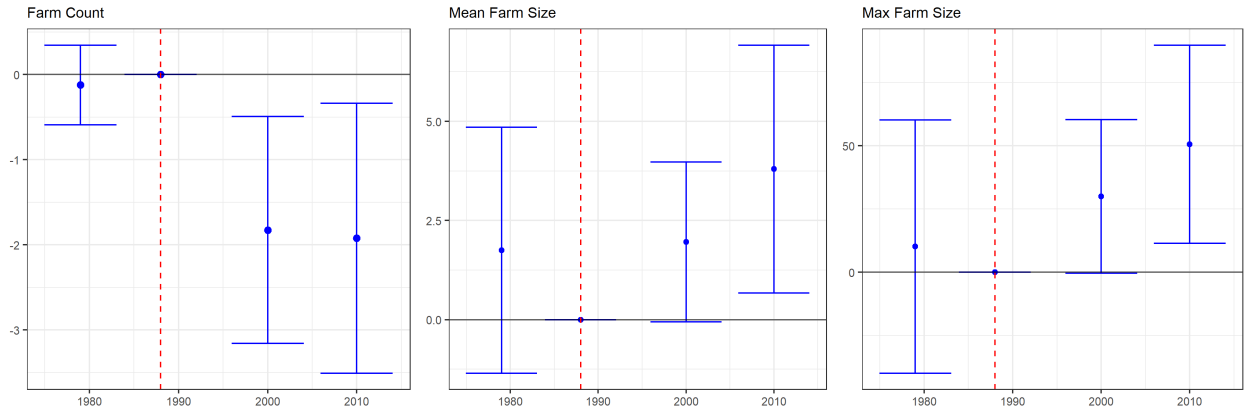


Figure 4: Village-Level Outcomes

Notes: This figure gives the results for our village-level event study. Outcomes are first-differenced. Farm count corresponds to the number of farms operating in the row crop market within the village, mean farm size corresponds to the unweighted average farm size within the village measured in ha, and the max farm size is the size in ha of the top centile of the within-village farm size distribution. The regression includes a series of controls set to their level in 1988 within the village, and interacted with a time-varying coefficients, as well as county-by-year fixed effects. Standard errors are clustered at the department level. The table associated to these results is Table A7.

As for the farm-level event study, the fact that exposure is non-responsive in the pre-period builds confidence for the design. The estimated effects in 2000 of moving from a share of 0 to 1 of farms with an above-median exposure are: an additional decrease in farm count of 1.8 (when the average decrease over that period is 7), an increase in the average farm size of 1.9ha, and an increase in maximum farm size by 30ha. We show results for additional outcomes in a difference-in-difference format in Table A15 and Table A16. Of interest, additional exposure also leads to an increase in total village agricultural area, and in the mean farm-level evenness of land allocation (though the village-level one remains constant). Because small farms are less even that large ones, this is likely indicative of exit at the bottom of the farm-size distribution.[34] We additionally show results for several points of the farm-size distribution.

[34] See Figure A5.

This set of results is useful to paint a picture of the very local reorganization of agricultural production following an additional exposure to the reform – and as shown in the previous sub-section – a decrease in farm-level profits. As expected, increased exposure leads to farm exit – and given both that average and maximum farm size increase, and that average farm crop evenness increases while village-level evenness remains constant – we can expect exit to happen at the bottom of the farm size distribution (larger farms have on average a higher evenness). Additional exposure does not, however, lead to a decrease in local agricultural activity as land use does not go down, but up. Combined with the previous farm-level results, these indicate that large decreases in subsidization have important effects on market dynamics, and because they impact farms heterogeneously according to their size, are also likely selective and likely to shift market characteristics for which small and large farms differ significantly. As most firm dynamics models will predict, larger establishments should be more productive than others on average. Depending on the relation between environmental externalities and productivity, this relative higher exposure of small farms to the reform could lead to significant changes in the prevalence of environmental externalities within the agricultural market. Our structural model precisely aims at recovering this relation. This section has also highlighted the impact of the reform on within-farm dynamics, and set changes in subsidies as a useful source of variation to identify the parameters of the farm problem.

4 Structural Analysis

Our structural analysis contains three sections. The first one is a presentation of our model – in turn demand for agricultural commodities, which we model as static, and supply: a combination of static input choices for the stage game, and dynamic entry, exit, capital accumulation and crop choice. The second section outlines our estimation strategy, the recovered parameters, and their implications. In the third and last section, we present our structural analysis of the MacSharry reform, and a study of subsidy scheme optimality under existing EU budget standards.

4.1 Model Outline

The goals of our structural analysis are two-folds. The first one is to recover the structure of production heterogeneity across French farms. This heterogeneity will inform the empirical relation between average costs of production, and environmental externalities. We argue in subsection B.2 that focusing on the chemical intensity of production is a relevant way to orient our analysis of environmental externalities within agriculture – we show there evidence for the correlation between farmland extension and waterway pollution from agricultural chemical use, heterogeneity in the chemical-intensity of production across French farms and suggestive evidence for chemical-using productivity growth. Our model will yield a relation between productivity and the chemical intensity of production. This relation will be recovered by the

combination of two elements: non-Hicksian productivity shocks shifting the farms' efficiency at using chemicals in production, and an elasticity of production relating land and chemicals. As first order conditions will show, the elasticity will sign the correlation between shocks and the production's chemical-to-land ratio, controlling for varying input prices and the crop-specific land and chemical input shares. This relation is key for our identification procedure, and follows the logic of Doraszelski and Jomandreu (2018). The relation for the chemical-to-output ratio is not as simple, but this second ratio follows a similarly positive relation with the non-Hicksian shocks under the substitutability of land and chemicals.

The second goal of the model – achieved by nesting supply within a partial equilibrium dynamic model – is to understand the extent to which the productivity-externality relation matters and relates to reallocations in production across farms. Dynamics are important here, because farm growth takes time and the extent and speed of that growth is informed by farm-specific productivity shocks.[35] Farm-size matters because larger farms will account for a larger share of production, and their capital stock will protect them from negative productivity shocks. Capital growth itself is endogenous to internal variations in productivity.

4.1.1 Static Demand

In our model, the zero profit condition will establish a price surface such that at any point on the surface potential entrants are indifferent between entering and remaining outside of the market. Output market clearing then allows us to reduce this surface to the point where supply equals demand. The demand system we use for this is purposefully simple, and does not model substitution in demand across crops. The set-up of our production function estimation requires aggregating crops into categories. We 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 follows:

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

Where e_{ct} is an idiosyncratic mean zero demand shock.

Annex subsection B.9 discusses an alternative modelling path using a linear quadratic demand for horizontally differentiated agricultural commodities themselves located within crop markets (for ex. 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 which will be the central element in our stage game – and parameters related to capital investment, entry and exit decisions.

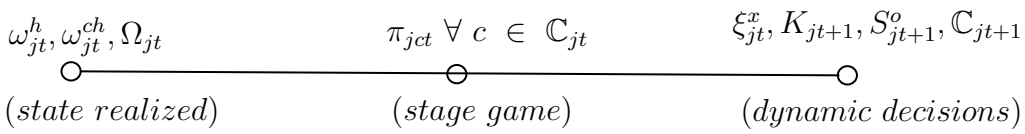
[35] Figure A20 shows the relation between farm size and farm age in the 2010 Census. For farms that entered post 1980, there is a clear positive and linear relation between tenure and size. We use the 2010 Census, as it is more likely indicative of trends within commercial farms, while previous Censuses will contain a large fraction of family farms with very different production and management behaviors.

Certain additional assumptions are only introduced after the estimation of the production function – to keep this first estimation as general as possible.

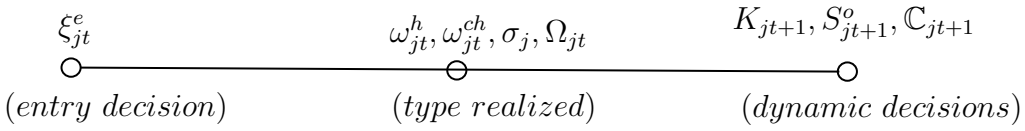
Agents & Timing: Farms act as single agents in a competitive row crop market with $\mathbb{C}$ differentiated crops. They have a type defined by the set: $\{\omega^h_{jt}, \omega^{ch}_{jt}, S^o_{jt}, K_{jt}, \sigma_j, \mathbb{C}_{jt}\}$, with resp. their TFP shock for their core competence, a non-Hicksian productivity shock, a land stock, a capital stock, a competence ladder, and a crop mix. The competence ladder structure allows us to recover crop-specific Hicksian shocks while observing most inputs at the farm-level. Specifically, farms have a competence ladder σ_j, which is a permutation of the crops in $\mathbb{C}$, and indicates how their unique TFP shock ω^h_{jt} becomes differentiated across products. We adopt a geometric structure for that competence ladder, such that the TFP for a given crop will be: $\omega^h_{jct} = \omega^h_{jt}\lambda^{\sum_{n=0}^{dim(\mathbb{C})-1} n1\{\sigma_j(n)=c\}}$.[36] Both ω^h_{jt} and ω^{ch}_{jt} are taken to be first-order Markov processes, exogenous to the farm's decisions. The state faced by a given farm is complemented by market states grouped in Ω_{jt}, which we will specify fully after describing the production process.

We assume the following timing for market incumbents and entrants:

Incumbent



Entrant



Incumbent Problem: The incumbent's problem can then be fully described by the following Bellman equation: with $\{X_{jct}\}_{\mathbb{C}_{jt}}$ the set of flexible inputs used in production, and

[36] This follows the approach of Mayer et al. (2014), where firms are endowed with a marginal cost type, and this marginal cost increases geometrically as firms add varieties to their mix. The addition of permutations σ accounts for the fact that we have separate markets which goods are not substitutable, while they model a unique market with horizontally differentiated varieties.

ξ^x_{jt} a dummy for the exit decision:

$$V(\omega^h_{jt}, \omega^{ch}_{jt}, S^o_{jt}, K_{jt}, \sigma_j, \mathbb{C}_{jt}; \Omega_{jt}) = \max_{\{X_{jct}\}_{\mathbb{C}_{jt}}, \xi^x_{jt}, K_{jt+1}, S^o_{jt+1}, \mathbb{C}_{jt+1}} \underbrace{\Pi_{jt}}_{\text{Static Profit}} + \underbrace{C(K_{jt}, K_{jt+1}, \Pi_{jt})}_{\text{Capital Adjustment}}$$
$$+ \xi^x_{jt} \underbrace{\left[(1-\delta_k)K_{jt}P^K_{jt} + P^s_{jt}S^o_{jt}\right]}_{\text{Scrap Value if Exit}} - (1-\xi^x_{jt}) \underbrace{f_k}_{\text{Cost of Incumbency}}$$
$$+ (1-\xi^x_{jt}) \underbrace{\left(\mathbb{E}_t \left[\frac{1}{1+r} V(\omega^h_{jt+1}, \omega^{ch}_{jt+1}, S^o_{jt+1}, K_{jt+1}, \sigma_j, \mathbb{C}_{jt+1}; \Omega_{jt+1}) | \omega^h_{jt}, \omega^{ch}_{jt}, \Omega_{jt} \right] \right)}_{\text{Expected Value of Continuation}} \tag{4}$$

We assume that land rental markets are perfect, that land owned is rented at the same price as the land used for production, and that there are no constraints to land use.[37] Crop choice is done at the end of the period for the next, and without facing any switching costs.[38] Static profit takes the same form as what was described in our reduced form section: the sum of rents from land rental, and profits coming from crop production:

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

The incumbent problem is key in that it sets the conditions for firm selection based on productivity. If productivity shocks are somewhat sticky, and a bad shock today makes a bad shock tomorrow more likely, negative shocks can induce farms to leave the market. This logic, and the role of productivity volatility in generating selection has been thoroughly studied by Collard-Wexler (2013) and Asker et al. (2014). The presence of a fixed cost of incumbency f_k, which we can interpret as the sum of fiscal and overhead costs unrelated to scale, makes that exit decision more likely. If $f_k = 0$, the cost of staying in one more period without producing, instead of leaving the market, boils down to the devaluation of one's capital stock which reduces the scrap value at exit. We will estimate f_k, which will be a key parameter to set the extent of the survival space for farms.

Entrant Problem: We assume that to operate, a farm needs to purchase an entry good that it then keeps throughout its lifetime. As such, potential entrants enter whenever the expected value of entry is larger that the price of this entry good. We write out the entrant's Bellman in subsection B.5. We assume that the entry good is produced by a

[37] When we later move to our numerical model, we will assume away the existence of owned land which can be rented. It is useful to keep it here, to clarify that capital corresponds to machinery and buildings, and that owned land is considered a separate source of dynamic farm value. We further assume that land rental markets are perfect, and that land is rented from and to others flexibly, without any aggregate land constraint, and at the same unique price.

[38] This is equivalent to saying crops are chosen at the start of the season prior to receiving the season-specific productivity shocks. The crop choice timing assumption will be key to address the selection issue introduced by only using single crop farms for the estimation of the production function – similarly to De Loecker et al. (2016).

competitive industry using labor, and that its production faces decreasing returns to scale, with $\alpha_e \in]0,1[$. Its price is f_t^e. We write:

$$Q_t^e = A_e L_t^{\alpha_e}$$

Firm dynamics models differ in their assumptions regarding entry costs, and Klenow and Li (2024) provide an overview of these.[39] In our case, the decreasing returns to scale assumption implies that the price of entry goods will be increasing convex in the mass of incumbents. The idea behind this modelling choice is that as the number of farms increase in France, the initial investment to open another one goes up. This captures the scarcity of the current stock of cleared agricultural land and built farm buildings. As more farms exist, new ones will likely have to settle in more peripheral areas where the initial investment to transform the land into functioning agricultural land are higher, and where farm buildings do not already exist and have to be built rather than purchased. Anecdotally, we show in Table A5 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 likely tracks with costs of entry into farming – and the fact that as total farmland decreases, it concentrates in relatively less mountainous and less remote regions. 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. In this case, the expected profit of entry is fully determined by the parameters of intervention and is orthogonal to the state of supply. The adjustment of the cost of the entry good will then provide a channel to pin down market size under intervention.

Production: Production decisions are made independently across production lines, with flexible inputs fully allocated across these lines.[40] 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 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.[41] We adopt the following specification, and use $\tilde{S}_{jct}$ to define the land-nest

[39] Klenow and Li (2024) discuss how some of the literature relies on fixed costs of entry (e.g. Hopenhayn (1992) and Hopenhayn and Rogerson (1993),Foster et al. (2008), Clementi and Palazzo (2016)), and models setting denominating fixed costs in units of labor – making them increase with growth (e.g. Grossman and Helpman (1991),Melitz (2003),Klette and Kortum (2004),Acemoglu et al. (2018)) – and discuss the consequences of both.

[40] 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). We discuss and estimate an alternative joint production framework where inputs are shared publicly across products – with a penalization – in Annex subsection B.6. We 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.

[41] 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, we consider that S_{jt}^o 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

aggregating land and chemicals into a unique effective input:

$$Q_{jct} = e^{\omega^h_{jct}+\varepsilon_{jct}} K_{jt}^{\alpha^c_k} L_{jct}^{\alpha^c_l} \Bigg[\underbrace{\Big\{\delta^c_s S^{\rho}_{jct} + (1-\delta^c_s)\Big(e^{\omega^{ch}_{jt}}\big\{\delta^c_p Pest^{\rho_2}_{jct} + (1-\delta^c_p)Fert^{\rho_2}_{jct}\big\}^{\frac{1}{\rho_2}}\Big)^{\rho}\Big\}^{\frac{1}{\rho}}}_{\equiv\ \tilde{S}_{jct}} \Bigg]^{\alpha^c_s} \quad (5)$$

We assume that $\alpha^c_k + \alpha^c_l + \alpha^c_s = 1\ ,\ \forall c$. In the case of production function estimation, the state variables in Ω_{jt} are the following: $\{Policy_t, \{P_{jct}\}_{\mathbb{C}}, \{P^x_{jt}\}_x, P^K_{jt}\}$ – the state of the CAP policy, a vector of crop prices, and a vector of input prices – all taken to be exogenous to the farm's decisions.[42] Output and input prices, as well as policy are assumed to follow Markov processes known to the farm and that we do not further specify.

Here ε_{jct} stands for a random productivity shock unobserved to all, ρ is the parameter of substitution between land and our chemical aggregator, ρ_2 is the parameter of substitution between pesticides and fertilizers, and the remaining α and δ parameters are crop-specific input shares. Elasticities follow from the usual formula: $\sigma = \frac{1}{1-\rho}$. We expect fertilizers and pesticides to be complements, and hence want to allow for ρ and ρ_2 to be different.[43] Land and chemicals will be substitutes if $\rho \in]0,1[$, and complements if $\rho < 0$. ω^h_{jct} is a TFP shock known to the farmer prior to making production decisions. ω^{ch}_{jt} is the non-Hicksian equivalent. An increase in its value will raise the effective amount of chemicals used in production – for a given amount of actual chemicals used. In a way $e^{\omega^{ch}_{jt}}\{\delta^c_p Pest^{\rho_2}_{jct} + (1-\delta^c_p)Fert^{\rho_2}_{jct}\}^{\frac{1}{\rho_2}}$ constitutes the farm's effort function for chemical use. We discuss later why our approach to the multi-product nature of farm production prevents us – under our current data environment – to model ω^{ch}_{jt} as product-specific. We note, however, that crops will still differ in their chemical-to-land ratio within a farm due to the crop-specific input shares. [44]

We use CES nests for land and chemicals, as they are a flexible way to allow for non-Hicksian shocks in a production function.[45] The outer Cobb-Douglas structure of our function relating capital, labor and the land aggregate simplifies the estimation, and gives us more stable coefficients than if we were relying on a third CES nest.

Our estimation of the production function will later rely on the combination of first-order condition coming from crop-specific profit maximization. When we combine the first

land, with a likely wedge in price between the two. This matches the ambiguous status of materials and of labor in Doraszelski and Jomandreu (2018), which are both computed using non-parametric aggregators of their static and dynamic components. When addressing the choice of instruments, we will be careful in only using lagged labor and land decisions to avoid any issues related to these dynamics from impacting estimation.

[42] Heterogeneity across farms in prices will then rely on exogenous factors like transportation costs, and for crop prices as well on fixed climate conditions that affect the quality of produced crops. We hence rule out monopsony power on the different input markets, and consider farms small enough that they face inelastic input supply curves. For input X, the farm faces price P^x_{jt}.

[43] This is highlighted in our discussion of contemporaneous technological changes in French agriculture in subsection 2.1.

[44] As mentioned above, our production function will vary across production lines. Different crops will have different input shares, on top of having different TFP shocks within the same farm. Farms will differ from each other in the production of the same crop depending on three internal elements: capital, TFP and chemical productivity. For the production function estimation, they will also exogenously face different input and output prices. All of these elements will play a role in defining how much chemicals a firm uses annually.

[45] See Doraszelski and Jomandreu (2018) for how CES provide a close approximation of non-hicksian productivity growth patterns. Both a Cobb-Douglas and a translog form would assume a degree of substitution across inputs, while we can recover it through estimation in the CES case.

order conditions related to optimal pesticide and land use in the maximization of each crop's flexible profit, we 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:

$$\begin{aligned}\rho\omega_{jt}^{ch} =& log\Big(\frac{P_{jt}^{p}}{P_{jt}^{s}}\Big) + log\Big(\frac{\delta_s^c}{\delta_p^c[1-\delta_s^c]}\Big) + (1-\rho_2)pest_{jct} - (1-\rho)s_{jct} \\ &+ (1-\frac{\rho}{\rho_2})log\Big[\delta_p^c Pest_{jct}^{\rho_2} + (1-\delta_p^c)Fert_{jct}^{\rho_2}\Big]\end{aligned} \tag{6}$$

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, Equation 6 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.[46] The MacSharry reform is likely to play a role along these three dimensions. The introduction of a land subsidy increased the relative price of chemicals, and the changing subsidization of crops also likely impacted farms' crop mix compositions. These are the two first elements of Equation 6. We saw in our farm-level event study that these led to a decrease in within farm chemical-to-land ratios. The third channel, if ω_{jt}^{ch} is indeed exogenous, will not respond to the reform directly, but reallocations and profit shocks related to the form are likely to favor the survival and market shares of farms with high ω_{jt}^{ch} at the time of the reform – hence increasing the average value of ω_{jt}^{ch} observed on the market. These joint selection and reallocation effects could counter-act the impact of the first two channels.

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. These assumptions are the following: we remove the possibility for farms to own land S_{jt}^{o} – and consider all land used is rented, we assume that all farms share the same competence ladder $\sigma_j = \sigma, \forall j$, and aggregate fertilizers and pesticides in a unique input which we 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 has the consequence of rendering the exogenous market state

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

Ω_t identical across farms.[47]

We introduce further notation, which will be useful to write the equation describing the evolution of the aggregate market state. The measure describing the allocation of farms across the state space is written μ_t. The incumbent and entrant problems give us two transition operators describing the evolution of a farm's type: $\phi_t(\omega', \omega^{ch'}, K', \mathbb{C}'|\omega, \omega^{ch}, K, \mathbb{C})$ describing the probability of evolution of a farm from $\{\omega, \omega^{ch}, K, \mathbb{C}\}$ to $\{\omega', \omega^{ch'}, K', \mathbb{C}'\}$ according to equilibrium policy decisions, and $\phi_{e,t}(\omega', \omega^{ch'}, K', \mathbb{C}'|\omega, \omega^{ch})$ for entrants after they draw their initial type.[48] We describe the value functions for both entrants and incumbents under our full set of assumptions, and write-out the expressions for these transition operators in Annex subsection B.5.[49] $G_1(.)$ 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:

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

The aggregate market state is then defined by $\Omega_t = \{\mu_t, N_t^e, f_t^e, \{P_{ct}\}_{\mathbb{C}}, \{P_t^x\}_x, P_t^K, Policy_t\}$ – with the last three elements taken as exogenous to the market. We denote by $Q_{ct}^S(\Omega_t)$ and $Q_{ct}^D(P_{ct})$ aggregate supply and demand for each crop. We denote by $\mathbb{M}_t$ the mass of firms implied by the measure μ_t. We define below our equilibrium concept, which is close to the one of Hopenhayn (1992).[50]

Equilibrium Concept: We will be looking 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. $Q_c^S(\Omega^\star) = Q_c^D(P_c^\star) \quad \forall c \in \mathbb{C}$. (**Output Market Clearing**)
- $f^{e,\star}$ s.t. $f^{e,\star} = Q^{e,\star}(\mathbb{M}(\mu^\star))^{-1}$. (**Entry Goods Market Clearing**)

[47] In this sense, there is no geography in our model, but for the possibility of chemical-intensity to reflect local soil and climate characteristics.

[48] The t under-script accounts for the fact that these transition dynamics can depend on the aggregate market state Ω_t: farmers of a same type will not make the same continuation, crop choice and investment decision under different price regimes.

[49] Our dynamic setting draws from the work of Hopenhayn and Rogerson (1993), and its extension to firms with capital as done by Clementi and Palazzo (2016).

[50] The algorithmic procedure used to identify the equilibrium is similar to that of Hopenhayn and Rogerson (1993), but for the fact that our fixed cost of entry adjusts with the mass of entrants. This breaks the linear homogeneity of the transition operator $T(.)$ used in Hopenhayn (1992) and Hopenhayn and Rogerson (1993). As a result, we directly look for a fixed point in the transition operator $T(.)$ for a given mass of entrants N^e and price pair that jointly satisfy market clearing and zero ex-ante profit. This fixed point is found conditional on a fixed cost of entry, and the monotonicity of f_e in the mass of entrants allows for an outer-loop which is a bisection-style search for that equilibrium cost of entry. Hopenhayn (1992) sets the conditions for a unique equilibrium with positive entry and exit to exist.

Role of Dynamics: Our farms make different dynamic decisions – capital adjustment, crop choice, and entry/exit decisions. They also face an exogenous varying state in the form of their TFP shock, over which they form expectations when it comes to making dynamic decisions. Entry and exit decisions are key to understanding selection based on type, as exit will be impacted by one's type and projection of future profits. Crop choice and capital stock matter for production – a farm's chemical use and chemical-to-land ratio will respond to its crop mix composition, while a farm's chemical-to-output ratio will decrease with its capital stock. Hence, these are key elements to account for, if we want to evaluate the impact of subsidy design on production externalities. Finally, capital and its relation to productivity shocks matter if we want 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.[51] As Collard-Wexler (2013) show, 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 and entry/exit – are then all key to recover how different subsidy design will interact with selection on productivity.

4.2 Estimation & Results

The estimation of the model is done for a market with two crop categories: $\mathbb{C} = \{wheat, other\}$.[52] 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.

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

[51] See Figure A20 showing the relation between farm size (in land use measured in ha) and their age in years of tenure. We 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. We measure size in land use here, as this is the only available relevant variable we observe in the Census.

[52] 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 our model to converge in a significant way.

4.2.1 Demand

We estimate two equations for demand, which take the following form – with lower case variables for logs:

$$\begin{cases} q_{wheat,t} = \alpha_{wheat} + \beta_{wheat} p_{wheat,t} + e_{wheat,t} \\ q_{other,t} = \alpha_{other} + \beta_{other} p_{other,t} + e_{other,t} \end{cases}$$

We estimate these equations using data from the FADN which we aggregate at the national level using sampling weights, we 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, we build measures of yearly local deviations from climate averages, that we 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. We also run a second design using only post-reform observations, and additionally using seasonal weather forecasts as instruments on top of realizations. We 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 A8.

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 we will be looking at the first-order conditions for crop-specific static profit maximization. Throughout this section, we will assume interior solutions to input choice decisions for profit maximization, and will hence remove FADN observations with zero flexible input use.

We will outline specific issues related to the FADN data structure and how they call for additional care when deriving estimating equations. We then proceed to the derivations giving us our two estimating equations, as well as the estimation procedure itself. The aim of decomposing the estimation in two steps is that we can first recover an equation where the non-Hicksian shock is the only endogenous unobservable – once recovered – we can then recover TFP shocks and remaining parameters in a second step with yet again a unique endogenous unobservable.[53]

Input Observability: Our analysis runs into the widespread issue of the level of input measurement in production panels. Besides land, all our inputs are measured at the

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

farm-level.[54] Our approach to this problem follows the method proposed by De Loecker et al. (2016), and we estimate our parameters using only data from single product farms. We hence bypass the input allocation question, and later use simple algebra and the problem's first order conditions – both described in Annex subsection B.7 – to recover the unobserved shocks for the multi-product farms. Contrary to De Loecker et al. (2016), and because we observe crop-specific land allocations, we can back-out these shocks while allowing for TFP shocks to be heterogeneous within a farm across its crops. We also note that using single product farms for estimation is likely to induce a selection issue – farms with more products are likely to be more productive.[55] Both Levinsohn and Petrin (2003) and De Loecker et al. (2016) argue that using an unbalanced panel is helpful in reducing this selection issue. For our product choice context, 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. We follow their dataset construction guidelines for all our estimation. We also extend their modification of the Olley and Pakes (1996) selection correction to our selection issue. This is done in subsection 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 our 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.[56] 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.[57] Given our context, we also assume that cross-farm differences in output quality will not be too large. Organic farming was less developed in the years surrounding the reform, and our general market is itself quite narrowly defined.

We observe both labor and land in volume, but only observe input bills for pesticides

[54] This implies that our 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 we do in Annex subsection 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)). Our case differs from these advances in that we 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.

[55] 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, we 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.

[56] See Foster et al. (2008) for an analysis of the difference between productivity measured using revenue or quantity data, and how both relate to the dynamics of firm selection.

[57] While the FADN contains both production, storage and sales data – we assume away the question of storage in our model.

and fertilizers.[58] We deflate these bills into volumes using input-specific Laspeyres price indices at the national level.[59] As we observe aggregate input bills for both our chemical categories, we 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, our 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: Our first stage relies on the expression given for the non-Hicksian shock in Equation 6, and follows the methodology of Doraszelski and Jomandreu (2018). This expression gives us a relation for the unobserved non-Hicksian shock depending only on parameters and observables. If we assume that the non-Hicksian shock follows a first-order Markov process, we can then create moment conditions on which can rely for estimation.

We write: $\omega^{ch}_{jt+1} = \mathbb{E}[\omega^{ch}_{jt+1}|\omega^{ch}_{jt}] + \zeta^{ch}_{jt+1} = g_{ch}(\omega^{ch}_{jt}) + \zeta^{ch}_{jt+1}$, and combine it with Equation 6. We can then approximate g_{ch} with a high-order polynomial. We build moments – with $A^{z,ch}_{jt}$ standing for our instrument matrix – and use them in a GMM estimation to recover the following parameters $\{\rho, \rho_2, \{\delta^c_s, \delta^c_p\}_{\mathbb{C}}\}$:

$$\mathbb{E}\left[\zeta^{ch}_{jt+1} A^{z,ch}_{jt}\right] = 0 \tag{7}$$

Second Stage: The estimating equation for the second stage relies on the first order condition for labor. It gives us a parametric function for the TFP shock that depends on observed variables and parameters that we will recover. We write it below. This expression is useful to build an estimate of the predicted part of the TFP process. [60]

$$\omega^h_{jct} = p^l_{jt} - p_{jct} - \mu - \alpha^c_k k_{jt} - log(\alpha^c_l) + (1 - \alpha^c_l) l_{jct} - \alpha^c_s \tilde{s}_{jct} \tag{8}$$

We would ideally like to follow an estimation procedure similar to the one used for our first step. That is, we would like to 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

[58] We observe respectively labor in hours, and land in hectares.

[59] These Laspeyres indices are themselves built using an agricultural price survey to which we 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 we 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.

[60] Production decisions are made in expectation of the random shock ε_{jct}, hence it is its expectation which enters the equation – with $\mathbb{E}[\varepsilon_{jct}] = \mu$

log of the production function and this Markov process) to form moments, where Equation 8 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 we 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 we use Ξ_{jct} to denote whether farm j only produces c in t, then the actual process we observe for the TFP in our 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 us to use an uncorrected specification of the non-Hicksian process. Here, we need a correction, and rely on a procedure which follows the lines of Olley and Pakes (1996) and De Loecker et al. (2016). We describe it in subsection B.8. With our additional control for selection, we denote $A^{z,h}_{jct}$ our matrix of instruments – we use the following moments to recover $\{\alpha^c_k, \alpha^c_l, \alpha^c_s\}$:[61]

$$\mathbb{E}\left[\left(\xi^h_{jct+1} + \varepsilon_{jct+1}\right) A^{z,h}_{jct}\right] = 0 \tag{9}$$

Estimation: For both steps, we recover our parameters using the generalized method of moments (GMM). We absorb the parameters related to the approximated functions $g_{ch}(.)$ and $g_h(.)$, following the approach of Wooldridge (2010) and Doraszelski and Jomandreu (2018). Our 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 our second stage, we 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.

We 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: We implement the procedure of Andrews et al.

[61] μ is absorbed within the intercept of the Markov process we use to model the TFP.

(2017) in subsection B.11 to shed some light on the respective role of our moments in driving the estimation of our production function parameters. We recover the sensitivity matrix associated with each of our estimation steps, which allows to understand the extent to which any moment-specific failure of exogeneity would impact our estimates. For our first step, the analysis shows our 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 our 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.

4.2.3 Costs of Capital Adjustment

We start by parametrizing the costs of capital adjustment. Most models of the adjustment costs literature focus on manufacturing, such as Cooper and Haltiwanger (2006), Asker et al. (2014) and Johnston (2020), and include non-convex costs of adjustment. We expect the nature of farm capital investment to significantly differ from those of plants. In the FADN, roughly 40% of farms' capital is machines, and 60% corresponds to buildings. Because this capital is not needed at every point of the production process, and because agricultural production is seasonal – we do not expect large capital investments to be disruptive of the production process – and hence do not model non-convexities in capital adjustment. Cooper and Haltiwanger (2006) discuss extensively the role of convex costs of adjustments in helping to replicate investment patterns. Convex costs will smooth investment over time. and justify farms investing in smaller more frequent amounts, rather than in a large discrete fashion. The less convex the costs, the more investment rates will correlate with productivity shock variation. We write:

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

We have four 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 us 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. We will rely on indirect inference for estimation. This means that we will 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 we have two categories of parameters, and hence will need two categories of moments. Given the already parameters we already estimated for our 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 us recover the effective cost of entry for that equilibrium, but not $\{A_e, \alpha_e\}$. To recover them, we 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.

We 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. We 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.

Our indirect inference procedure could then directly recover these four parameters, targeting the necessary moments for it. We however decide to use a two step procedures, which reduces the dimension of the search space and, given our model is slow to converge, significantly increases the precision we can reach with the estimation. We 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, we 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 $\{C_k^Q, f_k\}$. We can thus focus our 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 $f_e^{priceintervention}$, the market size under the land subsidy regime and the target we need to match for the change in market size going from price intervention to the land subsidy, we 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.

We will use the following routine for our first step: for each candidate vector of parameters, we will solve our model to its stationary equilibrium under the land subsidy policy regime, recover the policy functions and simulate a dataset akin to our actual farm-level data. We will then recover key moments in this simulated data and compare them to the observed data (for the period when the land subsidy was effective). Our estimates of the true parameters will 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.

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

Algorithm: For a potential vector of parameters Θ, we solve our dynamic game to its stationary equilibrium, and generate associated policy functions regarding capital investment, crop choice and market incumbency. We then create a sample of 1,000 farms, drawing their initial state at random from the equilibrium distribution of incumbents. We simulate their behavior for $T = 50$ periods, and compensate for exit by continuously drawing in new entrants to keep the sample size constant.[62] In order to address the impact of those random draws on the composition of the data, we bootstrap this exercise $B = 100$ times, and then compute average simulated moments over these farms, periods and bootstrap runs. We denote these moments estimate $\Phi^s(\Theta)$, and the observed moments in the FADN and the Census Φ^d. We then use a grid search to find the vector of parameters that solves:

$$\min_{\Theta} \left[\Phi^d - \Phi^s(\Theta)\right]' W \left[\Phi^d - \Phi^s(\Theta)\right] \tag{10}$$

Where W is a weighting matrix. Because our moments are similarly scaled, we use the identity matrix – following Asker et al. (2014).

Table 1: Moments for Dynamic Parameters

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
Ratio Incumbents 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.

[62] 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 our firms exit, and we have to compensate for this by drawing new firms – we cannot rely on the removal of early periods to address this bias. We 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 our case, as firms potentially always exit the market at some point. We address this by bootstrapping the simulation, rather than balancing the simulated dataset.

Identification: Indirect inference is a useful approach to recover some parameters of interest, when the likelihood function implied by the conditional distribution of the endogenous variables of one's econometric model is intractable (Gourieroux et al. (1993), Gallant and Tauchen (1996)). Indirect inference relies on the idea that while the likelihood function associated with a model might be intractable, it remains easy to simulate data draws using that same model. One can then specify some auxiliary parameters which are straightforward to estimate both in the true and in the simulated data, and infer the parameters of interest from the comparison of the auxiliary parameters. In our context, we use the four previously described moments as auxiliary parameters. These are easily estimated both in our FADN data, and in our simulated data – either as sample means or a sample coefficient of variation. Jointly, these auxiliary parameters and their estimation method yield a binding function which relates our parameters of interest Θ to the moments. With our notation, the binding function is precisely $\Phi(.)$, yielding $\Phi(\Theta)$ for any candidate vector of parameters. The simulation routine which aims at recovering $\Phi^s(\Theta)$ acts as a functional estimator of the binding function, and the distance minimization problem specified in Equation 10 is our calibration exercise which identifies which value of the parameter vector Θ gives the best match between the observed and simulated moments. Following Gourieroux et al. (1993), identification in our case will rely crucially on two elements. The first one is that our binding function 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 our 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.[63]

It is important for the moments we choose to be driven by the parameter vector Θ. Capital accumulation, crop choice and incumbency are the dynamic decision made by our farms. These decisions are interlinked, in that a farm with good prospects will be more likely to remain on the market, grow more crops, and to invest. A farm with bad prospects is more likely to exit, reduce its scope, and reduce or keep its current stock of capital. A firm's capital stock also shields it partially from future negative productivity shocks.

Relevance of moments: While our parameters are recovered jointly from the matching of all moments, it is useful to discuss the parameter-specific relevance of each moment. We show in Figure A22 the distribution of the three variables from which we 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

[63] Collard-Wexler (2013) notes that while maximum-likelihood estimation will be biased in the presence of simulation error (Pakes et al. (2007)), indirect inference will be consistent as simulation error will wash out in expectation when computing conditional means and moments.

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.

Our 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 we 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 our 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 we expect the zero profit condition to bind, we can also track entry costs by looking at expected profits of entry. In our 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 our effort to recover the cost of entry.

4.2.4 Results

We obtain the following set of results. We give in Table A9 the values we use for the additional set of model parameters that we directly calibrate.

Table 2: 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 Parameter Land-Chemicals	ρ	0.36	(0.22)
Substitution Parameter 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. We keep farms observed for at least three periods in a row[64]*, and that produce either only crops in the wheat group, or in the other crops group. We 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 we 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 which does not give use standard errors.*

Recovered Input Substitution: The two key parameters recovered with this estimation give us the elasticities of substitution across resp. land and chemicals, and fertilizers and pesticides. The first elasticity is $\sigma = 1.5684$, indicating that land and chemicals are substitutes, and the second is $\sigma_2 = .2988$, meaning that fertilizers and pesticides are complements. It is useful to compare these results to the ones frequently obtained in the agricultural economics literature.[65] Carpentier and Rainelli (1997) for France, McGuirk and Mundlak

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

[65] We note that there are debates surrounding how to specify production functions that include pesticides as inputs. Licht-

(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 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 our specification and the recovered value of σ_2 goes in this direction.

In terms of the substitution between land and chemicals, we obtain a relatively large value, which we can relate 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 our 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).[66]

4.3 Policy Analysis

With the estimated model, we can perform counterfactual analysis of different policy regimes and study their respective impacts on welfare.

Policy focus: We propose to evaluate the welfare consequences of three different policy regimes, and compare them to the baseline no-intervention case.

The first policy we evaluate is the price intervention regime, which was the main policy of the EU's Common Agriculural Policy umbrella between the the late 1960s and 1992. We

enberg and Zilberman (1986) for example argue for modelling pesticides as aiding for pest abatement rather than entering production functions as a direct input. Our approach aims at allowing for fertilizers and pesticides to potentially be complements, and for chemicals intensity to be associated with increases in productivity.

[66] 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 we 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 we 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 we 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. We 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 we obtain here.

specifically evaluate the impact as it existed in 1991, and approximate it as raising output prices 43% above the effective prices under the land subsidy regime, on both of our modelled crop markets. Beyond its historical interest, the price intervention regime is interesting to study as it has the largest consequences on firm survival. By setting a purchasing guarantee at very high price levels, the EU significantly reduced the force of selection with the agricultural market. It however did not change incentives regarding input choice, and hence any positive impact on pollution would come from the reduced selection of firms on their productivity, and in part on their chemical-using productivity. The intervention regime also changed farms' scale incentives, likely increasing production levels within a farm type. This scale effect has the possibility to raise the average input-to-output ratios of farms, if adjustments of capital do not track fully with the adjustment of scale induced by intervention.

The second policy we evaluate is the land subsidy. We model this land subsidy without historical set aside requirements, and in that sense we model a land subsidy as it existed in Europe from the early 2000s onwards. Specifically, we model a land subsidy which reduces the price of land by 35%. This policy is interesting to study, in that its mode of subsidization will distort the input choice incentives for all market incumbents. It might also reduce the strength of selective forces if its impact on lowering production costs is not fully compensated by reductions in output prices. The subsidy might also change scale decisions for firms.

Finally, we propose to evaluate a chemical tax combined with a lump sum payment, which compensates farmers for their expected taxation rate. We set the tax at a level such that the effect on the relative land-to-chemicals prices ratio is identical to that of the land subsidy. Chemical taxes are scarce in the current policy environment. In Europe, only Denmark, Sweden, Norway, Italy and France have pesticide taxes – and most have very low rates of taxation, while the shared approach to pesticide regulation has been to introduce an authorization procedure based on their risk assessment.[67] Fertilizers are regulated through the Nitrate Directive within the EU – which sets caps on the use of nitrogen application in at-risk areas. As such, the EU has so far taken a limited command-and-control approach to chemicals regulation in agriculture, setting infinite (in the case of pesticides), and finite spatially heterogeneous (in the case of fertilizers) quotas for their use. The relative advantages of quotas versus taxes have been heavily discussed since Weitzman (1974), who highlighted the role of the steepness of social cost and benefit curves in driving their relative comparative advantages. Here, we discuss the implementation of a chemicals tax, but do expect a non-linear cost curve from pollution and a general ability of farms to substitute away from chemicals in production. The chemical tax seems like a useful tool to mitigate the effect of agricultural pollution while limiting its consequences on efficiency. Because the tax will directly target pesticides, and not land through pesticides, we expect it to have a larger effect on pollution: the price of chemicals relative to all other inputs is affected. Because it favors

[67] The general mode through which pesticides have been regulated so far is through authorization procedures. Since 2006, the EU has had a policy called REACH for Registration, Evaluation, Authorisation and Restriction of Chemicals, which has generated many concerns when it to the role of lobbying in influencing authorisation procedures. In the United States, related policies have been implemented since the late 1990s. While these authorisation procedures have the capacity to ban pesticides, they do not change the incentives faced by farms once pesticides are authorised.

survival through the lump-sum, but does not encourage low-type survivors to produce, nor any incumbents to over-produce compared to the baseline – we expect its efficiency effects to be smaller.

Welfare Criterion: We evaluate the policy in terms of consumer outcomes, and consider them to be of two different types: the chemical-intensity of production embodied by the chemical-to-output ratio, and the effective price of agricultural commodities. This effective price includes the price of commodities, and the subsidy-per-unit which is paid through any given policy regime to producers, and ultimately would be levied on consumers through taxation. The different policies have a difference incidence on prices, and imply different levels of producer subsidization. In the case of price intervention, we consider the difference between the private price at which the market would actually clear under the intervention-induced level of output, and the level of the intervention price to be the subsidy-per-unit effectively paid out by consumers. This subsidy by definition is directly included in the price under intervention. This definition implicitly includes the social cost of over-production as the difference between the market clearing price and the policy-enforced one. Under the land subsidy regime, prices will fall as a consequence of reduced costs of production – but the effective price will be the sum of that price, and the amount of land subsidy received by farms per-unit produced. Finally, because we model the chemical tax as paired with a lump sum payment, this tax has no subsidy for producers, but it will likely raise prices as a consequence of increased costs of production linked with distorted input choice incentives.

We have aggregated chemicals into a single input. It is difficult to provide a social cost of aggregate chemical use at the farm level, which includes both the environmental damages stemming from pesticide and fertilizer use. We propose to use the valuations for pesticide-driven pollution implied by Denmark's and France's taxes on pesticides as approximations for the social cost of chemical pollution in agriculture. For Denmark, we only use the benchmark cost of 6.70 euros per kg of active ingredient used, and for France we use the mid-point cost of .9 euros per kg of active ingredient.

Welfare Impacts: We compare the three policy-induced equilibria and the no-intervention case. We express the outcomes in terms of their effect relative to the land subsidy regime, which is the current central regime of EU agricultural subsidization.[68] We plot the indifference curve implied by resp. the French and Danish pesticide tax. We specifically plot the indifference curves which intercept the land subsidy equilibrium. Any equilibrium located to the left of that curve will be welfare-improving, and any to the right will worsen welfare. We show values for each crop market.

[68] We translate these changes into changes in prices and chemical use using estimates for the average prices for each commodity group over 1995-2000, and chemical intensity of production under recommended agro-science practices. We use a baseline rate of .16kg of pesticide active ingredient per ton produced. See statistics from the Latvian government for that last estimate https://stat.gov.lv/en/statistics-themes/business-sectors/agriculture/press-releases/5773-use-pesticides-cereal-and-rape.

We first see that only the chemical tax is welfare improving compared to the land subsidy, and that both the no-intervention case and price intervention reduce welfare according to the valuation implied by the French and Danish taxes. In terms of prices first. The pre-1992 price intervention scheme by definition increased prices by 43% above the land intervention price. In terms of effective price, this corresponds to resp. a 27% and 42% increase. The chemical tax has a lower effective price for wheat, 1.3% lower than the land subsidy regime. However, the second crop being much more chemical-intensive, the tax's relative effect there is to raise the effective price by 5.6%. As could be expected, the no-intervention equilibrium has the lowest effective prices, resp. 2.1% and .7% lower than the ones of the land subsidy regime.

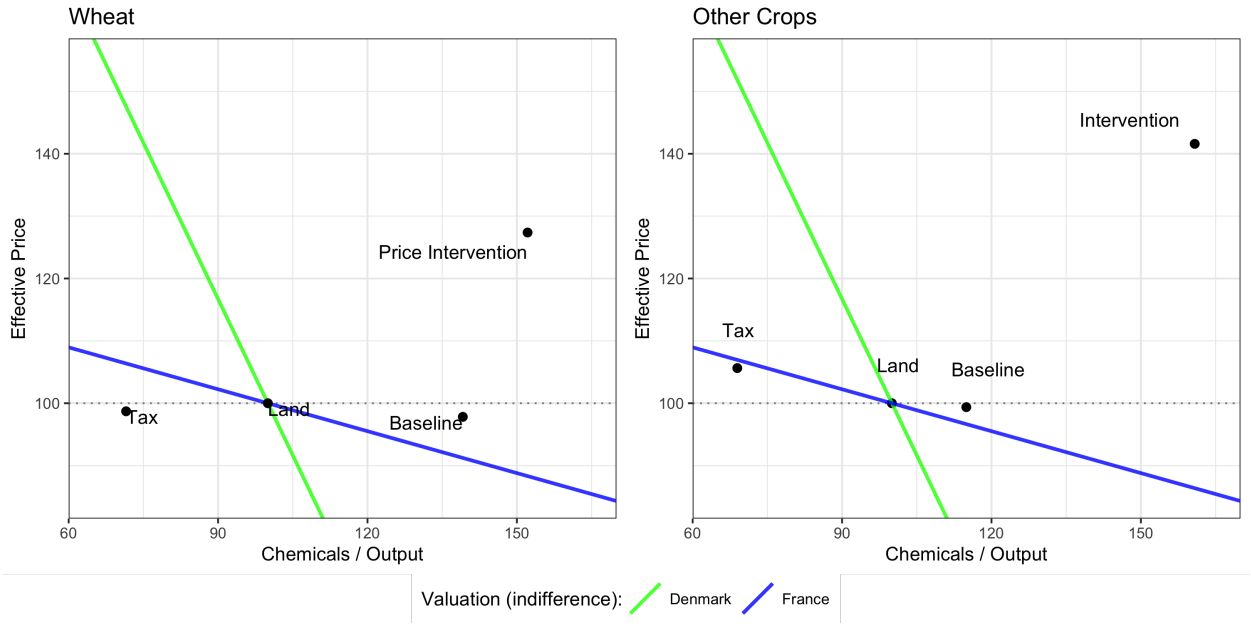


Figure 5: Welfare Consequences

Notes: We show the effective prices (crop price on the market plus subsidy per unit) and chemicals-to-output ratios linked with the different policy regimes, on both crop markets. Values are normalized to the land subsidy, which is the current CAP subsidy design. The blue and green lines describe the marginal valuation for pesticide pollution implied by existing Danish and French pesticide taxes, also expressed in percentage deviations from the effective price and pollution intensity existing under the land subsidy regime. Equilibria located to the left of each respective lines represent a welfare improvement compared to the land subsidy regime.

Output price intervention is also a poor instrument to address pollution, and it ranks last in terms of pollution. We will confirm this later looking at decompositions, but the gains from reduced selection are far-outweighed by the consequences of over-production. As farms increase their scale, capital stock does not follow through exactly, and respective flexible inputs-to-output ratios increase. Both the land subsidy and the chemical tax perform better

than the no-intervention case in terms of pollution, but the chemicals tax is by far the best policy in terms of pollution. Relative to the land subsidy, it decreases pollution intensity by 28.5% on the wheat market, and by 31% on the aggregate crop one. The full set of outcomes, detailed by market is in Table A1.

Decomposing Channels of Change: We can decompose the effects of policies as treatment and reallocation effects. Treatment will change the value of the outcome variable at the farm-crop level, while reallocation will change the allocation of production across farms. For the change in the average value of a given variable, which we denote ΔZ, and denoting the market share of farm of type j in equilibrium t – we can write:

$$\Delta Z = \sum_j \underbrace{s_{jt}\Delta Z_j}_{Treatment} + \underbrace{\Delta s_j \Delta Z_j + \Delta s_{jt} Z_{jt}}_{Reallocation}$$

We plot the decomposition graphically, *Within* indicating the impact of the policy only taking into account treatment. Here, we show the effect of removing price intervention on effective prices and pollution intensity. We show the figures for the other policies in Figure A23 and Figure A24.

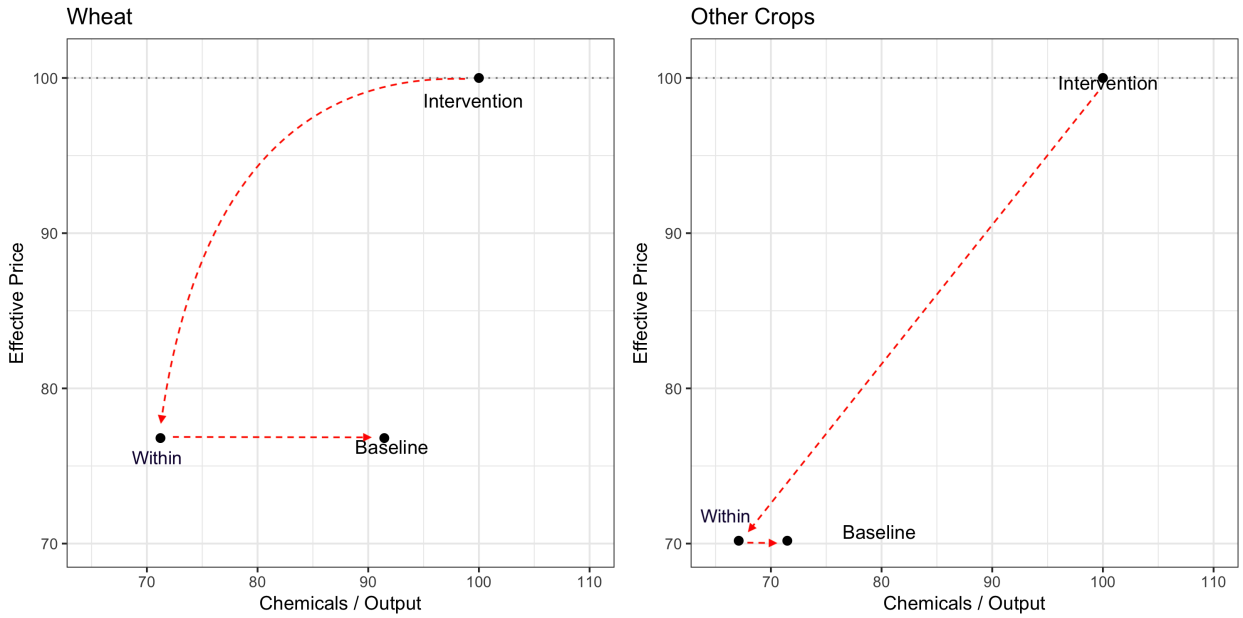


Figure 6: Impact of Ouput Price Intervention

Notes: We show the effective prices (crop price on the market plus subsidy per unit) and chemicals-to-output ratios linked with both price intervention and no policy intervention, on both crop markets. Values are expressed in percent of the effective price and chemicals-to-output ratio under price intervention. We decompose the change between price intervention and no-intervention using the Within point. Movements from *Intervention* to *Within* correspond to within farm-crop treatment effects of removing price intervention, while the movement from *Within* to *Baseline* describe the additional effect of reallocations (changing market shares).

This figure highlights very well the mechanisms at play when it comes to the effect of price intervention on pollution intensity. Removing price intervention decreases scale decisions, and the *Within* points for both crops are to the left of the baseline point. Reallocation from price intervention to baseline however increases pollution. This captures the fact that the reallocation of production towards more efficient firms indeed increases the average pollution intensity of the market, as these firms tend to pollute more. The other two policies have much smaller effects on the relative allocation of production, and mostly impact prices and pollution through within-establishment changes in production incentives.

Finally, we propose in subsection B.10 an extension of the decompositions used by De Loecker and Collard-Wexler (2015) to analyse the mechanisms that differentiate the chemicals tax and the land subsidy in terms of pollution effects.

5 Conclusion

In this paper, we study how the joint consideration of production efficiency and environmental externalities can challenge our understanding of the desirability of agricultural subsidies. This is an empirical question, which requires to know both about the relation between producer efficiency and their propensity to generate externalities, and about the distribution of producer heterogeneity in the market. Focusing on externalities stemming from the use of a polluting input, we model the efficiency-externality relation as driven by non-Hicksian shocks shifting the productivity of the polluting input, combined with a recovered input substitutability between this input and others.

In the case of the French row crop market, we find that more productive farms will also use more chemicals in production. In that context, we find that reducing selection based on non-Hicksian productivity reduces the pollution intensity of the market, but reducing pollution based on input-neutral productivity will raise it. Production subsidies can also increase the average scale of producers, likely decreasing their average input efficiency in production. We find that the pre-1990s output price subsidies of the EU ended up having a negative effect on pollution, as scale effects dominated the positive consequences of reduced farm selection. They also had large negative effects on efficiency. We further show that in the case of externalities coming from specific input use, subsidies directly targeting the price of its substitute can however generate large decreases in pollution, potentially higher than their welfare costs through lower market efficiency. The current EU land use subsidy dominates its prior output price intervention regime, and is potentially welfare improving compared to no-intervention. A specific tax for the polluting inputs, paired with a targeted lump sum payment however still fares better both in terms of efficiency and pollution. Lump sum payments are more desirable than input subsidies, as they only impact incumbency decisions, but not production ones. Directly targeting the polluting input rather than one of its substitutes is also more effective, as it shifts all the relevant relative input price ratios faced by the farm.

Our analysis has only considered pollution as damaging for welfare, because of welfare's direct valuation of public health and functioning ecosystems. Environmental externalities also matter for their dynamic consequences for production efficiencies. Climate change limits our future capacity for productivity gains, and so does species extinction and ecosystem disruptions. A future area of research for this agenda is to recast the question of the sources of agricultural productivity growth into a context where farms have dynamic interactions with the ecosystem and land within which they produce.

References

Acemoglu, D., U. Akcigit, H. Alp, N. Bloom, and W. Kerr (2018): "Innovation, Reallocation, and Growth," *American Economic Review*.

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*.

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*.

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*.

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*.

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*.

Femenia, F., A. Gohin, and A. Carpentier (2010): "The Decoupling of Farm Programs: Revisiting the Wealth Effect," *American Journal of Agricultural Economics*.

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*.

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

Gallant, A. R. and G. Tauchen (1996): "Which Moments to Match?" *Econometric Theory*.

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*.

Gourieroux, 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*.

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

Grossman, G. and E. Helpman (1991): "Quality Ladders in the Theory of Growth," *Review of Economic Studies*.

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

Hawkins-Pierot, J. T. and K. R. Wagner (2023): "Technology Lock-In and Costs of Delayed Climate Policy," *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*.

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. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*.

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*.

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*.

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

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*.

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*.

Weitzman, M. L. (1974): "Prices vs. Quantities," *Review of Economic Studies*.

Williamson, O. (1968): "Economies as an antitrust defense: the welfare tradeoffs," *American Economic Review*.

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

Appendix for online publication
TITLE

A Additional Figures and Tables

A.1 Figures

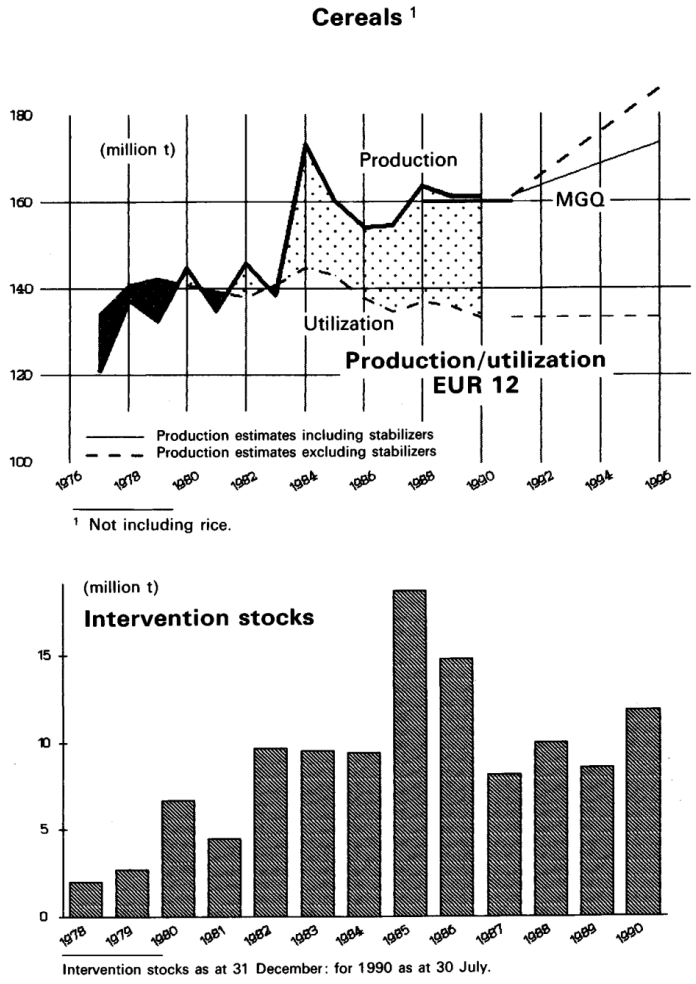


Figure 12

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 .

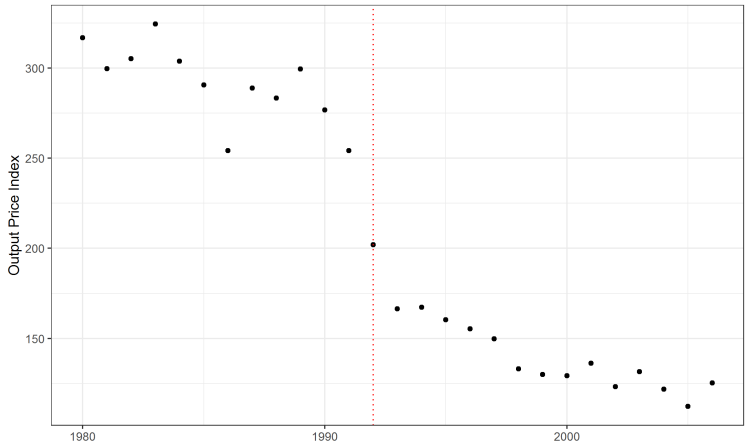


Figure A2: Average Farm Price

*Notes: We show a binscatter for a farm-level average output price, computed over the following crops **[LIST]**, using relative land shares as weights.*

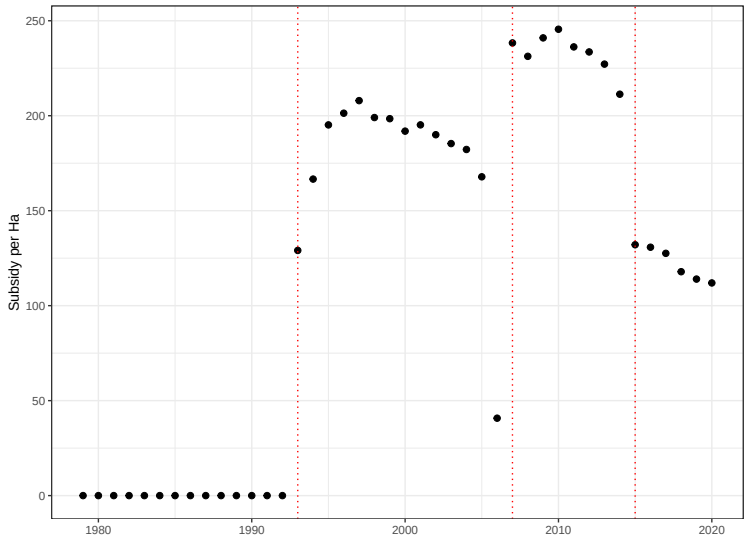


Figure A3: EU Land Subsidies under the CAP

Notes: We 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.

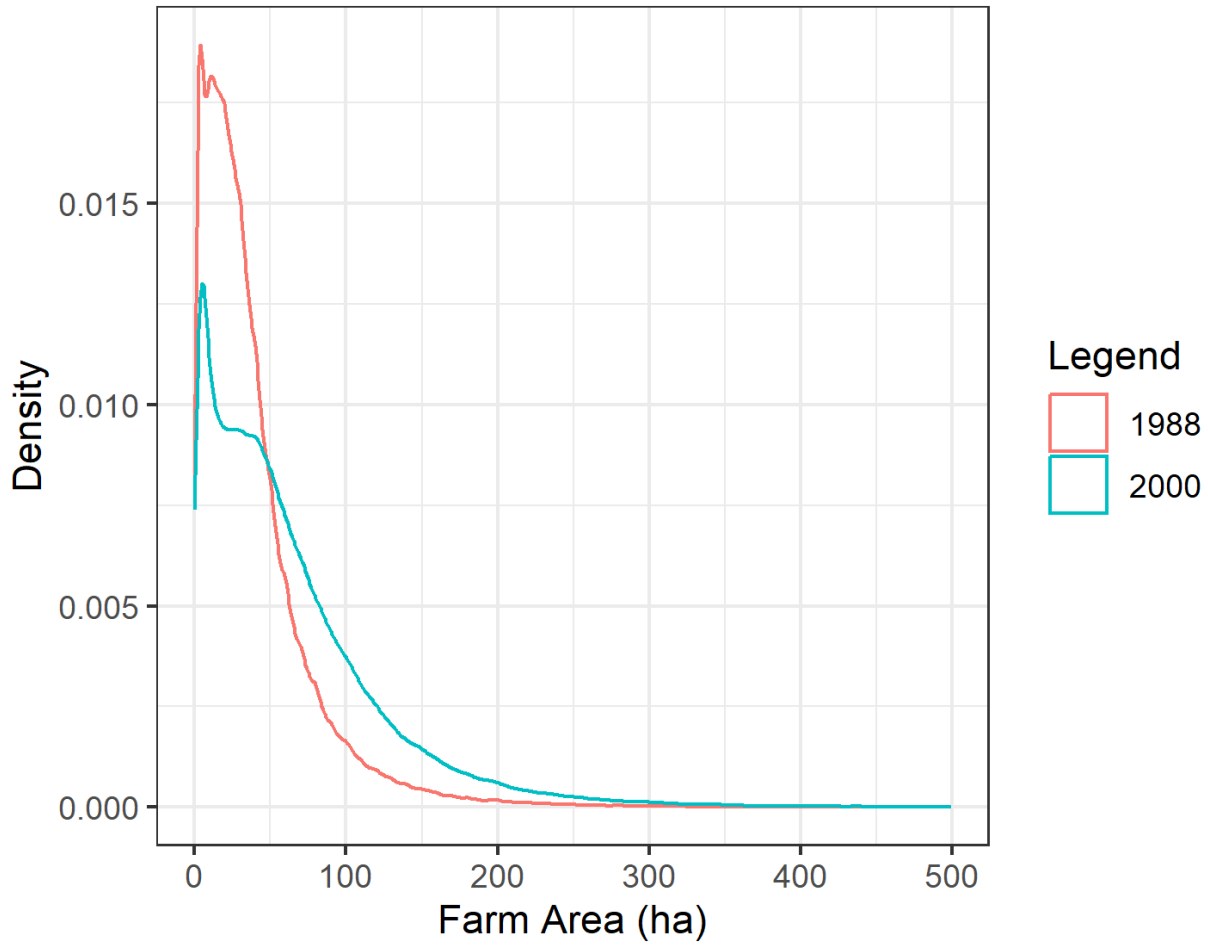


Figure A4: Farm Size Distribution in 1988 and 2000

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. We see that the distribution shifts significantly over time, and that density is reallocated in the right tail of the distribution. We know 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.

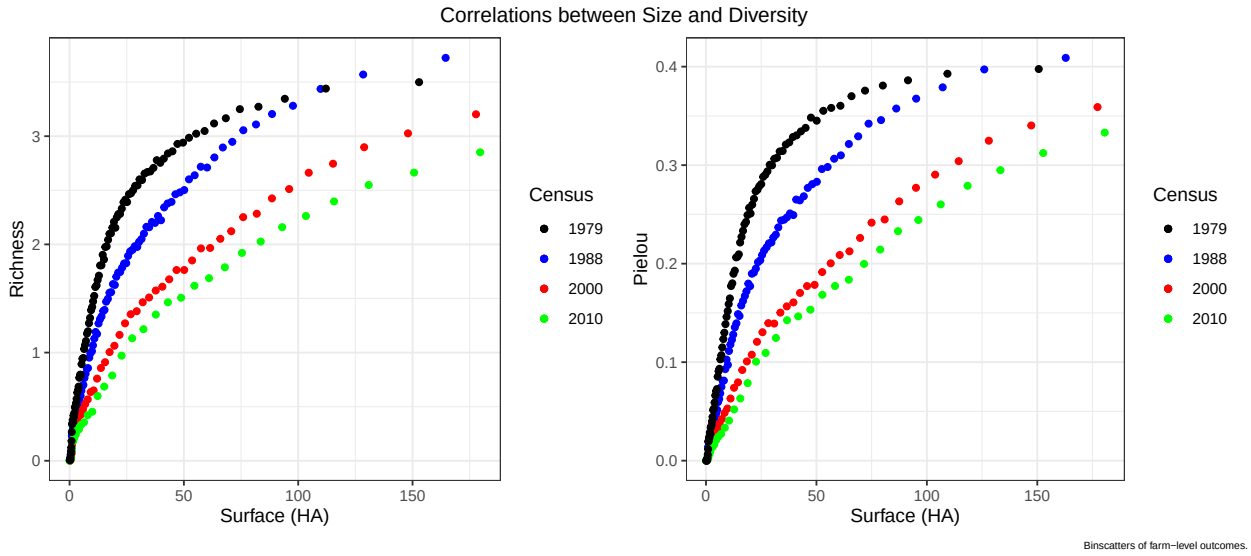


Figure A5: Evolution of Farm Crop Diversity

Notes: These binscatters are computed using data from the Agricultural Census. We take into account the maximum number of crops which we can track 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[69].

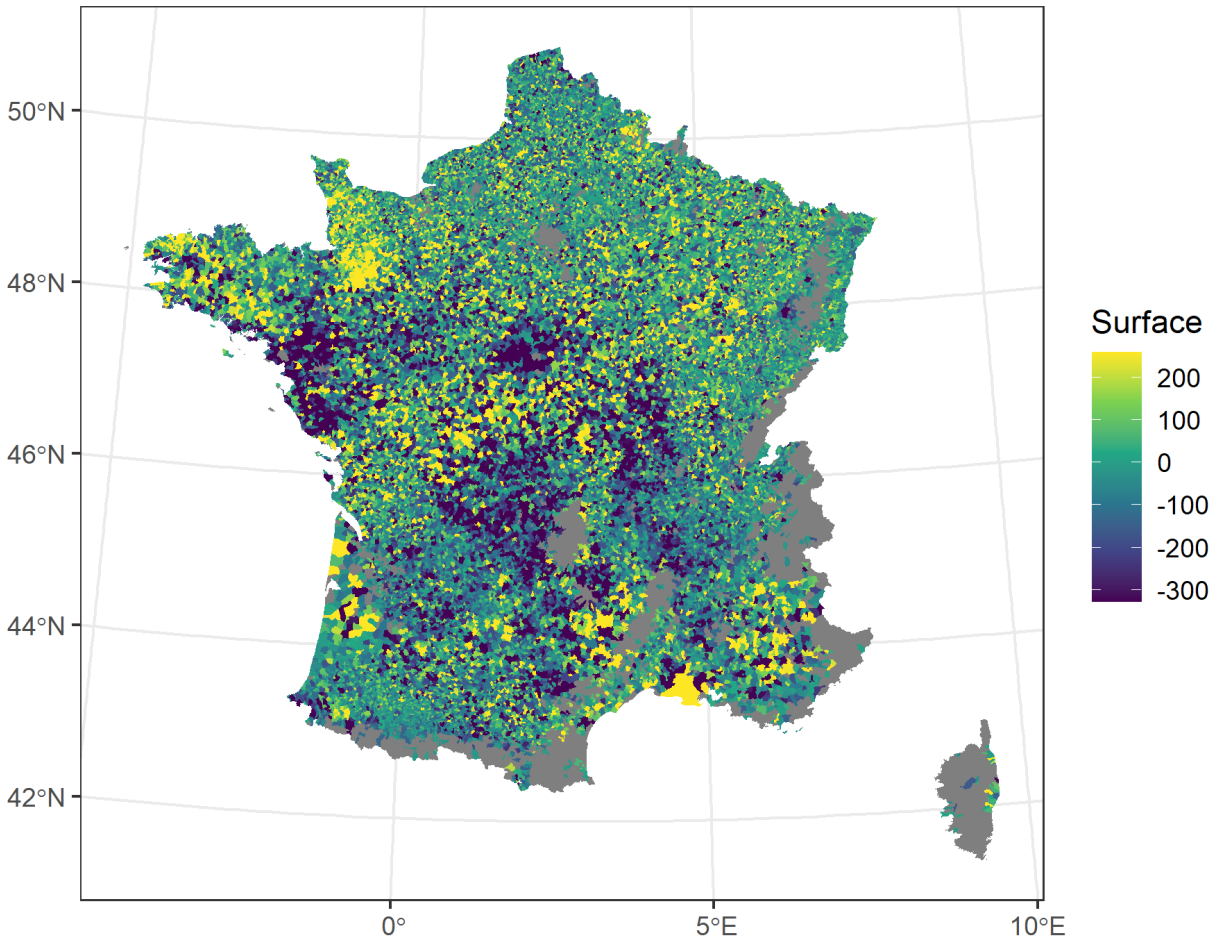


Figure A6: Evolution of Village-level Total Agricultural Area

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

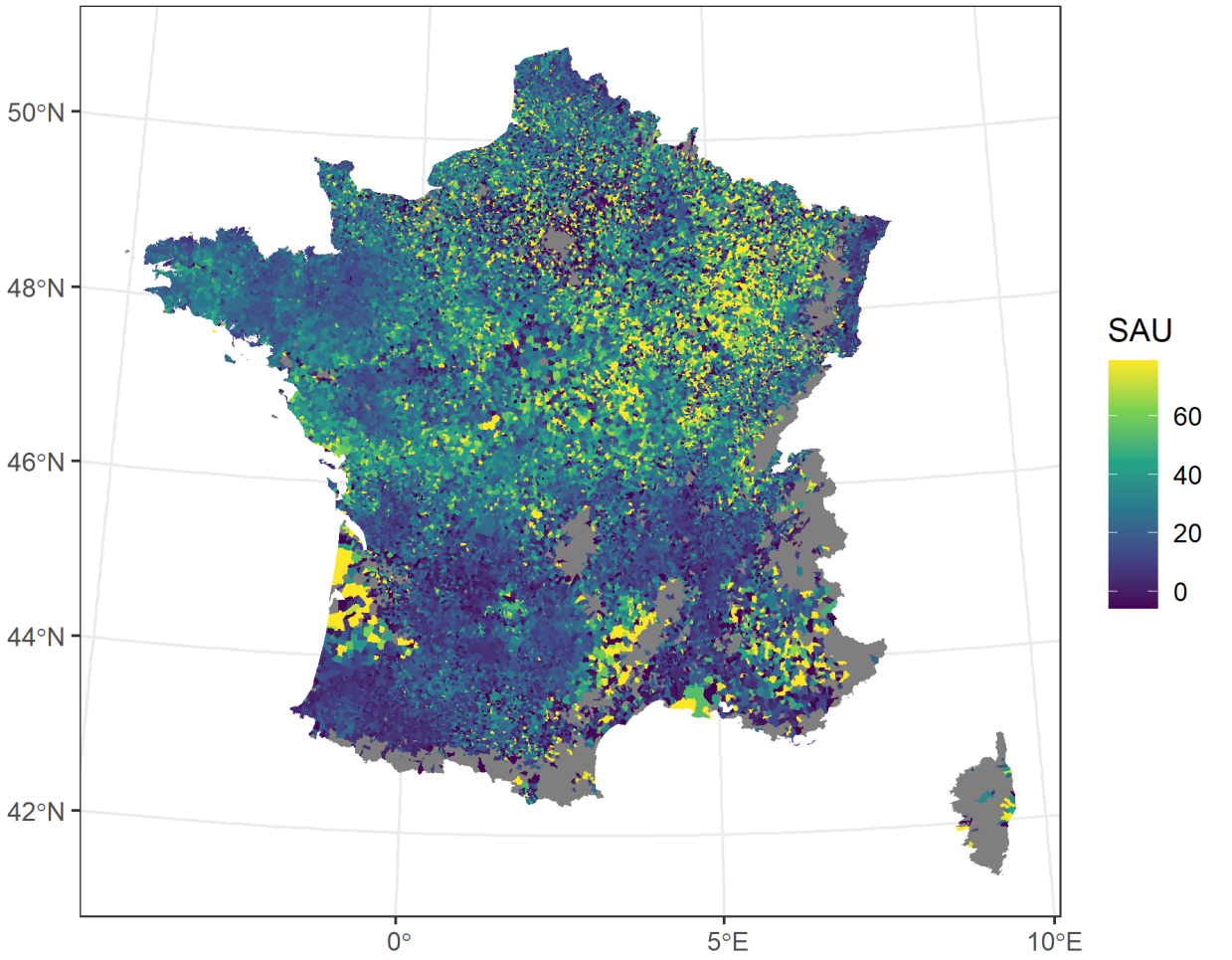


Figure A7: Evolution of Village-level Average Agricultural Area

Notes: This map shows the evolution of within-village average agricultural area between 1988 and 2000. We use the data from the relevant waves of the French Agricultural Census. We first note that virtually no village sees a decrease in average farm size, and then that the geographic distribution of farm growth is quite heterogeneous, with the largest increases happening in the area stretching between Paris and the Massif Central, which includes the row-crop intensive Beauce region.

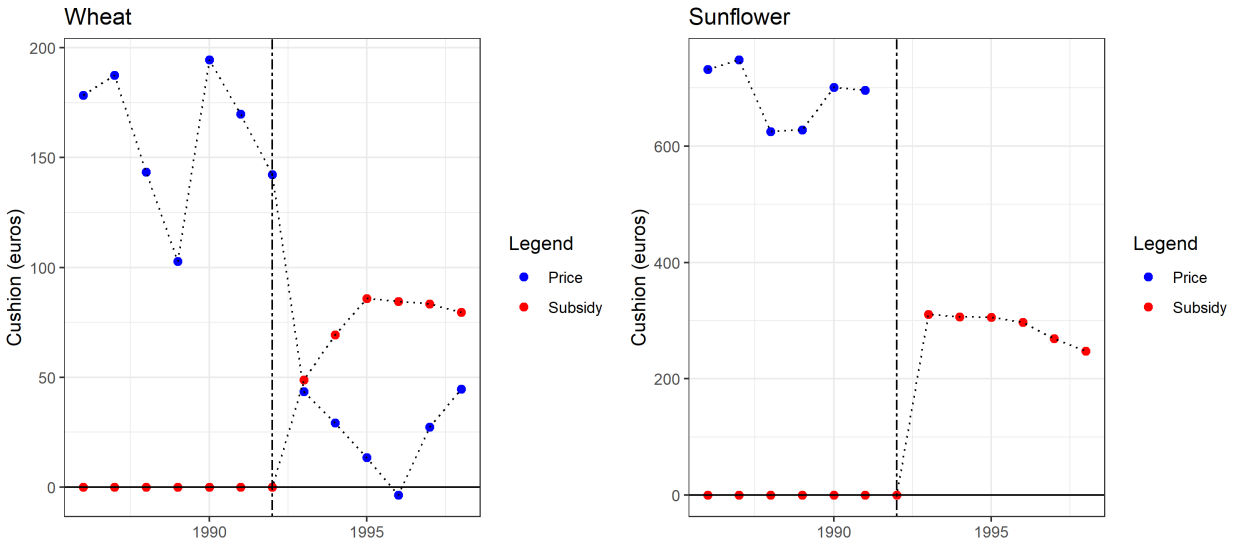


Figure A8: Variation in the Price Wedges Introduced by EU Policy

Notes: We illustrate the cross-crop variation in wedges introduced by EU policy which interact with the cross-farm heterogeneity in crop mix composition, to generate the variation used in our shift-share design. The price intervention line for sunflower stops in 1991 to illustrate the fact that the intervention price is fully removed from 1992 onwards. We collect intervention prices from yearly European regulations, and translate them to 2020 euros using appropriate exchange rates. Subsidies correspond to the median subsidy values observed in the FADN, and divided by average yields to be translated into subsidies per ton produced.

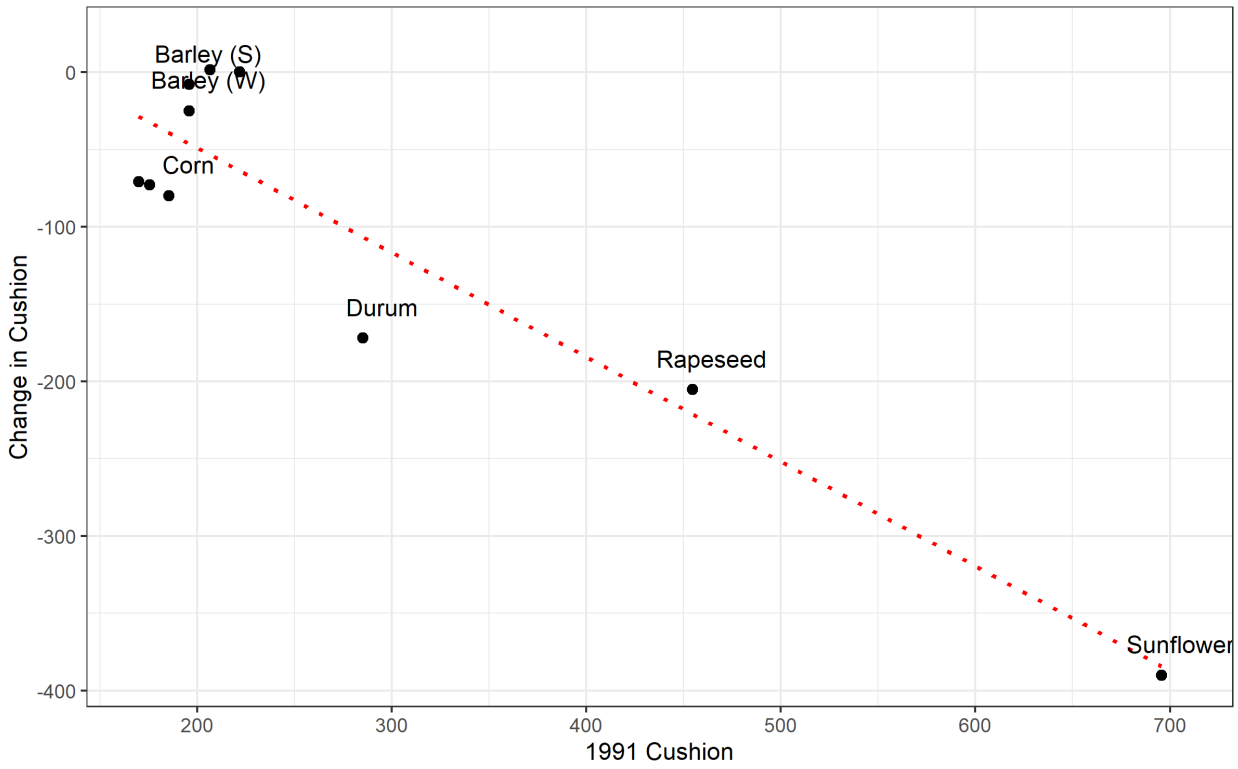


Figure A9: Convergence of Crop Cushions Post Reform

Notes: We illustrate 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.*

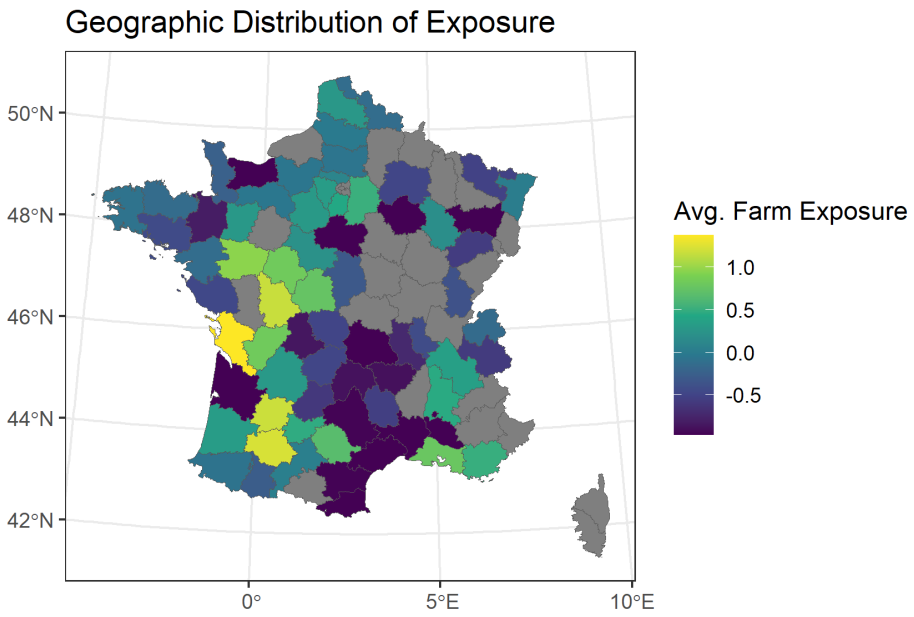


Figure A10: Department Weighted Averages of Farm-Level Exposure

Notes: We show 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. We show the geographic variation of the standardized exposure here.

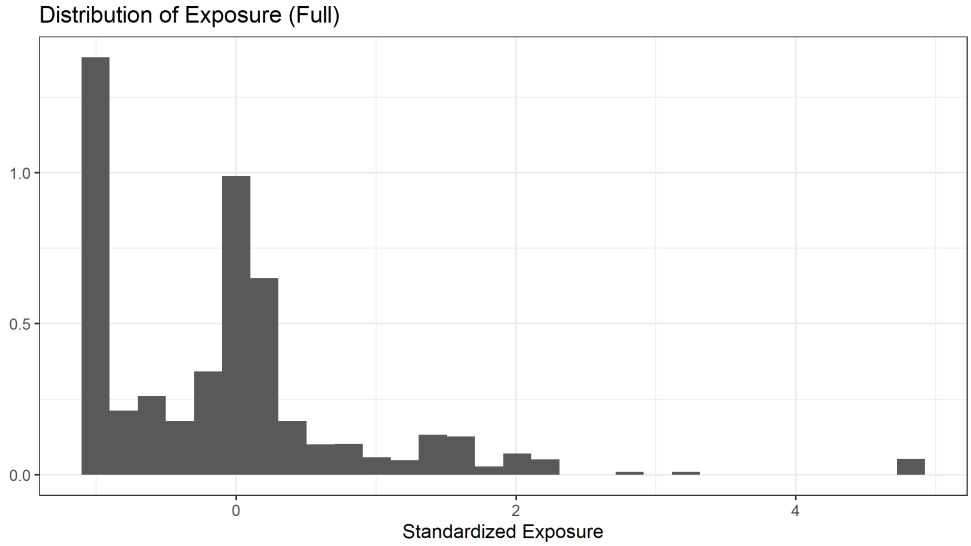


Figure A11: Distribution of Farm-Level Exposure

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

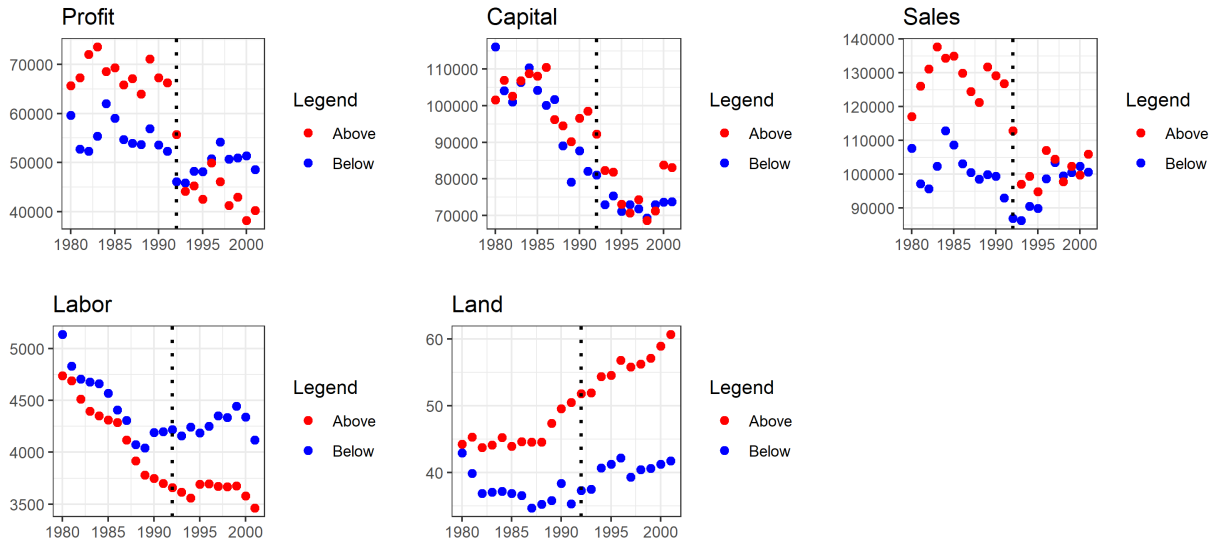
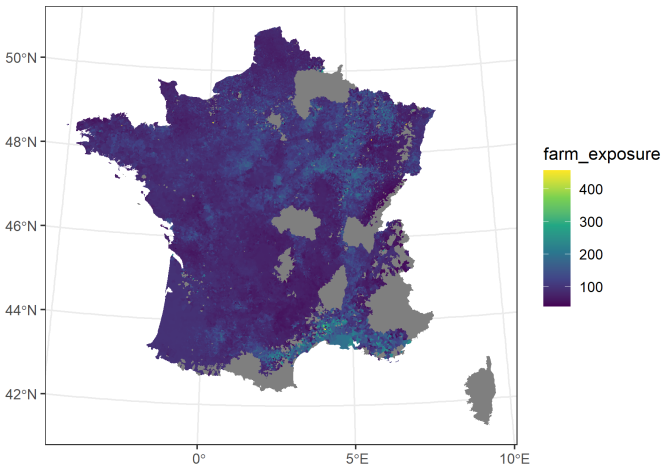
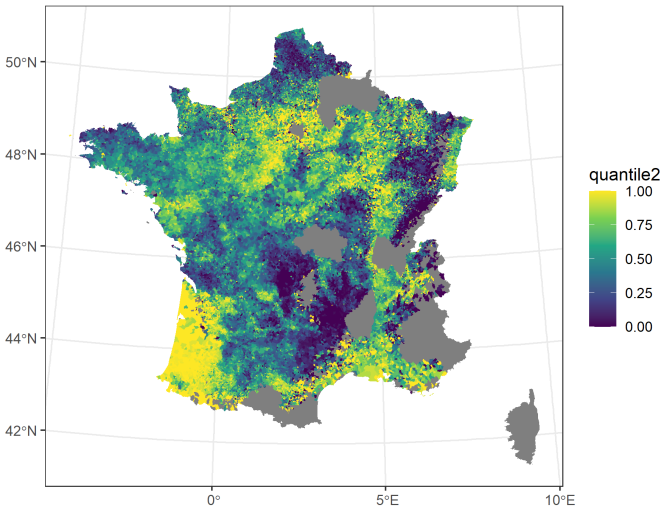


Figure A12: Balance in Levels and Trends

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.



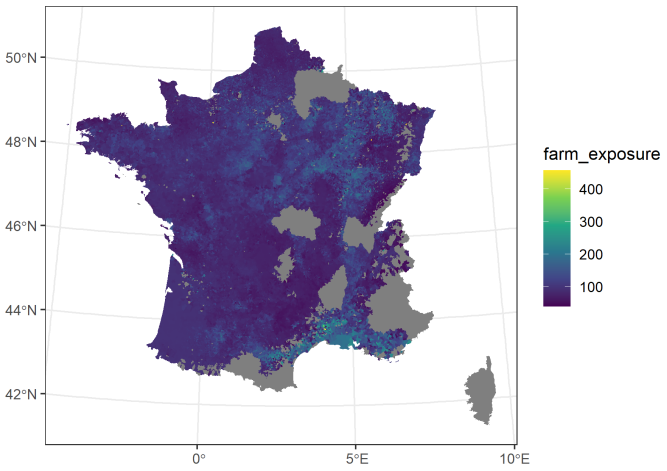
(a) Average Village Exposure



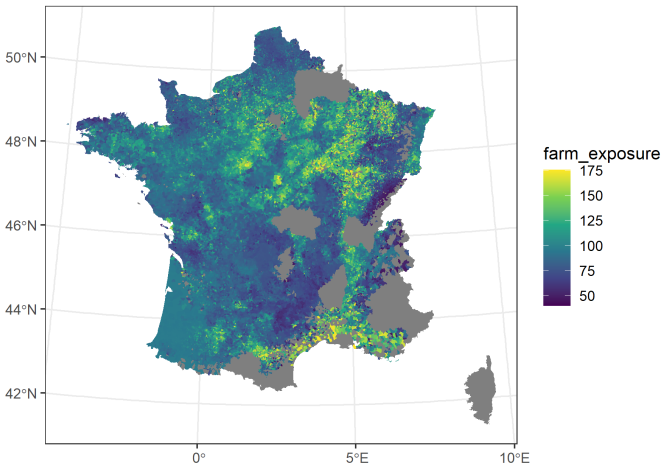
(b) Village Share Above Median Exposure

Figure A13: Geographic Distribution of Exposure

Notes: We plot graphically the distribution of 1988 average village exposure to the reform, and of the village share of farms situated above the median France-level exposure. We see that average farm exposure within a village has very large outliers along the Mediterranean coast. On both maps, areas with no data correspond to Corsica, the Paris urban area, and several mountainous regions of the country where large field agriculture is scarce.



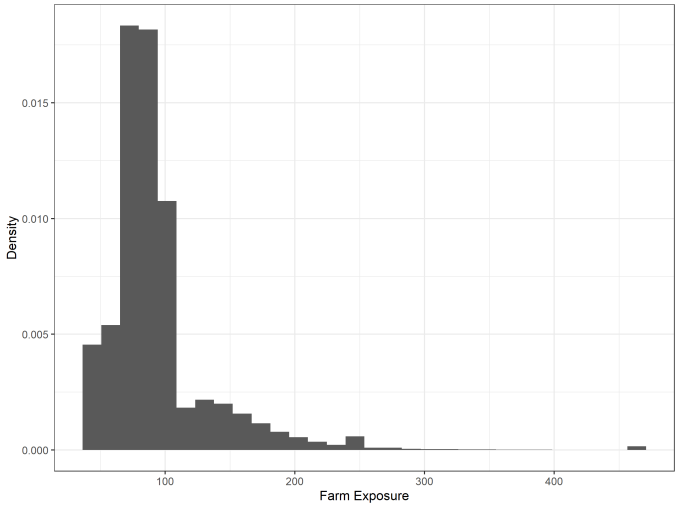
(a) Average Village Exposure



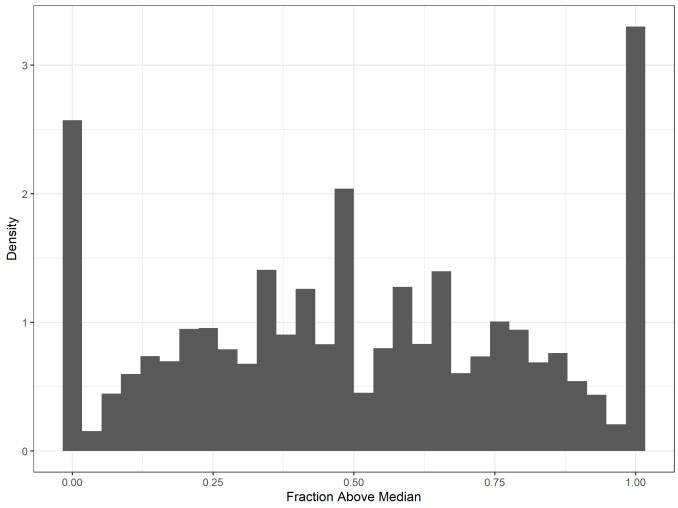
(b) Average Village Exposure - Truncated

Figure A14: Geographic Distribution of Exposure

Notes: We plot graphically the distribution of 1988 average village exposure to the reform, without and with a truncation of the top 5ppt of the distribution. We see how the distribution is heavily skewed, but also how this right tail in exposure is geographically concentrated along the Mediterranean coast.



(a) Average Village Exposure



(b) Village Share Above Median Exposure

Figure A15: Distribution of Exposure

Notes: We plot graphically the distribution of 1988 average village exposure to the reform, and of the village share of farms situated above the median France-level exposure.

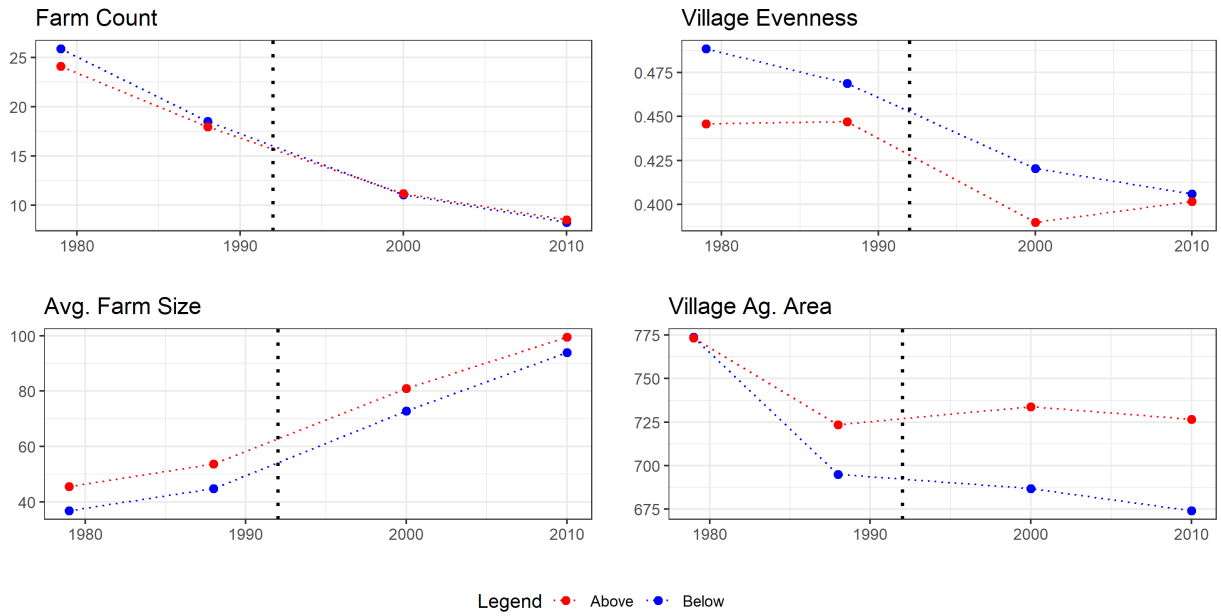


Figure A16: Balance Tests for the Village-Level Analysis

Notes: We plot the evolution of important farm-level characteristics across villages with a below and above median share of farms with an above median exposure to the reform in 1988.

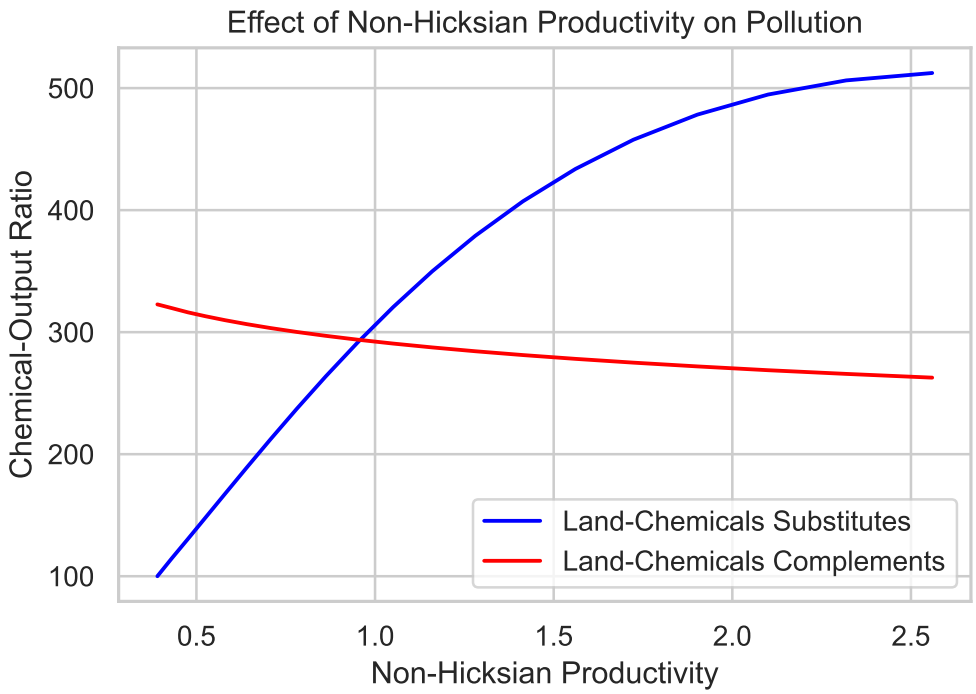
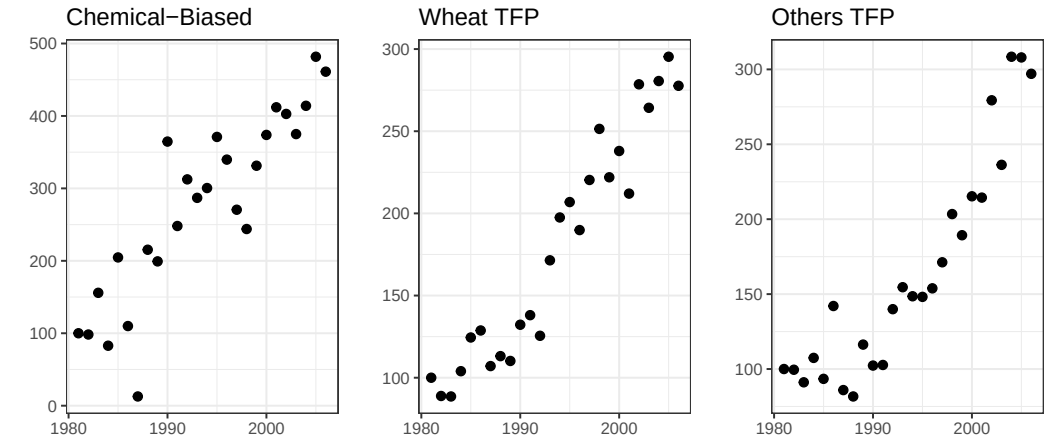
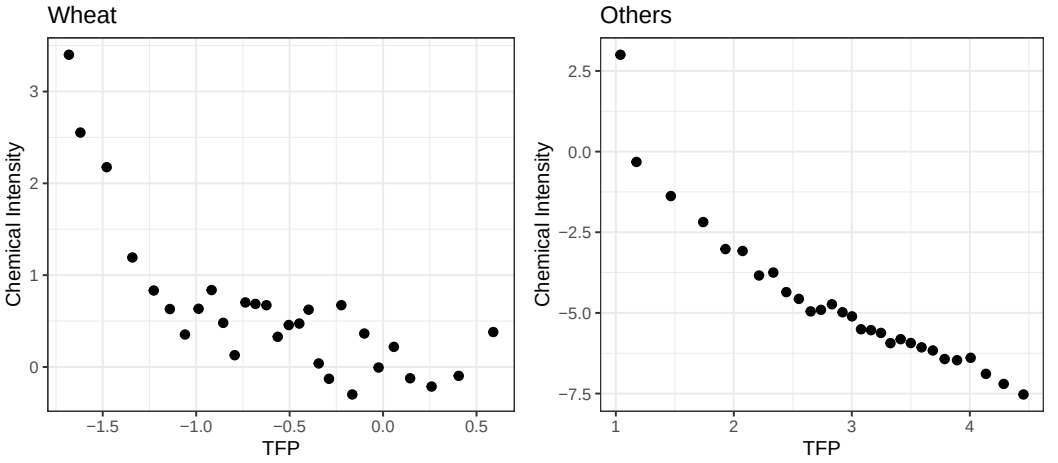


Figure A17: Relation Between Land-Chemicals Elasticity and non-Hicksian Productivity

Notes: This figure shows the effect of non-Hicksian productivity shocks on firms' chemical-to-output ratios at their optimal production level. We 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.



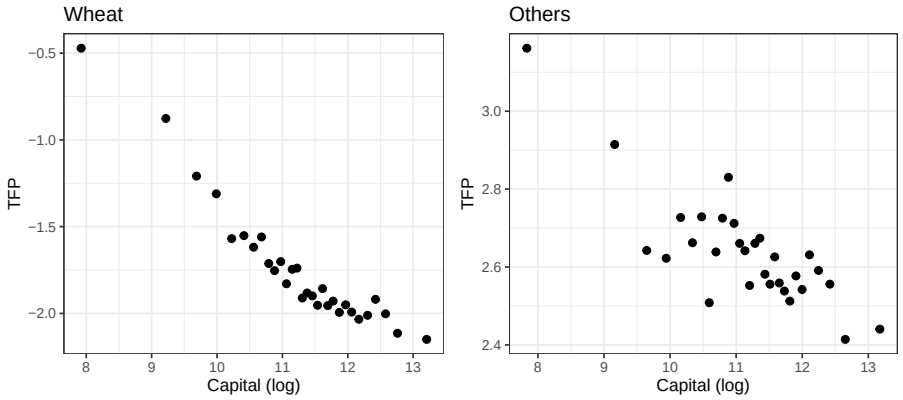
(a)



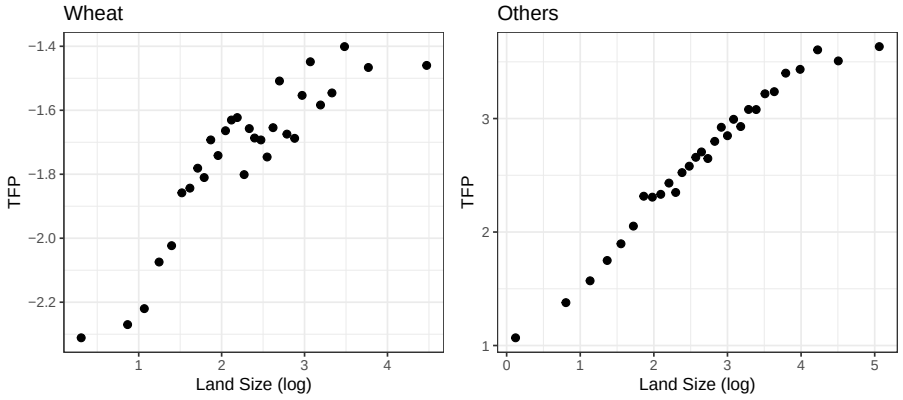
(b)

Figure A18: Recovered Productivity Shocks

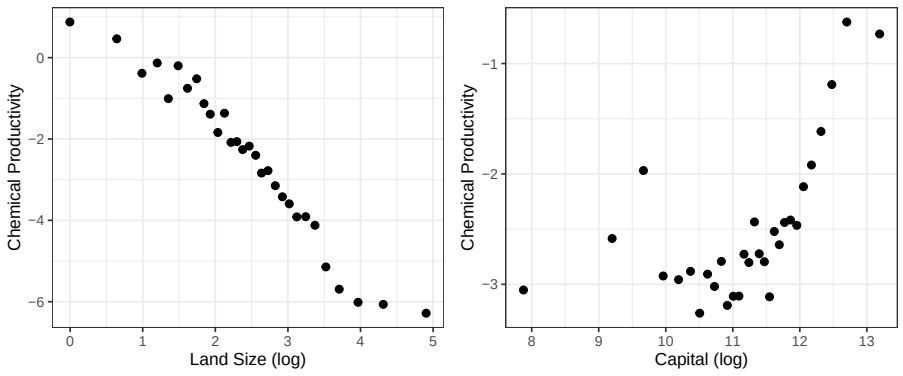
Notes: The first row of figures shows the evolution of productivity shocks as measured in the FADN, resp. the non-Hicksian shock which we introduce as farm-specific and shared across crops, and the two TFP shocks which are differentiated across crops within farms. Shocks for the first row of figures 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. We 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 A19: 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. We 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.*

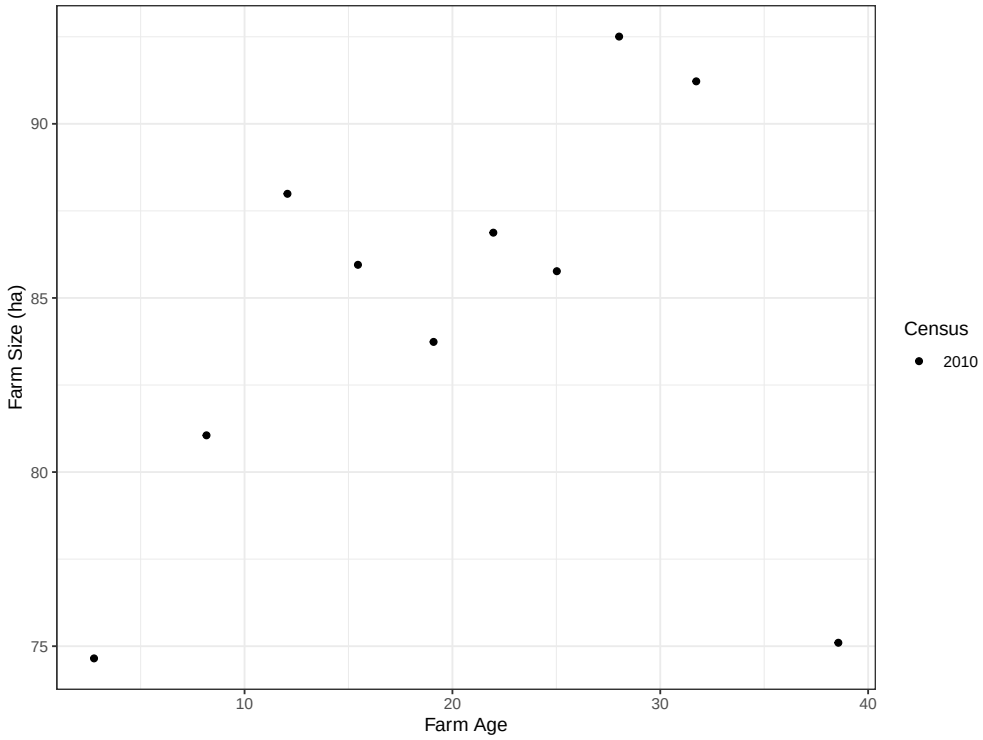
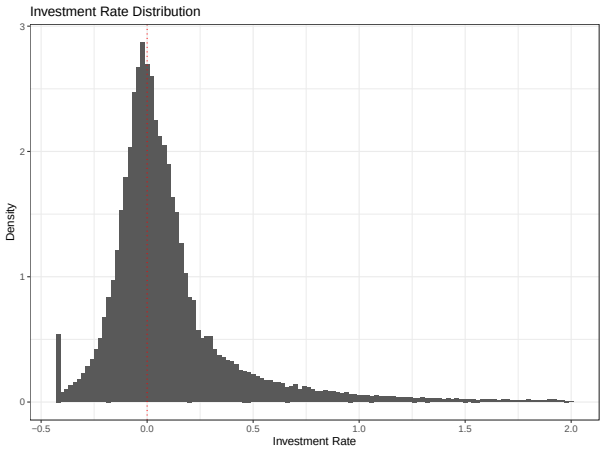
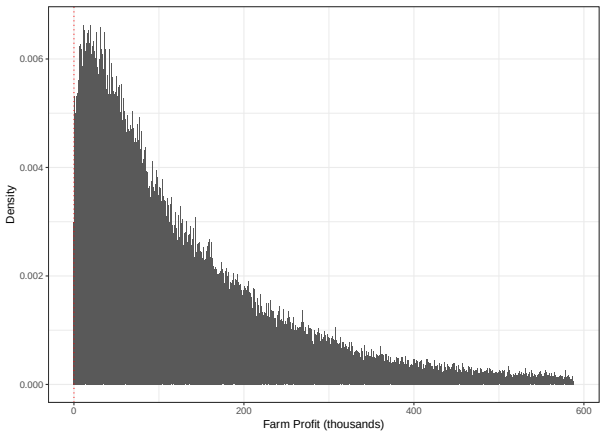


Figure A20: Relation between Farm Size and Farm Age in the 2010 Census

Notes: We plot a binscatter showing the relation between farm size and farm tenure, as observed in the 2010 French agricultural census. We use the 2010 Census to capture this relation, as the total number of farms ceases to decrease as quickly between 2000 and 2010, and we 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, we were more likely to be small non-commercial farms with different management styles.

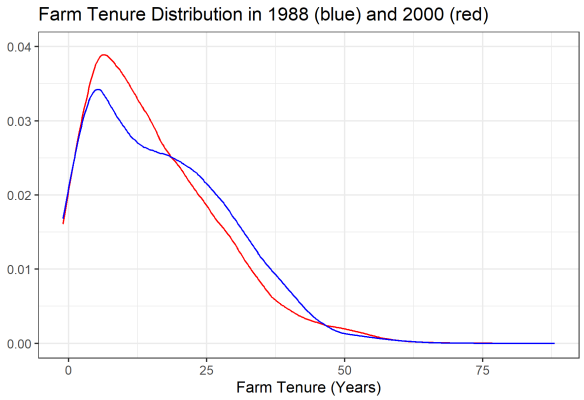


(a) Distribution of Investment Rates (FADN)



(b) Distribution of Profits (FADN)

Figure A21



(a) Distribution of Tenure (Census)

Figure A22

Notes: We plot the distributions from which we 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.

Table A1: Policy Outcomes

Panel A			**Wheat Market**						
Equilibrium	Subsidy Per Unit	Chemicals-Land	Chemicals-Output	Profit	Price	Effective Price	Output	Land	Chemicals
Land Subsidy	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Price Intervention	1087.16	163.49	152.15	127.93	143.00	127.37	167.75	156.15	255.30
Chem. Tax	0.00	99.36	71.53	99.77	110.82	98.71	90.02	64.85	64.43
Baseline	0.00	196.71	139.14	98.87	109.82	97.82	89.42	63.27	124.44
Panel B			**Other Crops**						
Equilibrium	Subsidy Per Unit	Chemicals-Land	Chemicals-Output	Profit	Price	Effective Price	Output	Land	Chemicals
Land Subsidy	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Price Intervention	13890.28	163.56	160.81	164.01	143.00	141.59	181.49	178.27	291.84
Chem. Tax	0.00	99.35	68.96	95.47	106.69	105.64	89.64	62.22	61.81
Baseline	0.00	196.71	114.95	101.27	100.36	99.37	86.48	50.55	99.41

Notes. We describe market-level outcomes under the different policy regimes we model using our counterfactual simulations. Values are expressed in percentage points of the levels of the land subsidy regime – which is the current subsidization form of agriculture within the EU. We show values for each of our two crop markets.

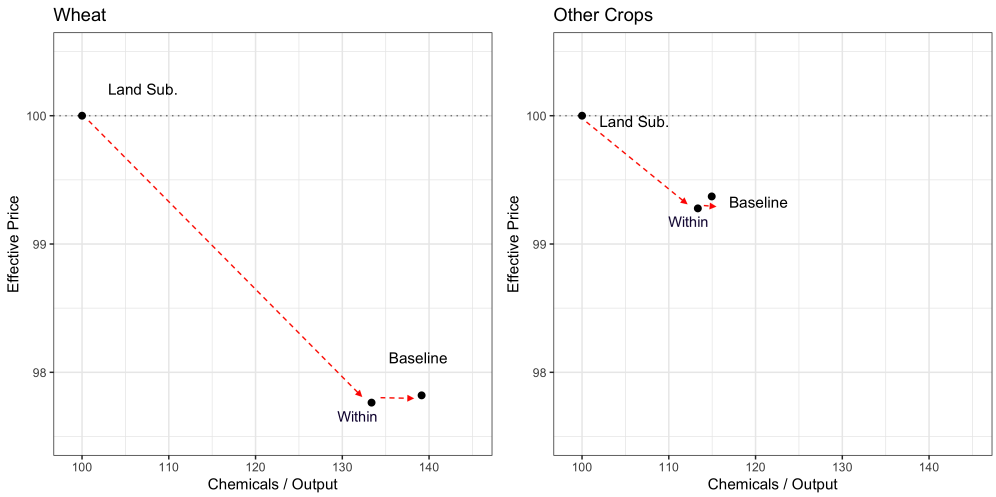


Figure A23: Impact of the Land Subsidy

Notes: We show the effective prices (crop price on the market plus subsidy per unit) and chemicals-to-output ratios linked with both the land subsidy and no policy intervention, on both crop markets. Values are expressed in percent of the effective price and chemicals-to-output ratio under the land subsidy. We decompose the change between land subsidy and no-intervention using the Within point. Movements from *Land Sub.* to *Within* correspond to within farm-crop treatment effects of removing the subsidy, while the movement from *Within* to *Baseline* describe the additional effect of reallocations (changing market shares).

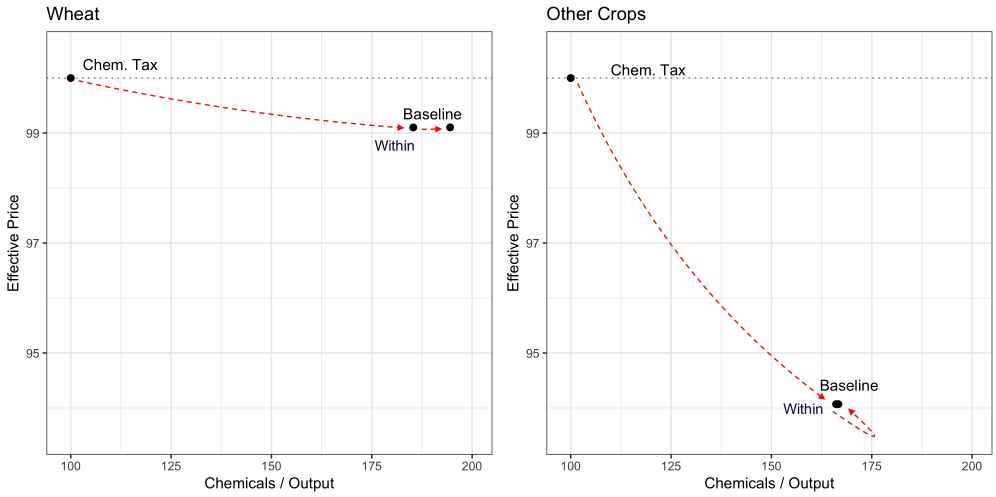


Figure A24: Impact of the Chemicals Tax

Notes: We show the effective prices (crop price on the market plus subsidy per unit) and chemicals-to-output ratios linked with both the chemicals tax and no policy intervention, on both crop markets. Values are expressed in percent of the effective price and chemicals-to-output ratio under the chemicals tax. We decompose the change between chemicals tax and no-intervention using the Within point. Movements from *Chem. Tax* to *Within* correspond to within farm-crop treatment effects of removing the tax, while the movement from *Within* to *Baseline* describe the additional effect of reallocations (changing market shares).

A.2 Tables

Table A2: 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 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)
	1970	1583279	29904735.3300		18.8878		0.5813	
	1979	1257168	29496571.6300		23.4627		0.5792	
	1988	1006120	28595798.5400		28.4219		0.5816	
	2000	653090	27856312.8900		42.6531		0.6130	
	2010	518925	27832907.9000		53.6357		0.6263	
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)
	1970	1034687	25984231.5400	9006112.4200	25.1131	8.7042	0.4916	0.6571
	1979	823432	25882409.8300	11568370.0200	31.4324	14.0490	0.4836	0.6578
	1988	602476	23827571.5600	11189216.8300	39.5494	18.5721	0.4771	0.6534
	2000	352098	22884643.2800	11105923.7700	64.9951	31.5421	0.4754	0.6295
	2010	260731	21827942.3500	10979980.2300	83.7182	42.1123	0.4608	0.6008

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 we 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: Evolution of Agricultural Land Ruggedness

	Year	Average	$10th$	$25th$	$75th$	$90th$	$99th$
1	1970	33.2264	13.3583	19.3155	41.8313	58.5489	107.5599
2	1979	33.1215	13.3772	19.3356	41.6604	58.3028	105.9168
3	1988	33.0303	13.3583	19.2822	41.5127	58.0048	105.8695
4	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. We 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 A6: Farm-Level Shift Share Results

Dependent Variables:	Output Price (log)	Sales (log)	Subsidies (log)	Profit (log)	Chemical-to-Land (log)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
$Exposure_j$l × 1984	-0.0718	0.0232	0.4192**	0.0209	0.2882**
	(0.0858)	(0.0610)	(0.1995)	(0.0509)	(0.1178)
$Exposure_j$ × 1985	-0.0048	0.0344	0.1238	0.0525	0.4619***
	(0.0470)	(0.0428)	(0.3202)	(0.0657)	(0.0730)
$Exposure_j$ × 1986	0.1386*	0.0099	-0.3486	0.0298	-0.1221
	(0.0806)	(0.0347)	(0.3351)	(0.0958)	(0.3536)
$Exposure_j$ × 1987	0.0921	0.0055	0.4313	0.0589	-0.0348
	(0.0596)	(0.0308)	(0.2699)	(0.0577)	(0.1622)
$Exposure_j$ × 1988	-0.0921	-0.0167	0.3283	-0.0184	0.1871*
	(0.0899)	(0.0386)	(0.2371)	(0.0749)	(0.1083)
$Exposure_j$ × 1989	-0.0267	0.0334	0.1405	0.0193	0.0515
	(0.0913)	(0.0298)	(0.1568)	(0.0453)	(0.1495)
$Exposure_j$ × 1990	-0.0363	-0.0226	0.2643	-0.0732	0.1115
	(0.0318)	(0.0501)	(0.2123)	(0.0498)	(0.1283)
$Exposure_j$ × 1992	-0.1109***	-0.0370	0.3724*	-0.1003	0.1060
	(0.0270)	(0.0249)	(0.1891)	(0.1041)	(0.1041)
$Exposure_j$ × 1993	-0.1896***	-0.2166***	0.6124	-0.3123***	-0.1561
	(0.0719)	(0.0243)	(0.3806)	(0.0756)	(0.1637)
$Exposure_j$ × 1994	-0.0879**	-0.2443***	0.7901**	-0.3157***	-0.6522***
	(0.0413)	(0.0640)	(0.3587)	(0.1002)	(0.2023)
$Exposure_j$ × 1995	-0.0862**	-0.2980***	0.9869***	-0.3966***	-0.5027***
	(0.0424)	(0.0595)	(0.2114)	(0.1243)	(0.1078)
$Exposure_j$ × 1996	-0.2202***	-0.2733***	0.6958**	-0.1065	-0.4212***
	(0.0456)	(0.0789)	(0.3147)	(0.1125)	(0.1266)
$Exposure_j$ × 1997	-0.2443***	-0.3559***	1.063***	-0.3426**	-0.3838***
	(0.0374)	(0.0861)	(0.1932)	(0.1512)	(0.1455)
$Exposure_j$ × 1998	-0.2066***	-0.2036***	1.001***	-0.3820***	-0.3931***
	(0.0568)	(0.0637)	(0.2050)	(0.0807)	(0.1197)
$Exposure_j$ × 1999	-0.1487	-0.3244***	1.178***	-0.0750	-0.2736**
	(0.0998)	(0.0647)	(0.2458)	(0.0790)	(0.1267)
$Exposure_j$ × 2000	-0.2042***	-0.3312***	1.078***	-0.1838*	-0.4259*
	(0.0328)	(0.0636)	(0.2563)	(0.1049)	(0.2217)
$Exposure_j$ × 2001	-0.0448	-0.2064***	0.8146***	-0.1429**	-0.1132
	(0.1253)	(0.0577)	(0.1595)	(0.0588)	(0.1533)
$Exposure_j$ × 2002	-0.1745**	-0.2771***	0.7911***	-0.0934	0.3113
	(0.0843)	(0.0725)	(0.1744)	(0.1373)	(0.1909)
Fixed-effects					
Region-Year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	2,204	3,856	2,757	3,856	3,391
R^2	0.79008	0.32947	0.65217	0.69880	0.32932

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

Notes. We 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. We control for the farms' 1991 total sales, capital stock, agricultural area and labor, allowing for a time-varying intercept for each of these controls. We add region-year fixed effects, and cluster the standard errors at the department level.

Table A7: Village-Level Shift Share Results

Dependent Variables:	Farm Count	Average Farm Size	Max Farm Size
Model:	(1)	(2)	(3)
Variables			
$Share_k \times 1979$	-0.1216	1.751	10.12
	(0.2387)	(1.581)	(25.52)
$Share_k \times 2000$	-1.827***	1.959*	30.00*
	(0.6799)	(1.028)	(15.51)
$Share_k \times 2010$	-1.922**	3.796**	50.62**
	(0.8092)	(1.595)	(20.00)
Fixed-effects			
County × Year	Yes	Yes	Yes
Fit statistics			
Observations	150,735	150,735	150,735
R^2	0.79013	0.52641	0.58401

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

Notes. We show the results of our shift-share event-study design at the village level. These coefficients correspond to the year-specific coefficient associated to farm exposure. Outcomes are first-differenced, ...

Table A8: 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. We obtain estimates of total quantity sold and average prices for France for our two crop categories using the FADN and its sampling weights. We then run regressions of sales on prices, instrumenting the prices with weather shocks in order which should shift supply but not demand. We 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 A9: 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: We use the recovered productivity shocks to estimate these other estimated parameters. Specifically, we 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.

B Appendix

B.1 Data Description

B.1.1 Intervention Prices

For the purpose of this paper, and to the best of our knowledge, we 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, we used the CAP-specific exchange rate which was then used to translate ECUs to each member state's currency. We similarly gathered data on commodity-specific land subsidies, which we 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. We 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:

$$i_t = i_{t-1}\left(\prod_i^n \frac{P_{i,t}}{P_{i,t-1}}\right)^{\frac{1}{n}} \tag{11}$$

We 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. We consider the hourly wage data provided by workers whose occupation is classified as that of an agricultural worker or farmer.

Finally, we 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. We 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 we 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.

We 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, we aggregate the hourly data into growing-season observations [70]: growing degree days (GDD) and heat degree days (HDD) for temperature[71], and total precipitation for rainfall. In order to match the forecast dataset, we 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 [72]. 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. We 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). We approximate this by bundling together forecasts produced one calendar month ago (our

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

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

[72] ECMWF provides an ensemble of 25 forecasts, which we average.

one month lead for the rest of the paper), produced two calendar months ago, up to five months ago. As such, the forecasts that we aggregate into growing season observations are not homogeneous in terms of lead value, but are the closest equivalent of it that we can get.

Similarly to the weather realization data, we 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, we filter Landsat data for images taken between June and August, which we crop for surface water and treat for clouds and cloud shadows. We then run the following function on each pixel:

$$Bloom = NIR - 1.03 * SWIR \tag{12}$$

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.[73]

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). We use the following workflow. We 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. We then use first a Gaussian filter, and then a Canny edge detection filter using four bands: red, blue, green and near infrared. We 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. We 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.

[73] 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 we think 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.

We 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. We 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[74], 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.

We 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

We 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.[75]

We 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.[76] 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.[77] 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)).

We 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.[78] We do so controlling

[74] By this we 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.

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

[76] 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.

[77] 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.

[78] We show in the annex, resp. Figure A6 and Figure A7, maps for the evolution at the village level of both total and average

for region-by-year fixed effects to further isolate the effect of agricultural growth. We 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.

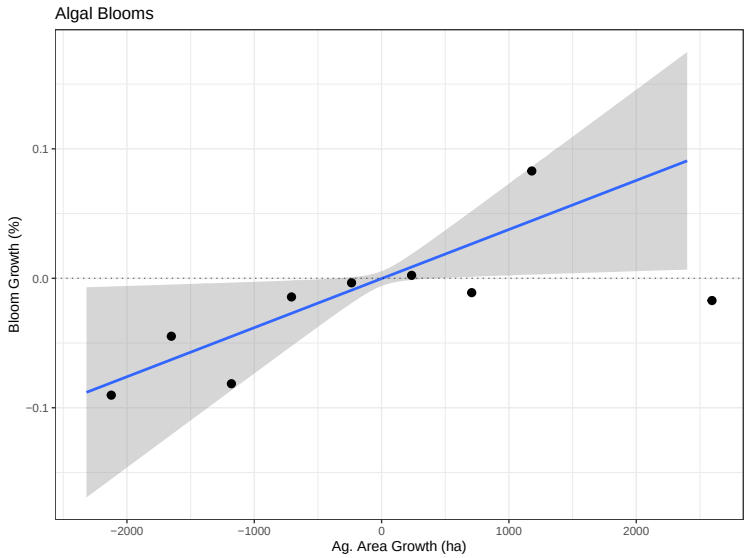


Figure A25: Correlates of Agricultural Area Expansion

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.

We 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 we 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. We observe a clear positive and linear relation between farm size and these two ratios. We 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,

farm size. While average area is increasing everywhere, there are clear 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.

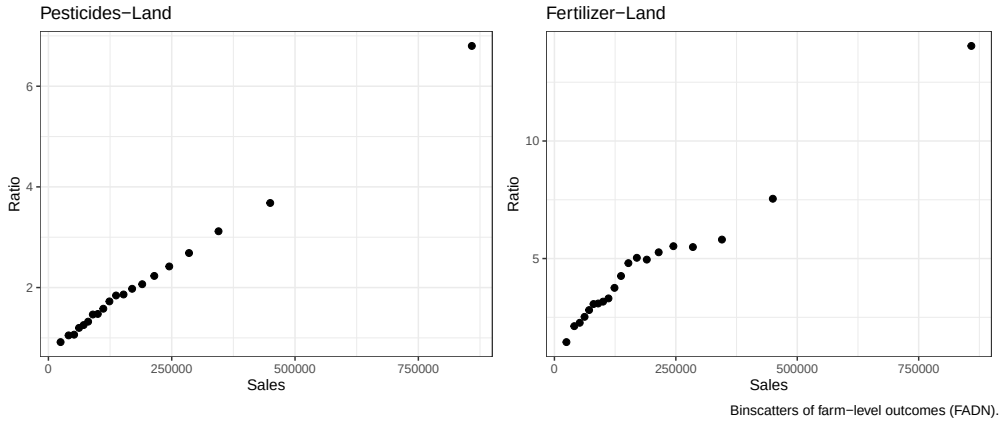


Figure A26: Chemical-to-Land Ratios in relation with Farm Size

Notes: We 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. We 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.2.3 A Mechanism: Suggestive Evidence for Chemical-Using Technological Change

We use farm-level data to compute fertilizer-to-land and phytosanitary-to-land input ratios as observed in the FADN. We further normalize the ratios 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. These graphs show that 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 is likely indicative of non-Hicksian technological change – in the form of chemical-enhancing productivity if land and chemicals are substitutes, or land-enhancing if they are complement. Because soil and climatic conditions can reasonably be considered as constant over such a short period of time – meaning that land productivity should remain constant – they are likely suggestive of chemical-intensive productivity change.

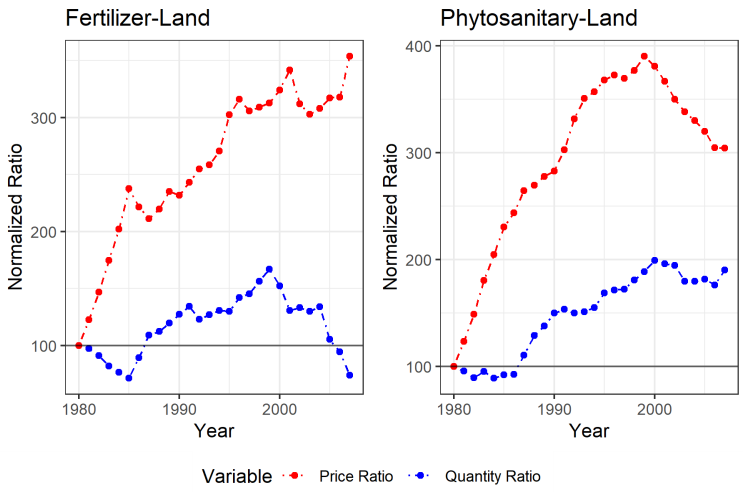


Figure A27: Suggestive Evidence for non-Hicksian Productivity Changes

Notes: These binscatters are computed using combined data from the FADN, the INSEE Ipampa price index series, and the Valeur Venale des Terres land price series. Land prices are corrected for the land subsidies in effect. We 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.

We note that a European Nitrate Directive started being implemented in France between 1997 and 2000, limiting the quantity of nitrogen which could be used by farms in at-risk area, which could explain part of the decrease observed from 1999 onwards.[79]

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

We 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. We 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. We then provide a more in-depth analysis of our conditional parallel-trends assumption driving our difference-in-difference design. We show trends in the same set of variables as shown in Figure A12, 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, we 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.

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

B.3.1 Rotemberg Weights

We 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 A10: Rotemberg Weights: With Controls

Panel A	Rotemberg Weights		
Crop	Weight	Shock	Share Variance
Sunflower	0.85	389.88	0.12
Wheat	0.12	70.72	0.031
Corn	0.057	79.63	0.083
Durum	0.015	171.69	0.015
Sorghum	0.0037	77.33	0.024
Oats	0.0011	-1.89	0.0057
Rye	0.00071	-0.57	0.010
Barley - Winter	-0.013	24.69	0.000096
Barley - Spring	-0.017	7.66	0.034
Colza	-0.016	205.17	0.010
Panel B	Correlation Matrix		
Weights	1.000	0.813	0.822
Shocks	0.813	1.000	0.655
Variance Shares	0.822	0.655	1.000

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.

As we can see, the weights vary in a quite striking way as we include the 1991 relative oil share. Most importantly we 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. We 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. We will investigate this later with a crop-by-crop iv strategy.

Table A11: Rotemberg Weights: Without Controls

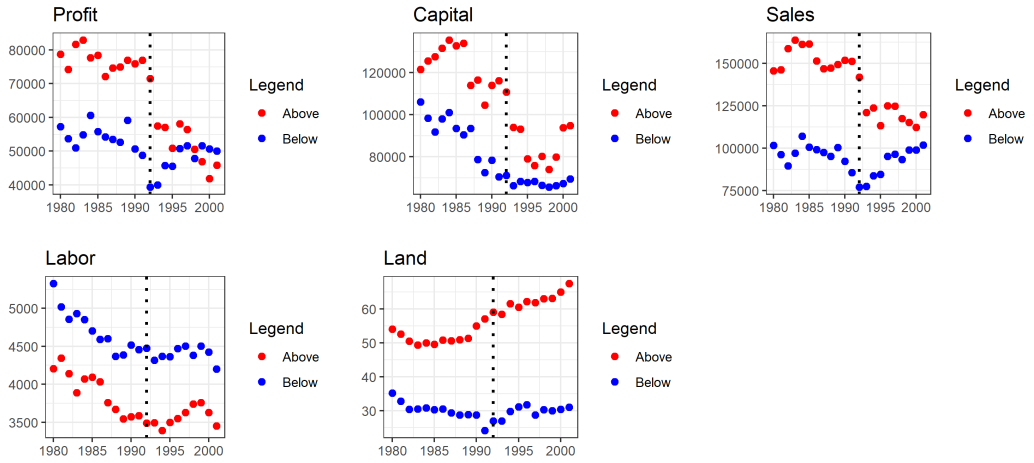
Panel A	Rotemberg Weights		
Crop	Weight	Shock	Share Variance
Sunflower	0.7529	389.8818	1.1710e-01
Wheat	0.0789	70.7184	3.0733e-02
Corn	0.0601	79.6289	8.2996e-02
Durum	0.0574	171.6932	1.5084e-02
Colza	0.0461	205.1690	2.3679e-02
Barley - Winter	0.0070	24.6915	5.7327e-03
Sorghum	0.0013	77.3312	1.0492e-02
Rye	0.0003	-0.5674	9.5837e-05
Oats	0.0002	-1.8912	3.3656e-02
Barley - Spring	-0.0044	7.6645	1.0507e-02
Panel B	Correlation Matrix		
Weights	1	0.8612	0.8208
Shocks	0.8612	1	0.6938
Variance Shares	0.8208	0.6938	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.

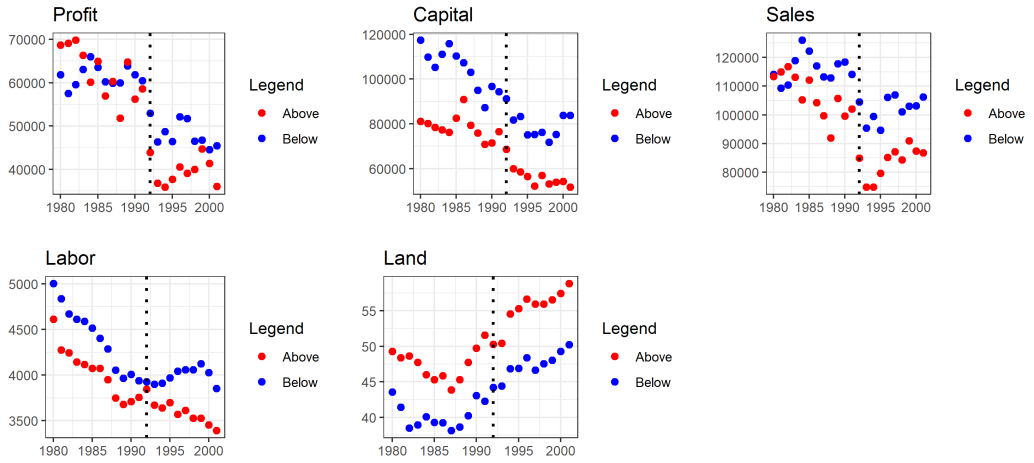
B.3.2 Balance Tests: Across Crop Shares

Next, we 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, we should keep two things in mind. The first thing is that everything else equal, we 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. We note that because we 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, we 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.



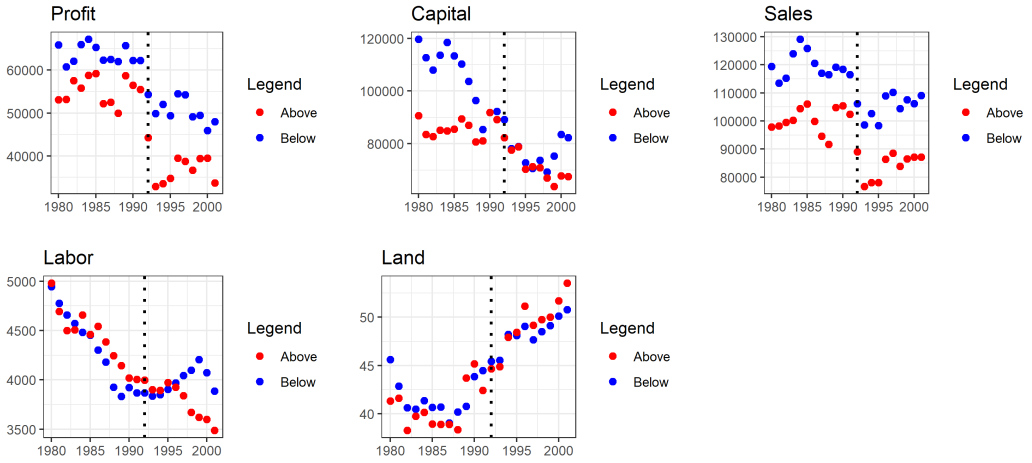
(a) Balance Tests - Wheat



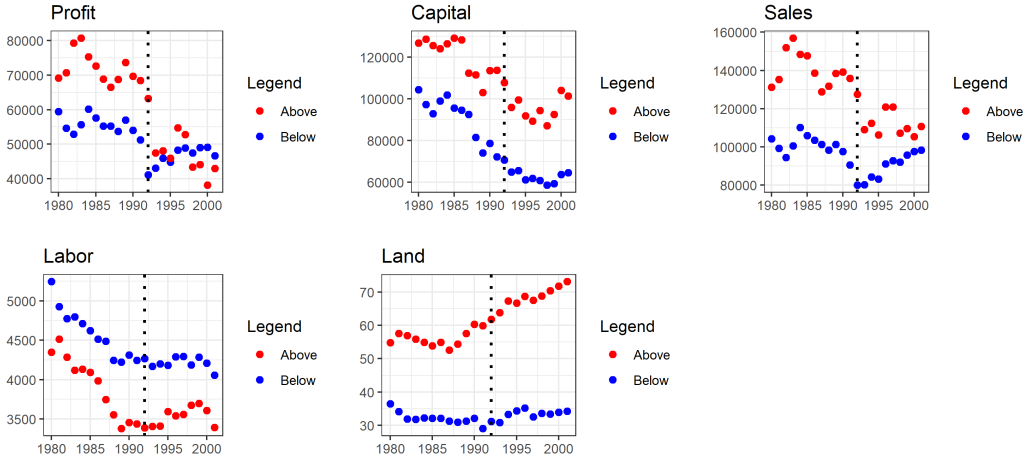
(b) Balance Tests - Sunflower

Figure A28: Trends per Category

Notes: We 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. We use FADN data for this exercise.



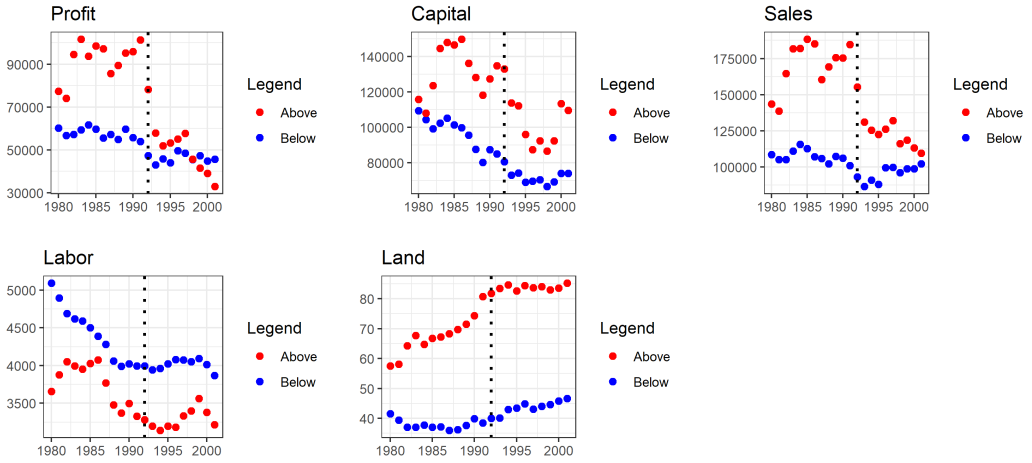
(a) Balance Tests - Corn



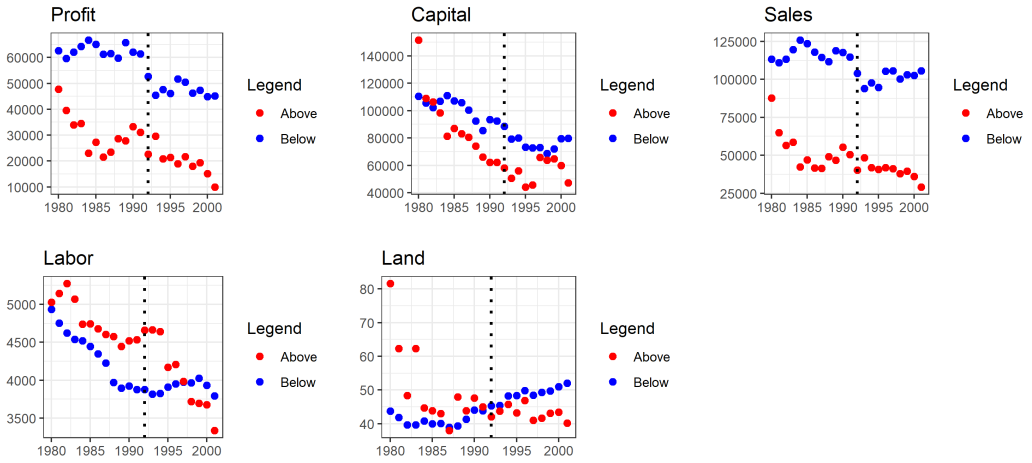
(b) Balance Tests - Barley

Figure A29: Trends per Category

Notes: We 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. We use FADN data for this exercise.



(a) Balance Tests - Colza



(b) Balance Tests - Rye

Figure A30: Trends per Category

Notes: We 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. We use FADN data for this exercise.

B.3.3 Additional Farm-Level Results

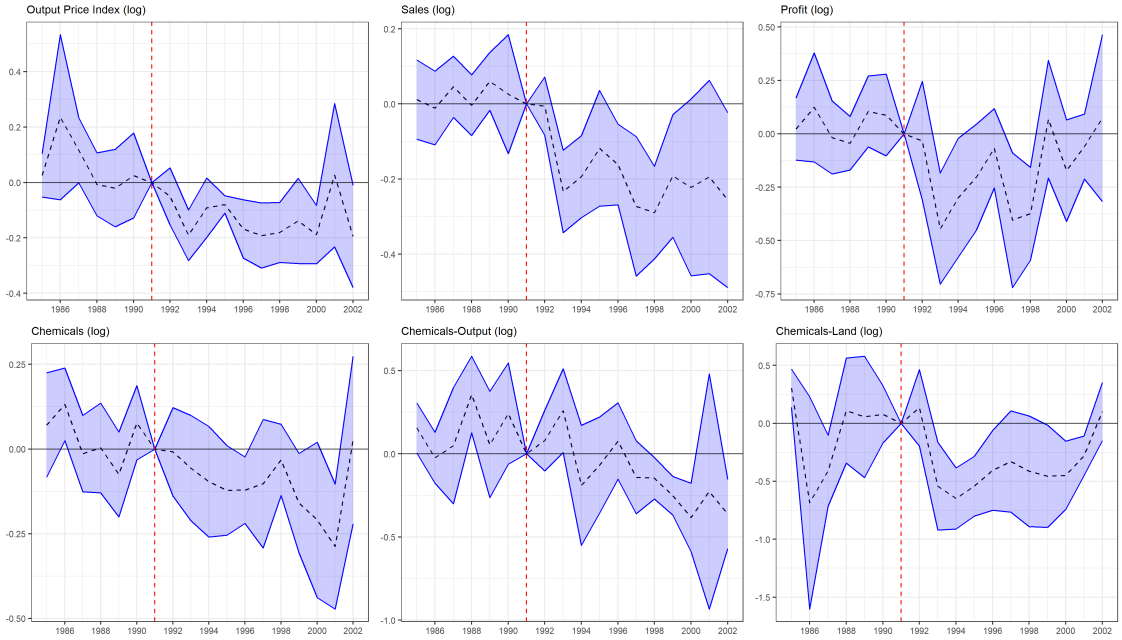


Figure A31: Farm-Level Event Study: Full Set of results

Notes: We plot graphically the result of our 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 sales, total subsidies received, profit measured as value added plus subsidies, their capital stock, their expenses in insurance, and the farms' chemical-to-land, fertilizer-to-land and phytosanitary-to-land ratios where chemicals and their sub-categories are always measured as the deflated farm-level input bill, and land is measured in ha. 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. We control for: 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, and the evenness of their crop mix in 1991. We also control for their 1983-1984 adoption trends in chemicals measured as their evolution of chemical-to-land and chemical-to-output ratio growth. We allow for a time-varying intercept for each of these controls. We add region-year fixed effects, and cluster the standard errors one-way at the region-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. We can do so using a modified estimating regression of the following form:

$$\Delta^{1991} Y_{jt} = Exp_j \beta^0 + Exp_j Post_t \beta^1 + X_j \Gamma^0 + X_j Post_t \Gamma^1 + \eta_{r(j)t} + \varepsilon_{jt} \tag{13}$$

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. We 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. We 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 A12: Difference-in-Difference Results

Panel A	Main Shift-Share Design (0/1)					
Dependent Variables:	Price (log)	Profit (log)	Sales (log)	Chemicals (log)	Chemicals-Output (log)	Chemicals-Land (log)
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
$Exposure_j$	0.0188	0.0048	-0.0307	0.0278	0.0833	-0.0033
	(0.0331)	(0.0419)	(0.0227)	(0.0269)	(0.0655)	(0.1025)
$Exposure_j \times Post_t$	-0.1627***	-0.1645**	-0.1732***	-0.1654***	-0.2460***	-0.3524***
	(0.0390)	(0.0646)	(0.0396)	(0.0401)	(0.0862)	(0.1226)
Mean Level	231.8	165,708.1	165,708.1	17,024.6	47.62	217.8
Fixed-effects						
Region-Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	1,612	2,251	2,251	2,251	2,088	2,141
R^2	0.78046	0.75309	0.39206	0.36372	0.41327	0.45596
Panel B	Alternative Design (1/1)					
Dependent Variables:	Price (log)	Profit (log)	Sales (log)	Chemicals (log)	Chemicals-Output (log)	Chemicals-Land (log)
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
$Exposure_j^{alt}$	-0.0176	0.0124	-0.0316	0.0199	0.0590	0.0405
	(0.0273)	(0.0357)	(0.0201)	(0.0228)	(0.0526)	(0.0817)
$Exposure_j^{alt} \times Post_t$	-0.1270***	-0.1413***	-0.1381***	-0.1246***	-0.2117***	-0.3068***
	(0.0316)	(0.0534)	(0.0320)	(0.0331)	(0.0692)	(0.0975)
Mean Level	231.8	165,708.1	165,708.1	17,024.6	47.62	217.8
Fixed-effects						
Region-Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	1,612	2,251	2,251	2,251	2,088	2,141
R^2	0.78479	0.75315	0.39398	0.36280	0.41473	0.45499
Within R^2	0.20777	0.09371	0.17121	0.21367	0.29322	0.30749
Clustered (Region-Year) standard-errors in parentheses						

Notes. We show the difference-in-difference results on a farm-level output price index, farm-level profit and total subsidies. 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. The output price index and profit are both differenced with respect to their 1991 value. We 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. We also control for their 1983-1984 adoption trends in chemicals measured as their evolution of chemical-to-land and chemical-to-output ratio growth. We allow for a time-varying intercept for each of these controls. We add region-year fixed effects, and cluster the standard errors at the region-year level.

We conclude this section by performing some heterogeneity analysis in Table A13, showing results for below and above median capital stock holders (as set in 1991). To simplify the comparison across the groups, we show the results for a difference-in-difference design where the coefficients associated resp. to $Exposure_j$ and $Exposure_j * Post_t$, resp. give the average effect of exposure pre-1991, and the relative additional effect of exposure post-1991. We write the regression as follows, and perform the regressions separately for each group:

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

These results are interesting as all the negative price and profit effects are concentrated among low capital farms. Coefficients for larger farms are both smaller in absolute values, and non-significant. This hints at the fact that the negative effects of the reform were concentrated among smaller farms, and that exit and reallocation should then be expected to happen in a way that shrinks the share of production allocated to these smaller establishments. If larger farms differ in terms of productivity and externalities from the smaller ones, these shifts will likely have a significant impact on market-level efficiency and pollution.

Table A13: Heterogeneity across Farm Size

Dependent Variables:	Sales (log)		Profit (log)	
Model:	(1)	(2)	(3)	(4)
Variables				
$Exposure_j$	0.0459	-0.0075	0.0836	0.0324
	(0.0317)	(0.0456)	(0.0627)	(0.0610)
$Exposure_j \times Post_t$	-0.2702***	-0.0021	-0.2264***	0.1338
	(0.0526)	(0.0682)	(0.0819)	(0.1226)
Mean Level	233.5	230.9	110,765.4	219,993.0
Fixed-effects				
Region-Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	1,106	1,145	1,106	1,145
R^2	0.50320	0.47801	0.80243	0.77368

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

Notes. We show the heterogeneity of our difference-in-difference results on a farm-level output price index, and farm-level profit. We show results for groups below and above the median capital stock in 1991. The output price index is computed as a weighted average of the farm's output prices, using the crops' relative land use as weights. $Exposure_j$ gives the pre-1992 effect of farm exposure on the outcome's growth rate, while $Exposure_j * Post_t$ gives the relative change in the effect of exposure after 1991. The output price index and profit are both differenced with respect to their 1991 value. We 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, and the evenness of their crop mix. We also control for their 1983-1984 adoption trends in chemicals measured as their evolution of chemical-to-land and chemical-to-output ratio growth. We allow for a time-varying intercept for each of these controls. We add region-year fixed effects, and cluster the standard errors one-way at the region-year level.

B.3.4 Independent Instruments

Finally, we decompose our 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. We only show the coefficients for the instruments interacted with our post-1991 dummy. These indicate the effect of crop-specific exposure to the reform after the reform happened. These regressions show significant heterogeneity in the results when it comes to sales – where decreases in sales volumes linked with exposure are specifically linked to the cultivation of sunflower, colza and durum – while sorghum and rye are associated with positive and

significant effects. These last two crops represent a small share of total row crop shares. Overall, the significant negative effects of the reform on sales are linked with the cultivation of oil crops, which we could also observe with the Rotemberg weights decomposition. In our regressions, we control for the share of output allocated to oil crops by farms in 1991, in order to control for the potential unobserved differences between farms growing oil crops and those that do not. These sales results translate to the profit ones, where the two significant coefficients are associated with oil crops. The decrease in the level of chemicals use is however more homogeneously negative, with negative coefficients for all the largest crops in terms of market share (wheat, barley, corn, sunflower and colza), and with all the significant coefficients being negative.

Table A14: Heterogeneity in Crop-Specific Effects

Dependent Variables:	Sales (log)	Profit (log)	Chemicals (log)
Model:	(1)	(2)	(3)
Variables			
Wheat × $Post_t$	-0.0035	0.0031	-0.0023
	(0.0022)	(0.0045)	(0.0021)
Barley (w) × $Post_t$	0.0031	0.0204	-0.0158**
	(0.0077)	(0.0156)	(0.0066)
Corn × $Post_t$	-0.0030	0.0020	-0.0043**
	(0.0021)	(0.0037)	(0.0021)
Rye × $Post_t$	0.7275**	-0.3893	0.1364
	(0.2915)	(0.6132)	(0.3092)
Sunflower × $Post_t$	-0.0031***	-0.0029*	-0.0029***
	(0.0009)	(0.0016)	(0.0009)
Colza × $Post_t$	-0.0054***	-0.0064**	-0.0033**
	(0.0015)	(0.0032)	(0.0014)
Durum × $Post_t$	-0.0143***	-0.0083	-0.0121***
	(0.0045)	(0.0051)	(0.0035)
Sorghum × $Post_t$	0.0627***	0.0104	0.0086
	(0.0155)	(0.0213)	(0.0147)
Barley (s) × $Post_t$	-0.0062	0.0334	-0.0144
	(0.0174)	(0.0369)	(0.0204)
Fixed-effects			
Region-Year	Yes	Yes	Yes
Fit statistics			
Observations	2,251	2,251	2,251
R^2	0.41908	0.75824	0.37582

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

Notes. We show the heterogeneity of our difference-in-difference results across crop-specific shocks. The instrument is decomposed in multiple crop-specific instruments. We include the crop shares in levels, but do not show the coefficients in the table. We 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. We also control for their 1983-1984 adoption trends in chemicals measured as their evolution of chemical-to-land and chemical-to-output ratio growth. We allow for a time-varying intercept for each of these controls. We add department-year fixed effects, and cluster one-way the standard errors at the region-year level.

B.4 Village-Level Results

Table A15: Village-Level Shift-Share Results

Dependent Variables:	Farms	Ag. Area	Max Size	Evenness (f)	Evenness (v)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
$Share_k \times 1988$	-0.4988	-0.6255	-1.836	-0.0057	0.0096
	(0.3251)	(8.394)	(2.243)	(0.0088)	(0.0107)
$Share_k \times 2000$	-1.053***	20.25***	4.462*	0.0371***	0.0044
	(0.2855)	(7.090)	(2.434)	(0.0067)	(0.0096)
Mean Change	-6.868	147.0	4,374.3	-0.0277	-0.0522
Mean Level	17.60	69,604.1	11,075.6	0.3242	0.4564
Fixed-effects					
County-Year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	88,671	88,671	88,671	88,671	88,671
R^2	0.61643	0.28472	0.26929	0.35634	0.42806

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

Notes. The shift-share regression is here run on the set of metropolitan French villages for which we can build a measure of exposure to the reform. By metropolitan, we mean that we exclude any French villages not located on the mainland French territory, excluding Corsica as well. Controls included in the regression but not shown are: the 1988 share of farms in the village with no exposure at all to the reform, the village's total agricultural area, the total area used for the cultivation of the crops we are interested in, and average farm size. All controls are interacted with time-varying coefficients. We also include county-by-year fixed effects. Standard errors are clustered at the department level.

Table A16: Village-Level Shift-Share for Ag. Land Results

Dependent Variables:	Ag. Area	Min	$25th$	Mean	$75th$	Max
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
$Share_k \times 1988$	-0.6255	-1.311**	-1.213*	-1.324	-1.003	-1.836
	(8.394)	(0.6492)	(0.7272)	(0.9010)	(1.172)	(2.243)
$Share_k \times 2000$	20.25***	2.730*	3.581**	3.727**	3.439**	4.462*
	(7.090)	(1.481)	(1.417)	(1.442)	(1.677)	(2.434)
Fixed-effects						
County-Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	88,671	88,671	88,671	88,671	88,671	88,671
R^2	0.28472	0.27428	0.30146	0.36108	0.33887	0.26929

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

Notes. The shift-share regression is here run on the set of metropolitan French villages for which we can build a measure of exposure to the reform. By metropolitan, we mean that we exclude any French villages not located on the mainland French territory, excluding Corsica as well. Controls included in the regression but not shown are: the 1988 share of farms in the village with no exposure at all to the reform, the village's total agricultural area, the total area used for the cultivation of the crops we are interested in, and average farm size. All controls are interacted with time-varying coefficients. We also include county-by-year fixed effects. Standard errors are clustered at the department level.

Table A17: Village-Level Shift-Share: Comparison of Exposure Aggregation

Dependent Variables:	Farm Count			Total Ag. Area		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
$Share_k \times 1988$	-0.4988			-0.6255		
	(0.3251)			(8.394)		
$Share_k \times 2000$	-1.053***			20.25***		
	(0.2855)			(7.090)		
Exposure (*average*) × 1988		0.0040			1.845	
		(0.0643)			(2.787)	
Exposure (*average*) × 2000		-0.0510			4.996***	
		(0.0755)			(1.778)	
Exposure (*max*) × 1988			0.0237			-3.596
			(0.0918)			(2.370)
Exposure (*max*) × 2000			-0.2249**			4.268***
			(0.0954)			(1.603)
Fixed-effects						
County-Year	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	88,671	88,671	88,671	88,671	88,671	88,671
R^2	0.61643	0.61627	0.61637	0.28472	0.28470	0.28476

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

Notes. This table shows the results of our village-level event-study using the design of Equation 3. We vary the way we aggregate the farm-level measure of exposure to the village-level, resp. using our preferred measure of the share of farms with an above-median exposure to the reform within the village in 1988, which we write as $Share_k$. This share is computed for farms with a positive exposure, hence active in the row crop market in 1988. This is the measure for which we show results in the paper. We then aggregate exposure using a simple unweighted average exposure, and the maximum exposure – both within farms with a positive exposure. We always control for the share of farms not active on the row-crop market in 1988. These regressions include the same set of controls as the village-level regressions shown in the paper.

B.5 Model Details

It is useful to re-write the incumbent problem under our additional set of assumptions, as well as outline the entrant's problem.

$$
\begin{aligned}
V(\omega^h_{jt}, \omega^{ch}_{jt}, K_{jt}, \mathbb{C}_{jt}; \Omega_t) = & \underset{\{X_{jct}\}_{\mathbb{C}_{jt}}, \xi^x_{jt}, K_{jt+1}, \mathbb{C}_{jt+1}}{max} \underbrace{\Pi_{jt}}_{\text{Static Profit}} + \underbrace{C(K_{jt}, K_{jt+1})}_{\text{Capital Adjustment}} \\
& + \xi^x_{jt} \underbrace{\left[(1-\delta_k)K_{jt}\right]}_{\text{Scrap Value}} - (1-\xi^x_{jt}) \underbrace{f_k}_{\text{Cost of Incumbency}} \\
& + (1-\xi^x_{jt}) \Bigg(\underbrace{\mathbb{E}_t \left[\frac{1}{1+r} V(\omega^h_{jt+1}, \omega^{ch}_{jt+1}, K_{jt+1}, \mathbb{C}_{jt+1}; \Omega_{t+1}) | \omega^h_{jt}, \omega^{ch}_{jt}, \Omega_t \right]}_{\text{Expected Value of Continuation}} \Bigg)
\end{aligned} \tag{14}
$$

We introduce two additional distributions, resp. $G_1(.)$ and $G_2(.)$ which are the distributions from which an entrant draws their initial TFP shock ω^h_{jt}, and their non-Hicksian productivity ω^{ch}_{jt}. We can then write the entrant's problem as follows:

$$
\begin{aligned}
V_e(\Omega_t) = & \underset{\xi^e_{jt}, \xi^x_{jt}, K_{jt+1}, \mathbb{C}_{jt+1}}{max} - K_{jt+1} \\
& + (1-\xi^x_{jt}) \Bigg(\underbrace{\mathbb{E}_t \left[\frac{1}{1+r} V(\omega^h_{jt+1}, \omega^{ch}_{jt+1}, K_{jt+1}, \mathbb{C}_{jt+1}; \Omega_{t+1}) | \omega^h_{jt}, \omega^{ch}_{jt}, \Omega_t \right]}_{\text{Expected Value of Continuation}} \Bigg)
\end{aligned} \tag{15}
$$

Embedded is the assumption that entrants do not face adjustment costs for their initial capital investment. The transition operators $\phi_t(\omega', \omega^{ch\prime}, K', \mathbb{C}' | \omega, \omega^{ch}, K, \mathbb{C})$ and $\phi_{e,t}(\omega', \omega^{ch\prime}, K', \mathbb{C}' | \omega, \omega^{ch})$ the require the additional introduction of the transition operator $J^h(.|.)$ and $J^{ch}(.|.)$ for the Hicksian shock, and of the policy functions κ^k for incumbent capital investment, $\kappa^{\mathbb{C}}$ for crop choice, and κ^k_e for entrant capital investment. These transition operators now write as:

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

We 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. And the analogue for entrants is $\phi_{e,t}(\omega', , \omega^{ch\prime}, K' | \omega, \omega^{ch})$:

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

B.6 Joint Production Estimation Framework

Here we describe an alternative joint production framework, for which we also provide estimates. This approach combines the work of Dhyne et al. (2022) and of Doraszelski and Jomandreu (2018). From the first, we 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. We 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, we 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 $\tilde{\omega}_{jct}$. We 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. We keep the functions g_{jct} fairly general for the moment, and parametrize them further later. One additional set of assumptions we 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 we 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. We 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. We 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 $\{L_{jt}, Fert_{jt}, Pest_{jt}\}$, we get the following expression:

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

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

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

Combining these two equations, we get:

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

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, we assume a specific shape for the g_{jct} function[80]. We take:

$$g_{jct} = K_{jt}^{\alpha_k} L_{jt}^{\alpha_L} \left\{ \delta_s^c S_{jct}^{\rho} + \delta_p^c \left[e^{\omega_{jt}^p Pest_{jt}}\right]^{\rho} + \delta_f^c \left[e^{\omega_{jt}^f Fert_{jt}}\right]^{\rho} \right\}^{\frac{\alpha_s}{\rho}}$$

Without relying on a translog specification, this is the simplest specification which allows us to model non-hicksian productivity shocks. We note that the heterogeneity in the g_{jct} across crops only relies on the $\{\delta_s^c, \delta_f^c, \delta_p^c\}_c$ parameters. Because we 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 we are interested. A more expansive heterogeneity could specify crop-specific elasticities ρ_c, which we assume away here.

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

$$p_{jt}^x - p_{jt}^s = \rho \omega_{jt}^x + (\rho - 1) x_{jt} + log\Big(\sum_{c \in \mathbb{C}_{jt}} \frac{\delta_x^c}{\delta_s^c} S_{jct}^{1-\rho} \Big)$$

This equation mirrors the first stage estimating equation from Doraszelski and Jaumandreu (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

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

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.

We propose to recover the $\{\delta^c_s, \delta^c_f, \delta^c_p\}_c$, and ρ in a first estimation step, and to recover the remaining parameters in a second step. To simplify the estimation process, we 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: We specify a structure for the two input-biased productivity shocks. For each of them, we assume an AR(1) structure such that, with ζ^x_{jt+1} an exogenous i.i.d. innovation shock:

$$\omega^x_{jt+1} = \mathbb{E}[\omega^x_{jt+1}|\omega^x_{jt}] + \zeta^x_{jt+1} = g^x(\omega^x_{jt}) + \zeta^x_{jt+1}$$

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

$$\mathbb{E}[(\zeta^p_{jt} + \zeta^f_{jt})A^1_{jt}] = 0$$

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, we 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 we can compute using the estimates from the first step. Now, we use crop-specific policy cushions, as well as land prices and agricultural wages as instruments denoted A^2_{jt} in a GMM estimation, with moments:

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

We 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. We 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). We also note that $\rho \in]0, 1[$, indicating that land, pesticides and fertilizers are substitutes in production, which was expected.

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

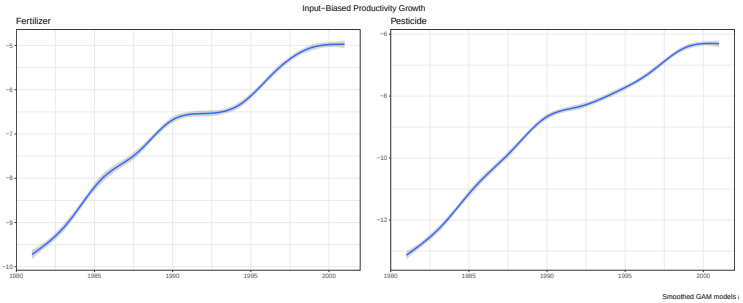


Figure A32: Input-Biased Productivities - Joint Production Framework

Notes: We 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, we estimate our production parameters on single crop-group firms, following De Loecker et al. (2016). Because we observe farm-crop level land allocations, knowing the parameters of our production specification, we 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, we can write the following relation between pesticides and fertilizers:

$$Fert^{1-\rho_2}_{jct} = \frac{P^P_{jt}}{P^F_{jt}} \frac{(1-\delta^c_p)}{\delta^c_p} Pest^{1-\rho_2}_{jct} \tag{16}$$

We 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, we 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 we have estimated the parameters of our production function, the right hand side of this expression is fully known. As we know total farm pesticide use, we can then recover the volumes of pesticide used for each of the crops that they grow. We can then use Equation 16 to recover the fertilizer allocations as well. With these, we 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

We 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, we want to build moments based on the following corrected process:

$$\omega^h_{jct+1} = \mathbb{E}[\omega^h_{jct+1} | \omega^h_{jct}, \Xi_{jct+1} = 1] + \xi^h_{jct+1} \tag{17}$$

With I_{jt} the farm's information set at the end of period t, we 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). We use a competence ladder, which means

that there will be a strict relation between the farm's TFP between c and c'. We 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 $\overline{\omega}^h_{jcc't+1}$.[81] We can then specify $g_h(.)$ as a polynomial of its two terms. While Equation 8 provides an expression for ω^h_{jct}, we also need one for the threshold $\overline{\omega}^h_{jcc't+1}$. For this we 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.[82] 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:

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

Equation 8 is used to move from the third to the fourth line. While we do not observe $\sigma_j^{-1}(c)$, we 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}, we can invert SP_{jt}, which is first recovered non-parametrically, to obtain a proxy for the threshold that we 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

[81] 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.

[82] 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.

of as homogeneous, exogenous local conditions can lead to similar crops having different 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.

We 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:

$$U_t = q^c_{0,t} + \alpha \int_{i\in\Omega} q^c_{i,t} di - \frac{1}{2}\gamma \int_{i\in\Omega} (q^c_{i,t})^2 di - \frac{1}{2}\eta \Big(\int_{i\in\Omega} q^c_{i,t} di \Big)^2 \tag{18}$$

$q_{0,t}$ 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, we note that two varieties can be two types of corn grown by two different farms, or wheat and corn grown by the same farm. We 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.

We 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:

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

We 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, we can also derive a choke price p_t^{max} above which farms will not face any positive demand.

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

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

We 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 we 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. We estimate the following regression using a two-stage least squares strategy, with exogenous shocks to firm supply as instruments. Specifically, we use local

realized and forecasted weather shocks as supply shocks.

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

We recover the following parameters. We 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: We 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. We 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:

$$U_t = q_{0,t}^c - \frac{1}{2}\gamma \int_{i\in\Omega} (q_{i,t}^c)^2 di - \frac{1}{2}\eta \Big(\int_{i\in\Omega} q_{i,t}^c di \Big)^2 + \sum_k \alpha_k \int_{i\in\Omega_k} q_{i,t}^c di \tag{19}$$

This corresponds to the simplest way in which we 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 we 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, we 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:

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

Where we 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 het-

erogeneity, 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. We write the entrant's Bellman as follows:

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

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

Again, we can follow the notation of Hopenhayn and Rogerson (1993) to write the law of motion which describes the evolution of the market. We 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. We 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 $\kappa(\delta, K, \sigma; \{p^{kt}_{max}\}_k)$ the investment policy function, and $\zeta(\delta, K, \sigma; \{p^{kt}_{max}\}_k)$ the incumbency policy function, we have:

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

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

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

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

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

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

We 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. We 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 $N^\star_e$ such that:

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

B.10 Decomposition

We 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 market-level outcomes when we introduce any given policy. We show effects of introducing a policy in a market when none previously existed, and will hence consider how the average of a given variable z moves from its value without policy z_t, to its value with policy z_{t+1}. It is useful to partition the incumbency space as follows: shared space pre- and post-policy introduction $\mathcal{B} = IncumbencySpace_t \bigcup IncumbencySpace_{t+1}$, the incumbency space only active pre-policy $\mathcal{A} = IncumbencySpace_t \setminus IncumbencySpace_{t+1}$, and finally the incumbency space only active with the policy $\mathcal{C} = IncumbencySpace_{t+1} \setminus IncumbencySpace_t$.

We can first study how the average within space B changes with the policy. We write $s^{\mathcal{B},t}_j$ the aggregate (both crops) market share of farm of type j in period t within the space $\mathcal{B}$, and $s^t_{j,c}$ the share of farm j's production allocated to crop c in period t. We can decompose:

$$\begin{aligned}
\Delta z|_{\mathcal{B}} = &\sum_{j \in \mathcal{B}} \underbrace{s^{\mathcal{B},t}_j \left[s^t_{j,c_1} \Delta z_{j,c_1} + s^t_{j,c_2} \Delta z_{j,c_2} \right]}_{\textit{Within-Farm Within-Crop}} \\
&+ \sum_{j \in \mathcal{B}} \underbrace{s^{\mathcal{B},t}_j \left[z^t_{j,c_1} \Delta s_{j,c_1} + z^t_{j,c_2} \Delta s_{j,c_2} + \Delta z_{j,c_1} \Delta s_{j,c_1} + \Delta z_{j,c_2} \Delta s_{j,c_2} \right]}_{\textit{Within-Farm Across-Crops}} \\
&+ \sum_{j \in \mathcal{B}} \underbrace{\Delta s^{\mathcal{B}}_j z^t_j + \Delta s^{\mathcal{B}}_j \Delta z_j}_{\textit{Across Farms - Within B}}
\end{aligned}$$

We can further study the effect of changing the boundaries of the incumbency space. We can decompose this effect into two channels: shrinkage away from $\mathcal{A}$, and extension to $\mathcal{C}$. We write for any two sets X^1 and X^2: $s^{t,\mathcal{X}^1 \cup \mathcal{X}^2}$ the market share held by farm j in period

t considering the entire space $\mathcal{X}^1 \bigcup \mathcal{X}^2$, and as previously by $s^{t,\mathcal{X}^1}$ its market share in the same period only considering the firms located in $\mathcal{X}^1$. We can write:

$$z^t|_{\mathcal{B}} - z^t|_{\mathcal{A}\bigcup\mathcal{B}} = \underbrace{\sum_{j\in\mathcal{B}} \left[s_j^{t,\mathcal{B}} - s_j^{t,\mathcal{B}\bigcup\mathcal{A}} \right] z_j^t}_{\textit{Gain in Market Share in } \mathcal{B}} - \underbrace{\sum_{j\in\mathcal{A}} s_j^{t,\mathcal{B}\bigcup\mathcal{A}} z_j^t}_{\textit{Loss in Value from Shrinkage}}$$

$$z^{t+1}|_{\mathcal{C}\bigcup\mathcal{B}} - z^{t+1}|_{\mathcal{B}} = \underbrace{\sum_{j\in\mathcal{B}} \left[s_j^{t+1,\mathcal{B}\bigcup\mathcal{C}} - s_j^{t+1,\mathcal{B}} \right] z_j^{t+1}}_{\textit{Loss in Market Share in } \mathcal{B}} + \underbrace{\sum_{j\in\mathcal{C}} s_j^{t+1,\mathcal{B}\bigcup\mathcal{C}} z_j^{t+1}}_{\textit{Gain in Value from Extension}}$$

We obtain the following results with the decomposition, looking in turn at the chemical-land and the chemical-to-output ratios.

Table A20: Decomposition of Environmental Effects

	Chem.-Land	Chem.-Q	Chem.-Land	Chem.-Q	Chem.-Land	Chem.-Q
	Common Space Evolution (on $\mathcal{B}$)					
	Price Intervention		Land Subsidy		Chemicals Tax	
Within-Farm Within-Crops	-0.00	41.68	-263.57	-14.13	-262.12	-40.62
Within-Farm Across-Crops	100.88	25.20	4.59	1.37	3.93	0.82
Across-Farms	1.29	0.67	-0.01	-0.00	-0.01	-0.01
	Shrinkage (away from $\mathcal{A}$)					
	Price Intervention		Land Subsidy		Chemicals Tax	
Gain of Market Share in $\mathcal{B}$	0.91	0.91	2.62	2.61	0.86	0.86
Shrinkage away from $\mathcal{A}$	-1.05	-0.99	-1.48	-1.90	-0.75	-0.81
	Extension (to $\mathcal{C}$)					
	Price Intervention		Land Subsidy		Chemicals Tax	
Loss of Market Share in $\mathcal{B}$	-5.39	-5.34	-1.83	-1.82	0.00	0.00
Extension to $\mathcal{C}$	3.28	4.15	1.13	1.37	0.00	0.00

Notes. We decompose the changes in market pollution outcomes from moving from a land subsidy to a chemicals tax. Changes are expressed in percentage of the land subsidy levels.

These decompositions confirm previous intuitions from Figure 5. On price intervention first. We see that selection – the joint movement from $\mathcal{A}\bigcup\mathcal{B}$ to $\mathcal{B}\bigcup\mathcal{C}$ has a positive effect on pollution – measured by either of our two ratios. These however happen to be quite small at the aggregate level. These small changes from selection are largely out-weighted by changes within farm. The chemical-to-land ratio at the aggregate level is very negatively affected by the increased share of crop 2. The chemical-to-output ratio is similarly impacted by this change. It is also further impacted by changes within-farm within-crop. The production function has constant returns to scale when we account for capital, but if capital does not adjust fully with scale choices, increases in scale will lead to increase in the input-to-output

ratio for all flexible inputs – leading to this 41% increase from within-within changes in the chemical-to-output ratio.

The decompositions for the land subsidy and the chemical tax show similar mechanisms at play. Selection effects are limited, and most of the effects are concentrated in the within-within channel. Incentives for input choices are now modified, leading to significant differences in pollution. Because the chemicals tax specifically targets chemicals, its within-within effect is more pronounced on the chemicals-to-output ratio than under the chemicals-to-land ratio.

Limitations: We want to note a few limitations specific to the comparison between the chemical tax and the land subsidy. First of all, because our model takes land to be fully flexible – while in reality land use among farms in France is a mix of land rental and land ownership – the model might over-state the substitutability of chemicals towards land. This is less of an issue in the case of the chemical tax, which changes the relative price ratio of chemicals will all inputs, and not just land. This comparison might then over-state the impact of the land subsidy on the chemical intensity of production, more than it over-states that of the chemical tax. A second limitation is that the chemical tax is paired with a lump sum for which informational requirements are large if we want to compensate farms for their exp-ante exposure to it. Specifically, the government would need to know farms' maximization problem down to their type and prices. Without the lump sum, the chemical tax will have a larger effect on selection, likely increasing its impacts on costs of production. We note that the land subsidy then has the clear advantage of being easy to implement.

B.11 Sensitivity Analysis for the Production Function Estimation

We 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. We 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.

We 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.

We 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. 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 we 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 $\hat{\theta}$. We present and discuss below our estimates of ΛK for each estimation step.

We 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, we 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 we can also assume to be

exogenous. Because we 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 A21: Sensitivity for First Step

	Constant	$log(Chem.Ratio_{t-1})$	$Capital_t$	$Wage_t$	$LandPrice_t$	$Land_{t-1}$	$FarmExposure_t$
$log(1-\rho)$	17.5560	-0.2420	0.0000	-2.0303	-0.0000	0.0007	-0.0003
ρ_2	-399.7498	9.6731	-0.0000	44.0966	0.0000	-0.0276	0.0046
δ_s (group 1)	-26.7153	0.3820	-0.0000	3.0439	0.0000	-0.0009	0.0004
δ_s (group 2)	6.8136	-0.1288	0.0000	-0.7720	-0.0000	0.0002	-0.0001
δ_p (group 1)	-40.6379	0.8927	0.0000	5.0727	0.0000	-0.0017	-0.0002
δ_p (group 2)	138.6659	-3.1918	0.0000	-15.6575	-0.0000	0.0072	-0.0017

Table A22: 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)	0.0091	0.0000	-0.0025	-0.0203	0.0009	-0.0000	0.0281	0.0039	0.0000	-0.0004	-0.0010
α_s (group 1)	-0.0042	-0.0000	0.0010	0.0078	-0.0002	0.0000	-0.0097	-0.0015	-0.0000	0.0002	0.0004
α_l (group 2)	-0.0021	-0.0002	0.0012	0.0067	-0.0004	0.0000	-0.0118	-0.0015	-0.0000	0.0001	0.0004
α_s (group 2)	0.0008	0.0000	-0.0003	-0.0017	0.0001	-0.0000	0.0026	0.0004	0.0000	-0.0000	-0.0001