THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 1 Capitole

Présentée et soutenue par

Anca VOIA

Le 23 novembre 2021

Efficacité et ciblage optimal des paiements pour services ecosystémiques

Ecole doctorale : **TSE - Toulouse Sciences Economiques**

Spécialité : **Sciences Economiques - Toulouse**

Unité de recherche :
UMR 5314 -TSE-R

Thèse dirigée par
Céline NAUGES et Sylvain CHABé-FERRET

Jury
Mme Jennifer Alix-Garcia, Rapporteure
M. Paul Ferraro, Rapporteur
M. François Salanié, Examinateur
M. Jean-Christophe Bureau, Examinateur
M. Vincent Marcus, Examinateur
Mme Julie Subervie, Examinatrice
Mme Anouch Missirian, Examinatrice
Mme Céline Nauges, Directrice de thèse

Toulouse School of Economics

PhD Thesis

Cost-Effectiveness and Optimal Design of Payments for Ecosystem Services

Anca Voia

November 2021

Acknowledgements

I would like to express my sincere gratitude to my thesis advisors and mentors, Céline Nauges and Sylvain Chabé-Ferret, for their continuous guidance and unwavering support. I especially thank Sylvain Chabé-Ferret for all the time and resources he invested in this thesis, for believing in me and for all his understanding and encouragement. I also wish to thank François Salanié for his availability and precious feedback and all members of the Environmental Group for their kindness and support. I must also thank Nour Meddahi and the doctoral school for their understanding and unfailing help.

I would like to extend my sincere thanks to the *Observatoire du Développement Rural*, especially to Gilles Allaire, Cedric Gendre, Pierre Cantelaube and Eric Cahuzac, for their generous sharing of time, expertise and data on the French Grassland Conservation Program. I also wish to thank Lionel Védrine for explaining the program details and Valentin Bellasen for sharing his knowledge on the environmental impacts of grassland.

I gratefully acknowledge outstanding research assistance from Stefania Pozzi, Marine Roux, Julie Klein and Bernardo Mayorga Salas. I am also grateful to Stefan Pollinger, Christian Bontemps, Phillipe Bontems, Christian Hellwig and Ulrich Hege for their insightful suggestions on the last chapter.

I would like to express my deepest appreciation to Jennifer Alix-Garcia, Paul Ferraro, Jean-Christophe Bureau, Vincent Marcus, Julie Subervie, Anouch Missirian and François Salanié for accepting to be part of the jury of this thesis.

I am grateful to my colleagues and friends for making my PhD journey better through their moral support and all the great moments we shared together. Special thanks to Sarah Lemaire, Paloma Carrillo, Yuting Yang and Eva Tène for always being there for me.

Finally, I would like to express my gratitude to my family for their tremendous understanding and encouragement in the past few years. I especially thank my husband Johnatan Vincent for his unconditional support and belief in me and our son Noam for giving me the energy to complete this PhD thesis.

Abstract

This thesis includes three independent papers on topics related to Payments for Ecosystem Services, namely on their cost-effectiveness and their optimal design. The first two chapters are co-authored with Sylvain Chabé-Ferret.

In the first chapter, we are interested in knowing whether Grassland Conservation Programs are a cost-effective way to fight climate change. Grassland, especially when extensively managed and when replacing cropland, stores carbon in the ground. As a result, Grassland Conservation Programs, that compensate farmers for maintaining grassland cover, might be an effective way to combat climate change, if they succeed in triggering an increase in grassland cover at the expense of cropland for a reasonable amount of money. In this chapter, we use a natural experiment to estimate the cost-effectiveness of the French Grassland Conservation Program. We exploit a change in the eligibility requirements for the program that generated a sizeable increase in the proportion of participants in the areas most affected by the reform. We find that the expansion of the program leads to a small increase in grassland area, mainly at the expense of croplands, which implies that the program expansion increased carbon storage. We also find that the elasticity of grassland provision is low, and that, as a result, the program has large windfall gains. To compute the benefit-cost ratio of the program, we combine our results with similar estimates from the literature using meta-analysis tools and we introduce the resulting parameter in a model of carbon storage in grassland. We find that, for a carbon price of 24 Euro/tCO_2eq, the climate benefits of the program are equal to 7±3% of its costs. When taking into account the other benefits brought about by grassland, we find the benefits of the program to be equal to 44±15% of its costs. We estimate that the program would break even for a carbon price of 194±122 Euro/tCO_2eq.

In the second chapter, we investigate the cost-effectiveness of Forest Conservation Programs and their potential for climate-change mitigation. Deforestation is a major contributor to the emission of greenhouse gases. Forest Conservation Programs that pay landowners for maintaining forest cover might thus be an effective way to fight climate

change as long as the benefits from avoided emissions exceed the cost of triggering the conservation of additional forest cover. In this paper, we use meta-analysis tools to estimate the benefit-cost ratio of Forest Conservation Programs implemented in developing countries. We combine 18 separate estimates of the additional forest cover conserved thanks to these programs with estimates of emissions from deforestation. We find that Forest Conservation Programs reduce the annual deforestation rate by 0.23±0.14 percentage points on average and thus provide climate benefits. Our results suggest that the value of the climate benefits of Forest Conservation Programs crucially depends on the permanence of their effects after the program stops. For a Social Cost of Carbon of 31 USD/tCO_2eq (or 24 Euro/tCO_2eq in 2007 exchange rate), we estimate that benefits are equal to 45±32% of the program costs if the impact of the program on deforestation stops just after the program ends, 78±56% of the program costs if the impact decreases progressively over 10 years, and 263±194% of the program costs if the impact persists forever. We estimate that Forest Conservation Programs would become cost-effective with a Social Cost of Carbon of 100 USD/tCO_2eq (or 77 Euro/tCO_2eq), even with no permanence. We find ample evidence of publication bias in the estimates of the impact of Forest Conservation Programs on deforestation, with certain of the most recent estimates over-estimating the impact of these programs by a factor of 10, artificially inflating their benefit-cost ratio above one.

In the third chapter, I investigate the optimal design of Payments for Ecosystem Services, with an application to Grassland Conservation Programs. To strike the right balance between agriculture and the environment, policymakers increasingly use Payments for Ecosystem Services programs. They are incentives offered to landowners conditional on the provision of an environmental service. However, information asymmetries may limit their effectiveness: due to differences in opportunity costs, offering a linear-uniform payment to all farmers increases the risk of windfall gains. Nonlinear payments are a way to decrease windfall gains by differentiating payments by the quantity offered. Another approach is to differentiate payments by geographic characteristics, a proxy for provision costs. In this chapter, I use a principal-agent model to provide insights on the optimality of different Payments for Ecosystem Services contract designs. I use data on the French Grassland Conservation Program contracts, and I exploit an exogenous change in the payment structure to identify and estimate nonparametrically the farmers' cost function and the distribution of their types. This allows me to select parametric specifications and to evaluate welfare for different contract designs. I find that the loss of using linear-uniform contracts instead of nonlinear ones is around 2.6% and that spatially-targeted linear-uniform contracts improve the welfare gain with respect to the linear-uniform contracts by 1.9%.

Moreover, I find a low cost of asymmetric information, with the surplus of nonlinear contracts being 87% of that under complete information.

Résumé

Cette thèse comprend trois articles indépendants concernant les programmes de paiements pour services écosystémiques, notamment sur leur efficacité et leur ciblage optimal. Les deux premiers chapitres sont co-écrits avec Sylvain Chabé-Ferret.

Dans le premier chapitre, nous souhaitons évaluer si les programmes de conservation des prairies sont des moyens efficaces de lutte contre le changement climatique. En effet, les prairies gérées de manière extensive et d'autant plus lorsqu'elles remplacent des terres cultivées, stockent une quantité plus importante de carbone dans le sol. C'est pourquoi, ces programmes qui rémunèrent les agriculteurs pour le maintien de la couverture des prairies pourraient être un moyen efficace de lutte contre le changement climatique, à condition d'induire une augmentation de cette couverture pour un coût raisonnable. Dans ce chapitre, nous utilisons une expérience naturelle pour estimer la rentabilité du Programme Français de Conservation des Prairies, le plus important au monde. Nous exploitons un changement dans les conditions d'éligibilité au programme qui a généré une augmentation importante de la proportion de participants dans les communes les plus impactés par la réforme. Nous constatons que l'expansion du programme a conduit à une légère augmentation de la superficie des prairies, principalement au détriment des terres cultivées, impliquant une augmentation des capacités de stockage de carbone. Nous constatons également que l'élasticité de l'offre de prairies est faible, ce qui démontre une capacité du programme à générer d'importants effets d'aubaine. Ensuite, afin de calculer le rapport coûts-bénéfices du programme, nous avons utilisé des outils de méta-analyse afin de combiner nos résultats avec les estimations similaires tirées de la littérature et nous introduisons le paramètre final dans un modèle d'évaluation des quantités de carbone stockés dans les prairies. Les résultats montrent que les bénéfices climatiques du programme sont égaux à 7$\pm$3% de ses coûts pour un prix du carbone de 24 Euro/tCO_2eq. En tenant compte des autres bénéfices apportés par les prairies, nous constatons que les bénéfices du programme sont estimés à 44$\pm$15% de ses coûts. Ainsi, afin que le programme atteigne l'équilibre, nous estimons que le prix du carbone devrait être à minima égal à 194$\pm$122 Euro/tCO_2eq.

Dans le deuxième chapitre, nous étudions l'efficacité des programmes de conservation des forêts et leur potentiel d'atténuation du changement climatique. La déforestation est un contributeur majeur de l'émission de gaz à effet de serre. Ainsi, les programmes de conservation des forêts qui rémunèrent les agriculteurs pour le maintien de la couverture forestière sembleraient être un moyen efficace de lutte contre le changement climatique tant que les bénéfices liés aux émissions évitées sont supérieurs au coût de conservation d'une couverture forestière supplémentaire. Dans cet article, nous utilisons des outils de méta-analyse pour estimer le rapport coûts-bénéfices des programmes de conservation des forêts. Nous combinons 18 estimations distinctes de la couverture forestière supplémentaire réalisée grâce aux programmes de conservation des forêts avec les estimations d'émissions dues à la déforestation. Ainsi, nous estimons que les programmes de conservation des forêts réduisent le taux de déforestation annuel de 0,23±0,14 points de pourcentage en moyenne et offrent ainsi des bénéfices climatiques. Nous constatons aussi que les avantages climatiques des programmes de conservation des forêts dépendent très fortement de la permanence de leurs effets après l'arrêt du programme. C'est pourquoi, pour un coût social du carbone de 31 USD/tCO_2eq (ou 24 Euro/tCO_2eq en taux de change 2007), nous estimons que les bénéfices sont égaux à 45±32% des coûts du programme lorsque les impacts du programme s'arrêtent dès la fin du programme, 78±56% des coûts du programme si l'impact diminue progressivement sur 10 ans, et enfin 263±194% des coûts du programme si l'impact persiste indéfiniment. Ces résultats nous permettent d'estimer le seuil de rentabilité des programmes, sans permanence, à un coût social du carbone de 100 USD/tCO_2eq (ou 77 Euro/tCO_2eq). Par ailleurs, nous trouvons de nombreuses preuves de biais de publication dans les estimations passées de l'impact des programmes. Certaines des récentes estimations pouvant surestimer l'impact des programmes par 10, surestiment artificiellement leur rapport coûts-bénéfices, le rendant ainsi positif.

Dans le troisième chapitre, j'étudie l'optimalité des paiements pour services écosystémiques via l'application aux programmes de conservation des prairies traité dans le premier chapitre. Aujourd'hui, afin de trouver le juste équilibre entre l'agriculture et l'environnement, les décideurs publics utilisent de plus en plus les paiements pour services écosystémiques (PSE). Les PSE sont des incitations offertes aux propriétaires fonciers sous réserve de la fourniture d'un service environnemental. Cependant, les asymétries d'informations peuvent limiter leur efficacité. En effet, offrir un paiement linéaire et uniforme à tous les agriculteurs augmente le risque d'arbitrage en raison des différences de coûts d'opportunité. Les paiements non linéaires sont un moyen de réduire les gains exceptionnels en différenciant les paiements par la quantité de service écosystémique offerte. Une autre approche consiste à différencier les paiements par caractéristiques géographiques, un proxy réaliste des coûts d'opportunité. Dans ce chapitre, j'utilise un modèle principal-agent pour fournir des informations sur l'optimalité des différentes conceptions de contrats de paiements pour services écosystémiques. J'utilise des données sur les contrats du Programme Français de Conservation des Prairies, et j'exploite un changement exogène de la structure de paiement pour identifier et estimer de manière non paramétrique la fonction de coût des agriculteurs et la distribution de leurs types. Cela me permet de sélectionner des caractéristiques paramétriques et d'évaluer le surplus total pour différents types de contrats. Je trouve ainsi que la diminution du surplus lié à l'utilisation d'un contrat linéaire uniforme versus un contrat non linéaire est d'environ 2,6%. En outre, les contrats linéaires uniformes ciblés spatialement améliorent le surplus total de 1,9% par rapport aux contrats linéaires uniformes classiques. Par ailleurs, j'estime que le coût des asymétries d'information est faible étant donné que le niveau de surplus obtenu via l'utilisation de contrats non linéaires est égal à 87% de celui obtenu dans un environnement d'informations pures et parfaites.

Contents

List of Figures

List of Tables

Introduction

Fighting climate change is one of the most important challenges facing mankind. After the burning of fossil fuels, deforestation and agricultural activity are the second and third major sources of climate change. Yet, they can also be a major part of the solution through public policy interventions such as taxes, subsidies or regulations aimed at triggering additional carbon sequestration. In my thesis, I focus on Payments for Ecosystem Services (PES) programs that pay voluntary landowners to adopt practices aimed at increasing carbon sequestration. These programs are widely used by governments around the world as a tool to incentivize landowners to adopt practices more favourable to the environment. In my work, I investigate whether PES are a cost-effective way to decrease emissions of greenhouse gases and I study the optimal design of such programs.

Payments for Ecosystem Services programs are voluntary agreements between a buyer (e.g. Government or private users) and a seller (e.g. landowner) in which a payment is given conditionally on an environmental service being adequately provided. The payment should compensate the landowner for the average compliance costs and for the forgone farming revenue associated with the adoption of greener practices. In general, a PES program targets at least one of the four environmental services among carbon sequestration, watershed services, biodiversity, and landscape beauty. The interest in studying Payments for Ecosystem Services lies in their role of achieving environmental protection goals, such as climate change mitigation and adaptation (depending on the environmental service they target). Indeed, PES programs aimed at carbon sequestration from grasslands and forests can help fighting climate change, as they store important quantities of carbon in the soil and aboveground (for forests only). Similarly, PES aimed at improved water quality might have natural adaptation benefits if, for example, more environmentally-friendly land uses are undertaken that could potentially decrease inhabitants' vulnerability to climate-related water problems.

Payments for Ecosystem Services programs, and especially Forest Conservation Programs, saw a big increase in terms of number and budget allocated after the 2015 Paris

Climate Agreement, which for the first time recognized forests as a key part of the solution to climate change. However, PES programs started to be implemented worldwide long before. For example, in the United States, the Conservation Reserve Program was introduced in 1985 with a yearly funding of around 2 billion USD. A similar budget was allocated to the Environmental Quality Incentives Program, that was implemented in 1996. In the European Union (EU), PES schemes were introduced first as accompanying measures of the 1992 Common Agricultural Policy (CAP) reform, but since 2000 they become core instruments of EU agricultural policies. The budget allocated to these programs increased from 76 million Euro in 1993 to almost 3 billion Euro in 2010. In China, the Sloping Land Conversion Program was introduced in 1999 and had a yearly budget of 360 million USD during the first years. China has the objective to increase the spending for this program up to 48 billion USD. Finally, in developing countries, programs aimed at Reducing Emissions from Deforestation and forest Degradation (REDD+) were introduced by Parties to the United Nations Framework Convention on Climate Change in 2007. Around 450 such programs were in place in developing countries in 2018, with a total yearly budget of 1 billion USD. In addition to REDD+, a growing number of national PES programs were implemented, Costa Rica introducing the first national-level forest PES in 1997.

One of the main strengths of Payments for Ecosystem Services compared to other public policies aimed at reducing carbon emissions (such as taxes and regulations) is its voluntary approach. However, this turns out to be also one of the main threats in terms of program efficiency. Indeed, farmers with the lowest costs of meeting the eligibility requirements are the most likely to enter the program. As a result, the program might end up paying some farmers for doing nothing differently from what they would have done without any payment. This in turn undermines the cost-effectiveness of the program. Most of the empirical literature has focused on estimating the additionality of PES programs. Additionality measures how much greener farmers practices have become thanks to the program and is a key parameter for evaluating the cost-effectiveness of a PES scheme. The general conclusion which emerges from this literature is that Payments for Ecosystem Services have a low level of additionality. Yet, only few recent papers take the analysis further and estimate the benefit-cost ratio of PES programs given the level of additionality estimated. The main contribution of this thesis is to add evidence on the cost-effectiveness of Payments for Ecosystem Services programs in terms of avoided greenhouse gases emissions. The results from the first two chapters suggest that PES programs are not cost-effective at current carbon prices. This is in contrast with the other recent estimates available that find a positive benefit-cost ratio for Forest Conservation Programs. As showed in chapter 2, the difference might be due to publication bias that is present in this literature. A potential solution to increase the additionality of PES programs that is widely advocated in the literature is to better target the population of interest in order to capture the heterogeneity in opportunity costs. I find, in chapter 3, that the welfare gains attributed to targeted payments compared to linear-uniform ones are actually quite small. Therefore, the reason why Payments for Ecosystem Services programs are not cost-effective and potential solutions for their improvement require further research.

All three chapters of this thesis therefore study the cost-effectiveness of Payments for Ecosystem Services and their optimal design.

Chapter 1 estimates the benefit-cost ratio of the French Grassland Conservation Program using a natural experiment. Grassland, especially when extensively managed and when replacing cropland, stores carbon in the soil. To this day, it is still unknown whether Grassland Conservation Programs trigger sufficient changes in grassland cover so as to be cost-effective ways to fight climate change and to improve the environment. This chapter tries to fill this gap. First, we estimate the additionality of the French Grassland Conservation Program using a change in eligibility criteria as a natural experiment in a difference-in-differences (DID) design. Our work thus contributes to the literature on the estimation of the additionality of Payments for Ecosystem Services using a natural experiment and to the one on the evaluation of the additionality of Grassland Conservation Programs. Second, we clarify the causal effect parameters we can identify with our natural experiment, using a theoretical model of participation in an environmental subsidy program characterized by a notch. Our approach delineates conditions under which one can use DID to identify the effect of such a reform. Third, we combine our additionality results with estimates from the agronomic literature on the dynamics of carbon storage after grassland is converted to cropland in order to build the benefit-cost ratio. Finally, we combine our estimates with the findings of Chabé-Ferret and Subervie (2009) and Gallic and Marcus (2019) in a meta-analysis to obtain a more precise benefit-cost ratio of the French Grassland Conservation Program.

Chapter 2 investigates the cost-effectiveness of Forest Conservation Programs implemented in developing countries using meta-analysis techniques. These programs got a big boost from the 2015 Paris Climate Agreement, which for the first time recognised forests as a key part of the solution to climate change mitigation. However, there is still no consensus among economists when it comes to their effectiveness on conserving forest cover. By summarizing 18 existing rigorous studies that evaluate forest conservation programs in a meta-analysis, this chapter tries to clarify the debate. We complement the existing literature reviews on this topic by estimating the benefit-cost ratio of Forest Conservation Programs implemented in developing countries. We compute the discounted benefits of the programs in terms of avoided emissions and compare them to the program costs. Moreover, we test and account for publication bias, that might occur when some studies are missing from the published record. Thus, this paper also contributes to the literature on the consequences of publication bias for estimates of crucial parameters in economics.

Chapter 3 provides insights on the optimality of different Payments for Ecosystem Services contract designs, using a principal-agent model. PES contracts are usually subject to asymmetric information between landowners and the service buyers that can limit their cost-effectiveness. The purpose of this chapter is to check the importance of hidden information in these types of programs and to quantify the loss in welfare associated with the use of linear-uniform contracts instead of nonlinear, optimal ones. To this end, using a principal-agent model, I exploit an exogenous change in the payment structure of the French Grassland Conservation Program to identify and estimate nonparametrically the farmers' cost function and the distribution of their types. This allows me to select parametric specifications and to evaluate welfare for different contract designs. The contribution of this chapter is to combine the almost exclusive theoretical literature studying the asymmetric information issue in the context of payment for ecosystem services and the emerging literature on nonparametric methods to identify and estimate agents' cost function and the distribution of their types.

Chapter 1

Are Grassland Conservation Programs a Cost-Effective Way to Fight Climate Change? Evidence from France

Sylvain Chabé-Ferret and Anca Voia

Abstract

Grassland, especially when extensively managed and when replacing cropland, stores carbon in the ground. As a result, Grassland Conservation Programs, that compensate farmers for maintaining grassland cover, might be an effective way to combat climate change, if they succeed in triggering an increase in grassland cover at the expense of cropland for a reasonable amount of money. In this paper, we use a natural experiment to estimate the cost-effectiveness of the French Grassland Conservation Program. We exploit a change in the eligibility requirements for the program that generated a sizeable increase in the proportion of participants in the areas most affected by the reform. We find that the expansion of the program leads to a small increase in grassland area, mainly at the expense of croplands, which implies that the program expansion increased carbon storage. We also find that the elasticity of grassland provision is low, and that, as a result, the program has large windfall gains. To compute the benefit-cost ratio of the program, we combine our results with similar estimates from the literature using meta-analysis tools and we introduce the resulting parameter in a model of carbon storage in grassland. We find that, for a carbon price of 24 Euro/tCO_2eq, the climate benefits of the program are equal to 7±3% of its costs. When taking into account the other benefits brought about by grassland, we find the benefits of the program to be equal to 44±15% of its costs. We estimate that the program would break even for a carbon price of 194±122 Euro/tCO_2eq.

1.1 Introduction

Fighting climate change is one of the most important challenges facing mankind. Comparing the cost-effectiveness of the various options available to decrease the emissions of greenhouse gases is critical for succeeding in this endeavor. In this paper, we estimate the benefit-cost ratio of one such option: Grassland Conservation Programs. Grassland, especially when extensively managed and when replacing cropland, stores carbon in the soil (Soussana et al., 2004). It also reduces water pollution (Agouridis et al., 2005) and increases biodiversity (Bretagnolle et al., 2012). As a result, Grassland Conservation Programs, that pay voluntary farmers for maintaining grassland cover, might be an effective way to protect the environment and to combat climate change. The key for these programs to be cost-effective is to trigger an increase in grassland cover at the expense of cropland for a reasonable amount of money. To this day, it is still unknown whether Grassland Conservation Programs trigger sufficient changes in grassland cover so as to be cost-effective ways to fight climate change and to improve the environment.

A key input to compute the benefit-cost ratio of a Grassland Conservation Program is its additionality (Chabé-Ferret and Subervie, 2013): how many additional hectares of grassland have been implanted or maintained thanks to the program. Additionality in turn depends on the elasticity of the supply of grassland. The more elastic (i.e. responsive to prices) the supply of grassland, the more cost-effective the program. In the limit, if the supply of grassland is fully inelastic, the program ends up paying farmers for doing nothing differently from what they would have done without any payment, and the effectiveness of the program is null.

Estimating additionality is no easy task because usual comparisons are very likely to be biased by confounding factors (Chabé-Ferret and Subervie, 2012). Comparing contracting farmers to non contracting farmers might overestimate the impact of the program. Indeed, contracting farmers take up the program not by chance but because they have lower costs of supplying grassland, and thus would have had a larger area of grassland than nonparticipants even in the absence of the program. The characteristics that make contracting farmers supply more grassland even in the absence of the program thus confound the effect of the program. Most of these characteristics are difficult to measure in usual datasets: the opportunity cost of grassland is mostly related to land quality, a difficult parameter to observe and summarize. The change of grassland area of contracting farmers after the implementation of the program might also be confounded by simultaneous changes in prices or in other policies.

In this paper, we estimate the benefit-cost ratio of the French Grassland Conservation Program, the largest Grassland Conservation Program in the world, using a natural experiment to solve for confounding bias.[1] The natural experiment that we exploit is a change in the eligibility requirements to the French Grassland Conservation Program that happened in 2000. Before 2000, contracting farmers had to have a ratio of grassland to agricultural usable area higher than 75% in order to be eligible to receive the payments. In 2000, new contracts were introduced that did not include this eligibility criteria. In the areas most affected by the reform, the proportion of beneficiaries of the Grassland Conservation Program doubled, increasing from 10% to 20%, while it remained stable around 15% in areas unaffected by the reform. Our identification strategy uses the change in grassland cover in areas where the proportion of beneficiaries remained constant as a counterfactual for the change in grassland area that would have occurred in treated areas in the absence of the reform. In a theoretical model, we show that this comparison identifies the effect of the program on farmers that entered after the 2000 reform under plausible assumptions. The main assumption that we are making is that treated and control areas do not differ in their changes in grassland area absent the reform (an assumption generally called the Parallel Trends Assumption). We find support for the Parallel Trends Assumption in our pre-reform data. To account for possible effects of the program on non contracting farmers, we conduct our analysis at the commune level.[2] Contracting farmers might indeed decide to buy, rent or exchange grassland with non-contracting farmers, a phenomenon called "leakage effect". In the presence of leakage effects, the program might not add a single hectare of grassland in a commune but still appear to increase the area in grassland at the level of the individual participating farm. Working with commune-level data preserves our analysis from this issue.

Our results show that grassland cover increases in the communes most affected by the reform, and that this increase comes mostly at the expense of cropland. As a consequence, the 2000 reform helped store carbon in the ground, which suggests that the program brings positive environmental benefits. Unfortunately, the changes in grassland cover that we find are small compared to their monetary cost. The loosening of the eligibility criteria in 2000 lead to a substantial inflow of money in affected communes (around 5,000$\pm$513 Euro over five years, or an increase of 42.46$\pm$6.21%), but to a comparatively small increase

[1]Natural experiments leverage quasi-experimental variation in exposure to a policy in order to neutralize the effect of confounding factors (Chabé-Ferret and Subervie 2012, Dominici et al., 2014).

[2]Communes are the smallest administrative unit in France. There are approximately 36,000 communes in France. The average size of a French commune is around 7 sq.mile, which is a little less than half of the average size of a US Census Block Group. Leakage effects, that act through the functioning of the land market, are likely to be the most important at the commune level.

in grassland area (3.73±7.31 hectares per commune, or an increase of 0.76±1.49%). We thus estimate that the elasticity of the supply of grassland is low (around 0.02±0.04), meaning that an increase in prices by 10% would only increase the supply of grassland by 0.2%.

To compute the benefit-cost ratio of the program, we combine our additionality estimates to estimates of how carbon storage changes when grassland is converted into cropland. With a Social Cost of Carbon of 24 Euro/tCO_2eq, we estimate that the value of CO_2 emissions avoided thanks to the program is equal to 12±24% of its costs. When taking into account the other benefits brought about by grassland, we find the benefits of the program to be equal to 72±141% of its costs. We estimate that the program's benefits would equal its costs for a Social Cost of Carbon of 80±389 Euro/tCO_2eq. Our estimates are not precise enough to decide whether the benefits of the French Grassland Conservation Program are superior to their costs. To increase the precision of our estimates, we combine our results with similar estimates of the impact of the French Grassland Conservation Program using meta-analysis tools. Chabé-Ferret and Subervie (2009) use DID-matching to estimate the additionality of the Grassland Conservation Program in 2005. Gallic and Marcus (2019) use a change in the eligibility rules of the French Grassland Conservation Program in 2015 in order to estimate its additionality. Both papers find results very similar to ours (even if somewhat smaller on the additionality side), despite using individual level data and different identification strategies. Combining their results with ours using a meta-regression, we find that the climate benefits of the program are equal to 7±3% of its costs, its total environmental benefits to 44±15% of its costs and that it would break even for a carbon price of 194±122 Euro/tCO_2eq.

Overall, our results suggest that the increase in the number of beneficiaries of the French Grassland Conservation Program that resulted from the relaxation of the eligibility requirements in 2000 did not provide environmental benefits large enough to cover the costs of the reform. Although our results are by nature local and thus valid only for the group of farmers affected by the 2000 reform, they suggest that the French Grassland Conservation Program, and by extension Grassland Conservation Programs in general, are probably not cost-effective ways to fight climate change. There are indeed good reasons to believe that the areas most affected by the 2000 reform are the ones where grassland cover was more at risk of conversion into cropland. The share of grassland was already low and decreasing and cropland was expanding. It is thus likely that additionality is lower in the zones that benefited from the program prior to the reform. Results in Gallic and Marcus (2019) that cover different areas than the ones we do here find very similar results to ours. Similar programs elsewhere in Europe and the US will probably be characterized by much the same levels of additionality and thus reach similar benefit-cost ratios. Similar programs might have better benefit-cost ratios if they have larger additionality or if grassland stores more carbon than in France. But it seems unlikely that Grassland Conservation Programs can achieve similar climate benefits and benefit-cost ratios to the ones of Forest Conservation Programs. Indeed, forests store the same quantities of carbon in the ground as grassland does, but also, unlike grassland, store carbon above ground. Unless the elasticity of forest provision is much smaller than that of grassland, Forest Conservation Programs will have better benefit-cost ratios than grassland. Recent estimates suggest that the climate benefits of Forest Conservation Programs exceed their costs by a factor of two (Jayachadran et al., 2017; Simonet et al. 2018). A recent meta-analysis finds more modest climate benefits of Forest Conservation Programs, around 53$\pm$38% of their costs, and estimate that they would break even for a carbon price of 100 USD/tCO_2eq (Chabé-Ferret and Voia, 2021). Even with these more modest estimates, Forest Conservation Programs thus appear to be more promising that Grassland Conservation Programs to fight climate change. Nevertheless, if and when the Social Cost of Carbon reaches 200 Euro/tCO_2eq, Grassland Conservation Programs will become cost-effective, according to our estimates.

Our paper contributes to several literatures. First, we contribute to the literature estimating the additionality of Payments for Ecosystem Services by using a natural experiment. Most of the previous evaluations of the additionality of Payments for Ecosystem Services use observational methods, namely a combination of matching with difference-in-differences (DID) (see e.g. Chabé-Ferret and Subervie (2013)). Of the few studies of the impacts of Payments for Ecosystem Services that rely on natural experiments, most evaluate rather small scale programs. Kuhfuss and Subervie (2018) look at the additionality of the French Payments for Ecosystem Services aimed at pesticide reduction in the Languedoc-Roussillon region using the exogenous variation in the timing of the implementation of the program as a natural experiment. Simonet et al. (2018) use the introduction of a Brazilian forest conservation program to estimate its additional effects in the Para state. Apart from our paper, only Alix-Garcia et al. (2012, 2015) and Gallic and Marcus (2019) use a natural experiment to evaluate a nationwide Payment for Ecosystem Services Program (the Mexican Forest Conservation Program and the French Grassland Conservation Program respectively).[3] Second, we contribute to the evaluation of the additionality of Grassland Conservation Programs. Very few evaluations focus on Grassland Conservation Programs, although they are a major component of the EU Payments for Ecosystem Services. Arata

[3]It is interesting to note that Gallic and Marcus came to their idea for their paper after seeing a presentation of an earlier version of our own work. We see this cross-pollination as a testament to how research can influence work done by policy-makers.

and Sckokai (2016) identify a statistically significant increase in the share of grassland for participant farmers in all EU Payments for Ecosystem Services in five member states. Pufahl and Weiss (2009) apply a DID-matching approach to a non-representative subsample of German farms to show that the whole EU program of Payments for Ecosystem Services is likely to increase both the grassland area and the area under cultivation. Only Chabé-Ferret and Subervie (2009) and Gallic and Marcus (2019) perform a similar estimate to the one we are trying to achieve. Third, we clarify the causal effect parameters we can identify with our natural experiment, using a theoretical model of participation in a environmental subsidy program characterized by a notch (Kleven and Waseem, 2013) that is removed after some date. Our approach delineates conditions under which one can use DID to identify the effect of such a reform. Our model extends the one in Chabé-Ferret and Subervie (2013) to account for the existence of bunching at the eligibility threshold and encompasses the use of data on aggregated units such as communes as a way to eschew the problems posed by leakage effects. Pollinger (2021) develops an alternative approach to compute the benefit-cost ratio of an environmental program characterized by a kinked incentive, based on a density discontinuity estimator. Fourth, we combine our estimates with similar estimates from Chabé-Ferret and Subervie (2009) and Gallic and Marcus (2019) in a meta-analysis to obtain more precise estimates of the benefit-cost ratio of the French Grassland Conservation Program. Meta-analysis are increasingly used to synthetize estimates from a broad range of literature. Chabé-Ferret and Voia (2021) for example conduct an extensive meta-analysis of the additionality estimates of Forest Conservation Programs (as do Samii et al. (2014) and Snilsveit et al. (2019)). Fifth, we combine our additionality estimates with estimates from the agronomic literature on the dynamics of carbon storage after grassland is converted to cropland. We derive closed form solutions for the climate and environmental benefits of a program affecting land use changes that are of separate usefulness. We include uncertainty estimates on the parameters of the dynamics of carbon storage into our benefit-cost ratios estimates using the Delta method. Sixth, we contribute to the literature estimating the most cost-effective ways to fight climate change. Jayachadran et al. (2017) estimate the benefit-cost ratio of a Forest Conservation Program. Chabé-Ferret and Voia (2021) derive the benefit-cost ratio of Forest Conservation Programs using a meta-analysis. Christensen et al. (2021) decompose the discrepancy between the impacts of weatherization programs predicted by theoretical models and their realized benefits. A complete comparison of the benefit-cost ratios of all the policy options for fighting climate change is still out of reach, but we hope that this paper, along with others cited here, gets us closer to that goal. Gillingham and Stock (2018) provide estimates of the cost of actions that can be taken to fight climate change but do not include Grassland Conservation Programs nor an explicit meta-analysis.

The remainder of the paper is structured as follows: Section 1.2 describes the French Grassland Conservation Program; Section 1.3 exposes our empirical strategy; Section 1.4 introduces the data used in this paper; Section 1.5 presents the results and the robustness checks; Section 1.6 presents the cost-benefit analysis; Section 1.7 concludes.

1.2 The French Grassland Conservation Program

The French Grassland Conservation Program is the largest Grassland Conservation Program in the world. Over the period 2003-2007, a yearly budget of around 350 million Euro was allocated to subsidize 4.6 million hectares of grassland, covering 60% of the total grassland area in France (CNASEA, 2008). The program was created in 1993 as part of a broader set of Payments for Environmental Services introduced in the European Union as accompanying measures of the 1992 Common Agricultural Policy (CAP) reform. Since 2000, Payments for Ecosystem Services have become a core instrument of EU agricultural policies as part of the second pillar of the CAP. Subsidies for grassland conservation were included in the agri-environmental programs of several European countries, such as the German Cultural Landscape Program (KULAP), the Austrian Agri-environmental Program (OPUL), the United Kingdom's Environmental Stewardship Scheme or the Irish Rural Environment Protection Scheme (REPS) (Institut de l'Elevage, 2007). The yearly budget allocated to these programs varies from 100 million Euro for the German KULAP to 283 million Euro for the Austrian OPUL. Similarly, in the United States grassland conservation measures were in place since 2002 through the Grassland Reserve Program[4], with a funding of 38 million dollars yearly (USDA-NRCS, 2010).

The French Grassland Conservation Program is designed as a Payments for Ecosystem Services scheme. Payments for Ecosystem Services are voluntary agreements between a seller (a landowner) and a buyer (the Government or private users) in which a payment is given conditional on an environmental service being adequately provided (Alston et al., 2013). The payment is computed so as to compensate the landowner for the average compliance costs and for the forgone farming revenue associated with the adoption of greener practices or so as to reflect the value of the environmental service provided. In general, a Payments for Ecosystem Services program targets at least one of the four environmental services among carbon sequestration, watershed services, biodiversity and scenic beauty.

[4]Since 2014 the program is called the Conservation Reserve Program-Grasslands

The French Grassland Conservation Program was created in 1993 with the goal of stopping the decrease in grassland cover (from 43% of the agricultural area in 1970 to 36% in 1988 and only 27% in 2010). It was first called "*Prime au Maintien des Systemes d'Elevage Extensifs*" (PMSEE). PMSEE was a five-year contract during which farmers committed to keep the grassland on the same plots. In exchange, they were paid 35 to 46 Euro per hectare of grassland if they met two criteria: *(i)* a specialization rate (share of permanent and temporary grassland in the total usable agricultural area) higher than 75% and *(ii)* a loading ratio (density of livestock units (LU) per hectare of forage area) below 1.4. In 1998, PMSEE was renewed for another five years and an eligibility requirement related to the use of fertilisers was introduced: farmers were not allowed to exceed 70 kilograms of nitrogen per hectare of grassland. PMSEE was replaced in 2003 by a new extensive grazing scheme called "*Prime Herbagère Agro-Environnementale*" (PHAE). The eligibility criteria for PHAE were similar to those for PMSEE with three main exceptions. First, the thresholds for eligibility in terms of share of grassland and density of livestock units varied at department[5] level. Some departments kept the same thresholds as for PMSEE, while others chose a threshold for the specialization rate smaller than 75%, but never smaller than 50%. Also, some departments set the loading ratio higher than 1.4 LU/ha, but never larger than 1.8. Second, additional requirements were introduced, especially in order to limit the use of phytosanitary products and fertilizers on the plots. Finally, the payments were increased to 76 Euro per hectare of conserved grassland.

PMSEE and PHAE were two national programs that specifically target grassland conservation. However, starting in 2000, France launched an ambitious new Payments for Ecosystem Services program as part of the National Plan for Rural Development (NPRD). It was first called "*Contrat Territorial d'Exploitation*" (CTE) and was replaced in 2003 by "*Contrat d'Agriculture Durable*" (CAD). Among all the new Payments for Ecosystem Services that this program instituted, two broad categories were actually subsidies to grassland conservation: the measures 19 and 20. The measure 19 subsidized the maintenance of grassland opening where it was colonized by scrubs and trees, while the measure 20 subsidized extensive grassland management through mowing and/or pasture. The eligibility requirements for measures 19 and 20 were mainly that fertilization was limited on the field (in general, below 60 kilograms of nitrogen per hectare of grassland). The main difference is that measures 19 and 20 did not have any requirements on the specialization rate. As a consequence, these measures were taken also by farmers who were in general not eligible for PMSEE or PHAE due to a small share of grassland. Thus, measures 19 and 20 generated

[5]There are 95 departments in France.

a new influx of farmers into the French Grassland Conservation Program. The timelime of the French Grassland Conservation Program is presented in Figure 1.1, while a detailed description of the eligibility requirements is given in Figure 1.2.

1.3 Empirical Strategy

In this section, we delineate our empirical strategy. In order to do so, we first present a simple model of how farmers react to the incentives triggered by the grassland program before and after the 2000 reform. We then detail the sources of identification of our parameter estimates. Finally, we present our econometric strategy.

1.3.1 Farmers' reaction to the Grassland Conservation Program

We posit that there is a continuum of farmers of unit size characterized by their technical effectiveness at generating income from grassland θ. θ is distributed according to the cumulative distribution function F_θ over the interval $\left[\underline{\theta}, \overline{\theta}\right]$. Returns from grassland are given by the function $R(q,\theta)$ where $q \in [0,1]$ is the share of the total area of the farm allocated to grassland. R is a continuous, twice differentiable function of both arguments. It is concave, with $R_{qq} < 0$, $R_\theta > 0$ and $R_{q,\theta} > 0$. Before 2000, farmers' response to the program can be described as follows:[6]

$$\max_q R(q,\theta) + tq\mathbb{1}[q \geq \bar{q}]. \tag{1.1}$$

In this optimization problem, when farmers cross the $\bar{q}$ threshold, they receive a compensation t for each of their additional units of grassland beyond $\bar{q}$ but also for all the $\bar{q}$ inframarginal units. This large discontinuity in the incentives faced by the farmers around $\bar{q}$ is called a notch (Kleven and Waseem, 2013). The optimal response by farmers include some bunching at $\bar{q}$. In order to understand why, let us solve the farmers' problem as presented in equation (1.1).

Let us first define $\pi(t,\theta) = \max_q R(q,\theta) + tq$ as the farmers' problem without the participation constraint. π is increasing in both t and θ.[7] The interior solution for the optimal supply of grassland without the notch constraint is $q(t,\theta)$ and is defined as the solution to the first order equation $R_q(q,\theta) + t = 0$. By the implicit function theorem,

[6] $\mathbb{1}[A]$ is equal to one if A is true and to zero otherwise.

[7] By the envelope theorem, $\pi_t = q \geq 0$. We also have that $\pi_\theta = R_\theta > 0$.

$q(t,\theta)$ exists, is unique and is increasing in θ and in t.[8] There are also two corner solutions: $q = 0$ and $q = 1$. The condition $q(t,\theta) = 0$ defines $\underline{\theta}_0(t)$ such that all farmers with θ below $\underline{\theta}_0(t)$ have $q = 0$. The condition $q(t,\theta) = 1$ defines $\overline{\theta}_1(t)$ such that all farmers with θ above $\overline{\theta}_1(t)$ have $q = 1$. Let $q^*(t,\theta)$ summarize the supply function in the farmers' problem without participation constraint. It is equal to $q(t,\theta)$ when θ is between $\underline{\theta}_0(t)$ and $\overline{\theta}_1(t)$, to zero below $\underline{\theta}_0(t)$ and to one above $\overline{\theta}_1(t)$.

Let us now come back to the farmers' problem with the participation constraint as presented in equation (1.1). The constraint defines two new thresholds, $\underline{\theta}^*(t,\bar{q})$ and $\overline{\theta}^*(t,\bar{q})$ such that farmers that have $\theta \leq \underline{\theta}^*(t,\bar{q})$ will choose to supply $q^*(0,\theta)$ units of grassland, farmers with θ between $\underline{\theta}^*(t,\bar{q})$ and $\overline{\theta}^*(t,\bar{q})$ will choose to supply exactly $\bar{q}$ units of grassland and farmers with θ above $\overline{\theta}^*(t,\bar{q})$ will chose to supply $q^*(t,\theta)$ units of grassland. The reason why this is so is as follows. Because $q(t,\theta)$ is increasing in θ, the condition $q(t,\theta) = \bar{q}$ defines a threshold $\overline{\theta}^*(t,\bar{q})$ such that all farmers with $\theta \geq \overline{\theta}^*(t,\bar{q})$ choose to participate in the program, since they comply with the participation constraint even when it is not required. This is because π is increasing in t and thus farmers always prefer to benefit from the subsidy if it is not constraining for them. Farmers just below $\overline{\theta}^*(t,\bar{q})$ face two possibilities. They can increase their supply of grassland in order to reach $\bar{q}$ and be eligible for the payment. Their profit would then be $\pi_c(\bar{q},t,\theta) = R(\bar{q},\theta) + t\bar{q}$. Or they can choose to not benefit from the subsidy and supply $q(0,\theta)$ units of grassland, for a profit of $\pi(0,\theta)$. For a farmer such that $q(0,\theta) + dq = \bar{q}$, the loss incurred by bunching at $\bar{q}$ is equal to $(R_q(q,\theta) + t)dq$ and the gain is equal to the whole difference from $\pi(0,\theta)$ to $\pi_c(\bar{q},t,\theta)$. The first term is much smaller than the second, so that a farmer very close to $\overline{\theta}^*(t,\bar{q})$ chooses to bunch at $\bar{q}$. But as θ decreases, the cost of providing unprofitable amounts of grassland increases while the benefit from doing so decreases. At $\underline{\theta}^*(t,\bar{q})$, farmers are indifferent between participating and not participating. As a consequence, all farmers with θ below $\underline{\theta}^*(t,\bar{q})$ do not participate in the program.

Figure 1.3 summarizes the results of the theoretical model. The rightmost curve is the supply curve when farmers face a subsidy $t = 0$. The second and third curves are defined for prices of grassland of $t = t_0$ and $t = t_1 > t_0$ respectively. These curves represent $q^*(t,\theta)$, that is the level of supply in the absence of the participation constraint. When the subsidy increases, farmers supply more grassland at each level of θ. Let us now examine what happens when the participation constraint is active. Let us focus on the case where $t = t_0$, that is on the two rightmost curves. In the presence of the participation

[8] Since $R_{qq} < 0$, the sign of the derivatives of q with respect to its arguments are those of the first order condition with respect to each argument. The result follows since $\frac{\partial R_q(q,\theta)+t}{\partial t} = 1$ and $\frac{\partial R_q(q,\theta)+t}{\partial \theta} = R_{q,\theta}$.

constraint, farmers below $\underline{\theta}^*(t_0, \bar{q})$ have no incentive to participate in the program, since the benefits of reaching the participation constraint are smaller than the costs of complying with the constraint. As a consequence, these farmers do not participate in the program and thus supply $q^*(0, \theta)$. This is materialized by the fact that the curve with $t = 0$ is drawn in a continuous line for these farmers. Farmers between $\underline{\theta}^*(t_0, \bar{q})$ and $\overline{\theta}^*(t_0, \bar{q})$ prefer to participate in the program, but choose to bunch at $\bar{q}$, because their optimal supply with a subsidy of t_0 would be lower than $\bar{q}$. This is manifested by the fact that the second curve is below $\bar{q}$ for these farmers and is drawn in a interrupted line, to show that their supply absent the participation constraint is unobserved (as is their supply absent the subsidy). Finally, farmers above $\overline{\theta}^*(t_0, \bar{q})$ participate in the program and supply the same amount of grassland they would have supplied absent the participation constraint ($q^*(t_0, \theta)$). This is shown on the plot by the fact that the curve $q^*(t_0, \theta)$ is drawn in a continuous line above $\overline{\theta}^*(t_0, \bar{q})$. Among those farmers, those that are below $\overline{\theta}_1(t_0)$ supply the quantity $q(t_0, \theta)$ predicted by the solution to the first order condition to the unconstrained problem. The farmers that are above $\overline{\theta}_1(t_0)$ bunch at $q = 1$ and affect all of their area to grassland.

1.3.2 The 2000 reform and sources of identification

The 2000 reform of the Grassland Conservation Program can be modelled as taking off the $\mathbb{1}[q \geq \bar{q}]$ constraint from equation (1.1) and increasing the subsidy from t_0 to t_1. The resulting supply curve is the leftmost curve on Figure 1.3. Now, almost everyone participates in the program, apart from the farmers whose grassland supply is zero (the ones with θ below $\underline{\theta}_0(t_1)$). No farmers bunch at $\bar{q}$ anymore since the participation constraint has been lifted. Finally, all farmers have increased their supply of grassland compared with the pre-reform state.[9]

Farmers can be separated in several groups on the basis of how they react to the 2000 reform. These groups are the basis of our identification strategy. We use the random variable T (for *type*) to denote the different groups formally. Farmers with $\theta \leq \underline{\theta}_0(t_1)$ supply zero proportion of grassland at all prices ($t = 0$, $t = t_0$ and $t = t_1$). We denote them with $T = b_{000}$. Farmers with $\underline{\theta}_0(t_1) < \theta \leq \underline{\theta}_0(t_0)$ supply zero proportion of grassland at prices $t = 0$ and $t = t_0$. We denote them with $T = b_{00}$. Farmers with $\underline{\theta}_0(t_0) < \theta \leq \underline{\theta}_0(0)$ supply zero proportion of grassland only when $t = 0$. We denote them with $T = b_0$. Farmers with $\underline{\theta}_0(0) < \theta \leq \underline{\theta}^*(t_0, \bar{q})$ do not bunch neither at zero nor at $\bar{q}$. They move from

[9] Note that this is not a general result, since bunching farmers might supply less than when forced to bunch at $\bar{q}$. The actual price increase from t_0 to t_1 was in practice probably large enough to avoid this type of countervailing effects.

not receiving the program in 2000 to receiving the program in 2005. We call them *compliers* and denote them with $T = c$. Farmers with $\underline{\theta}^*(t_0, \bar{q}) < \theta \leq \overline{\theta}^*(t_0, \bar{q})$ bunch at $\bar{q}$ when $t = t_0$. We call them *bunchers* and denote them with $T = b$. Farmers with $\overline{\theta}^*(t_0, \bar{q}) < \theta \leq \overline{\theta}_1(t_1)$ do not bunch neither at $\bar{q}$ nor at one. We call them *always takers* and denote them with $T = at$. Farmers with $\overline{\theta}_1(t_1) < \theta \leq \overline{\theta}_1(t_0)$ supply a proportion of grassland of one when $t = t_1$. We denote them with $T = b_1$. Farmers with $\overline{\theta}_1(t_0) < \theta \leq \overline{\theta}_1(0)$ supply a proportion of grassland of one when $t = t_1$ and $t = t_0$. We denote them with $T = b_{11}$. Farmers with $\theta > \overline{\theta}_1(0)$ supply a proportion of grassland of one at all prices. We denote them with $T = b_{111}$. We also add a last category of farmers: the ones who supply a positive proportion of grassland at all prices and do not participate in the program. We call them *never takers* and denote them with $T = nt$. The existence of never takers is not compatible with our model but is a feature of the dataset. It is easy to rationalize their existence, either by adding a heterogeneous fixed cost of participating in the program (measuring the hassle it takes to apply for the new contracts under the CTE/CAD program) or by introducing a second eligibility requirement, such as a sufficiently small loading ratio for example.

In order to understand the idea behind our identification strategy, we are going to focus on the first two groups: the *compliers* and the *always takers*. The first group of farmers moves from $t = t_0$ to $t = t_1$ between 2000 and 2005 and is our control group, while the second group moves from $t = 0$ to $t = t_1$ and is our treated group. They both are characterized by an absence of bunching, which simplifies identification. Our identification strategy consists in comparing what happens to the *compliers* who enter the program after the reform to what happens to the *always takers* who stay with the program all along. In order to state our identification strategy rigorously, we need some additional notation. First, let $q_{2000}(t, \theta)$ and $q_{2005}(t, \theta)$ denote the supply of grassland in 2000 and 2005 respectively, for farmers with a grassland productivity level of θ and facing the level of subsidy t. These two supply functions differ from $q^*(t, \theta)$ by the fact that they contain the effect of exogenous shocks to grassland supply. These shocks are of two types: annual shocks common to all farmers, such as the ones affecting the price of crops, meat, milk, farm inputs, etc; and farmer-specific idiosyncratic shocks that we assume are i.i.d. across farmers and across time and independent from θ.

Our identification strategy is based on a Difference-In-Difference estimator comparing the grassland supply of the *compliers* to the grassland supply of the *always takers*:

$$DID_q^{at} = \mathbb{E}[q_{2005}(t_1, \theta) - q_{2000}(0, \theta) | T = c] - \mathbb{E}[q_{2005}(t_1, \theta) - q_{2000}(t_0, \theta) | T = at]. \quad (1.2)$$

Identification of the causal effect of the grassland subsidy follows from the following assumption:

Assumption 1 (Parallel trends, always takers) *We assume that grassland supply would have followed parallel trends among compliers and always takers if they had been exposed to the same price change:*

$$\mathbb{E}[q_{2005}(t_1,\theta) - q_{2000}(t_0,\theta)|T = c] = \mathbb{E}[q_{2005}(t_1,\theta) - q_{2000}(t_0,\theta)|T = at].$$

Assumption 1 actually encompasses two separate assumptions. For one, it requires that influences other than the grassland subsidy have the same average impact among *compliers* and *always takers* between 2000 and 2005. That means that crop prices and input prices would have influenced the share of grassland in both groups in the same way. Moreover, Assumption 1 also requires that the impact of the change in the grassland subsidy from t_0 to t_1 would have been the same in the two groups. Under Assumption 1, our DID^{at} estimator identifies $LATE_{q_{2000}}$, the causal effect of moving from $t = 0$ to $t = t_0$ in 2000 for the group of compliers:

$$DID_q^{at} = \mathbb{E}[q_{2000}(t_0,\theta) - q_{2000}(0,\theta)|T = c] = LATE_{q_{2000}}. \tag{1.3}$$

Another route to identification would have been to use the *never takers* as a source of comparison for the change in grassland supply of the *compliers* in the absence of the reform:

$$DID_q^{nt} = \mathbb{E}[q_{2005}(t_1,\theta) - q_{2000}(0,\theta)|T = c] - \mathbb{E}[q_{2005}(0,\theta) - q_{2000}(0,\theta)|T = nt]. \tag{1.4}$$

Identification of the causal effect of the grassland subsidy follows from the assumption that both *never takers* and *compliers* would have had parallel trends in the absence of the reform:

Assumption 2 (Parallel trends, never takers) *We assume that grassland supply would have followed parallel trends among compliers and never takers in the absence of the reform:*

$$\mathbb{E}[q_{2005}(0,\theta) - q_{2000}(0,\theta)|T = c] = \mathbb{E}[q_{2005}(0,\theta) - q_{2000}(0,\theta)|T = nt].$$

Under Assumption 2, our DID^{nt} estimator identifies $LATE_{q_{2005}}$, the causal effect of moving from $t = 0$ to $t = t_1$ in 2005 for the group of compliers:

$$DID_q^{nt} = \mathbb{E}[q_{2005}(t_1, \theta) - q_{2005}(0, \theta) | T = c] = LATE_{q_{2005}}. \tag{1.5}$$

1.3.3 Estimating the effect of the reform at the commune level

In practice, we choose to perform our analysis not at the individual level, but at the commune level. We choose this approach because it helps solve several issues we have encountered when trying to take our identification strategy to the data. We face two main issues when trying to operationalize the estimators presented in equations (1.2) and (1.4). Let us examine each of them in turn and explain why using data aggregated at the commune level helps solve them.

The first issue is that we do not observe the groups of *compliers*, *always takers*, and *never takers*. Our data does not allow us to identify with enough certainty the policy recipients in the outcome data. We are actually missing a lot of beneficiaries of the grassland program in the pre-2000 data. This is because the identifiers used in the surveys where outcomes are measured differ from the identifiers used in the administrative data where beneficiaries of the Grassland Conservation Program are listed. We have tried to do our best at matching the two sources but our matching rate is far from satisfactory. The problem is that the measurement error changes over time: we fail to identify a lot of beneficiaries before 2000, but the successful matching rate increases steeply after 2000. As a consequence, at the individual level, we wrongly allocate farmers that are *always takers* into the group of *compliers*. This biases our estimator downwards perhaps severely.

The second issue is that we have made the implicit assumption that there are no leakage effects of the program. Leakage would occur if contracting farmers exchanged land with non contracting farmers because of the policy, the former renting or buying grassland from the latter, and the latter renting or buying cropland from the former. Leakage is a plausible reaction to the program, since contracting farmers receive a subsidy per hectare of grassland, they now value grassland more relative to cropland than non contracting farmers do. A comparison between contracting and non contracting farmers at the individual level would confound the leakage effects with a true additional effect of the program and would thus overestimate the total effect of the program. The overall effect of the subsidy on grassland area could very well be null but our individual level DID estimator would estimate it to be positive. Performing our treatment effect estimate at the commune level enables us to account for possible leakage effects of the policy. We posit that most leakage, if it exists, takes place at the commune level, between geographically close farmers. This is a credible assumption since land markets are mostly local. As a consequence, with our approach, any transfer of land between farmers residing in the same commune that does not alter the overall land use within the commune is not counted as additional.

We thus estimate our main regressions at the commune level. At the commune level, we have access to accurate data on the number of beneficiaries of the Grassland Conservation Program and to accurate data on the proportion of grassland in the usable agricultural area. We compute the growth rate in the total number of beneficiaries of grassland contracts per commune between 2000 and 2005. If the growth rate is positive, the commune belongs to the treated group, while if the growth rate is equal to zero, the commune is used as a control. We denote communes in the treated group with the random variable $D = 1$ and communes in the control group with the random variable $D = 0$. Our main outcome variable of interest is $Q_{c,y}$, the proportion of grassland in the usable agricultural area of commune c in year y. $Q_{c,y}$ is the average of the proportion of grassland in the usable agricultural area of all farms in commune c in year y, weighted by $w_{i,c}$ the share of each farm in the total usable agricultural area of commune c:

$$Q_{c,y} = \sum_{i=1}^{N_c} w_{i,c} q_{i,y}(t_{i,y}, \theta_i), \tag{1.6}$$

with N_c the number of farms in commune c, $t_{i,y}$ the value of the subsidy received by farmer i in year y and $q_{i,y}$ the functions $q_{2000}(t, \theta)$ and $q_{2005}(t, \theta)$ as defined above. Our commune-level DID estimator can then be defined as follows (omitting the c index for brevity):

$$DID_Q = \mathbb{E}[Q_{2005} - Q_{2000}|D = 1] - \mathbb{E}[Q_{2005} - Q_{2000}|D = 0]. \tag{1.7}$$

The main theoretical result of this section is that, under a mild set of assumptions, DID_Q is equal to a weighted average of farm-level LATEs as defined in equations (1.3) and (1.5):

Proposition 1 *Under a set of conditions made precise in Appendix 1.8.1, there exists strictly positive scalars α and β with $\alpha + \beta = \Pr(T = c|D = 1)$ such that:*

$$DID_Q = \alpha LATE_{q_{2000}} + \beta LATE_{q_{2005}}.$$

Proof: See Appendix 1.8.1. ■

Finally, in order to compute elasticity estimates and benefit-cost ratios, we also compute the impact of the reform on the monetary transfers received at the commune level using the same DID estimator as in equation (1.7):

$$DID_M = \mathbb{E}[M_{2005} - M_{2000}|D = 1] - \mathbb{E}[M_{2005} - M_{2000}|D = 0], \tag{1.8}$$

with $M_{c,y}$ the monetary transfer received by farmers in commune c in year y as part of the Grassland Conservation Program. The following proposition shows that this DID estimator identifies the weighed average of the transfers received by compliers in 2000 and in 2005 multiplied by N, the average number of farms in a commune:

Proposition 2 *Under a set of conditions made precise in Appendix 1.8.1, there exists strictly positive scalars α and β with $\alpha + \beta = \Pr(T = c|D = 1)$ such that:*

$$DID_M = N(\alpha\mathbb{E}[t_0 q_{2000}|T = c] + \beta\mathbb{E}[t_1 q_{2005}|T = c]).$$

Proof: See Appendix 1.8.1. ■

Propositions 1 and 2 are the core of our empirical strategies. They show that, under plausible assumptions, our identification strategy relying on commune-level data identifies meaningful treatment effect parameters. First, these parameters are computed on the subpopulation of *compliers*, the farmers that enter the program after the 2000 reform. Second, the DID estimate of the effect of the reform on the proportion of grassland in the usable agricultural area DID_Q is equal to a weighted average of two impacts of the reform on compliers: the one moving them from a subsidy of 0 to t_0 in 2000 and the one moving them from 0 to t_1 in 2005. The dual nature of our treatment effect parameter stems from the dual nature of the comparison groups that we use to proxy the trends of compliers absent the reform: *always takers* and *never takers*. *Always takers* benefit from the program both in 2000 and in 2005. As a consequence, they proxy for the change that *compliers* would have experienced if the requirement of a specialization rate higher than 75% had been cancelled before 2000 and *compliers* had been allowed to enter the program with a subsidy rate of t_0. *Never takers* do not benefit from the program in 2000 nor in 2005. As a consequence, they proxy for the change that compliers would have experienced if they had not been allowed to enter the program after 2000. Third, under the same assumptions, the DID estimate at the commune level of the effect of the reform on the transfers received by farmers enrolled in the Grassland Conservation Program, DID_M, is also a weighted average of two transfers. The first transfer is the average amount of money that would have been received by *compliers* if they would have been allowed to enter the Grassland Conservation Program before 2000. The second transfer is the average amount of money received by *compliers* once they have been allowed to enter the Grassland Conservation Program after 2000. Fourth, the weights involved in computing the treatment effect parameters identified by DID_Q and DID_M are the same: α weighs the treatment effects defined in 2000 and β the ones defined in 2005. α is equal to the difference in the proportion of *always takers* between the control and treated communes while β is equal to the difference in the proportion of *never takers* between the control and treated communes. Under our assumptions, $\alpha + \beta$ is equal to the proportion of *compliers*. DID_Q and DID_M thus identify the sum of two Intention to Treat Effects (ITE): the effect on compliers multiplied by the proportion of compliers. Fifth, when we compute the elasticity of grassland supply, we compare the change in grassland area (obtained by multiplying our proportion estimate DID_Q by the average agricultural area at the commune level and dividing it by the average 2000 level) to the change in monetary transfers estimated by DID_M (divided by the average amount of transfers at the commune level in 2000). Even though this estimate is not the average of the two separate elasticities of the 2000 and 2005 impacts taken separately, it still is a valid elasticity of the average response of the compliers to two different transfers. Sixth, all of these interpretations of DID_Q and DID_M rest on several assumptions, among which the most important is the absence of diffusion effects. Nevertheless, our approach is robust to a relaxation of this assumption. If the diffusion effects are limited to the commune level (which is highly likely since most diffusion effects take place on the land market and thus are concentrated within a commune), our estimators include the response of both *always takers* and *never takers* to the reform. They thus estimate the total effect of the reform, net of any indirect impacts on *never takers* and *always takers*. Seventh and finally, another critical assumption for the valid interpretation of our estimator is that *bunchers* are in the same proportion in treated and control communes, so that their fate does not influence our estimator. We believe this assumption is well-justified since the proportion of *compliers* is small. If this assumption was to be wrong, our resulting estimates would be biased upwards. Indeed, *bunchers* experience a less intense response to the reform since they were already bunching too high with respect to the unconstrained incentive. As a consequence, their change in grassland area between 2000 and 2005 is less steep than the one that would have been experienced by *compliers* if they had been allowed into the program in 2000. If *bunchers* are not in the same proportion in both treatment and control groups, our estimate thus provides an upper bound on the effect of the reform on *compliers*.

1.3.4 Estimation

Our data is a commune-year panel over four periods. We estimate a two-way fixed effects model, which is an extension of the simple DID to more than two periods. The baseline equation is given by:

$$Y_{ct} = \widetilde{\alpha} D_{ct} + \widetilde{\beta} X_{ct} + \widetilde{\eta_c} + \widetilde{\xi_t} + \widetilde{\epsilon_{ct}} \tag{1.9}$$

where Y_{ct} is the aggregated outcome variable (for example the share of permanent grassland area in commune c at time t), D_{ct} is a dummy taking a value of one starting in 2003 for communes where the number of beneficiaries increased after the reform, X_{ct} is the vector of aggregated control variables (for example the number of small farms in commune c at time t), $\widetilde{\eta_c}$ and $\widetilde{\xi_t}$ represent the commune and year fixed effects. The fixed effects control for time-invariant unobserved commune characteristics (e.g. altitude, slope) and for effects that are common to all communes at one point in time (e.g. changes in CAP policies that affect every farmer in the same way). ϵ_{ct} is the error term and includes unobserved variables such as managerial ability, environmental preferences and prices. We also include department-specific yearly effects in our main specification. The estimated standard errors are robust to heteroskedasticity and are clustered at the commune level to account for serial correlation in the outcome variables (Bertrand et al., 2003).

The parameter of interest, $\widetilde{\alpha}$, captures the average causal effect of the program expansion that followed the change in eligibility criteria. This estimate captures the full impact of the reform, on both beneficiaries and non-beneficiaries located in the same commune. For this parameter to be a consistent estimate of the impact of the reform, the parallel trends assumption must hold, meaning that there should be no systematic differences in outcome trends between treated and control communes before the reform. We test this assumption by comparing trends in outcomes between treated and control communes before the reform.

To check the robustness of the DID specification we re-estimate the intention-to-treat effect using the changes-in-changes (CIC) model proposed by Athey and Imbens (2006). The CIC model is a nonlinear generalization of the DID model to the entire distribution of potential outcomes. The estimated treatment effect is given by the difference between the actual and the counterfactual distribution of the outcome variable in the treated communes. In turn, this difference is given by the difference between the outcome variable of the control communes with the same rank (i.e. in the same quantile) before and after the reform.[10] The key identifying assumption of the CIC method is the time invariance within groups assumption. It is the counterpart of the parallel trends assumption in the DID case and it requires that the population of agents within groups does not change over time.

[10] Specifically, a treated group with a level Y of the outcome variable in the pre-treatment period is matched with a control commune with the same level of the outcome in the same period. Then, this control commune is matched to a control commune with the same rank in the post-treatment period.

However, it has been rarely used in practice so far as the existing statistical tools used for its implementation are quite limited.[11]

1.4 Data

We construct our database at commune level using two types of data. First, we use administrative data from France's *Agence de Services et de Paiements* (ASP) provided to us by the *Observatoire du Développement Rural* (ODR). This data contains information on all beneficiaries of grassland programs from 1999 to 2006.[12] To build our treated and control groups we count the number of beneficiaries in each commune and we compute the growth rate in the number of beneficiaries from before to after the reform.

Second, in order to estimate the outcome and control variables, we resort to farm level data provided by the French Ministry of Agriculture. More specifically, we use the 2000 agricultural census and the farm structure surveys from 1993 to 2007. These surveys are conducted every two years between censuses on 10% of the population of farmers. To construct our variables of interest, we first weight the farm level data using the sampling weights provided in the survey and then sum the weighted data at commune level.

Our main outcomes are the share of permanent grassland, crops and fodder in the total utilised agricultural area, the specialization rate (% of permanent and temporary grassland in the total utilised agricultural area) and the loading ratio (the ratio of livestock units to the forage area). To obtain a better understanding of the potential land use changes triggered by the grassland program, we also look at variables such as the share of total usable agricultural area, the share of forest area and the share of nonproductive land in the total farm area within a commune. Except for the loading ratio, which is transformed applying the inverse hyperbolic sine,[13] we express all our outcome variables as shares in order to account for size differences between communes. Our control variables include the number of farms for each type of crop orientation and for each economic size and the total number of farms in each commune. A detailed definition of all these variables is given in Appendix 1.8.2.

[11] In R, we use the single available command, "CiC" from the "qte" package, which only allows for one pre-treatment period and one post-treatment period and does not allow for the inclusion of covariates.

[12] That dataset contains information such as the commune of residence, the years in which the farmers were enrolled in a grassland program, the number of hectares enrolled and the payment they received every year.

[13] We apply the inverse hyperbolic sine (IHS) transformation to the loading ratio to correct for its highly skewed distribution with a mass point at zero and to ensure equivalence in the unit of measure and interpretation of results with the other outcome variables. IHS is defined as $log(Y_i + (Y_i^2 + 1)^{\frac{1}{2}})$. It is defined at zero and can be interpreted similarly to a log-linear specification.

Our final dataset includes only farmers having at least one hectare of utilised agricultural area and only those communes where at least one farmer has received a subsidy for grassland conservation over the period 1999 to 2006. The sample constraint on communes enables us to build treatment groups with more similar characteristics than if we would have included also communes with no beneficiary of the Grassland Conservation Program over the analysed period. We work with two balanced panels: one from 1993 to 1997 and one from 2000 to 2007. The reason why we decided to split the data into two periods is that survey identifiers are erased after each census. In our case this happens in 2000, so having a coherent balanced panel over the whole period is impossible. We thus use a balanced panel of 9,998 communes from 1993 to 1997 to perform the placebo test and a balanced panel of 10,468 communes from 2000 to 2007 to recover the treatment effect. Among these, 7,808 communes are common between the two periods.[14] We choose the time window 1993-2007 to avoid possible complications due to the fact that there was no Grassland Conservation Program before 1993 and that the new scheme starting in 2007 had many changes compared to the previous one.

Table 1.1 reports the mean and standard deviation of our outcome variables, by treatment group and sample. Recall that our control communes are those in which farmers are benefiting from the grassland subsidy for the whole 1993-2007 period. Thus, as a consequence of the program requirements, they have a higher share of permanent grassland and specialization rate and a lower loading ratio than the treated communes, where farmers became beneficiaries only after the 2000-2003 reform. The control communes have also a higher share of forest and nonproductive land and a bigger part of the agricultural area that is owned. Conversely, the farms located in treated communes have a higher share of crops, fodder and utilised agricultural area and have more rented land than farmers in control communes. This selection in levels does not create any problems for our identification strategy since the DID methodology removes permanent differences between the treated and control groups.

1.5 Results

In this section we start by presenting the magnitude of the effect of the 2000 reform on the number of contracting farmers and the amount of transfers received as part of the

[14] We also build a balanced panel of the 7,808 communes over the whole period, but we observe a huge drop in all our outcome variables between 1997 and 2000 that we cannot explain otherwise than by a change in the weighting system starting with the 2000 census. We thus choose to split the sample into two periods in order to avoid capturing this decrease in the treatment effect estimation.

Grassland Conservation Program in the communes affected by the program expansion. We then show the results of the main regressions estimating the impact of the reform on outcomes based on our baseline equation 1.9. Finally we present some robustness checks of the main results.

1.5.1 The Size of the Program Expansion in Treated Communes

Figure 1.4 shows the total number of beneficiaries of grassland conservation contracts over time, as a function of the treatment status of the commune. As expected and by construction, the treated communes see a sharp increase in the number of participants starting after 2000 and especially marked from 2002 to 2003. The number of beneficiaries in treated communes jumps from slightly above 20,000 in 2000 to slightly above 35,000 in 2003, or an increase of about 75%. In the control communes, the number of beneficiaries is almost constant over time. Figure 1.6 shows that the proportion of farmers benefiting from the Grassland Conservation Program also rises sharply after 2000 in treated communes, while it remains stable in control communes. Formally, we estimate the impact of the reform on the share of beneficiaries in treated communes to be 10.7 ± 0.35 p.p. (Table 1.3), which represents a near doubling of the proportion of contracting farmers in treated communes. The map of France in Figure 1.5 shows that both treated and control communes are quite heterogeneously dispersed throughout the country, which is good for our identification strategy since it suggests that they are rather similar at least in their location and thus in the opportunity cost of grassland. The only two areas not covered are the Paris basin where there is no grassland and Corsica that we exclude from the analysis.

The key insight behind the change in the proportion of participants on which our identification strategy rests is that this increase in the number of beneficiaries stems from the entry of the *compliers* in the program. The *compliers* are farmers that were ineligible to the program before 2000 because their specialization rate was too low, but that are free to enter the program after 2000 once the requirement on the specialization rate is relaxed. In order to test this part of our model, we define *potential compliers* as farmers who have a specialization rate strictly lower than 75% in 2000 and we regress this indicator on the treatment dummy (which is defined at the commune level). Our assumption is that we will see more *potential compliers* in 2000 in treated communes (where a lot of new entrants will appear after the 2000 reform). Hopefully, the proportion of *potential compliers* in 2000 will be higher in treated communes by the same amount as the proportion of compliers that we have estimated in Table 1.3 (roughly 10%). The results from this regression are presented in Table 1.4. We find that the proportion of *potential compliers* is higher in treated communes than in control communes by 7.5 to 10.3 p.p., which is very close to our estimate of the proportion of *compliers*. As a consequence, our theory that the increase in the proportion of participants in treated communes comes mainly from farmers ineligible to the program before the 2000 reform is vindicated.

The amount of monetary transfers as part of the French Grassland Conservation Program increased markedly in treated communes, as shown in Figure 1.7. We estimate that the program expansion increased the total amount of grassland subsidies in treatment communes by 5,000±513 Euro (Table 1.3), or a 42% increase. Figure 1.7 shows that the amount of subsidies increased in control communes as well, because of the increase in the per hectare payment that accompanied the introduction of the new programs, but this increase is of smaller magnitude. Note finally that the average increase of transfers in treated communes is very close to the increase received by the average *complier*. Indeed, the average monetary transfer to *compliers* is equal to the average transfer at the commune level divided by the proportion of compliers (roughly 0.10) and by the average number of farmers per commune (roughly 10). These two operations approximately cancel out, which implies that the average monetary impact of the reform at the commune level is roughly equal to the average monetary impact at the complier level.

1.5.2 The Impact of the Program Expansion on Outcomes

We present both graphical evidence and regression results of the effect of the 2000 reform of the Grassland Conservation Program on our outcomes of interest. As a general description of the graphical evidence, the first column of plots in each figure, denoted by (a), represents the placebo test on the 1993-1997 sample of communes. The second one, denoted by (b), shows the treatment effect of the program on the sample of communes from 2000 to 2007. The first line of plots presents the trends in average outcome variables by treatment status, while the second line shows the yearly coefficients on the difference between treated and controls. These coefficients can be interpreted as an estimate of the impact of being treated on the outcome variable in a given year relative to the reference year. The effect is statistically significant if zero is not included in the 95% confidence interval, represented by dashed lines. We present regression results for different specifications with and without additional control variables and with and without department-year fixed effects. The results are consistent across specifications even though the point estimates slightly change with the introduction of controls or additional fixed effects. Our preferred specification is the one that accounts for both commune characteristics and yearly, department specific shocks.

Figure 1.8 shows that, graphically, there is no difference in the share of permanent grassland between treated and control communes from 1993 to 1997, as the coefficients of the interaction term fluctuate around zero before 2000. Between 2000 and 2007 the wedge opens up, suggesting a small positive impact of the Grassland Conservation Program on the share of permanent grassland area. Figure 1.9 shows that there is a small increase in crop area in treated communes compared to control communes from 1995 to 1997, while after 2000 the difference becomes negative. The share of fodder area does not appear to be affected by the change in eligibility requirements, as the yearly coefficients swing around zero both before and after 2000 (Figure 1.10). In Figure 1.11 we can observe that the specialization rate is stable before 2000 and increases afterwards, indicating a positive effect of the grassland program on this outcome. Finally, in Figure 1.12 it seems that there is a slight decrease in the loading ratio between 1993 and 1997 in the treated communes compared to control communes, while after 2000 there is no difference in the loading ratio of the two groups. All in all, the visual evidence suggests that the grassland program leads to a small increase in the share of permanent grassland area and the specialization rate, a decrease in the share of crops and no change in the share of fodder area and the loading ratio.

Table 1.2 presents the results of the fixed effects regressions. The estimated coefficients confirm the conclusions of the graphical evidence, but are in general not statistically different from zero. Nevertheless, we find that the share of permanent grassland area increases after the reform by 0.28±0.55 p.p. in treated communes compared to control communes. Likewise, the specialization rate increases by 0.45±0.49 p.p. At the same time, the share of crop area decreases by a similar amount, -0.40±0.39 p.p., while there is no difference in the share of fodder area and loading ratio between the two groups of communes. An interesting pattern that arises from these results is a potential switch from crops to grassland in the treated communes from the pre- to the post-treatment period.

Apart from croplands, the additional grassland area that we find after 2003 might also come from forest or nonproductive land. Figure 1.13 shows that the share of utilised agricultural area in total farm area slightly decreases in treated communes with respect to control communes after 2000, while before there was no difference between the two groups. Contrariwise, as shown in Figure 1.14, the share of forest area increases in the post-treatment period. Figure 1.15 indicates that the difference in the share of nonproductive land between the comparison groups was slightly positive in the pre-treatment period and it became almost null afterwards. The regression results from Table 1.5 suggest that the share of utilised agricultural area in total farm area remains rather stable over the whole period between the treated and control communes. Moreover, the share of forest area increases over time, from -0.25±0.43 p.p. to 0.10±0.35 p.p., while the share of nonproductive land decreases by almost the same amount, from 0.23±0.33 p.p. to 0.00±0.29 p.p. Thus, since the share of utilised agricultural area does not change over time and the decrease in nonproductive land is compensated by the increase in forest area, we argue that the increase in the share of grassland comes mainly from the decrease in the share of crops.

Putting everything together, our interpretation of the results is that the policy reform induced some farmers living in the treated communes to keep more grassland on their farms mainly at the expense of croplands.

1.5.3 Robustness Checks

Changes-in-changes. Our identification strategy relies on the parallel trends assumption. However, for some of our outcome variables we acknowledge the existence of pre-treatment trends that, even though not statistically significant, might invalidate our methodology. For this reason we perform a robustness check using the non-parametric equivalent of the DID method, the CIC strategy. Due to difficulty in practical implementation, the CIC regressions do not include fixed effects or additional controls. Table 1.6 shows that this method yields very similar results to our preferred specification including both control variables and commune, year and department-year fixed effects.

Different samples. Our sample is composed of two balanced panels, one from 1993-1997 and one from 2000-2007. To test the sensitivity of our results to this choice, we re-estimate the model using two unbalanced panels from 1993-1997 and 2000-2007 and a balanced panel restricted to the same communes for the whole 1993-2007 period. The results are summarized in Table 1.7 and Table 1.8. Even though the precision and magnitude of the estimated coefficients vary slightly with the sample size (i.e. the bigger the sample size, the more precise the estimate), in all cases the qualitative findings remain similar to the ones estimated on the balanced sample of different communes between the two periods.

1.6 Elasticity Estimates and Cost-Benefit Analysis

In this section we start by computing the elasticity of the additional permanent grassland supply with respect to the amount of subsidies. Next, we build a cost-benefit analysis by comparing the additional costs of the program due to the eligibility criteria change with its additional benefits, quantified using values taken from the literature. Throughout this section we present mean estimates along with their 95% confidence intervals that we build using transformed standard errors through the Delta Method.[15]

1.6.1 Elasticity Estimate

The impact we measure of the French Grassland Conservation Program's reform on commune level outcomes is not statistically different from zero. However, what matters for policymakers is the relative size of the impact compared with the amount of money spent. We find evidence that the policy reform was accompanied by a substantial inflow of money in treated communes compared to control communes, of around 5,000±513 Euro per hectare over the 5 years of grassland contracts, corresponding to an increase of 42.46±6.21%.[16] This amount of additional subsidies corresponds to a comparatively small increase in grassland area of 3.73±7.31[17] hectares per treated commune, or an increase of 0.76±1.49%[18] in grassland area. Therefore, we estimate a low elasticity of the supply of grassland with respect to the amount of the subsidy of 0.02±0.04.[19] These elasticity estimates are summarized in Table 1.9.

Our results imply that the cost per hectare of additional permanent grassland over the 5 years of contracts is 1,340±2,628 Euro,[20] which is almost three times bigger than the actual subsidy per hectare over the same period of time, of 450 Euro.[21] Dividing the additional spending due to the reform by the actual subsidy per hectare of grassland gives an estimate of the increase in the subsidized area at the commune level. We find that the reform has increased the amount of subsidized area by 11 hectares per treated commune. Given that the corresponding increase in grassland area is 3.73 hectares per commune, we estimate a low additionality ratio of 34%.[22]

[15] See Appendix 1.8.1 for a description of the Delta Method.

[16] The percentage change is computed as the ratio between the estimate of the additional amount of subsidies and the counterfactual mean of the amount of subsidies in treated communes after the reform (i.e. (5,000 Euro /11,775 Euro) × 100).

[17] The additional hectares of grassland are computed by multiplying the estimate of the share of permanent grassland area with the sample mean of the total utilised agricultural area in treated communes after the reform (i.e. 0.28p.p./100 × 1,333 ha).

[18] The percentage change is computed as the ratio between the estimate of the share of permanent grassland area and the counterfactual mean of the share of permanent grassland area in treated communes after the reform (i.e. (0.28p.p./37.02%) × 100).

[19] The elasticity of the supply of grassland is computed as the ratio between the percentage change in grassland area and the percentage change is the amount of subsidies (i.e. 0.76/42.46%).

[20] The cost per additional hectare of grassland is obtained by dividing the estimated additional cost to the additional hectares of grassland (i.e. 5,000 Euro/3,73 ha).

[21] The subsidy per hectare of grassland for PHAE and CTE/CAD together was about 90 Euro.

[22] The additionality ratio is as the ratio between the additional subsidized hectares and the additional hectares of grassland (i.e. (3.73 ha/11 ha) x 100)

1.6.2 Cost-Benefit Analysis

In this section, we perform a cost-benefit analysis of the reform. To estimate the benefits of the reform, we model the emissions per hectare in the presence of the reform and in its absence. We choose to model the dynamics of carbon stored in the soil after a change in soil usage using the saturated exponential function that Arrouay et al. (2002) propose for France:

$$F_{s,u}(t) = \Delta_{s,u}(1 - e^{-k_{s,u}t}), \tag{1.10}$$

where $F_{s,u}(t)$ is the cumulated flow of carbon into the soil t years after converting the soil from use s to use u in tons of carbon per hectare (tC/ha), $\Delta_{s,u}$ is the long run difference in carbon storage between soil use u and soil use s and $k_{s,u}$ is the speed at which carbon flows after conversion. Figure 1.16 shows the flows of carbon after the conversion from grassland to cropland and from cropland to grassland using the parametrizations proposed by Arrouay et al. (2002). In the long run, grassland stores 25tC/ha more than cropland on average in France. The conversions between grassland and cropland are not symmetric: while carbon is depleted very fast when grassland is converted to cropland ($k_{g,c} = 0.07\text{year}^{-1}$, implying that 7.4tC are lost in the first 5 years after conversion of grassland (g) to cropland (c)), it takes a lot of time to rebuild the carbon content in the soil after conversion of grassland ($k_{c,g} = 0.025\text{year}^{-1}$, implying that 2.9tC are stored in the first five years after cropland is converted to grassland).

To estimate the benefits from the program, we estimate the value of a hectare of grassland saved by the program. In the absence of the program, grassland is converted into cropland at $t = 0$ and starts emitting immediately. Emissions per unit of time (here per year) in the absence of the program, $E^0(t)$, can be computed as the negative of the first derivative of the cumulated carbon flow into the soil after conversion of grassland to cropland:

$$\begin{aligned} E^0(t) &= -3.66F'_{g,c}(t) \\ &= -3.66\Delta_{g,c}k_{g,c}e^{-k_{g,c}t} \end{aligned} \tag{1.11}$$

where 3.66 is the constant of conversion from tons of carbon into the soil to tons of CO_2 equivalent, so that emissions are expressed in $tCO_2eq/ha/year$. In the presence of the program, depending on how fast the effect of the program stops, emissions start at $t = x$. In our main specification, we assume that $x = 5$, meaning that the program has no permanence: the area in grassland saved by the program is converted to cropland as soon as the payments stop. As a consequence, we have:

$$E^1(t,x) = \begin{cases} 0 & \text{if } t \leq x \\ -3.66\Delta_{g,c}k_{g,c}e^{-k_{g,c}(t-x)} & \text{if } t > x. \end{cases} \tag{1.12}$$

In order to compute the value of the program, we first compute the value of one hectare of grassland saved by the program. We assume that, absent the program, this hectare would have been converted into cropland at year $t = 0$ and would have emitted $E^0(t)$ tons of CO_2 equivalent each year. We also assume that, under the program, this hectare would have been conserved as grassland until year $t = x$ and would have emitted $E^1(t)$ tons of CO_2 equivalent each year. The climate benefits of one hectare of grassland saved by the program until year x is thus:

$$B_c(x) = -\int_0^\infty \left(E^1(t) - E^0(t)\right) SCC_t e^{-rt} dt, \tag{1.13}$$

with SCC_t the Social Cost of Carbon at time t and r the discount rate. Assuming a constant Social Cost of Carbon, we show in Appendix 1.8.1 that the climate benefits from preventing the conversion of one hectare of grassland until date x is:

$$B_c(x) = \frac{-3.66\Delta_{g,c}SCC}{1 + \frac{r}{k_{g,c}}} \left(1 - e^{-xr}\right). \tag{1.14}$$

The intuition for the formula for $B_c(x)$ is as follows. The ratio in the first part of the formula measures the discounted benefit of keeping one hectare of grassland from converting to cropland forever. The numerator measures the social value of all the carbon stored in the ground under one hectare of grassland instead of one hectare of cropland. This is the social value of 25 tons of carbon, or 91.5 tCO_2eq. Using a Social Cost of Carbon of 24 Euro as proposed by the U.S. Environmental Protection Agency (EPA),[23]the social value of the carbon stored in the ground under one hectare of grassland versus one hectare of cropland is 2,196 Euro. The denominator serves to discount the stock of carbon by the time it takes for it to be released after conversion. The carbon is indeed not released all at once after conversion to cropland. What drives the amount of discounting is the ratio $\frac{r}{k_{g,c}}$. When $k_{g,c}$, the speed of extraction of carbon from the ground, is low relative to r,

[23] The EPA middle estimate (i.e. using a discount rate of 3%) for the SCC in 2010 is $31 (in 2007 USD) per ton of averted CO_2. Using the USD-EUR exchange rate of 2007 (i.e. 1 USD = 0.77 EUR), the SCC equals approximately 24 Euro.

a lot of emissions occur far in the future and the discounting is important. When $k_{g,c}$ is large, a lot of emissions happen very soon after conversion and the discounting is small. With $k_{g,c} = 0.07$ and $r = 0.02$, the value of the total stock of carbon into the ground under grassland is discounted by 77%. The last part of the formula accounts for the fact that the program only displaces emissions over time. As expected, when $x \to \infty$, this term tends to one and there is no discounting. When $x = 5$ years, the discounting is equal to 9.5%, meaning that the program only saves the equivalent of 9.5% of the total value of carbon stored in the soil. With the parameter values selected up to now, the climate value of preventing one hectare of grassland from converting to cropland for 5 years is equal to 162.54 Euro.

Grassland also brings benefits beyond reducing carbon emissions (cleaner water, pollination services, hunting and landscape). We assume that the value of these services is B_a Euro/ha/year and that they disappear instantaneously when grassland is converted into cropland. Adding these services to the climate benefits brings the following formula for computing the total climate benefits from grassland:

$$B(x) = \left(\frac{-3.66\Delta_{g,c}SCC}{1 + \frac{r}{k_{g,c}}} + \frac{B_a}{r} \right) \left(1 - e^{-xr}\right). \tag{1.15}$$

The proof of this result is in Appendix 1.8.1. Puydarrieux and Devaux (2013) estimate the values of the services brought by grassland as 44 Euro/ha/year for water quality, 60 Euro/ha/year for pollination, 4 Euro/ha/year for hunting,[24] and 60 Euro/ha/year for landscape amenities. In total, these additional benefits bring 168 Euro/ha/year. The discounted value of these benefits over 5 years is equal to 799.36 Euro/ha. Thus, the total benefit of preventing the conversion of 1 ha of grassland to cropland for five years is equal to 961.9 Euro.

Let us now compute the total benefit from the program and its benefit-cost ratio using our estimates of the impact of the reform on additionality and on transfers. We estimate that the program reform has increased grassland area at the commune level by 3.73±7.31 ha for a cost of 5,000±513 Euro. Assuming that these benefits last for five years only, and that grassland is converted to cropland as soon as the payments stop, the total value generated by the program is equal to 3.73*961.9=3,587.88 Euro, which implies a benefit-cost ratio of 0.72±1.41. The climate benefits of the program are equal to 3.73*162.54=606.27 Euro, which implies a climate benefit-cost ratio of 0.12±0.24. Assuming instead that the benefits of the

[24] Here we consider the hunting as a supply activity and not as a leisure activity. Thus we value it at the market price of the prey.

program last forever, even if the payments stop after 5 years (a very optimistic assumption which yields to an upper bound on the benefit estimates), we find that the total value generated by the program would be equal to 10,108*3.73=37,702.84 Euro, and thus that the program would have a benefit-cost ratio of 7.54±14.8. Under the assumption of full permanence of the program impacts after 5 years, the climate benefits of the program would be equal to 3.73*1,708=6,370.84, and its benefit-cost ratio to 1.27±2.53. Our estimates enable us to compute two additional critical values: the degree of permanence of the program effects that would enable the program to break even and the Social Cost of Carbon that would make the program break even. Considering only climate benefits, the effects of the program have to persist for 72 years after payments stop for the program to break even. When taking into account both climate benefits and the other benefits from grassland, the effects of the program have to persist for 2 years and 2 months for the program to break even. In the absence of any effect of the program beyond five years, the Social Cost of Carbon that would make the program break even on climate benefits alone is equal to 198±392 Euro/tCO_2eq. Under the same assumption, but including all the other benefits that grassland provides, the program would break even for a carbon price of 80±389 Euro/tCO_2eq. The summary of the cost-benefit analysis in presented in Table 1.10.

To improve the precision and validity of our benefit-cost analysis, we combine our own estimates of the additionality of the program with similar estimates obtained in the literature. Two other works have estimated the additionality of the French Grassland Conservation Program. Chabé-Ferret and Subervie (2009) use DID-matching to estimate the additionality of the Grassland Conservation Program in 2005 and find that it has increased the specialization rate of treated farms by 2±4 p.p., or 1.4±2.7 ha, for an additional cost of 3,500 Euro. Gallic and Marcus (2019) use a change in the eligibility rules of the French Grassland Conservation Program in 2015 in order to estimate its additionality. They use two changes as natural experiments: the end of grassland subsidies for farmers located outside of Less Favoured Areas and the opening of grassland subsidies to some farmers inside Less Favoured Areas that were not eligible before. Since Gallic and Marcus (2019) have access to data on all French farmers, their estimates are much more precise than ours.[25] There are several points worthy of notice in Gallic and Marcus (2019). First, they estimate that the program has no permanence: farmers leaving the program immediately decrease their proportion of grassland by 2.47±0.39 p.p., and do not move further in the subsequent years. Second, farmers entering the program experience a similar increase in their proportion of

[25] We have tried to access the same data as Gallic and Marcus (2019) but their access is reserved to members of the statistical services of the French Ministry of Agriculture.

grassland area: 2.48±0.43 p.p.[26] Both of these estimates yield an impact of the Grassland Conservation Program of 1.2±0.35 additional hectares of grassland for each treated farm, for a cost of 2,622 Euro per farm.[27] The benefit-cost ratios obtained using Chabé-Ferret and Subervie (2009) estimates is equal to 1.4*961.9/3,500=0.38±0.74 for the total benefits and to 1.4*162.54/3,500=0.07±0.13 for the climate benefits alone. The benefit-cost ratios obtained using Gallic and Marcus (2019) estimates is equal to 1.2*961.9/2,622=0.44±0.16 for the total benefits and to 1.2*162.54/2,622=0.07±0.03 for the climate benefits alone. Combining these three estimates of the benefit-cost ratio of the French Grassland Conservation Program into one using a meta-regression, we find a climate benefit-cost ratio of 0.07±0.03 and a total benefit-cost ratio of 0.44±0.15 (Figure 1.17). We also estimate that the program would break even for a carbon price of 194±122 Euro/tCO_2eq.

1.7 Conclusion

Payments for Ecosystem Services are being increasingly used in the context of development and environmental policies around the world. Yet, the empirical analysis of their effectiveness remains somewhat sparse. In this paper we provide an evaluation of a major nationwide Payments for Ecosystem Services program, the French Grassland Conservation Program, the largest of such programs in the world. Grassland Conservation Programs, that pay farmers for maintaining grassland cover, might be an effective way to combat climate change, if they succeed in triggering an increase in grassland cover at the expense of cropland for a reasonable amount of money. Unlike most of the previous literature evaluating the effect of Payments for Ecosystem Services, our approach does not rely on matching beneficiaries with similar non-beneficiaries. Instead, we use an exogenous change in the eligibility criteria for participating in a grassland program as a natural experiment. We perform our analysis at the aggregated, commune level in order to account for potential leakage effects within communes and we exploit the natural experiment in a difference-in-differences design: we compare changes in outcomes both over time and between areas where the number of grassland beneficiaries increased after the policy change and areas where the number of beneficiaries remained the same. We show in a theoretical model that our estimator recovers a policy-relevant treatment effect under plausible assumptions.

Our results suggest that the reform of the French Grassland Conservation Program

[26] This is the average of the additionality impacts estimated by Gallic and Marcus on cattle growers and on crop growers weighted by their respective proportion in the treated population.

[27] Amounts computed using Figure 8 in Gallic and Marcus (2019) in a DID design and weighting the results by the proportion of cattle growers and crop growers among the treated.

did increase the amount of transfers in the communes most affected by the reform (by 5,000±513 Euro, or 42.46±6.21%). The reform also managed to induce beneficiaries located in treated communes to increase the grassland area on their farm mainly at the expense of croplands. As such, the reform has generated positive environmental benefits. However, we find that the additionality of the program is low as the subsidized area increased by 11 hectares per commune, while the permanent grassland area only increased by 3.73 hectares (or 0.76±1.49%). As a consequence, we estimate that the elasticity of the supply of grassland is low (0.02±0.04). To estimate the benefit-cost ratio of the reform, we combine our additionality estimate with a model of the dynamics of carbon storage in grassland and estimates of the value of the various ecosystem services provided by grassland. We find that the reform of the Grassland Conservation Program has provided climate benefits equal to 12±24% of its costs, and total environmental benefits equal to 72±141% of its costs. In order to improve the precision of our estimates, we combine them with other estimates of the additionality of the French Grassland Conservation Program using a meta-regression. These estimates are similar in size, even if somewhat smaller than ours, and, together with ours, imply that the climate benefits of the French Grassland Conservation Program are equal to 7±3% of its costs and its total benefits to 44±15% of its costs. We estimate that the carbon price that would make the benefits of the program equal to its cost is 194±122 Euro/tCO_2eq.

Our study contributes to the current increase in policymakers' demand for evidence based analysis of public policies. Several issues deserve attention in future research. First, the cost-effectiveness of the program might be increased if we use an estimate of the true cost for a farmer to participate in a Payment for Ecosystem Service program instead of the government transfers to the farmers. Because participation in Payments for Ecosystem Services is voluntary, farmers' costs of adopting the greener practices are lower than the transfer they receive. Estimating these true costs is still an area for further research. Second, explicitly estimating the heterogeneity across space in both costs and treatment effects would potentially demonstrate the advantage of spatially targeting grassland subsidies.

1.8 Appendix

1.8.1 Proofs

Proof of Proposition 1

In what follows, in order to save on notation and to simplify the derivations, we assume that all farms are of the same size and all communes are of the same size (in practice, we weigh each farm by its usable agricultural area in our commune-level regressions). As a consequence, we assume that each commune has the same number of farms. We also assume an absence of diffusion effects, so that the Stable Unit Treatment Value Assumption is valid. That means that the treatment status of farm i only affects the outcome of farm i and no other. This is not a mild assumption and the main text discusses what happens to our estimator when it is relaxed. Under these simplifying assumptions, the area of grassland among treated and control communes can be written as the sum of the area of grassland in each type of farm weighted by their respective proportions in each type of commune:

$$\mathbb{E}[Q_y|D=d] = \sum_{\tau\in\Omega} \mathbb{E}[q_{i,y}(t_{i,y},\theta_i)|D=d, T_i=\tau] \Pr(T_i=\tau|D=d), \tag{1.16}$$

for $d \in \{0,1\}$ and $\Omega = \{b_{000}, b_{00}, b_0, c, b, at, b_1, b_{11}, b_{111}, nt\}$.

We can now write the commune-level DID_Q estimator as a function of the changes in types:

$$\begin{aligned} DID_Q = \sum_{\tau\in\Omega} & (\mathbb{E}[q_{i,2005}(t_{i,2005},\theta_i) - q_{i,2000}(t_{i,2000},\theta_i)|D=1, T_i=\tau] \Pr(T_i=\tau|D=1) \\ & -\mathbb{E}[q_{i,2005}(t_{i,2005},\theta_i) - q_{i,2000}(t_{i,2000},\theta_i)|D=0, T_i=\tau] \Pr(T_i=\tau|D=0)) . \end{aligned} \tag{1.17}$$

We now assume that the average changes of grassland area over time are the same for each type of farms in both treated and control communes:

Assumption 3 (Same trends by type) *We assume that, $\forall \tau \in \Omega$:*

$$\begin{aligned} & \mathbb{E}[q_{i,2005}(t_{i,2005},\theta_i) - q_{i,2000}(t_{i,2000},\theta_i)|D=1, T_i=\tau] \\ & = \mathbb{E}[q_{i,2005}(t_{i,2005},\theta_i) - q_{i,2000}(t_{i,2000},\theta_i)|D=0, T_i=\tau] = \delta^\tau. \end{aligned}$$

Assumption 3 is mild in that it is highly plausible that farms of the same type react in the same way to the same price changes. Under Assumption 3, we have:

$$DID_Q = \sum_{\tau \in \Omega} \delta^{\tau} \left(\Pr(T_i = \tau | D = 1) - \Pr(T_i = \tau | D = 0) \right). \tag{1.18}$$

Finally, let us assume the following on the proportion of each types:

Assumption 4 (Proportion of types) *We assume that:*

1. $\forall \tau \in \{b_{000}, b_{00}, b_0, b, b_1, b_{11}, b_{111}\}$, $\Pr(T_i = \tau | D = 1) = \Pr(T_i = \tau | D = 0)$,

2. $\Pr(T_i = c | D = 0) = 0$.

Item 1 in Assumption 4 implies that the proportion of *bunchers* in treated and control communes is the same. This is a strong assumption. In general, it mostly means that we disregard the behaviour of *bunchers* in our estimator. This is warranted since they represent a tiny fraction of the farmers. Item 2 in Assumption 4 implies that the proportion of *compliers* in control communes is zero. It means that the reason why these communes see a stability in the number of participants over time is because there are no new entrants in the Grassland Conservation Program.

A consequence of Assumption 4 is that the proportion of *compliers* in the treated group is equal to a fraction of the proportion of *always takers* and of *never takers* from the control group. In order to see this, note that the sum of the proportions of all of the types conditional on the treatment indicator is equal to one (T_i is a partition): $\forall d \in \{0,1\}$, $\sum_{\tau \in \Omega} \Pr(T_i = \tau | D = d) = 1$. Since the proportion of *bunchers* is the same in both treated and control groups (item 1 in Assumption 4), we have that $\Pr(T_i = c|D = 1) + \Pr(T_i = at|D = 1) + \Pr(T_i = nt|D = 1) = \Pr(T_i = at|D = 0) + \Pr(T_i = nt|D = 0)$. As a consequence, we have: $\Pr(T_i = c|D = 1) = \Pr(T_i = at|D = 0) - \Pr(T_i = at|D = 1) + \Pr(T_i = nt|D = 0) - \Pr(T_i = nt|D = 1) = \alpha + \beta$.

Under Assumption 4, equation (1.18) becomes:

$$\begin{aligned}
DID_Q &= \sum_{\tau \in \{at, nt, c\}} \delta^{\tau} \left(\Pr(T_i = \tau | D = 1) - \Pr(T_i = \tau | D = 0) \right) && (1.19) \\
&= -\alpha \delta^{at} - \beta \delta^{nt} + (\alpha + \beta) \delta^{c} && (1.20) \\
&= \alpha(\delta^{c} - \delta^{at}) + \beta(\delta^{c} - \delta^{nt}) && (1.21) \\
&= \alpha LATE_{q_{2000}} + \beta LATE_{q_{2005}}, && (1.22)
\end{aligned}$$

where the first equality uses item 1 in Assumption 4, the second and third equalities use the implication of Assumption 4 and the last equality uses Assumptions 1 and 2.

Proof of Proposition 2

We use the same set of simplifications used in Section 1.8.1. All farms are of the same size and all communes are of the same size. As a consequence, each commune has the same number of farms. We also assume an absence of diffusion effects, so that the Stable Unit Treatment Value Assumption is valid. Under these simplifying assumptions, the transfers received by treatment and control communes are the sum of the transfers received by each type of farm weighted by their respective proportions in each type of commune multiplied by N, the number of farms in each commune:

$$\mathbb{E}[M_y|D=d] = N \sum_{\tau\in\Omega} \mathbb{E}[t_{i,y}q_{i,y}(t_{i,y},\theta_i)|D=d, T_i=\tau]\Pr(T_i=\tau|D=d), \tag{1.23}$$

for $d \in \{0,1\}$ and $\Omega = \{b_{000}, b_{00}, b_0, c, b, at, b_1, b_{11}, b_{111}, nt\}$.

We can now write the commune-level DID_M estimator as a function of the changes in types:

$$\begin{aligned} DID_M = N \sum_{\tau\in\Omega} &(\mathbb{E}[t_{i,2005}q_{i,2005}(t_{i,2005},\theta_i) - t_{i,2000}q_{i,2000}(t_{i,2000},\theta_i)|D=1, T_i=\tau]\Pr(T_i=\tau|D=1) \\ &-\mathbb{E}[t_{i,2005}q_{i,2005}(t_{i,2005},\theta_i) - t_{i,2000}q_{i,2000}(t_{i,2000},\theta_i)|D=0, T_i=\tau]\Pr(T_i=\tau|D=0)). \end{aligned} \tag{1.24}$$

Under Assumption 3, the change in transfers received by the *always takers* and the various types of *bunchers* is the same in treated and control communes. Under Assumption 4, the contribution of the *bunchers* to DID_M becomes zero. The contributions of *never takers* to DID_M is also zero by construction (they receive no transfers). We thus have:

$$\begin{aligned} DID_M = N\,&(\mathbb{E}[t_1q_{i,2005}(t_1,\theta_i) - t_0q_{i,2000}(t_0,\theta_i)|T_i=at](\Pr(T_i=at|D=1) - \Pr(T_i=at|D=0)) \\ &+\mathbb{E}[t_1q_{i,2005}(t_1,\theta_i)|D=1, T_i=c]\Pr(T_i=c|D=1)). \end{aligned} \tag{1.25}$$

Under Assumption 1, we also have:

$$\mathbb{E}[t_1q_{2005}(t_1,\theta) - t_0q_{2000}(t_0,\theta)|T=c] = \mathbb{E}[t_1q_{2005}(t_1,\theta) - t_0q_{2000}(t_0,\theta)|T=at].$$

Using the fact that $\Pr(T_i=c|D=1) = \alpha+\beta$ and $\Pr(T_i=at|D=1) - \Pr(T_i=at|D=0) =$

$-\alpha$, we have:

$$DID_M = N\left(\alpha \mathbb{E}[t_0 q_{i,2000}(t_0,\theta_i)|T_i = at] + \beta \mathbb{E}[t_1 q_{i,2005}(t_1,\theta_i)|T_i = c]\right). \tag{1.26}$$

Closed form solutions for the discounted benefits of grassland

Let us start with the formula for climate benefits:

$$B_c(x) = -\int_0^\infty \left(E^1(t) - E^0(t)\right) SCC_t e^{-rt} dt,$$

The second part of the expression is the simplest to start with:

$$\begin{aligned}
B_c^0(x) &= \int_0^\infty E^0(t) SCC_t e^{-rt} dt \\
&= -3.66\Delta_{g,c} k_{g,c} SCC \int_0^\infty e^{-(k_{g,c}+r)t} dt \\
&= -3.66\Delta_{g,c} k_{g,c} SCC \left[-\frac{1}{k_{g,c}+r} e^{-(k_{g,c}+r)t}\right]_0^\infty \\
&= -3.66\Delta_{g,c} k_{g,c} SCC \left[\frac{1}{k_{g,c}+r}\right] \\
&= -\frac{3.66\Delta_{g,c} SCC}{1+\frac{r}{k_{g,c}}},
\end{aligned}$$

where the second equality stems from assuming that SCC_t is constant over time and uses the formula for E_t^0, the third equality uses the formula for the integral of an exponential, the fourth equality the fact that $\lim_{t\to\infty} e^{-(k_{g,c}+r)t} = 0$ and $e^0 = 1$.

The second part of the expression requires a change of variable $y = t - x$:

$$\begin{aligned}
B_c^1(x) &= \int_0^\infty E^1(t) SCC_t e^{-rt} dt \\
&= -3.66\Delta_{g,c} k_{g,c} SCC \int_x^\infty e^{-k_{g,c}(t-x)} e^{-rt} dt \\
&= -3.66\Delta_{g,c} k_{g,c} SCC \int_0^\infty e^{-k_{g,c}y} e^{-r(y+x)} dy \\
&= -3.66\Delta_{g,c} k_{g,c} SCC e^{-rx} \int_0^\infty e^{-(k_{g,c}+r)y} dy \\
&= -\frac{3.66\Delta_{g,c} SCC e^{-rx}}{1+\frac{r}{k_{g,c}}},
\end{aligned}$$

where the second equality stems from assuming that SCC_t is constant over time and using the formula for E_t^1, the second equality uses the change of variable $y = t - x$ and the last equality uses the formula for the integral of an exponential.

Let us now examine the closed form formula for the discounted benefits from a stream of yearly services B_a lasting x years:

$$\begin{aligned} B_a(x) &= \int_0^x B_a e^{-rt} dt \\ &= B_a \int_0^x e^{-rt} dt \\ &= \frac{B_a}{r}(1 - e^{-rx}), \end{aligned}$$

where the last equality stems from the formula for the integral of an exponential function.

The Delta Method

Transformation of one variable. We denote by ω^2 the asymptotic variance of the estimated coefficient $\widetilde{\alpha}$. Then, for the regression coefficient holds $\sqrt{n}(\widetilde{\alpha}\text{-}\alpha) \xrightarrow{d} N(0, \omega^2)$. The statement of the Delta Method says that if we transform an estimator by a function g, the following property holds:

$\sqrt{n}(g(\widetilde{\alpha})\text{-}g(\alpha)) \xrightarrow{d} N(0, \omega^2 g'(\alpha)^2)$, where g' denotes the first derivative of g.

This implies that the variance of the transformed estimator is given by:

$V[g(\widetilde{\alpha})] = V[\widetilde{\alpha}] \times g'(\widetilde{\alpha})^2$.

Transformation of two variables. To approximate the variance of some multi-variable function $G = G(\widetilde{\alpha}_x, \widetilde{\alpha}_y)$, we:

- take the vector of partial derivatives of the function G with respect to each parameter in turn : $\frac{\partial G}{\partial \widetilde{\alpha}_x}$ and $\frac{\partial G}{\partial \widetilde{\alpha}_y}$;
- right-multiply this vector by the variance-covariance matrix, $\Sigma = \begin{bmatrix} Var(\widetilde{\alpha}_x) & Cov(\widetilde{\alpha}_x, \widetilde{\alpha}_y) \\ Cov(\widetilde{\alpha}_x, \widetilde{\alpha}_y) & Var(\widetilde{\alpha}_y) \end{bmatrix}$
- right-multiply the resulting product by the transpose of the original vector of partial derivatives, G^T.

What we are interested in here is the standard error of the transformed variables, which equals the square root of the estimated variance. We apply the Delta Method transformation of one variable to obtain the standard error of the additional hectares of permanent grassland area and of the total benefits in Euro and the standard error of the percentage changes in grassland and money. We also use the Delta Method transformation of two variables to compute the standard errors of the elasticity estimates and the benefit-cost ratios, the standard error of the cost per additional hectare of grassland ratio and the cost per unit of averted CO_2 emission. We performed the computations in R using the "deltamethod" command from the "msm" package.

1.8.2 Data

Outcome variables:

- share of permanent grassland area (% of total utilised agricultural area) = share of natural grassland or pastures having more than 6 years on the same plot and low productivity grassland area;
- share of crop area (% of total utilised agricultural area) = share of cereals, industrial crops, pulses and protein crops;
- share of fodder area (% of total utilised agricultural area) = share of corn forage and silage, forage root crops and other annual forages;
- specialization rate (%) = the share of temporary and permanent grassland in the total utilised agricultural area;
- loading ratio = density of livestock units (cattle, equines, goats and sheep expressed in cattle units) in the forage area (permanent grassland and fodder area without corn forage);
- share of utilised agricultural area (% of total farm area) = share of annual crops, permanent crops and temporary and permanent grassland;
- share of forest area (% of total farm area) = share of timber and logging forests;
- share of nonproductive land (% of total farm area) = share of nonproductive heath, wasteland and non-agricultural area;
- share of owned land (% of total utilised agricultural area);
- share of permanently rented land (% of total utilised agricultural area);
- share of temporary rented land (% of total utilised agricultural area) = share of temporary rented land and land in sharecropping.

Control variables:

- type of crop orientation = cereals and protein crops, general crops, vegetable crops, flowers and horticulture, designated viticulture, other type of viticulture, fruits and other permanent crops, milk cattle, beef cattle, milk-beef cattle, other herbivorous, granivorous, mixed crops, poly-elevation herbivorous orientation, poly-elevation granivorous orientation, field crops and herbivorous;
- economic size = less than 4 ESU[28], between 4 and 8 ESU, between 8 and 16 ESU, between 16 and 40 ESU, between 40 and 100 ESU and more than 100 ESU ;
- number of farms = weighted number of farms.

[28] European Size Unit is a standard gross margin of 1200 Euro that is used to express the economic size of a farm (Eurostat:Statistics Explained).

1.8.3 Figures

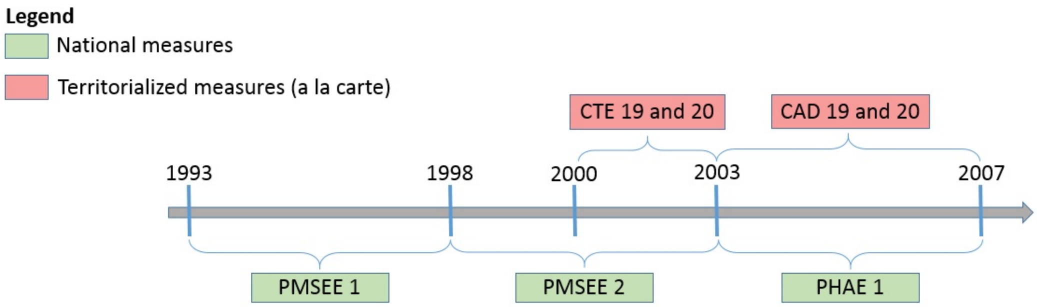


Figure 1.1: Timeline of the reforms of the French Grassland Conservation Program.

Eligibility Criteria \ Measure	PMSEE 2	PHAE 1	CTE/CAD 19 AND 20
Farmer's age	≤ 60 years	≤ 60 years	—
Farm size	≥ 3 ha UAA and ≥ 3 LU	—	—
Specialization Rate (Grassland/Utilised Agricultural Area)	≥ 75%	≥ 50% - ≥ 75% department dependent	—
Loading Ratio (Livestock Units/Fodder Area)	≤ 1.4	≤ 1.4 - ≤ 1.8 department dependent	≤ 1.4 - ≤ 1.8 department dependent
Fertiliser use (Units of Azote/ha of Grassland)	≤ 70	≤ 60	≤ 60
Max amount of subsidy / ha of grassland	**46€**	**76€**	**91€**

Figure 1.2: Eligibility rules of the French Grassland Conservation Program.

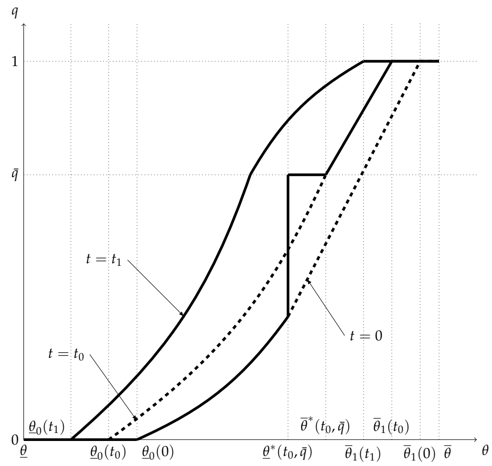


Figure 1.3: Theoretical model of grassland supply.

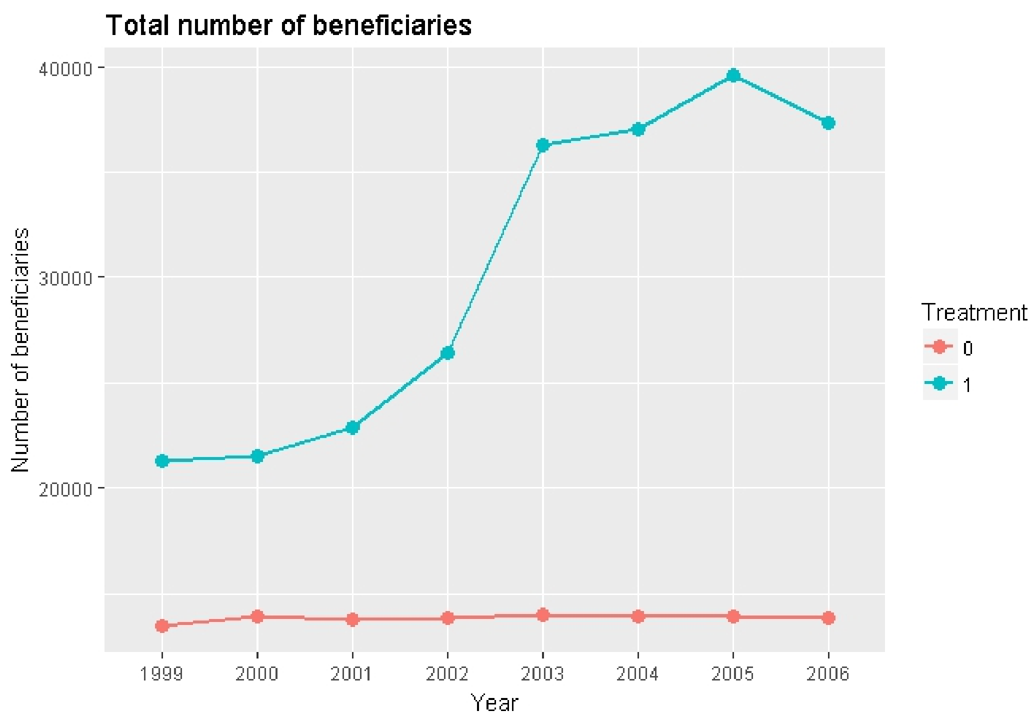


Figure 1.4: Total number of beneficiaries of grassland conservation schemes from 1999 to 2006, by treatment status.

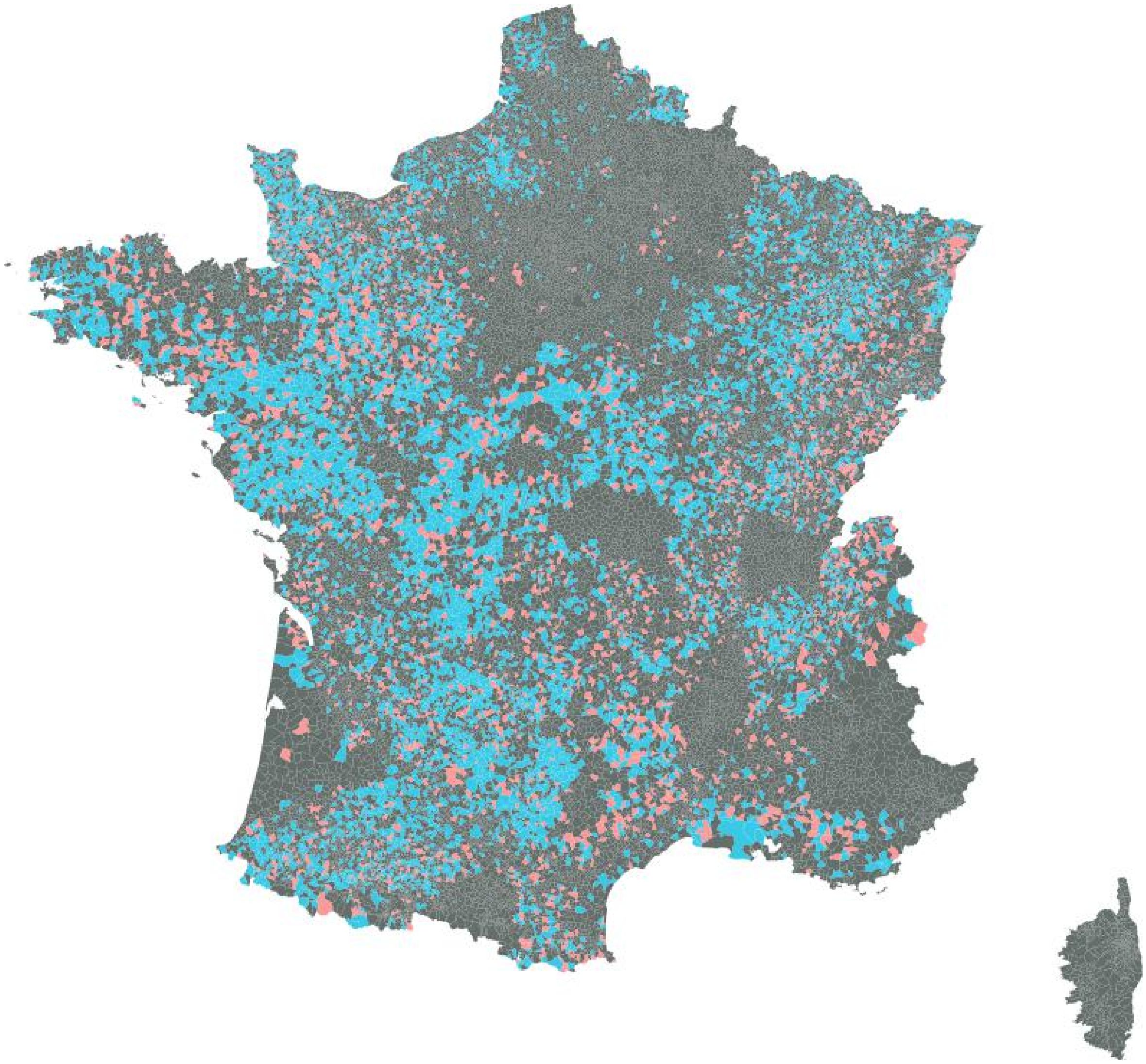


Figure 1.5: Map of France showing the treated communes (in blue) and the control communes (in pink).

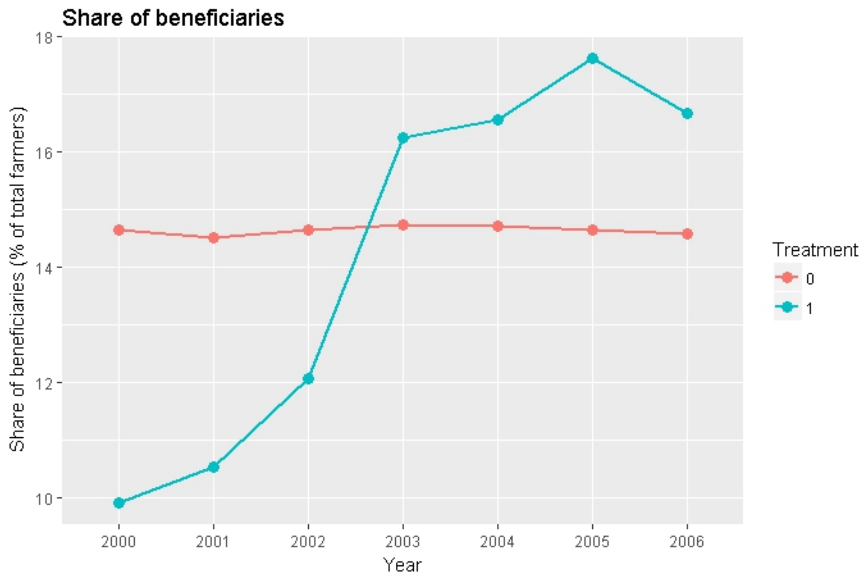


Figure 1.6: Share of beneficiaries of grassland conservation schemes in total farmers from 2000 to 2006, by treatment status.

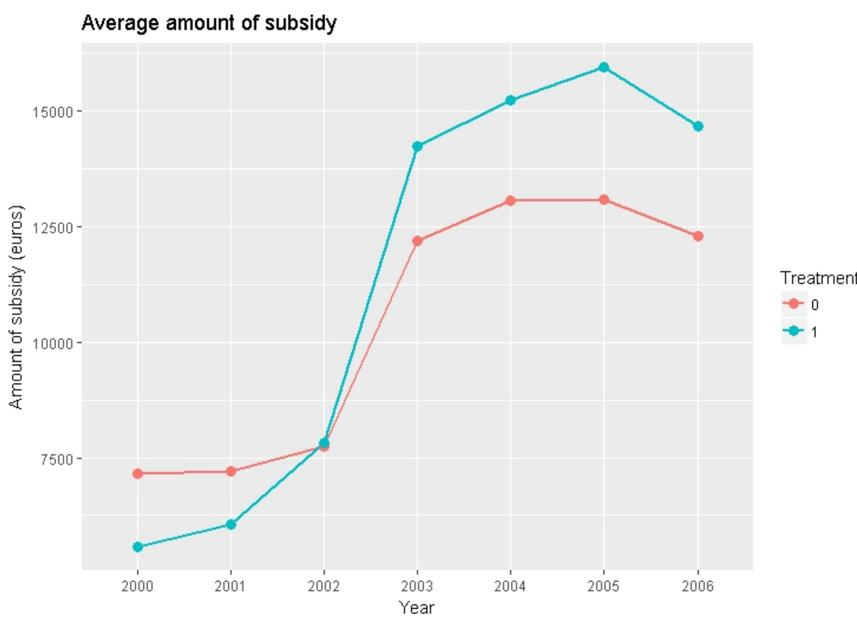
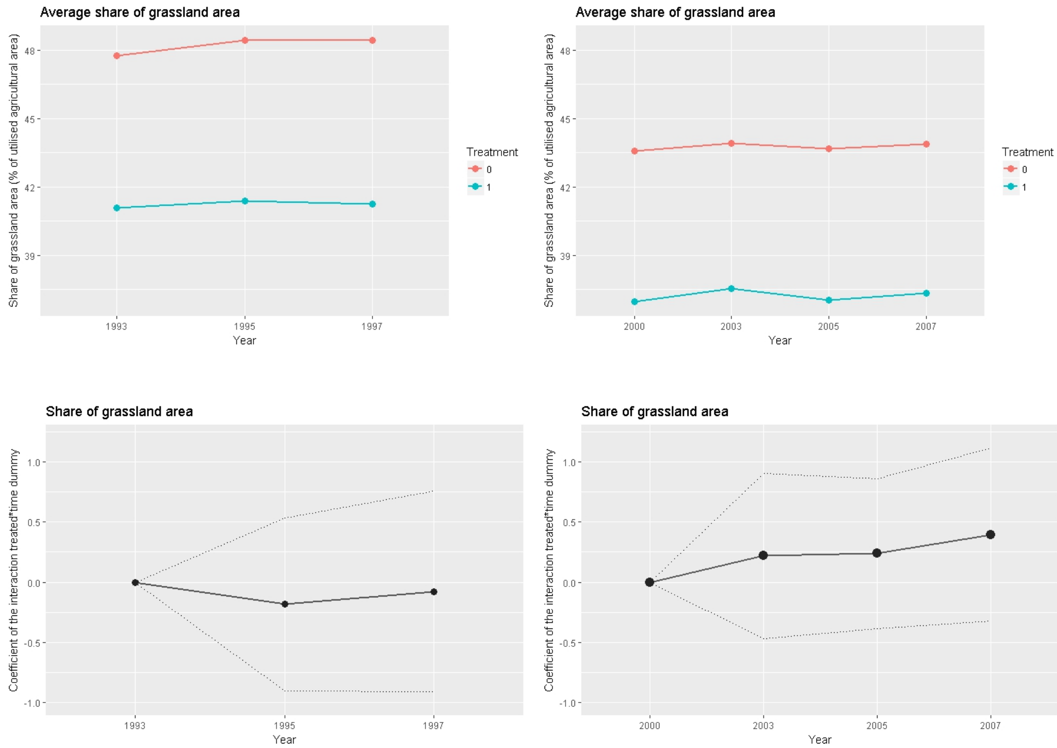


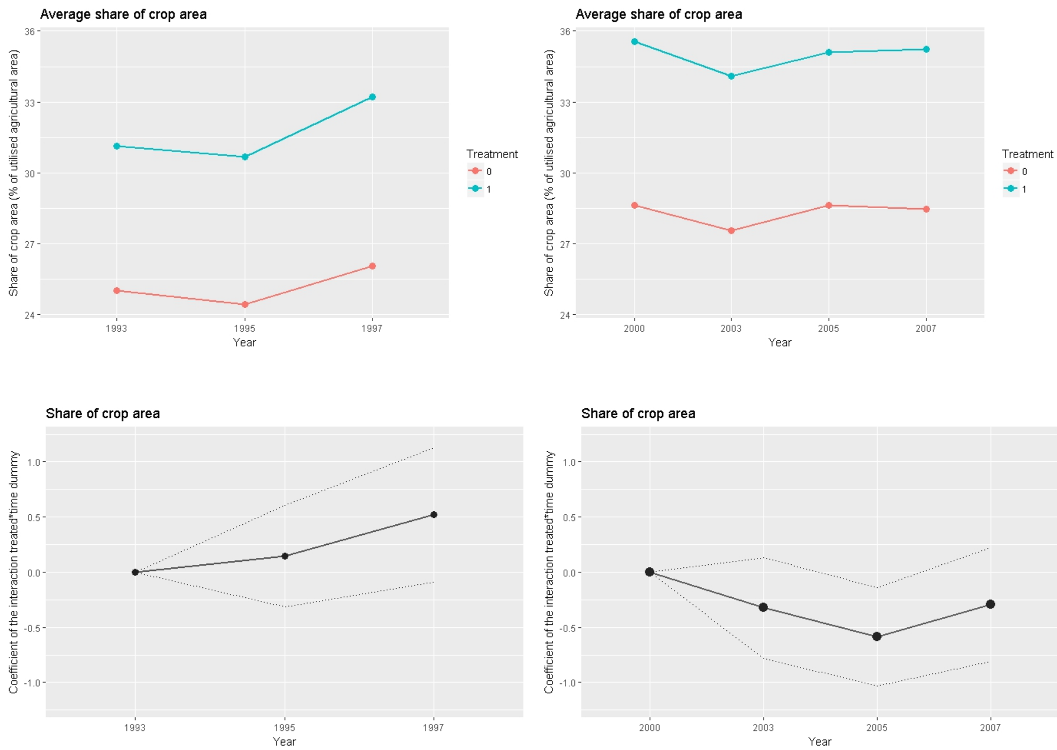
Figure 1.7: Average amount of subsidies (in Euro) paid to beneficiaries between 2000 and 2006, by treatment status.



(a) Placebo Test (1993-1997)

(b) Treatment Effect (2000-2007)

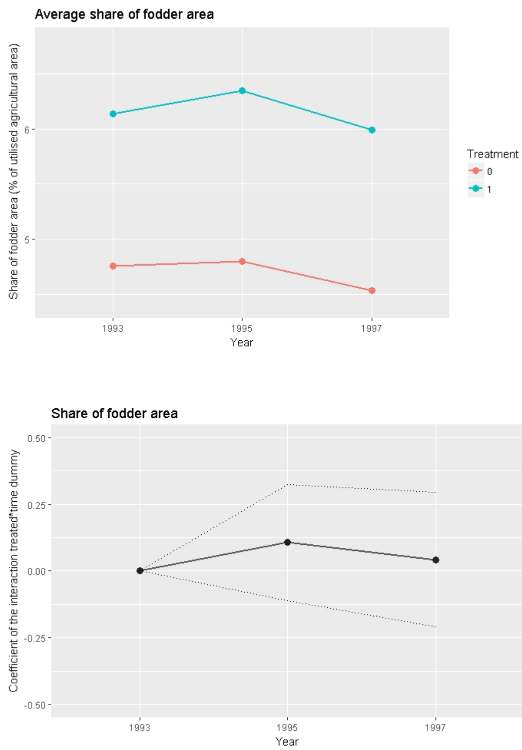
Figure 1.8: (i) Trends in the average share of permanent grassland area in total utilised agricultural area by treatment status and (ii) Estimated coefficients of the interaction treated*time dummy on the share of permanent grassland area and their 95% confidence interval (represented by dashed lines).



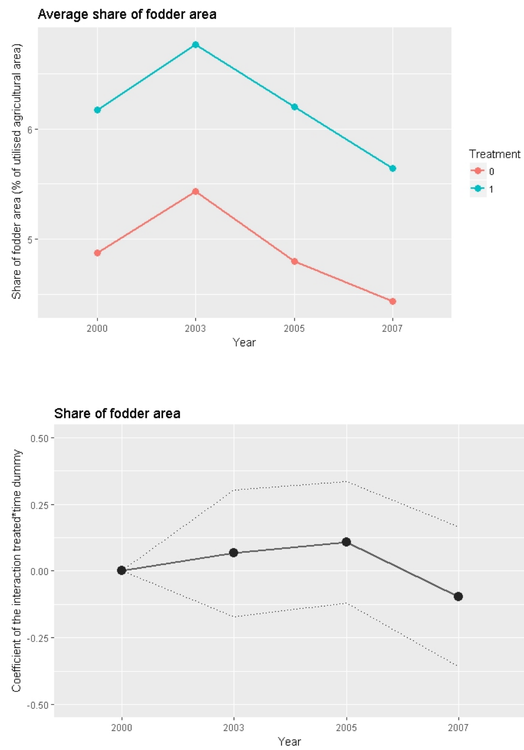
(a) Placebo Test (1993-1997)

(b) Treatment Effect (2000-2007)

Figure 1.9: (i) Trends in the average share of crop area in total utilised agricultural area by treatment status and (ii) Estimated coefficients of the interaction treated*time dummy on the share of crop area and their 95% confidence interval (represented by dashed lines).

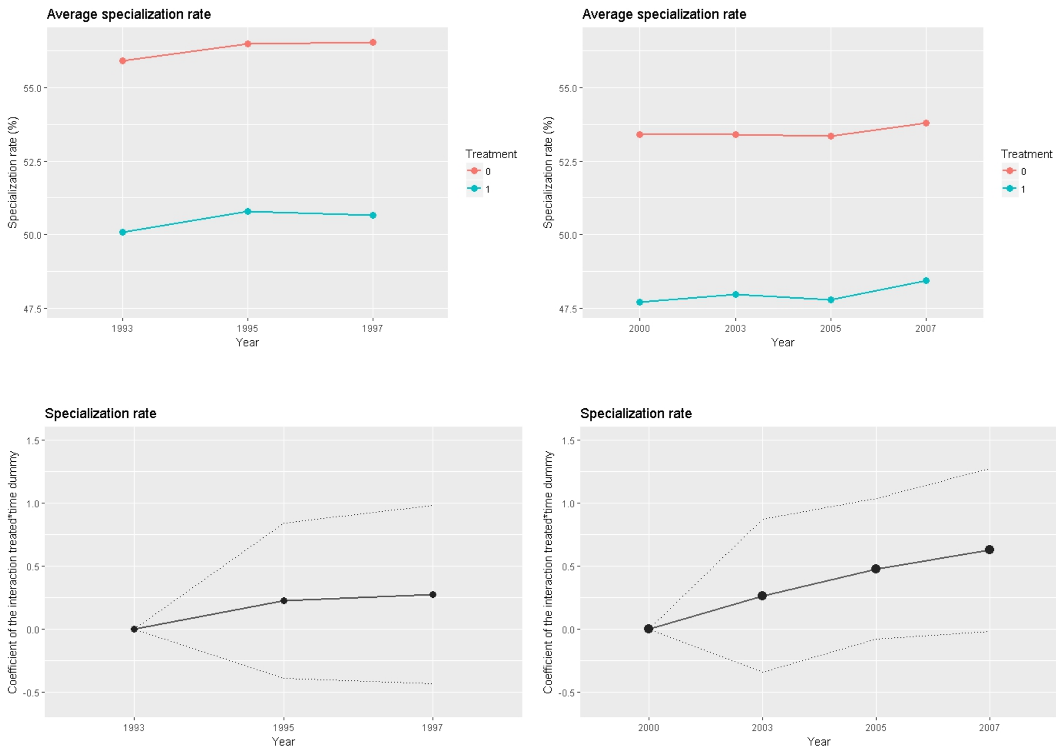


(a) Placebo Test (1993-1997)



(b) Treatment Effect (2000-2007)

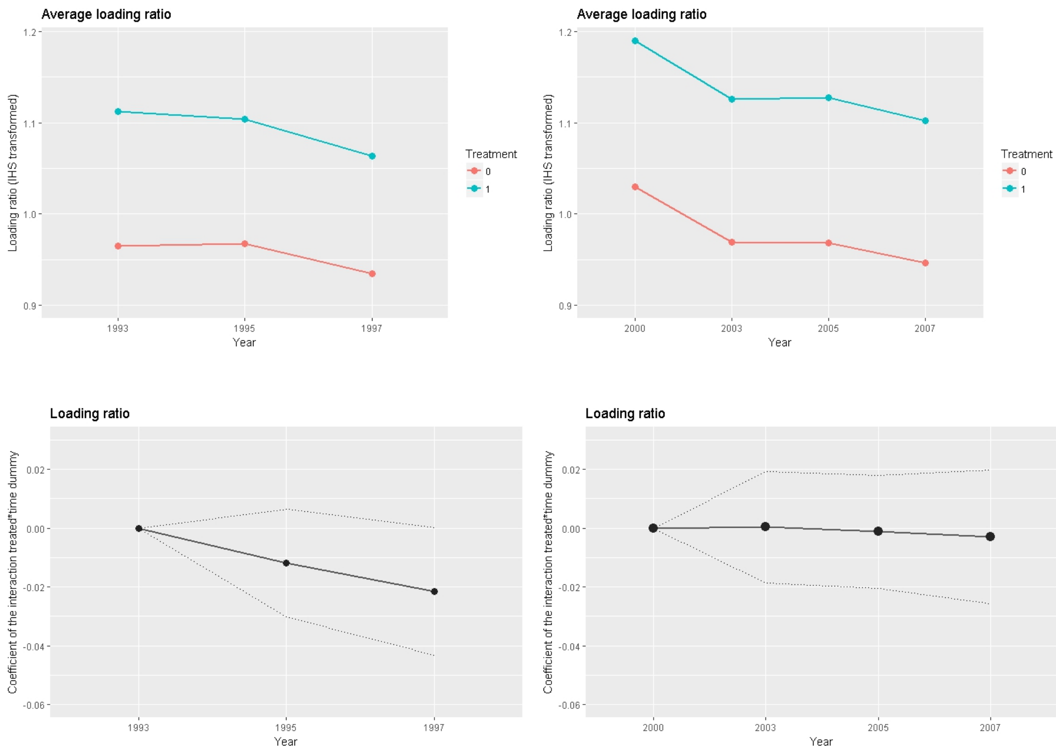
Figure 1.10: (i) Trends in the average share of fodder area in total utilised agricultural area by treatment status and (ii) Estimated coefficients of the interaction treated*time dummy on the share of fodder area and their 95% confidence interval (represented by dashed lines).



(a) Placebo Test (1993-1997)

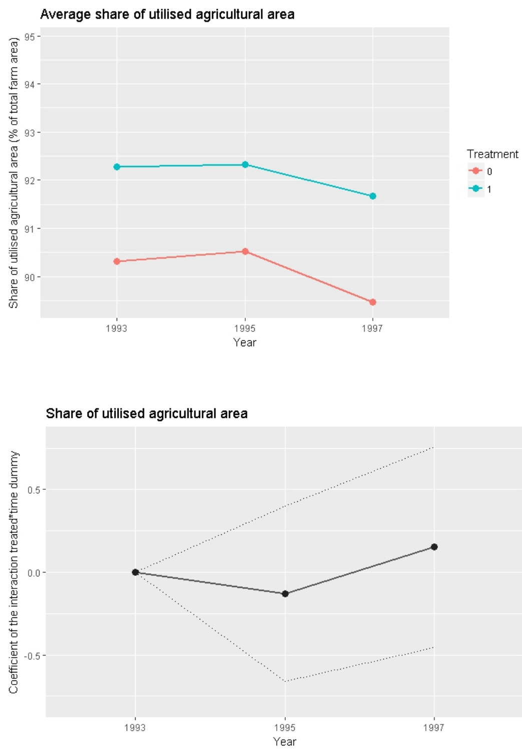
(b) Treatment Effect (2000-2007)

Figure 1.11: (i) Trends in the average specialization rate by treatment status and (ii) Estimated coefficients of the interaction treated*time dummy on the specialization rate and their 95% confidence interval (represented by dashed lines).

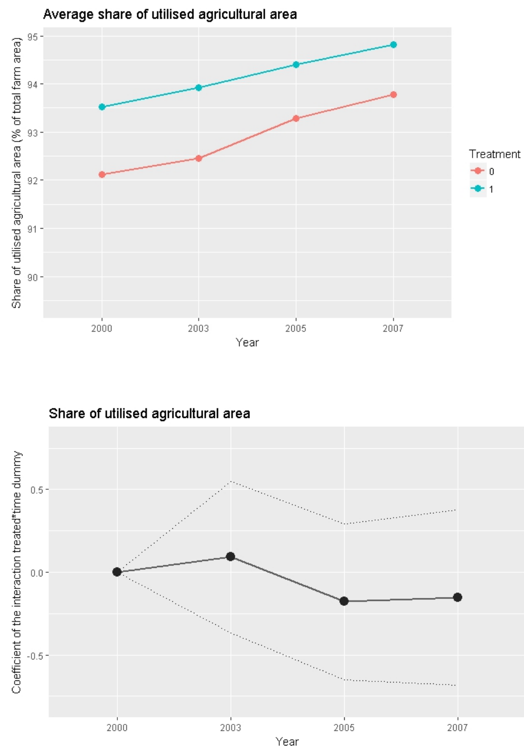


(a) Placebo Test (1993-1997) (b) Treatment Effect (2000-2007)

Figure 1.12: (i) Trends in the average loading ratio by treatment status and (ii) Estimated coefficients of the interaction treated*time dummy on the loading ratio and their 95% confidence interval (represented by dashed lines).

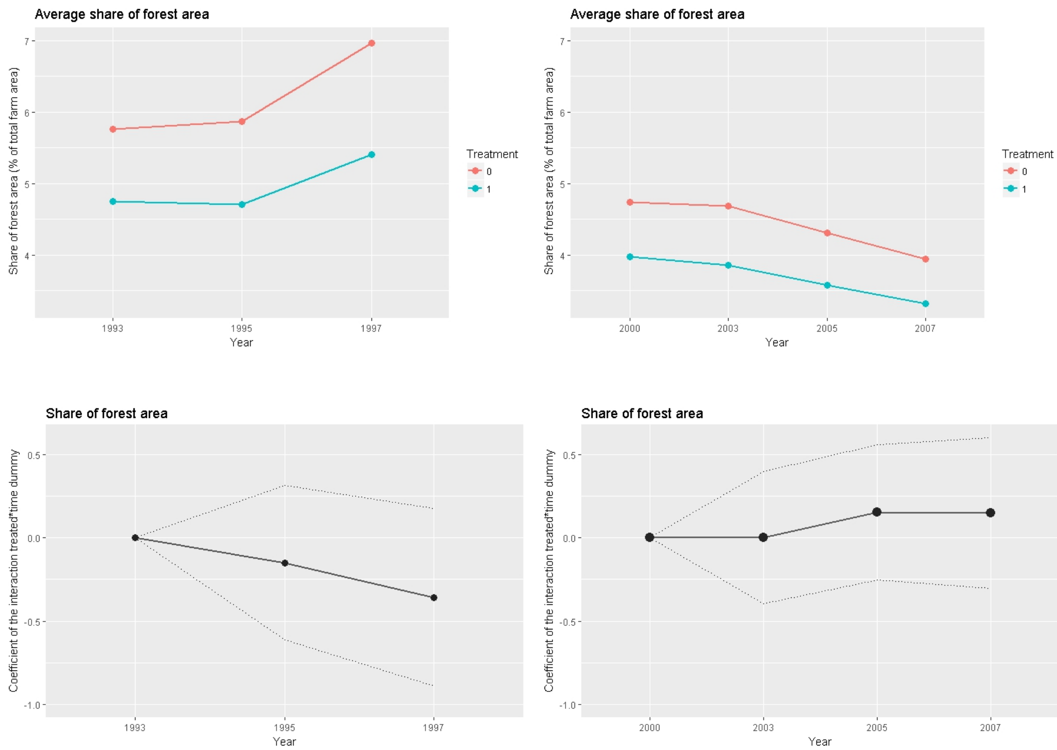


(a) Placebo Test (1993-1997)



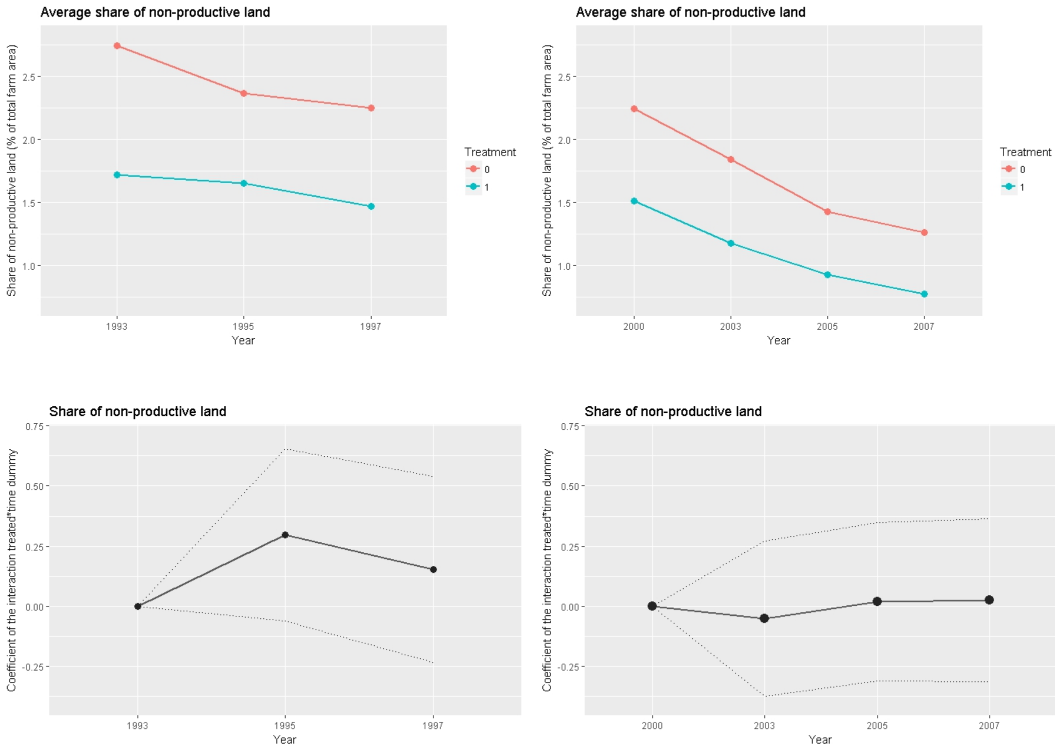
(b) Treatment Effect (2000-2007)

Figure 1.13: (i) Trends in the average share of utilised agricultural area in total farm area by treatment status and (ii) Estimated coefficients of the interaction treated*time dummy on the share of utilised agricultural area and their 95% confidence interval (represented by dashed lines).



(a) Placebo Test (1993-1997) (b) Treatment Effect (2000-2007)

Figure 1.14: (i) Trends in the average share of forest area in total farm area by treatment status and (ii) Estimated coefficients of the interaction treated*time dummy on the share of forest area and their 95% confidence interval (represented by dashed lines).



(a) Placebo Test (1993-1997)

(b) Treatment Effect (2000-2007)

Figure 1.15: (i) Trends in the average share of nonproductive land in total farm area by treatment status and (ii) Estimated coefficients of the interaction treated*time dummy on the share of nonproductive land and their 95% confidence interval (represented by dashed lines).

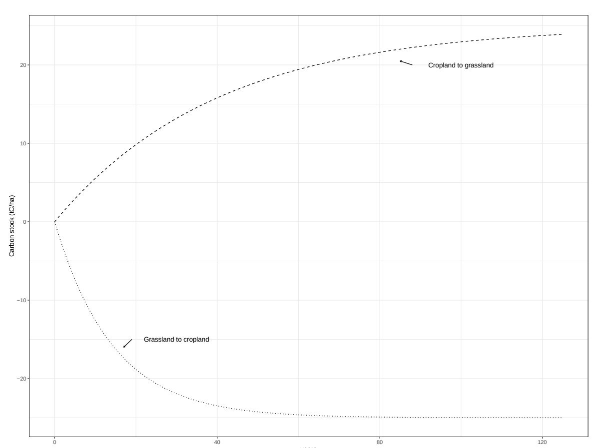
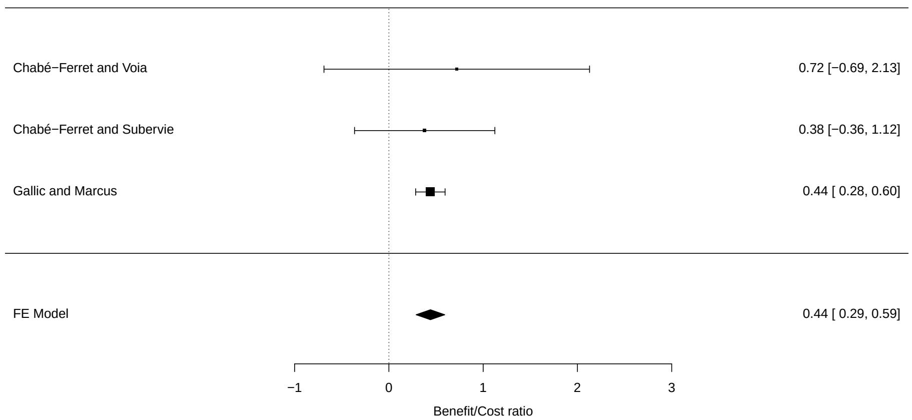
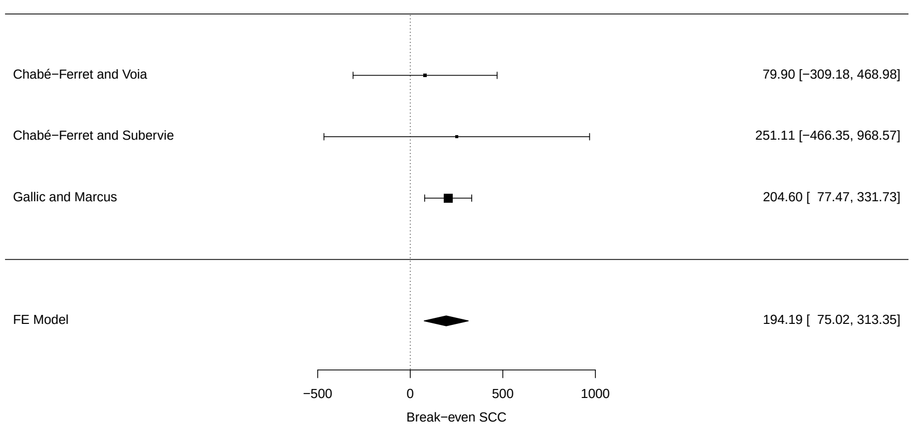


Figure 1.16: Evolution of the stock of Carbon in the soil when land use changes.



(a) Benefit-cost ratio



(b) Break-even SCC

Figure 1.17: Meta-analysis of the benefit-cost ratio and break-even SCC of the French Grassland Conservation Program.

1.8.4 Tables

Table 1.1: Mean and standard deviation of outcome variables, by treatment group and by sample

	1993-1997		**2000-2007**	
	Treated group	**Control group**	**Treated group**	**Control group**
Panel A				
Share of permanent grassland area	41.24	48.20	37.22	43.76
	(31.87)	(34.66)	(30.41)	(34.41)
Share of crop area	31.67	25.18	35.00	28.33
	(26.97)	(26.49)	(27.62)	(27.94)
Share of fodder area	6.15	4.69	6.19	4.89
	(8.63)	(8.01)	(7.96)	(7.81)
Specialization rate	50.52	56.32	47.97	53.49
	(31.97)	(34.32)	(31.35)	(34.60)
Loading ratio	1.68	1.42	1.73	1.47
	(3.07)	(2.76)	(4.41)	(2.96)
Panel B				
Share of utilised agricultural area	92.09	90.13	94.17	92.91
	(13.36)	(16.09)	(10.75)	(13.42)
Share of forest area	4.96	6.20	3.69	4.42
	(10.77)	(12.57)	(9.06)	(10.66)
Share of nonproductive land	1.61	2.45	1.10	1.69
	(6.22)	(8.42)	(4.32)	(6.85)
Observations	**6,827**	**3,171**	**7,243**	**3,225**

Table 1.2: DID-FE Results: Panel A

	Placebo Test (1993-1997)				**Treatment Effect (2000-2007)**			
	No DEPxTIME FE		With DEPxTIME FE		No DEPxTIME FE		With DEPxTIME FE	
	No controls	With controls	No controls	With controls	No controls	With controls	No controls	With controls
Outcome Variables								
Share of permanent grassland area	−0.44	−0.38	−0.17	−0.13	0.09	0.24	0.16	0.28
	(0.34)	(0.34)	(0.36)	(0.35)	(0.27)	(0.27)	(0.28)	(0.28)
Share of crop area	0.58	0.59	0.35	0.33	−0.33	−0.33	−0.38	−0.40
	(0.24)	(0.23)	(0.25)	(0.24)	(0.19)	(0.19)	(0.20)	(0.20)
Share of fodder area	0.12	0.11	0.08	0.07	0.01	−0.01	0.04	0.03
	(0.10)	(0.10)	(0.11)	(0.11)	(0.10)	(0.10)	(0.10)	(0.10)
Specialization rate	0.06	0.12	0.21	0.25	0.26	0.40	0.35	0.45
	(0.29)	(0.29)	(0.31)	(0.30)	(0.25)	(0.25)	(0.26)	(0.25)
Loading ratio	−0.02	−0.02	−0.02	−0.02	−0.00	−0.01	0.00	−0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	**9,998**	**9,998**	**9,998**	**9,998**	**10,468**	**10,468**	**10,468**	**10,468**

Note: Year and commune fixed effects estimation. Robust standard errors clustered at commune level in parenthesis. *p $<$ 0.1; **p $<$0.05; ***p $<$ 0.01.

Table 1.3: First Stage Results

	Treatment Effect (2000-2007)
Outcome Variables	
Share of beneficiaries (%)	10.71
	(0.18)
Total subsidies (Euro)	4,994.86
	(261.93)
Observations	**10,468**

Note: Year, commune and department-year fixed effects estimation. All regressions include the full set of control variables. Robust standard errors clustered at commune level in parenthesis.

Table 1.4: Testing the identification strategy

	(1)	(2)	(3)
Treated commune	0.075	0.103	0.078
	(0.007)	(0.007)	(0.007)
Constant	0.538	0.461	0.369
	(0.006)	(0.006)	(0.006)
Observations	10,468	10,435	10,075

Note: Estimates of the impact of the treatment at the commune level on the proportion of farmers ineligible to the program in 2000. Column (1) considers all farmers with a specialization rate below 75% in 2000 to be ineligible. Column (2) considers all farmers with a specialization rate below 75% and strictly positive in 2000 to be ineligible. Column (2) considers all farmers with a specialization rate below 75% and strictly positive in 2000 and with a loading ratio between 0.3 and 1.8 to be ineligible. Standard errors are in parenthesis.

Table 1.5: DID-FE Results: Panel B

	Placebo Test (1993-1997)				**Treatment Effect (2000-2007)**			
	No DEPxTIME FE		With DEPxTIME FE		No DEPxTIME FE		With DEPxTIME FE	
	No controls	With controls	No controls	With controls	No controls	With controls	No controls	With controls
Outcome Variables								
Share of utilised agricultural area	0.04	0.04	0.02	0.01	−0.19	−0.17	−0.06	−0.08
	(0.25)	(0.25)	(0.25)	(0.25)	(0.21)	(0.21)	(0.21)	(0.21)
Share of forest area	−0.34	−0.34	−0.25	−0.25	0.03	0.03	0.09	0.10
	(0.22)	(0.22)	(0.22)	(0.22)	(0.17)	(0.17)	(0.18)	(0.18)
Share of nonproductive land	0.28	0.27	0.22	0.23	0.18	0.16	−0.01	0.00
	(0.17)	(0.17)	(0.17)	(0.17)	(0.15)	(0.15)	(0.15)	(0.15)
Observations	**9,998**	**9,998**	**9,998**	**9,998**	**10,468**	**10,468**	**10,468**	**10,468**

Note: Year and commune fixed effects estimation. Robust standard errors clustered at commune level in parenthesis. *p < 0.1; **p <0.05; ***p < 0.01.

Table 1.6: CIC Results

	Placebo Test (1993-1997)	**Treatment Effect (2000-2007)**
Outcome Variables: Panel A		
Share of permanent grassland area	−0.12	0.28
	(0.34)	(0.32)
Share of crop area	0.29	−0.43
	(0.26)	(0.25)
Share of fodder area	0.15	0.00
	(0.12)	(0.13)
Specialization rate	0.23	0.46
	(0.28)	(0.30)
Loading ratio	−0.02	−0.01
	(0.01)	(0.01)
Outcome Variables: Panel B		
Share of utilised agricultural area	−0.09	−0.13
	(0.20)	(0.19)
Share of forest area	−0.32	0.04
	(0.18)	(0.16)
Share of nonproductive land	0.20	0.04
	(0.11)	(0.09)
Observations	**9,998**	**10,468**

Note: Changes-in-changes estimation. Regressions do not include fixed effects and control variables. Bootstrapped standard errors in parenthesis. *p < 0.1; **p <0.05; ***p < 0.01.

Table 1.7: DID-FE Results: Unbalanced Panel

	Placebo Test (1993-1997)				**Treatment Effect (2000-2007)**			
	No DEPxTIME FE		With DEPxTIME FE		No DEPxTIME FE		With DEPxTIME FE	
	No controls	With controls	No controls	With controls	No controls	With controls	No controls	With controls
Outcome Variables: Panel A								
Share of permanent grassland area	−0.35	−0.29	−0.08	−0.04	−0.02	0.14	0.08	0.21
	(0.35)	(0.34)	(0.36)	(0.35)	(0.27)	(0.27)	(0.28)	(0.28)
Share of crop area	0.56	0.57	0.33	0.31	−0.22	−0.23	−0.27	−0.30
	(0.24)	(0.23)	(0.25)	(0.24)	(0.20)	(0.19)	(0.21)	(0.20)
Share of fodder area	0.12	0.11	0.08	0.07	−0.01	−0.03	0.03	0.01
	(0.10)	(0.10)	(0.11)	(0.11)	(0.10)	(0.10)	(0.10)	(0.10)
Specialization rate	0.14	0.19	0.28	0.32	0.19	0.33	0.29	0.41
	(0.30)	(0.30)	(0.31)	(0.31)	(0.25)	(0.25)	(0.26)	(0.25)
Loading ratio	−0.02	−0.02	−0.02	−0.02	−0.01	−0.01	−0.00	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Outcome Variables: Panel B								
Share of utilised agricultural area	−0.10	−0.10	−0.13	−0.13	−0.23	−0.22	−0.06	−0.09
	(0.25)	(0.25)	(0.26)	(0.26)	(0.22)	(0.22)	(0.22)	(0.22)
Share of forest area	−0.28	−0.27	−0.22	−0.22	0.04	0.04	0.09	0.11
	(0.22)	(0.22)	(0.23)	(0.23)	(0.18)	(0.18)	(0.18)	(0.18)
Share of nonproductive land	0.29	0.29	0.26	0.27	0.20	0.19	−0.01	−0.01
	(0.17)	(0.17)	(0.17)	(0.17)	(0.16)	(0.16)	(0.16)	(0.16)
Observations	**10,599**	**10,599**	**10,599**	**10,599**	**11,463**	**11,463**	**11,463**	**11,463**

Note: Year and commune fixed effects estimation. Robust standard errors clustered at commune level in parenthesis. *p < 0.1; **p <0.05; ***p < 0.01.

Table 1.8: DID-FE Results: Same Sample of Communes

	Placebo Test (1993-1997)				**Treatment Effect (2000-2007)**			
	No DEPxTIME FE		With DEPxTIME FE		No DEPxTIME FE		With DEPxTIME FE	
	No controls	With controls	No controls	With controls	No controls	With controls	No controls	With controls
Outcome Variables: Panel A								
Share of permanent grassland area	0.03	0.06	0.18	0.19	0.25	0.38	0.31	0.44
	(0.39)	(0.38)	(0.40)	(0.39)	(0.32)	(0.31)	(0.32)	(0.31)
Share of crop area	0.45	0.48	0.22	0.21	−0.32	−0.31	−0.36	−0.38
	(0.28)	(0.27)	(0.29)	(0.28)	(0.23)	(0.22)	(0.24)	(0.23)
Share of fodder area	0.11	0.10	0.08	0.07	0.09	0.06	0.13	0.11
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Specialization rate	0.50	0.54	0.58	0.60	0.18	0.29	0.26	0.36
	(0.34)	(0.33)	(0.34)	(0.34)	(0.30)	(0.29)	(0.30)	(0.30)
Loading ratio	−0.01	−0.01	−0.01	−0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Outcome Variables: Panel B								
Share of utilised agricultural area	0.15	0.16	0.07	0.07	−0.31	−0.30	−0.23	−0.24
	(0.27)	(0.27)	(0.28)	(0.28)	(0.23)	(0.23)	(0.23)	(0.23)
Share of forest area	−0.19	−0.20	−0.09	−0.10	0.07	0.06	0.13	0.14
	(0.24)	(0.24)	(0.25)	(0.25)	(0.20)	(0.20)	(0.20)	(0.20)
Share of nonproductive land	−0.01	0.00	−0.04	−0.03	0.32	0.31	0.18	0.19
	(0.18)	(0.18)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
Observations	**7,808**	**7,808**	**7,808**	**7,808**	**7,808**	**7,808**	**7,808**	**7,808**

Note: Year and commune fixed effects estimation. Robust standard errors clustered at commune level in parenthesis. *p < 0.1; **p <0.05; ***p < 0.01.

Table 1.9: Elasticity Estimate

Outcome	**ITT estimate**	**% change**	**Elasticity**
Additional hectares of grassland	3.73±7.31	0.76%±1.49%	**0.02±0.04**
Additional monetary transfers (in Euro)	5,000±513	42.46%±6.21%	

Note: Estimate of the elasticity of the additional supply of grassland with respect to the additional amount of the subsidy per treated communes as a result of the French Grassland Conservation reform in 2000. The confidence interval around the estimated values is given by the formula: point estimate ± value from the standard normal distribution for the selected confidence level (i.e. 1.96) x standard error of the point estimate (computed using the Delta Method).

Table 1.10: Cost-Benefit Analysis

Study	Benefits			Costs	Benefit-Cost Ratio	Break-even SCC
	ITT estimate (ha)	Benefits per ha (Euro/ha)	Total (Euro)	ITT estimate (Euro)		(Euro/tCO_2eq)
Chabé-Ferret and Voia						
Climate benefits only	3.73±7.31	162.54±45.51	606±1,200	5,000±513	**0.12±0.24**	198±392
All benefits	3.73±7.31	961.90±45.51	3,588±7,033	5,000±513	**0.72±1.41**	80±389
Chabé-Ferret and Subervie						
Climate benefits only	1.4±2.7	162.54±45.51	228±444	3,500±513	**0.07±0.13**	369±721
All benefits	1.4±2.7	961.90±45.51	1,347±2,598	3,500±513	**0.38±0.74**	251±717
Gallic and Marcus						
Climate benefits only	1.2±0.35	162.54±45.51	195±79	2,622±513	**0.07±0.03**	323±145
All benefits	1.2±0.35	961.90±45.51	1,154±341	2,622±513	**0.44±0.16**	205±127

Note: The costs of the Grassland Conservation Program reform compared with the social benefits at commune level. The confidence interval around the estimated values is given by the formula: point estimate ± value from the standard normal distribution for the selected confidence level (i.e. 1.96) x standard error of the point estimate (computed using the Delta Method). The literature estimates come from Arrouay et al. (2002) for the climate benefits and from Puydarrieux and Devaux (2013) for the other ecosystem services.

Chapter 2

Are Forest Conservation Programs a Cost-Effective Way to Fight Climate Change? A Meta-Analysis

Sylvain Chabé-Ferret and Anca Voia

Abstract

Deforestation is a major contributor to the emission of greenhouse gases. Forest Conservation Programs that pay landowners for maintaining forest cover might thus be an effective way to fight climate change as long as the benefits from avoided emissions exceed the cost of triggering the conservation of additional forest cover. In this paper, we use meta-analysis tools to estimate the benefit-cost ratio of Forest Conservation Programs implemented in developing countries. We combine 18 separate estimates of the additional forest cover conserved thanks to these programs with estimates of emissions from deforestation. We find that Forest Conservation Programs reduce the annual deforestation rate by 0.23±0.14 percentage points on average and thus provide climate benefits. Our results suggest that the value of the climate benefits of Forest Conservation Programs crucially depends on the permanence of their effects after the program stops. For a Social Cost of Carbon of 31 USD/tCO_2eq, we estimate that benefits are equal to 45±32% of the program costs if the impact of the program on deforestation stops just after the program ends, 78±56% of the program costs if the impact decreases progressively over 10 years, and 263±194% of the program costs if the impact persists forever. We estimate that Forest Conservation Programs would become cost-effective with a Social Cost of Carbon of 100 USD/tCO_2eq, even with no permanence. We find ample evidence of publication bias in the estimates of the impact of Forest Conservation Programs on deforestation, with certain of the most recent estimates over-estimating the impact of these programs by a factor of 10, artificially inflating their benefit-cost ratio above one.

2.1 Introduction

Deforestation, defined as land-use change from forest to other land uses, is the second-major source of climate change, after burning fossil fuels, and it accounts for nearly 20% of the total emissions of greenhouse gases. Therefore, in order to limit the increase in temperature at below 2^{o}C before 2030, reducing emissions from the forest sector needs to play an important role in climate policies. Forest Conservation Programs that pay landowners for maintaining forest cover might thus be an effective way to fight climate change as long as the benefits from avoided emissions exceed the cost of triggering the conservation of additional forest cover (Chabé-Ferret et al., 2015). Forest Conservation Programs got a big boost from the 2015 Paris Climate Agreement, which for the first time recognised forests as a key part of the solution to climate change mitigation. As of May 2018, there were 426 REDD+ (Reducing Emissions from Deforestation and forest Degradation) projects and programs in 57 developing countries (Simonet et al. 2018). These programs offer financial rewards to developing countries for reducing emission from forested land. In addition to REDD+, other payments for ecosystem services (PES) targeted at forest conservation are increasingly implemented at national and sub-national levels.

In this paper, we use meta-analysis tools to estimate the benefit-cost ratio of Forest Conservation Programs implemented in developing countries. We combine 18 separate estimates of deforestation avoided thanks to Forest Conservation Programs obtained using either randomized controlled trials or observational methods with estimates of emissions from deforestation. We test and account for publication bias, that might occur when some studies are missing from the published record. Finally, we compute the discounted benefits of the programs in terms of avoided emissions and compare them to the program costs.

We find that Forest Conservation Programs reduce the annual deforestation rate by 0.23$\pm$0.14 percentage points on average. We also find strong evidence of publication bias in the literature. We estimate that publication bias can explain why some recent influential studies have found impacts of Forest Conservation Programs 10 times larger than the ones we find.

We find that the value of the climate benefits of Forest Conservation Programs crucially depends on the permanence of their effects after the program stops. For a Social Cost of Carbon of 31 USD/tCO_2eq, we estimate that benefits are equal to 45$\pm$32% of the program costs if the impact of the program on deforestation stops just after the program ends, while they are equal to 78$\pm$56% of the program costs if the impact decreases progressively over 10 years, and 263$\pm$194% of the program costs if the impact persists forever. We estimate that

Forest Conservation Programs would become cost-effective with a Social Cost of Carbon of 100 USD/tCO_2eq, even with very low permanence. We believe further research is needed in order to estimate the permanence of the effects of Forest Conservation Programs. We urge researchers, funders and policy-makers to commit to publishing all the results from these future studies, whatever the statistical significance of their results. Doing so will be critical to ensure that we obtain unbiased estimates of the benefits of Forest Conservation Programs.

Our paper contributes to several streams of literature. We first contribute to the literature on the evaluation of the effects of Forest Conservation Programs. Several qualitative literature reviews of this research question exist (Pattanayak et al., 2010, Alix-Garcia and Wolff, 2014) and find that forest conservation programs yield mixed results in terms of avoided deforestation. Two meta-analysis of the effects of Forest Conservation Programs have been conducted, one by Samii et al. (2014) and the other by Snilsveit et al. (2019). Both find results very similar to ours in terms of the decrease in deforestation due to Forest Conservation Programs: 0.21$\pm$0.18 p.p. for Samii et al. (2014) and 0.12$\pm$0.07 standardized mean difference for Snilsveit et al. (2019). Both papers do not perform an extensive benefit-cost analysis nor perform tests and corrections for publication bias. Recent influential results find impacts larger by one order of magnitude for programs in Brazil (Simonet et al., 2018) and in Uganda (Jayachandran et al., 2017), and thus find benefit-cost ratios larger than one. Our findings suggest that these results are inflated by publication bias. We also contribute to the literature on the consequences of publication bias for estimates of crucial parameters in economics. Ioannidis et al. (2017) estimate the extent of publication bias in 159 areas of research using similar meta-analysis tools as the one we are using here and find that 80% of the reported effects are exaggerated by a factor of two, with one third inflated by a factor of four or more. Andrews and Kasy (2019) find evidence of publication bias in experimental economics and in the minimum wage literature. Using a similar method, Hendren and Sprung-Keyser (2020) detect publication bias in the literature on the impacts of social programs. Gechert et al. (2021) find that the elasticity of substitution between capital and labor is overestimated by a factor of three in the published literature, with half of this overestimate stemming from publication bias. Nemati and Penn (2020) find major signs of publication bias in the literature on conservation nudges. DellaVigna and Linos (2021) find that the impact of nudges in the published literature is six times larger than in subsequent scale ups by government agency, and that 77% of that overestimation can be explained by publication bias. Jackson and Mackevicius (2021) find signs of publication bias in the literature on the impact of school spending, but, as in our case, they find that it does not affect the meta-analytic estimate substantially.

The rest of the paper is structured as follows: Section 2.2 describes what Forest Conservation Programs are and their scope; Section 2.3 presents our methodological approach; Section 2.4 presents our results and Section 2.5 concludes.

2.2 Forest Conservation Programs

Deforestation is the second-major source of climate change, accounting for nearly 20% of the total emissions of greenhouse gases. According to FAO (2015), carbon emissions from deforestation decreased by about 25% over the period 2001-2015, from 3.9 to 2.9 gigatons of carbon dioxide (CO2) per year. However, the most recent data for 2017 suggests that average annual emissions from 2015 to 2017 were 63 percent higher than the average in the preceding 14 years, rising at 4.9 gigatons per year (Wolosin and Harris, 2018). In addition to carbon sequestration, forests provide locally many other environmental services, such as biodiversity, watershed protection and climate regulation.

The global perception is that reducing emissions from deforestation, reforestation or forest restoration may be a less expensive option compared to other climate change mitigation policies, such as hybrid and electric car subsidies. The reason why this may be the case is that developing countries have the highest deforestation rates (Alix-Garcia and Wolff, 2014). The possibility of low-cost and potentially effective carbon sequestration options in developing countries led to the development of programs aiming at Reducing Emissions from Deforestation and forest Degradation (REDD+). REDD+, introduced in 2007 by Parties to the United Nations Framework Convention on Climate Change, is a mechanism that offers financial incentives to developing countries for reducing emissions from forested lands. Beyond deforestation and forest degradation, REDD+ includes the role of conservation, sustainable management of forests and enhancement of forest carbon stocks (UN-REDD, About REDD+). As of May 2018, there were 426 REDD+ projects and programs in 57 developing countries (Simonet et al. 2018). REDD+ can be seen as an international Payment for Ecosystem Services (PES), as governments receive payments conditionally on reducing forest-related emissions (Pattanayak et al., 2010). The REDD+ programs are financed through bilateral and multilateral agreements between countries and institutions. Simula (2010) estimated that 7 billion dollars were committed between 2008 and 2010. Multi-donor trust funds represent around 40% of the total, the most important being UN-REDD (318 million dollars), the Amazon Fund (721 million dollars), the Forest Carbon Partnership Facility (1.3 billion dollars) and the Forest Investment Program

(570 million dollars). In addition to REDD+, a growing number of PES programs targeting forest conservation are implemented at national and sub-national level in developing countries. One of the first national-level forest PES was introduced in Costa Rica in 1997. According to Ecosystem Marketplace, by 2019 there were 312 active avoided forest conversion payment schemes.

In general, the Forest Conservation Programs are designed as PES-like programs. Wunder (2005) defines PES as a voluntary transaction between at least one buyer and one seller in which a payment is given conditionally on the provision of a well-defined environmental service (or a land use likely to secure that service). Specifically, Forest Conservation Programs are 5 to 20-years voluntary contracts between public or private buyers (e.g. national governments and international organizations) and individual landowners or small communities in which a yearly payment is given conditionally on maintaining all or a part of the forested area owned. In general, the environmental services associated with forest conservation can be local or national, such as hydrological services and erosion control or global such as carbon sequestration and biodiversity conservation.

Forest Conservation Programs usually pay landowners per hectare of forest conserved. The payment should cover their opportunity cost of leaving trees standing and eventually their effort in protecting the forest against external threats. In our sample, the payment ranges from 2 USD/ha (collective forest conservation programs in Cambodia) to 43 USD/ha (Costa Rica's national program). Wunder (2008) shows that landowners will chose to participate in such a program if their opportunity cost is lower that the payment received, if they have sufficient capacity for protecting the forest, if they have property rights on eligible forested areas and if they trust the purchaser. Some of these features favor the poor, while others favor the rich. Therefore, the characteristics of participants versus nonparticipants in forest-PES programs largely depend on the context. For example, in Costa Rica participants tend to be richer, while in Mexico participants are found in areas with higher levels of poverty; in Uganda, there is more take-up from households that are credit-constrained, while in Cambodia participants need more capital assets to be able to participate.

In practice, due to information asymmetries on the opportunity cost of landowners, the payments tend to overcompensate the participants. In this case all or part of the conservation subsidy becomes a net income gain for the participant and the cost-effectiveness of the program gets diminished. The same happens when landowners put under PES forest land that they wouldn't have deforested even in the absence of payments. However, households who own land with low risk of deforestation tend to be poor. Thus, the net income transfer may help alleviate poverty. Targeting environmental effectiveness versus poverty alleviation is a trade-off that government-coordinated forest-PES programs have to deal with. Indeed, Alix-Garcia et al. (2012) found that larger avoided deforestation happened in less-poor areas. The little empirical evidence that exists on the interaction of the environmental and poverty objectives suggests that forest PES should be considered a ”win-neutral” rather than a ”win-win” strategy (Alix-Garcia et al., 2015).

2.3 Methods

In this section, we present our methodological approach. We start by delineating a simple model of deforestation and carbon emissions that enables us to define additionality, climate benefits and the benefit-cost ratio of Forest Conservation Programs. We then explain how we searched the literature for relevant papers and which data we extracted from the papers we ended up selecting. Finally, we present our econometric methodology.

2.3.1 A simple model of deforestation and carbon emissions

To be able to derive estimates of the additionality of Forest Conservation Programs and their benefit-cost ratios, we delineate a simple model of deforestation and carbon emissions. We start at date $t = 0$ with one plot of forested land of unit size (for example one hectare). We assume that, in the absence of the Forest Conservation Program, the plot is going to be deforested at a constant yearly rate d_0. In the absence of the program, forest cover thus decreases over time following a geometric process:

$$F_0(t) = (1 - d_0)^t. \tag{2.1}$$

We assume that a Forest Conservation Program starts at $t = 0$ and lasts K years. The program reduces the deforestation rate by θ percentage points (p.p.) for all the duration of the program: $d_1(t) = d_0 - \theta$, for $t \leq K$. We generally do not observe what happens after the program stops. We thus have to resort to various assumptions in order to bound the possible impacts of the program. We simulate three possibilities for the deforestation rate after the program stops. As a lower bound on the impacts of the program, we assume that the deforestation rate returns to d_0 as soon as the program stops. In that case, $d_1^l(t) = d_0$ for $t > K$. As an upper bound on the impacts of the program, we assume that the deforestation rate remains lower by θ p.p. forever: $d_1^u(t) = d_0 - \theta$ for $t > K$. Finally, as an intermediary assumption, we posit that the deforestation rate catches up with d_0 κ years after the end of the program, at a linear rate. In that case, $d_1^r(t) = d_0 - \theta + (t-K)\frac{\theta}{\kappa}$ when $K < t \leq K+\kappa$ and $d_1^r(t) = d_0$ for $t > K+\kappa$. Figure 2.1a plots the evolution of forest cover under these various assumptions. In what follows, we denote forest cover in the presence of the program as $F_1(t)$, except when we want to make one of these assumptions more salient.

To derive the benefits of the program, we need to transform the changes in forest cover into carbon emissions. We first assume that each forest stores G tons of equivalent CO_2 above ground.[1] Carbon storage thus evolves as $G_d(t) = GF_d(t)$ for $d \in \{0,1\}$. We also assume that carbon from deforested areas reaches the atmosphere after a delay of T years because decomposition processes transforming organic carbon in CO_2 take time. Emissions are thus defined as $E_d(t) = G_d(t-T-1) - G_d(t-T)$, for $t > T$. We also posit that emissions are zero for $t < T$. Figure 2.1b shows the path of emissions stemming from the model under these various assumptions under reasonable parametrizations.

We define the social benefits of a Forest Conservation Program as follows:

$$B = \sum_{t=0}^{\infty} \beta^t SCC_t(E_1(t) - E_0(t)), \tag{2.2}$$

with β the discount rate and SCC_t the Social Cost of Carbon. In practice, we use a constant estimate of the Social Cost of Carbon of 31 USD/tCO_2eq. We approximate the infinite stream of benefits by a finite sum over a sufficient amount of time so that estimates do not change by adding more periods.[2]

To compute the benefit-cost ratio of Forest Conservation Programs, we need in addition an estimate of the program true costs. In cost-benefit analysis, the true costs of a program are not the transfers received by the participants but the deadweight loss generated by the taxation needed to generate these transfers and the change in profits due to the participation constraints. In practice, both the deadweight loss from taxation and the profit loss from the constraints of the program are difficult to observe. We know that the profit loss from the constraints of the program has to be smaller than the transfers received by the participants, otherwise they would not enter the program. We make the assumption that the true costs of the program (both deadweight loss and profit loss) are well approximated by the transfers received by the participants, C. Even though this assumption seems reasonable, it clearly requires further research. Another interpretation of our approach is that of a decision-maker that has committed to spend C and who is trying to select the program that will deliver the larger bang for her buck.

[1] We do not consider for now the role of carbon storage below ground.

[2] In current work, we are trying to derive closed form solutions for the benefits of the program so that they can be computed more easily.

As a conclusion to this section, here are the parameters we need in order to be able to derive a benefit-cost ratio for Forest Conservation Programs. We first need an estimate of the impact of Forest Conservation Programs on the deforestation rate, in p.p.: θ. We then need an estimate of the counterfactual deforestation rate, d_0, in order to be able to compute counterfactual emissions. We also need an estimate of program duration K, of the transfers received by participants C and of the stock of carbon in the standing forest G. We extract these estimates from a systematic search of the literature evaluating the impact of Forest Conservation Programs.

2.3.2 Literature search and data extraction

In order to identify the relevant studies to be included in this meta-analysis, we first did a hand search of cited papers in the previous reviews of the literature. We then searched Web of Science and Econlit online databases using key words related (i) to payments for ecosystem services (e.g. ”payments for ecosystem services” OR ”payments for environmental services”), (ii) to the outcome of interest (e.g. ”forest” OR ”deforestation”) and (iii) to the program evaluation method employed (e.g. ”differences-in-differences” OR ”matching” OR ”fixed effects” OR ”first difference” OR ”regression discontinuity”, ”instrumental variable”, etc.). This search, done in May-June 2018, resulted in a list of 421 published and 4 unpublished articles, all written in English language. After title and abstract screening, we were left with 32 studies relevant to our topic, among which we retained 18 for the analysis. The PRISMA flow diagram in Figure 2.2 details the identification and selection process.

We used two inclusion criteria in the selection process. First, we included only articles that were evaluating a Forest Conservation PES-like Program on forest-related outcomes (e.g. deforestation rate, forest cover, NDVI index). Second, we selected only studies that reported a quantitative treatment effect causally identified thought robust program evaluation techniques. The complex characteristics of forest-PES programs make it difficult to evaluate their environmental effectiveness from simple before-after or with-without comparisons. Thus, researchers were pushed to adopt counterfactual-based evaluation techniques that explicitly compare treated and control areas. When program enrollment is randomized, a simple comparison between the two groups is enough to reliably estimate the program's impact. However, in most cases, the enrollment is not randomized, but rather based on different administrative and political criteria. Together with the voluntary participation aspect, these two characteristics lead to systematic differences between participants and nonparticipants. Therefore, robust program evaluation techniques have to be used to correct for the selection bias that arises in these cases (Borner et al., 2017). The most common such quasi-experimental methods are: matching, differences-in-differences, regression discontinuity designs and fixed effects regressions.

We excluded one relevant study because we were unable (even after contacting the authors) to retrieve the standard errors associated with the estimated treatment effect. Thus, after the screening process we were left with 18 independent studies, 15 published in peer-reviewed journals and 3 still working papers. The list of studies included in this meta-analysis and their characteristics is given in Table 2.1. These studies represent 10 forest conservation programs in 7 developing countries. Figure 2.3 shows the countries present in our sample. Latin American countries are the most represented (5 out of 7), while there are only two studies in Asia and one in Africa. There are only two papers that use randomized control trials (RCTs) to evaluate the PES program, while all the other studies use some form of matching, half of them combining matching with a difference-in-differences design. In general, the matching technique used is a one-to-one nearest-neighbor matching, either using the Mahalanobis distance or the Propensity Score to define "closeness" and applying the bias-correction (i.e. adjusting by covariates after matching) proposed by Abadie and Imbens (2006). All papers performed matching on, or included in the regression analysis in the case of RCTs, covariates related to land quality, socio-economic conditions (with very few exceptions) and accessibility to markets, which are believed to jointly determine the participation in the program and forest-use decisions. In all studies the outcome variable was measured using satellite images. From each study we choose only one treatment effect that best controls for different types of biases, resulting in 18 independent treatment effects.

2.3.3 Measures/Emissions

Outcome variable

Our outcome variable representing the environmental effectiveness of Forest Conservation Programs is the annual avoided deforestation (or the increase in forest cover) expressed in percentage points. Not all studies report the same forest-related outcome: more than half of the studies report deforestation rates, while the rest report changes in forest cover. Due to the small number of studies, we choose to pool both types of forest outcomes when we estimate the mean effect. However, given that the forest cover measure includes not only deforestation (i.e. forest loss) but also forest gain, we expect the effect sizes from the studies using this outcome variable to be larger. Indeed, Samii et al. (2014) found that this was actually the case in their sample. Therefore, we account for these differences in the heterogeneity analysis. Furthermore, not all studies reported the outcome variable in percentage points, but rather in hectares. In such cases, we converted the outcome using the following formula:

$$\hat{\theta}_i = \frac{\text{Outcome (ha)}_i}{\text{Hectares enrolled in the program}_i} \times 100 \tag{2.3}$$

We express all outcome variables in annual terms to facilitate the comparison, but we account for the difference in the number of intervention years analyzed in the subgroup analysis. When various different treatment effects are presented in a study, we chose the one that best controls for different types of biases, which in general represents the authors' preferred specification. Finally, we transform the reported standard errors accordingly using the Delta Method in order to recover an estimate of the precision of the treatment effect in p.p. $\hat{\sigma}_i$.[3]

Moderator variables

The independent variables that might influence the between-study heterogeneity and that we consider in our analysis can be grouped in two categories: (i) study design characteristics, that are chosen by the authors of the studies and (ii) forest conservation program characteristics, that are independent of the authors' will. For simplicity, we code all these as dummy variables.

Study design characteristics that we are interested in testing for their influence on the treatment effect is the use of experimental versus quasi-experimental methods (*RCT*), whether leakage effects are taken into account in the main estimate or not (*Leakage*), whether the evaluation of the program occurs at national versus sub-national level (*National*), whether the number of years of the program (i.e. number of years in which the landowners receive payments) evaluated are less or more than 5 (*Duration*) and whether studies use deforestation or forest cover as an outcome variable (*Deforestation*).

In terms of program characteristics that might explain the variation between studies, we compare programs funded by the Government (both national or local) versus those funded by conservation funds, NGOs (national or international) or other sources of private funding (e.g. tourists, consumers) (*Government*), programs in which there is a possibility of a collective payment (or other non-monetary benefits) to those where payments are only attributed to private households with clear land tenure rights (*Collective PES*), programs

[3]When standard errors where not reported in the paper, we first tried to recover them using the information available in the paper and if this was not possible, we contacted the authors by email.

that are implemented within, or very close to, protected areas versus those implemented away of protected areas (*Within PA*), programs that pay a lower than average (in the sample) subsidy[4] and those that pay a higher than average subsidy (*Subsidy*) and programs implemented in Central and South America or elsewhere in the world (here, Asia or Africa) (*Central-South America*).

Cost and benefit variables

From each study, we extract the payment per hectare of conserved forest. We assume that the payments represent the program's cost, as, with some few exceptions, we have no information on other administrative costs.

We obtain, for each country or region, the average carbon stored in trees per hectare of forest from the World Resource Institute (more specifically, from Global Forest Watch - Climate). The value of the average carbon stored in trees is based on satellite estimates and captures both the above-ground and below-ground storage. However, the latest value available is from year 2000 and only for 30% canopy density. Thus, we might be underestimating the carbon stored per hectare of tree cover. We transform the carbon stock in tCO_2eq by multiplying it by 3.67.[5]

2.3.4 Meta-analysis and correction for publication bias

Meta-analysis is a statistical method to synthesize data coming from separate primary studies that answer the same empirical question. The main goal of a meta-analysis is to compute a summary effect for the treatment effect (i.e. effect size), which in general has a higher statistical power than what can be achieved by individual studies. When the effect varies from one study to the next, meta-analysis also allows to assess the reasons for the dispersion. Finally, a meta-analysis also helps to test for the presence of publication bias, a phenomenon that might inflate our estimates of the impact of Forest Conservation Programs.

We apply meta-analysis to the set of estimates of the impact of Forest Conservation Programs on deforestation rates and their estimated standard errors $(\hat{\theta}_i, \hat{\sigma}_i)$, for $i = 1 \dots N$, with N the number of studies included in the systematic review.[6] The most basic meta-

[4]We transform each subsidy in its 2005 USD equivalent and then divide it by the Consumer Price Index (CPI) in each country in 2005 to obtain the subsidy in 2005 real prices.

[5]In this version of the paper we are not using this information but rather we use the estimate of Carbon content from Jayachandran et al.(2017) for simplicity purposes. This information will be used when we will build cost-benefit ratios for each individual study.

[6]In future work, we will also meta-analyse the benefit-cost ratios and cost-effectiveness parameters.

analysis regression is as follows:

$$\hat{\theta}_i = \theta + \epsilon_i + \nu_i, \tag{2.4}$$

where θ is the average treatment effect in the population, $\epsilon_i \sim \mathcal{N}(0, \hat{\sigma}_i^2)$ is sampling noise and $\nu_i \sim \mathcal{N}(0, \tau^2)$ accounts for the heterogeneity in treatment effects across sites. Equation (2.4) can be complemented with a set of covariates that try to explain part of the heterogeneity in treatment effects across sites.

A key distinction in meta-analysis is between fixed effects and random effects meta-analysis. Fixed effects meta-analysis is characterized by $\tau = 0$: there is no heterogeneity of effect sizes across sites. Random effects meta-analysis is characterized by $\tau > 0$: there is heterogeneity between effect sizes across sites. Obtaining an estimate of the value of τ is thus crucial to distinguish between homogeneous and heterogeneous treatment effects. Finding an accurate estimate of τ is also important since a consistent estimator of θ in Equation (2.4):

$$\hat{\theta} = \sum_{i=1}^{N} \omega_i \hat{\theta}_i \text{ with } \omega_i = \frac{\frac{1}{\hat{\sigma}_i^2 + \hat{\tau}^2}}{\sum_{i=1}^{N} \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}^2}}. \tag{2.5}$$

The weights ω_i give more importance to precise results (with smaller $\hat{\sigma}_i^2$). The importance given to more precise results is less strong when there is strong heterogeneity in treatment effects (when $\hat{\tau}^2$ is large). Several estimators have been proposed for τ^2 and they do not vary much in practice. In what follows, we will use the restricted maximum likelihood estimator (REML) which has been shown in simulations to perform better than the alternatives (Viechtbauer, 2005). When fitting a random-effects model, τ^2 is estimated and treated as a known constant, ignoring thus the uncertainty in the estimate. This can lead to test statistics that are too large and confidence intervals that are too narrow, especially when the number of studies is small and there is substantial heterogeneity (as it is the case here). To correct for this problem, we use the Knapp and Hartung (2003) adjustment. We assess the heterogeneity of the pooled treatment effect using several common measures proposed in the literature (Borenstein et al., 2009). Cochran's Q-statistic is the weighted sum of squares on a standardized scale and reflects the excess variance, which is the part that is attributed to differences in the true effects from study to study. However, the Q-statistic and its significance highly depend on the number of studies included in the meta-analysis. Contrarily, the between-study variance estimated in the first step of the random-effects model is insensitive to the number of studies, but it is based on the metric of the effect size. A measure that is invariant to both the number of studies and the scale is the I^2 statistic (Higgins et al., 2003) which gives the proportion of observed dispersion that is real (i.e. not due to the sampling error). However, even I^2 highly depends on the precision of the included studies. To overcome this limitation, we further compute prediction intervals (Harrer et al., 2019). While the confidence interval quantifies the precision of the mean effect size, the prediction interval shows the actual dispersion of effect sizes around the mean.

Meta-analysis is severely complicated by the existence of publication bias (Stanley and Doucouliagos, 2012). Publication bias occurs when only statistically significant results get reported in academic journals. As a consequence of publication bias, smaller effects are missing from the published record, which biases the meta-analysis estimator in equation (2.5) upwards. Publication bias has been shown to affect many literatures in economics (Ioannidis et al., 2017; Andrews and Kasy, 2019). Several methods have been proposed to test and correct for publication bias in practice. Kvarven et al. (2019) test the ability of several of these methods to correct for publication bias in meta-analysis by comparing their results to that of pre-registered massive replications in psychology. They find that the FAT-PET-PEESE method of Stanley and Doucouliagos (2012) performs the best at reproducing the results of pre-registred replications (it is on average unbiased). We thus use this approach to detect and correct for publication bias.

In practice, Stanley and Doucouliagos (2012)'s FAT-PET-PEESE method is a sequential three-step procedure. The first step is a Funnel Asymmetry Test (FAT). This test is based on the intuition that publication bias generates a negative correlation between precision (as measured by the standard error $\hat{\sigma}_i$) and effect size (here measured by $\hat{\theta}$). In practice, the Funnel Asymmetry Test is performed using the following regression:

$$\hat{\theta}_i = \alpha_0 + \alpha_1 \hat{\sigma}_i + \epsilon_i + \nu_i \tag{2.6}$$

and testing for the null hypothesis that $\alpha_1 = 0$. If the null hypothesis of $\alpha_1 = 0$ is rejected in favor of $\alpha_1 < 0$, we have evidence of publication bias. The second step of Stanley and Doucouliagos (2012)'s procedure is the Precision-Effect Test (PET). It consists in testing whether there is a non-zero treatment effect when estimating equation (2.6) using the null hypothesis that $\alpha_0 = 0$. Rejecting the null is taken as evidence that there is a non-zero treatment effect. In that case, Stanley and Doucouliagos (2012) propose to estimate the true bias-corrected treatment effect using a Precision-Effect Estimate with Standard Error (or PEESE). This estimator accounts for the fact that the link between the standard error and the treatment effect is non-linear but rather looks like a quadratic. In practice, PEESE estimates the bias-corrected treatment effect using $\hat{\beta}_0$ estimated using the following regression:

$$\hat{\theta}_k = \beta_0 + \beta_1 \hat{\sigma}_k^2 + \epsilon_k + \nu_k. \tag{2.7}$$

A final issue in estimating equations (2.6) and (2.7) is that of the proper estimation method to use. A fixed effects analysis does not account for the possible existence of treatment effect heterogeneity across sites and might therefore overestimate the precision of the estimated treatment effect parameter. A random effects analysis accounts for possible heterogeneity, but it might also confound heterogeneity with publication bias and worsen the impact of publication bias by increasing the weights of imprecise studies. Stanley and Doucouliagos (2015) propose to eschew both fixed and random effects meta-analysis and to replace them by an Unrestricted Weighted Least Squares approach. They show that the Unrestricted Weighted Least Squares estimator point estimate is equal to the fixed effects estimator, but has better coverage properties than the fixed effect estimator when there is treatment effect heterogeneity and is less biased than the random effects estimator when there is publication bias. In our estimations, we thus use Stanley and Doucouliagos (2015)'s Unrestricted Weighted Least Squares estimator.

2.4 Results

In this section, we present our main results. We first start with a quick description of our data. We then present our estimates of the additionality of Forest Conservation Programs. We focus specifically on the results of tests for publication bias and on the bias-corrected estimates. We end with a derivation of the benefit-cost ratio of Forest Conservation Programs.

2.4.1 Descriptive statistics

Table 2.2 presents the descriptive statistics for the outcome variable, for the moderator variables and for the variables used in the cost-benefit analysis. In our sample of individual studies, the subgroups are balanced for the *Leakage*, *Duration* and *Collective PES* variables. 67% of studies use deforestation as an outcome variable. 61% of studies are about programs that are publicly-funded. 33% of studies evaluate a Forest Conservation Program at the national level and the same proportion evaluate a program that is located within a protected area. 55% of programs give a subsidy higher than the sample average. 83% of studies come from Central and South America and only 11% use RCT as evaluation method.

Table 2.2 shows that the mean additionality estimate in our sample is of 0.79 p.p., meaning that on average, Forest Conservation Programs decrease deforestation by 0.8 p.p. each year. This estimate masks strong heterogeneity. Figure 2.4a plots the distribution of the 18 treatment effect estimates in our sample. The distribution is characterized by a double hump: a first set of estimates of program effects are clustered just above zero, while a second set is clustered around 2 p.p. Figure 2.4b plots each individual estimate along with its confidence interval. In the figure, the most precise estimates seem to be closer to zero, while the largest estimates seem to be the most imprecise ones. Finally, Figure 2.4c shows a funnel plot, that is how treatment effect estimates relate to their precision (here measured by their standard error). This figure confirms what Figure 2.4b suggested: there is a clear positive correlation between effect size and standard error (and thus imprecision): the most precise studies are also the ones that yield the smallest effect sizes.

2.4.2 Additionality of Forest Conservation Programs

Table 2.3 shows our estimates of the meta-regressions (2.6) and (2.7) using the Unrestricted Weighted Least Squares Estimator. The first result is that the coefficient α_1 in equation (2.6) is significantly different from zero, meaning that there is a positive correlation between the imprecision of a result (measured by its standard error) and its effect size: the most precise results are also the smallest. This is a clear sign of publication bias. This is made apparent on Figure 2.4c by the black line *FATPETWLS* which has a positive slope. The second result from the first column of Table 2.3 is that the estimate of α_0 obtained using Unrestricted Weighted Least Squares is not statistically different from zero. It is equal to 0.036±0.20, meaning that the PET estimate of the average effect of Forest Conservation Programs is not statistically distinguishable from zero.[7] The PEESE estimate is nevertheless larger and slightly more precise: 0.23±0.14. This estimate is materialized in Figure 2.4c by the location where the discontinuous line meets the y-axis.

Figure 2.4c also reports the results of estimating equation (2.4) using the Unrestricted Weighted Least Squares Fixed Effects estimator (*MetaAnalysisWLS*) and the Restricted Maximum Likelihood Random Effects estimator (*MetaAnalysisRE*). The Fixed Effects estimator is almost identical to the PEESE estimator at 0.26±0.14. The Random Effects estimator is larger at 0.61±0.33.[8]

Figure 2.5 summarizes our results. Our preferred estimates are the ones obtained

[7] We report the precision of each estimate using the half-width of its confidence interval.
[8] The precision of the Random Effects estimator appears on Figure 2.4b.

using the Unrestricted Weighted Least Squares estimator, either using the PEESE formulation or the simplest formulation that ignores publication bias. Both estimators yield very similar point estimates of a reduction in deforestation by 0.23±0.14 p.p. each year that the program is in place. The PET estimator is less optimistic while the Random Effects estimator is more optimistic. Both of these estimators are also less precise.

Figure 2.5 reports the two most recent and most influential estimates of the effect of Forest Conservation Programs on f orest cover (Jayachandran et al. (2017) and Simonet et al. (2018)). Both of these papers report estimates that are one order of magnitude above the ones we have estimated (around 2 p.p. per year). Note nevertheless that the precision of these results is much lower than ours (2±2 p.p.), and thus that our results are contained in their confidence intervals. We interpret the results in these two papers as being driven by publication bias. Indeed, they appear in Figure 2.4c to be very close to the PEESE curve (they are indicated by the two yellow dots). In what follows, we use the mean estimate from these two studies (2±2 p.p.) to illustrate the consequences of publication bias and of ignoring the importance of sampling noise when computing estimates of the benefit-cost ratio.[9,10]

2.4.3 Cost-Benefit Analysis

In this section, we report our estimates of the benefit-cost ratio of Forest Conservation Programs. One way to build such an estimate would be to compute a benefit-cost ratio for each program separately and then to aggregate these estimates using a meta-regression. This approach requires a separate estimate of precision for each benefit-cost ratio, which in turn requires a closed-form formula for both benefits and costs in order to be able to apply the Delta Method. We are working on obtaining such closed-form formulas in order to implement this approach.

What we are doing in this section is much simpler. We derive the benefit-cost ratio of the average Forest Conservation Program in our sample incorporating our estimates of treatment effects from the previous section in equation (2.2). Since equation (2.2) is not a closed form formula, we derive standard errors and confidence intervals using Monte-Carlo simulations. We use a horizon of 1000 years to compute the discounted climate

[9]In no way are we singling out the researchers involved in these papers. Publication bias is a systemic problem that is due the incentives faced by researchers and to the behavior of editors and referees that show marked preferences for statistically significant results.

[10]Note that we might be attributing to publication bias the impact of another confounding source. For example, it could be that the most precise results are obtained where the deforestation rate is already low, implying a negative correlation between precision and effect size. We are in the process of testing that possibility by collecting data on the baseline deforestation rate.

benefits. In agreement with the average program in our data, we set the counterfactual deforestation rate to $d_0 = 0.02$, the program duration to $K = 5$ years, the delay between deforestation and emissions to $T = 10$ years, the delay for the deforestation rate to reach the pre-treatment deforestation rate after the program stops at $\kappa = 10$ years and the stock of Carbon in the forest to $G = 565 MtCO_2eq$.[11] In our simulations, we set program costs to three levels: 11, 31 and 43 USD/year, corresponding respectively to the 25th, 50th and 70th percentiles of the distribution of costs in our sample. We compute the benefit-cost ratio using two distinct estimates of the impact of Forest Conservation Programs on deforestation. The first estimate corresponds to our preferred meta-analytical estimate: $\theta = 0.002$. The second estimate is one order of magnitude larger, in order to correspond to the most recent published estimates of the impact of Forest Conservation Programs: $\theta = 0.02$.

Figure 2.6 presents our estimates of the benefit-cost ratios of Forest Conservation Programs under various assumptions. Figure 2.6a shows the benefit-cost ratio of a Forest Conservation Program with our basic set of assumptions, as a function of the size of the impact of the program on deforestation (0.2 p.p. *vs* 2 p.p.) and of our assumptions on the post-program trajectory of the deforestation rate. With the impact we have estimated in this paper (0.2 p.p. reduction in the deforestation rate), Forest Conservation Programs do not appear to be cost-effective. Under the unfavorable assumption that the deforestation rate immediately returns to the pre-treatment rate when the program ends (*CatchingUp=Lower*), the benefit-cost ratio of Forest Conservation Programs is 0.45 ± 0.32. Under the assumption that the deforestation rate catches up linearly towards the pre-treatment rate (*CatchingUp=Linear*), the benefit-cost ratio of Forest Conservation Programs is 0.78 ± 0.56. It is only under the most favorable assumption that the impacts of the program last forever (*CatchingUp=Upper*) that the program benefits are larger than its costs (the benefit-cost ratio is equal to 2.63 ± 1.94 in that case).[12] With the much more optimistic impacts estimated in the most recent papers (2 p.p. reduction in the deforestation rate), Forest Conservation Programs are cost-effective no matter the assumption on the trajectory of deforestation after the program ends. Under the pessimistic assumption that the effects of the program cease right after the program stops, the benefit-cost ratio of Forest Conservation Programs is of 3.19. It is even much larger (5.7 and 33.9) under more optimistic assumptions.

Figures 2.6c and 2.6d show what happens to our benefit-cost ratio estimates when the cost of the program decreases to 11 USD/year or when the Social Cost of Carbon

[11] We use the estimate of Carbon content in the biomass above ground from Jayachandran et al (2017): 307 MT of biomass per hectare, or 154 MT of carbon per hectare, multiplied by 3.67 to obtain $MtCO_2eq$.

[12] This last estimate is not shown on the graph in order not to dwarf all the estimates.

increases to 100 USD/tCO_2eq. In both of these cases, the benefit-cost ratio of Forest Conservation Programs moves above one with our estimates of program impacts. It is the case even under the less optimistic assumption that the program impacts on the deforestation rate stops as soon as the program stops. Forest Conservation Programs can thus become cost-effective ways to fight climate change when their costs are in the lower end of the spectrum or if the estimates of the impacts of climate change become larger. Figure 2.6b shows what happens to our estimates of the benefit cost ratio of Forest Conservation Programs when the cost of the program increases to 43 USD/year. In that case, the benefit cost ratio of Forest Conservation Programs decreases to 0.32 and 0.56 under the assumptions of no permanence (*CatchingUp=Lower*) and moderate permanence (*CatchingUp=Linear*) respectively.

Figure 2.6a also shows that the comparison of benefit-cost ratios with our estimates of program impacts and the more optimistic recent estimates in the literature are missing a crucial element: precision. The results of our meta-analysis are indeed much more precise than the more optimistic recent estimates in the literature. The standard error of our additionality estimate is 0.07 p.p. while that of the most recent results in the literature are around 1 p.p. As a consequence, when computing confidence intervals using Monte Carlo simulations, our estimates of the benefit-cost ratio are much more precise than the ones using the less precise but more optimistic recent estimates in the literature. Even more important, some of our estimates of the benefit-cost ratio are precise enough to exclude one from their 95% confidence interval: we indeed estimate a benefit-cost ratio of 0.45 ± 0.32 under the pessimistic assumption that the deforestation rate returns to its pre-treatment level right after the program stops. The more optimistic estimates hide tremendous variability: they cannot exclude that the benefit-cost ratio is lower than one. They also cannot exclude that the benefit-cost ratio is equal to our most pessimistic estimate of 0.45 either. Figure 2.6a exemplifies the worrying consequences of publication bias: when there is publication bias, the true value of an imprecisely estimated parameter is generally in the bottom part of the 95% confidence interval of the most imprecise published estimates. In our case, the consequences of publication bias are even more stark: publication bias makes the benefit-cost ratio move above one.

2.5 Conclusion

In this paper, we use meta-analysis to estimate the impact of Forest Conservation Programs on deforestation. We find that Forest Conservation Programs decrease deforestation by

0.23±0.14 p.p. on average. We find major signs of publication bias in the literature, with the more precise results exhibiting the lowest impacts. We find that some of the most recent estimates of the impact of Forest Conservation Programs are overestimated by a factor of 10, artificially pushing their benefit-cost ratio above one.

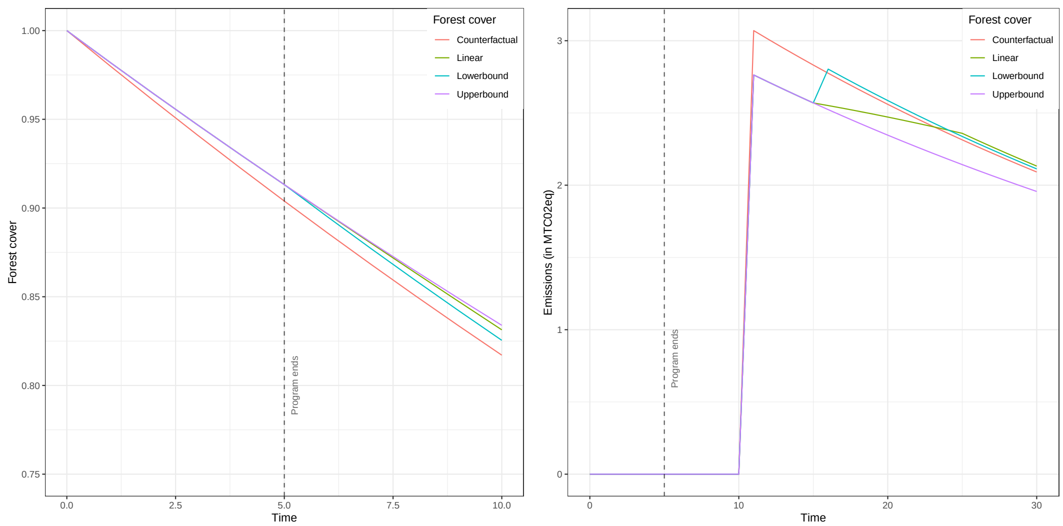
We use our estimates to generate the climate benefit-cost ratio of Forest Conservation Programs. We find that the value of the climate benefits of Forest Conservation Programs crucially depends on the permanence of their effects after the program stops. For a Social Cost of Carbon of 31 USD/tCO_2eq, we estimate that benefits are equal to 45±32% of the program costs if the impacts of the program on deforestation stop just after the program ends, while they are equal to 78±56% of the program costs if the impact progressively decreases over 10 years, and 263±194% of the program costs if the impact persists forever. We estimate that Forest Conservation Programs would become cost-effective with a Social Cost of Carbon of 100 USD/tCO_2eq, even with very low permanence.

In view of our results, we do not believe that Forest Conservation Programs pass a cost-benefit test at current carbon prices. There is nevertheless enough uncertainty around our estimates to warrant further investigation. By far the most important degree of uncertainty is the one surrounding the permanence of the effects of Forest Conservation Programs on deforestation. Even modest levels of permanence of effects (like a progressive fading out of impacts over a 10-year period) are enough to almost double benefits. Further research needs to investigate the permanence of the effects of Forest Conservation Programs once the payments have stopped. A second crucial area of further research is the estimation of the true cost of forest conservation efforts for landowners. In this paper, as in the rest of the literature, we have assumed that program costs are equal to the transfers received by the participants. In practice, the costs to participants are probably lower than the transfers they receive, thereby increasing the benefit-cost ratio of Forest Conservation Programs.

In future research efforts, we urge researchers, funders and policy-makers to take extremely seriously the challenge of publication bias. Our results suggest that recent influential estimates of the effects of Forest Conservation Programs on deforestation are probably overestimated by a factor of 10. We recommend that all future analysis aiming at estimating the impact of Forest Conservation Programs be pre-registered and commits to a publication of its results, no matter their size and statistical significance. This would be made possible for example by using the format of registered reports. A related possibility would be to launch a worldwide initiative aiming at estimating the impact of these programs, where the results of each individual estimate are included in an eventual meta-analysis.

2.6 Appendix

2.6.1 Figures



(a) Forest cover

(b) Emissions

Figure 2.1: Evolution of forest cover and emissions according to several assumptions of post-program permanence.

Note: Counterfactual depicts forest cover in the absence of the program; *Linear* shows forest cover when the deforestation rate returns linearly to the pre-treatment deforestation rate over κ years; *Lowerbound* graphs forest cover when the deforestation rate returns to the pre-treatment deforestation rate as soon as the program stops; *Upperbound* plots forest cover under the assumption that the effects of the program on the deforestation rate remain forever. In the simulations shown on the picture, we have set $d_0 = 0.02$, $\theta = 0.002$, $K = 5$, $\kappa = 10$, $T = 10$ and $G = 565MtCO_2eq$.

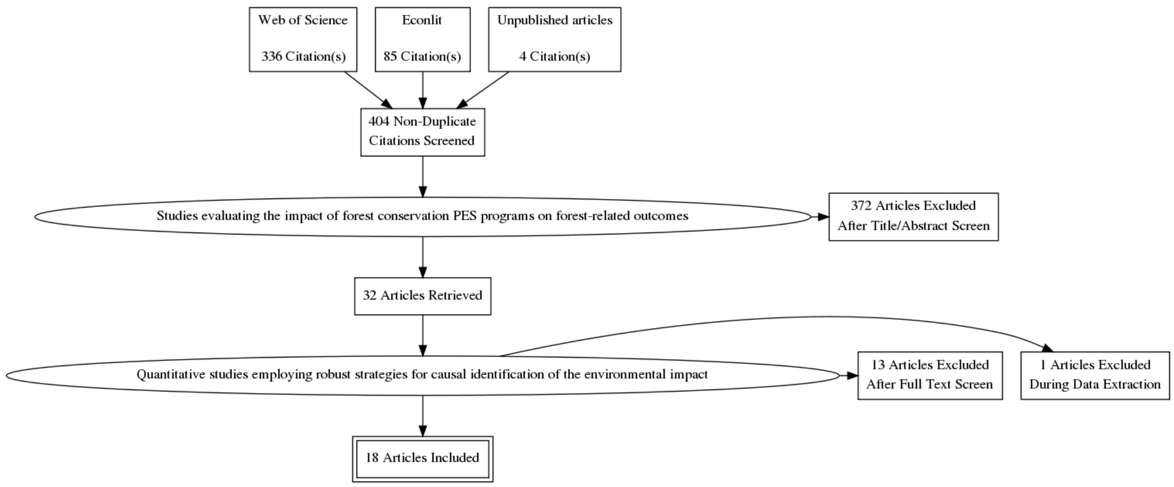
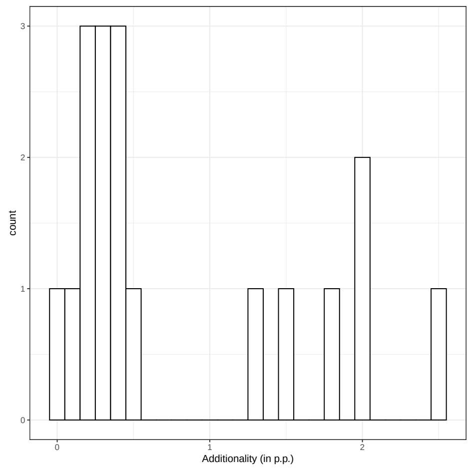
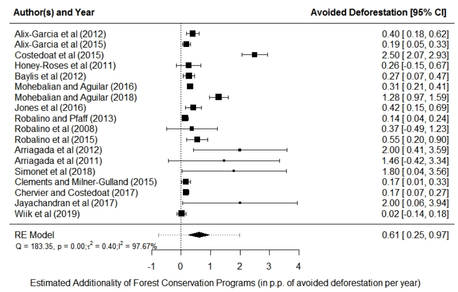


Figure 2.2: PRISMA flow chart for the identification and selection of studies included in the meta-analysis.

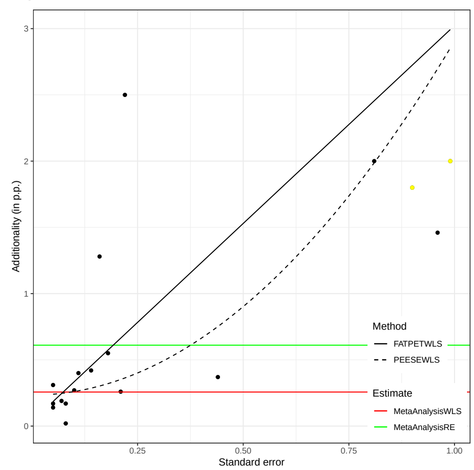
Figure 2.3: Map of the countries with estimates of the impact of Forest Conservation Programs on deforestation included in our sample.



(a) Histogram



(b) Forest plot



(c) Funnel plot

Figure 2.4: Distribution of individual level estimates of the impact of Forest Conservation Programs on Deforestation.

Note: The histogram plots the empirical distribution of the treatment effect estimates across studies included in our meta-analysis. The forest plot shows each individual estimate from each study included in our meta-analysis along with its 95% confidence interval. The forest plot also reports an estimate of equation (2.4) estimated using the Restricted Maximum Likelihood Random Effects Estimator along with its 95% confidence interval. The forest plot also reports an estimate of τ^2, of Cochran's Q-statistic and of the I^2 statistic. p is the p-value for a test of null that $\tau^2 = 0$. The funnel plot shows how each effect size relates to its precision (as estimated by its standard deviation). The funnel plot also reports the treatment effect estimated using the Restricted Maximum Likelihood Random Effects Estimator (*MetaAnalysisRE*) and the Unrestricted Weighted Least Squares Estimator (*MetaAnalysisWLS*). The funnel plot also shows the fitted FAT and PEESE curves obtained estimating equations (2.6) and (2.7) using the Unrestricted Weighted Least Squares Estimator (*FAT-PETWLS* and *PEESEWLS* respectively). The yellow dots on the funnel plot single out the individual estimates from Jayachandran et al. (2017) and Simonet et al. (2018).

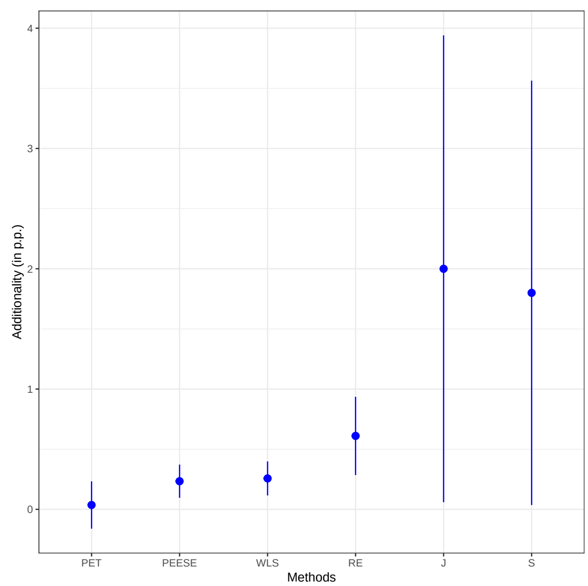
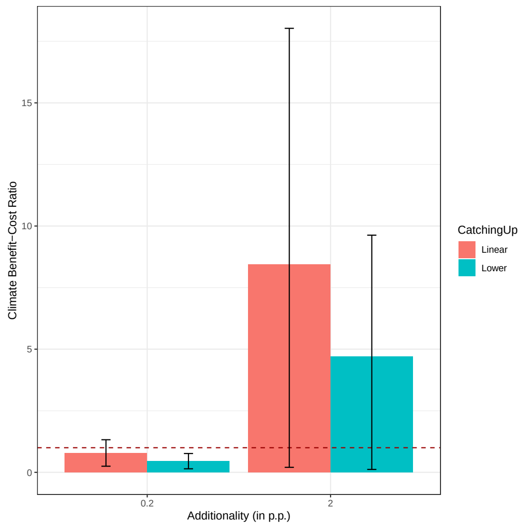
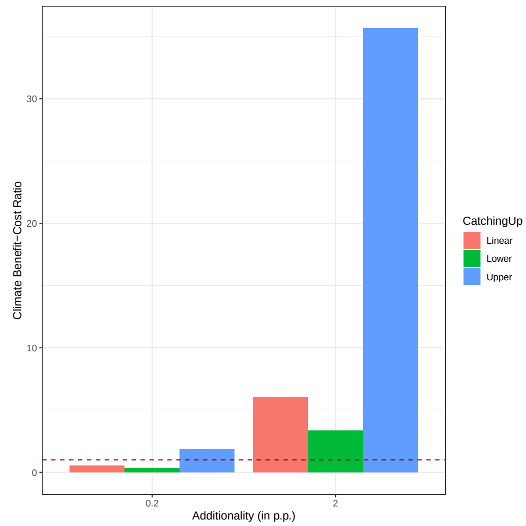


Figure 2.5: Summary of the results of the meta-analysis of the impact of Forest Conservation Programs.

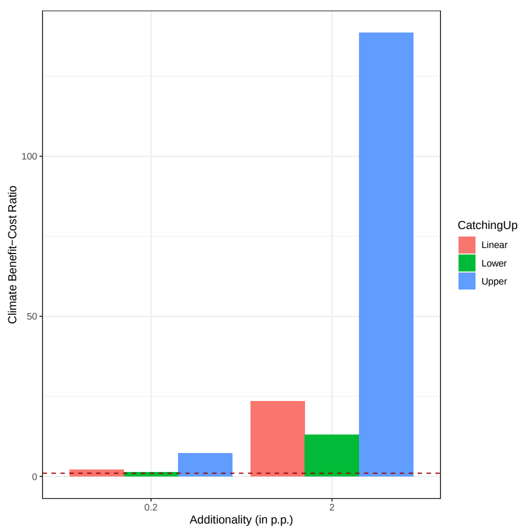
Note: Each dot is a point estimate of the impact of Forest Conservation Programs on forest cover. Confidence bands materialize the 95% confidence intervals. *PET* and *PEESE* are obtained estimating equations (2.6) and (2.7) using the Unrestricted Weighted Least Squares Estimator. *WLS* and *RE* are obtained estimating equation 2.4 using the Unrestricted Weighted Least Squares Estimator and the Restricted Maximum Likelihood Estimator respectively. *J* and *S* are the individual estimates from Jayachandran et al. (2017) and Simonet et al. (2018).



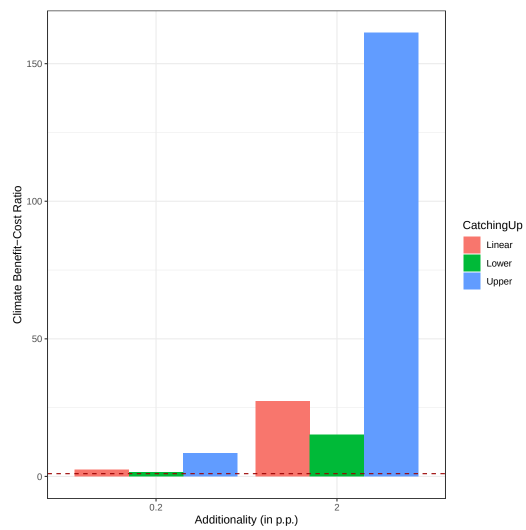
(a) $SCC = 31$, $C = 31$.



(b) $SCC = 31$, $C = 43$.



(c) $SCC = 31$, $C = 11$.



(d) $SCC = 100$, $C = 31$.

Figure 2.6: Benefit-cost ratio of Forest Conservation Programs under various sets of assumptions.

Note: Benefit-cost ratios estimated using the formula delineated in equation (2.2). We set $d_0 = 0.02$, $K = 5$, $T = 10$, $\kappa = 10$, $\beta = 0.98$ and $G = 565$ in all our simulations. d_0 is the counterfactual deforestation rate, K is program duration (in years), T is the delay from deforestation to emissions (in years), κ is the time it takes for the deforestation rate to converge to the pre-treatment deforestation rate after the program stops (in years), SCC is the social cost of carbon (in USD/tCO_2eq), β is the yearly discount rate, G is the carbon content of forest (in tCO_2eq/ha) and C is the yearly costs of the program (in USD/year). *Additionality* is the impact of Forest Conservation Programs on deforestation (in p.p. of reduced deforestation per year). *CatchingUp* shows the impact on our benefit-cost estimates of three assumptions about how the deforestation rate converges to the pre-treatment deforestation rate once the program stops. *Linear* assumes that the catching up occurs at a linear rate over κ years. *Lower* assumes that the deforestation rate returns to its pre-treatment value right after the program ends. *Upper* assumes that the program impact on the deforestation rate never fades away. The horizontal discontinuous line materializes the point at which the benefit-cost ratio is equal to one. Confidence bands materialize the Monte Carlo generated 95% confidence-intervals.

2.6.2 Tables

Table 2.1: List of included papers (grouped by program)

Program	Author(s) and Year	Control group	Time period	Outcome variable	Evaluation method	Carbon storage data
Mexico PSAH	Alix-Garcia et al. (2012)	Matched from the rejected applicants and future recipients	2003-2006	% of deforested area	Matching + Tobit	National
Mexico PSAH	Alix-Garcia et al. (2015)	Matched from rejected applicants	2003-2011	Mean dry season NDVI (normalized difference vegetation index)	Matching + Fixed Effects	National
Mexico PSA-CABSA	Costedoat et al. (2015)	Matched non-treated ejidos	2007-2013	Ha of forest cover	Matching + DID	State (Chiapas)
Mexico MBCF	Honey-Roses et al. (2011)	Matched polygons not subject to PES	1993-2009	% of avoided deforestation	Spatial matching	Mean of 2 states (Michoacan and Mexico)
Mexico MBCF	Baylis et al. (2012)	Matched grid cells without PES	1993-2009	Forest cover dummy	Spatial matching + DID	Mean of 2 states (Michoacan and Mexico)
Ecuador PSB	Mohebalian and Aguilar (2016)	Matched pixels from non-enrolled forests	2008-2014	Deforestation dummy	Matching	Province (Pichincha)
Ecuador PSB	Mohebalian and Aguilar (2018)	Matched pixels from non-enrolled forests	2008-2014	p.p. of avoided deforestation	Matching	Mean of 2 provinces (Napo and Orellana)
Ecuador PSB	Jones et al. (2016)	Matched non-participant households from cooperatives where at least 1 HH participated	2004-2013	p.p. of deforestation rate	Matching + Fixed Effects	Province (Sucumbios)
Costa Rica PSA	Robalino and Pfaff (2013)	Matched untreated parcels	1997-2000	p.p. of deforestation rate	Matching	National
Costa Rica PSA	Robalino et al. (2008)	Matched untreated parcels	2000-2005	p.p. of deforestation rate	Matching	National
Costa Rica PSA	Robalino et al. (2015)	matched parcels away from parks	2000-2005	Deforestation rate dummy	Matching	National
Costa Rica PSA	Arriagada et al. (2012)	Matched non-participant farms	1992-2005	Ha of change in forest cover	Matching + DID	Province (Heredia)
Costa Rica PSA	Arriagada et al. (2011)	Matched eligible, but non-participant census tracts	1997-2005	Ha of net deforestation	Matching	National
Brazil PAS	Simonet et al. (2018)	Households in communities that didn't receive PES	2010-2014	% of forest cover	DID	State (Para)
Cambodia IbisRice/Ecotourism	Clements and Milner-Gulland (2015)	Matched grid cells from non-treated villages	2005-2010	Ha of deforestation rate	Matching + DID	Province (Preah Vihear)
Cambodia CA	Chervier and Costedoat (2017)	Matched grid cells from non-treated communes	2005-2012	p.p. of forest cover loss	Matching + DID	Mean of 2 states (Koh Kong and Pursat)
Uganda RCT	Jayachandran et al. (2017)	Randomized control villages	2011-2013	Ha of tree cover	OLS (RCT)	Mean of 2 districs (Hoima and Kibaale)
Bolivia Watershared	Wiik et al. (2019)	Randomized control communities	2000-2016	% of deforestation	Generalized additive models (RCT)	Department (Santa Cruz)

Table 2.2: Descriptive statistics

	Mean	Std. Dev.	Min	Max
Outcome variable				
Avoided deforestation (p.p.)	0.79	0.80	0.02	2.50
Moderator variables				
RCT	0.11	0.32	0	1
Leakage	0.50	0.51	0	1
National	0.33	0.49	0	1
Duration	0.50	0.51	0	1
Deforestation	0.67	0.49	0	1
Government	0.61	0.50	0	1
Collective PES	0.50	0.51	0	1
Within PA	0.33	0.49	0	1
Subsidy	0.55	0.51	0	1
Central and South America	0.83	0.38	0	1

Table 2.3: FAT-PET-PEESE with Unrestricted Weighted Least Squares

	FAT-PET	PEESE
Intercept	0.036	0.234***
	(0.100)	(0.070)
Standard error	2.986**	
	(1.078)	
Variance		2.672
		(1.648)
Num.Obs.	18	18
R2	0.324	0.141

Note: Results of estimating equations (2.6) and (2.7) using the Unrestricted Weighted Least Squares Estimator (FAT-PET and PEESE respectively). Standard errors in parenthesis. *, ** and *** stand for coefficients statistically significantly different from zero at 10%, 5% and 1% respectively.

Chapter 3

Optimal Design of Payments for Ecosystem Services: Evidence from France

Anca Voia

Abstract

To strike the right balance between agriculture and the environment, policymakers increasingly use Payments for Ecosystem Services (PES) programs. They are incentives offered to landowners conditional on the provision of an environmental service. However, information asymmetries may limit their effectiveness: due to differences in opportunity costs, offering a linear-uniform payment to all farmers increases the risk of windfall gains. Nonlinear payments are a way to decrease windfall gains by differentiating payments by the quantity offered. Another approach is to differentiate payments by geographic characteristics, a proxy for provision costs. In this paper, I use a principal-agent model to provide insights on the optimality of different Payments for Ecosystem Services contract designs. I use data on the French Grassland Conservation Program contracts, and I exploit an exogenous change in the payment structure to identify and estimate nonparametrically the farmers' cost function and the distribution of their types. This allows me to select parametric specifications and to evaluate welfare for different contract designs. I find that the loss of using linear-uniform contracts instead of nonlinear ones is around 2.6% and that spatially-targeted linear-uniform contracts improve the welfare gain with respect to the linear-uniform contracts by 1.9%. Moreover, I find a low cost of asymmetric information, with the surplus of nonlinear contracts being 87% of that under complete information.

3.1 Introduction

To strike the right balance between agriculture and the environment, policymakers in both developed and developing countries increasingly use Payments for Ecosystem Services (PES). PES are voluntary agreements between a buyer (e.g. Government or private users) and a seller (e.g. landowner) in which a payment is given conditionally on an environmental service being adequately provided (Alston et al., 2013). The payment is computed so as to compensate the landowner for the average compliance costs and for the forgone farming revenue associated with the adoption of greener practices. In general, a PES program targets at least one of the four environmental services among carbon sequestration, watershed services, biodiversity, and scenic beauty. Depending on the environmental service targeted, these programs can be an effective tool for climate change mitigation or adaptation. For example, PES programs aimed at carbon sequestration from forests and grasslands can help mitigating climate change, while PES programs aimed at improved water quality can have natural adaptation co-benefits. This can happen if more environmentally-friendly land uses are promoted in the entire watershed, which will decrease the inhabitants' vulnerability to climate-related water problems (Wertz-Kanounnikoff et al., 2011).

PES contracts are usually subject to asymmetric information between landowners and the service buyers that can limit their cost-effectiveness (Ferraro, 2008). Landowners have private information about their opportunity costs of supplying the environmental services that can be used to secure higher payments than the minimum necessary to induce participation in the contract. This consequently limits the program's cost-effectiveness, as these informational rents are financed through taxes and thus are socially costly.

When there is hidden information, most of the relevant literature advocates the use of a screening mechanism (Bourgeon et al., 1995; Fraser, 1995; Wu and Babcock, 1996; Moxey et al., 1999; Ozanne et al., 2001; and many others in the context of PES), which consists of a menu of contracts that induces farmers to choose the contract that corresponds to their opportunity cost of providing the environmental service. One can then find inside this class an optimal contract, whose payments often turn out to be nonlinear with respect to the quantity of the service which is provided. Nevertheless, it is fair to say that these contracts are seldom used in practice, maybe because computing and implementing these payments is too complicated. What is often proposed in practice are linear-uniform contracts, that pay the same amount per hectare to all landowners targeted by the policy, regardless of location. These contracts are easier to implement and become particularly interesting when political interests are at play. As Boyer and Laffont (1999) state, long-term political objectives and the presence of multiple interest groups may lead to an excessive fluctuation of policies that could be restricted by favouring simple instruments instead of separating incentive mechanisms. However, these contracts have the disadvantage of being less flexible and more costly for the regulator, as landowners can secure higher informational rents, especially when there is substantial heterogeneity in their opportunity costs. Part of this opportunity cost is private information, while part of it is public information and depends on observable local variables, such as geographical and climatic characteristics. The latter information can be used to reduce the informational rents, for example by designing spatially-targeted contracts based on these observable variables.

In this paper, I use a principal-agent model to provide insights on the relative performance of different PES contract designs: nonlinear, linear- uniform and linear spatially-targeted. To check the importance of hidden information, I compare welfare levels under these three contract designs to the first-best welfare that could be hypothetically obtained if the regulator could use the optimal contract under complete information. For this purpose, I use data on the French Grassland Conservation Program contracts created in 1993 with the goal of stopping the decrease in grassland cover. This program consisted in five-years contracts between farmers and the Government in which farmers received a linear-uniform payment for each hectare of grassland conserved. The main focus is on the reform that happened in 2003, when "*Prime au Maintien des Systémes d'Elevage Extensif*" (PMSEE) was replaced by "*Prime herbagére Agro-Environnementale*" (PHAE). Under PHAE the eligibility criteria were loosened and established at department level (contrary to national level for PMSEE) and the payment per hectare of grassland was increased from 46 to 76 Euro. To derive the results of interest, I exploit this change in the payment. I argue in Section 3.3 that this change is exogenous and not linked to a change in farmers' opportunity costs. This assumption is key for the nonparametric identification and estimation of the cost function and the distribution of types.

I find that the loss associated with the use of linear-uniform contracts instead of the nonlinear ones is small, of only 2.6%. This result suggests that linear-uniform contracts are almost efficient, in addition to being easy to implement. This conclusion contrasts the general idea that simple contracts are inefficient, but is in line with results from other studies. D'Haultfoeuille and Février (2020) find a loss of 16% in the case of contracts between the French National Institute of Statistics and Economics and the interviewers they hire, while Pollinger (2021) finds that the optimal subsidy for solar panels in Germany is close to linear, as the loss with respect to the actual piecewise linear subsidy is of only 0.016%. I also find that spatially-targeted contracts at the regional level can improve the optimality with respect to linear-uniform contracts offered at the national level by 1.9%. Overall, I find a low cost of asymmetric information, the surplus under nonlinear contracts being 87% of the one under complete information.

There are many papers that have already studied the asymmetric information issue in the context of Payment for Ecosystem Services contracts from a theoretical point of view (Wu and Babcock 1996, Ozanne et al. 2001, Crépin 2005, Arguedas and Van Soest 2011). Some of them also provide an application, such as numerical illustrations or simulations (Bourgeon et al. 1995, Fraser 1995, Moxey et al. 1999, Gren 2004, Canton et al. 2009). However, very few of them are providing a way of estimating farmers' marginal cost and the distribution of their types by imposing parametric restrictions (Sheriff 2009, Mason and Plantinga 2013). The general conclusion coming from this literature is that the more targeted a contract, the more efficient. In other words, nonlinear contracts are the most effective. For example, Mason and Plantinga (2013) find that using nonlinear contracts instead of linear-uniform ones reduces government expenditures by 60%. However, in the case of a hypothetical land retirement program and a wetland creation program, Sheriff (2009) and Crépin (2005) find that linear-uniform subsidies achieve efficiency with a small sacrifice in terms of cost relative to the nonlinear contract. My paper contributes to this literature by following a more recent approach based on nonparametric methods to identify and estimate farmers' cost function and the distribution of their types. For the identification of the principal-agent model, Perrigne and Vuong (2011, 2012) and Bontemps and Martimort (2014) assume that the policy being implemented is already optimal, while Abito (2017), D'Haultfoeuille and Février (2020) and Pollinger (2021) exploit changes in the actual policy and recover the functions of interest from the agents' program only. This paper follows closely D'Haultfoeuille and Février's (2020) approach as it is the one that fits best the policy change exploited in this paper and has the advantage of being robust to the presence of selection issues, namely whether the increase in the payments attracted different agents than the ones already in the program.

The remainder of the paper is organized as follows: Section 3.2 presents the French Grassland Conservation Program and the data I use in the empirical part. Section 3.3 sets up the theoretical model. Section 3.4 presents the identification strategy and estimates the parameters of interest. Section 3.5 compares the welfare obtained under different contract designs. Finally, Section 3.6 concludes.

3.2 The French Grassland Conservation Program

The French Grassland Conservation Program is part of a broader set of PES schemes introduced in the European Union (EU) as accompanying measures of the 1992 Common Agricultural Policy (CAP) reform. Since 2000, they became a core instrument of EU agricultural policies as part of the second pillar of the CAP. The EU budget allocated to these schemes increased from 76 million Euro in 1993 to 3.03 billion Euro in 2010 (Arata and Skokai, 2016). Subsidies for grassland conservation were included in the agri-environmental programs of several European countries, such as the German Cultural Landscape Program (KULAP), the Austrian Agri-environmental Program (OPUL), the United Kingdom's Environmental Stewardship Scheme or the Irish Rural Environment Protection Scheme (REPS) (*Institut de l'Elevage*, 2007).

In France, support to grassland conservation was created in 1993 with the goal of stopping the decrease in grassland cover (from 43% of the agricultural area in 1970 to 36% in 1988 and only 27% in 2010). The program was first called "*Prime au Maintien des Systemes d'Elevage Extensifs*" (PMSEE) and it offered five-year contracts during which farmers committed to keeping permanent grassland on the same plots. In exchange, they were paid up to 46 Euro per hectare of grassland if they met two criteria: *(i)* a specialization rate (share of permanent and temporary grassland in the total usable agricultural area) higher than 75% and *(ii)* a loading ratio (density of livestock units (LU) per hectare of forage area) below 1.4. In 1998, PMSEE was renewed for another five years, and an eligibility requirement related to the use of fertilisers was introduced: farmers were not allowed to use more than 70 kilograms of nitrogen per hectare of grassland. PMSEE was replaced in 2003 by a new grassland conservation scheme called "*Prime Herbagère Agro-Environnementale*" (PHAE). The eligibility criteria for PHAE were similar to those for PMSEE with three main exceptions. First, the thresholds for eligibility in terms of share of grassland and density of livestock units were allowed to vary at department[1] level. Some departments kept the same eligibility criteria as for PMSEE[2], while others loosened the thresholds on the specialization rate (between 50% and 75%) and on the loading ratio (between 1.4 and 1.8 LU/ha) in order to increase farmers' participation. Second, additional requirements were introduced, especially in order to limit the use of phytosanitary products and fertilizers on the plots. Finally, the payments were increased to 76 Euro per hectare of conserved grassland. It is this change in the payment scheme that I will exploit to derive the results of interest.

[1] There are 95 departments in France.

[2] 22 departments, representing 1/4 of departments with PHAE, kept the same eligibility criteria as for PMSEE.

3.2.1 Data

I use two types of data provided by the *Observatoire du Développement Rural* (ODR). First, I use administrative data from France's *Agence de Services et de Paiements* that contains information on every beneficiary of the French Grassland Conservation Program from 1999 to 2006. This dataset includes the commune of residence, the years in which farmers were enrolled in a grassland program, the number of hectares enrolled and the payment they received every year. For the analysis, only farmers that benefited from both PMSEE and PHAE and that lived in departments that kept the same eligibility criteria between the two programs were selected. The reason why these restrictions on the sample were made is to avoid the issue of self-selection of farmers into the new program and to make sure that the payment is the only change in contract conditions between the two programs.

Second, I use farm level data coming from the 2000 Agricultural Census and the 2005 Farm Structure Survey[3] conducted by the French Ministry of Agriculture. For each beneficiary was selected its total agricultural area, total hectares of grassland, total hectares of crops and number of livestock. Thus, my dataset is a panel that follows 725 beneficiaries of the French Grassland Conservation Program in 2000 and 2005. I choose to use year 2005 as the post-reform period in the analysis, as by this time all transitions from PMSEE to PHAE have been made.[4]

The descriptive statistics in Appendix 3.7.2 show that the payment received by farmers increased from PMSEE to PHAE, whereas the quantity of grassland put under contract decreased. This is counter-intuitive, given that the eligibility criteria were loosened and the amount of the subsidy increased. What actually happened is that many farmers put fewer hectares of grassland under the PHAE contract in order to keep some flexibility on other plots (Ministry of Agriculture, 2006). However, the total hectares of grassland increased between 2000 and 2005. It is this quantity that will be used to perform the estimations, under the assumption that the increase in the price of grassland under contract translates into an increase in the price of grassland not under contract. Indeed, according to Eurostat, the price of permanent grassland was rather constant around 2,500€ per hectare between 1998 and 2002 and it increased to around 4,600€ per hectare for all the 2003-2007 period.

[3]This survey is conducted on 10% of the population of farmers.

[4]In 2003, there were still some farmers that were beneficiaries of PMSEE, as they entered the 5-years contract in 1999.

3.3 Model

In this section I first model the farmers' decision of conserving grassland. The information recovered on the optimal quantity provided will then be used in Section 3.4 to identify and estimate nonparametrically the opportunity cost function and the distribution of farmers' types. Second, I model the regulator's program under different types of contract designs, both implemented without asymmetric information and under asymmetric information - nonlinear, linear-uniform and spatially-targeted contracts -. The detailed computations are found in Appendix 3.7.1. These models will then be used in Section 3.5 to perform the welfare analysis.

3.3.1 The farmers' program

A regulator is interested in the provision of grassland conservation and offers a subsidy t for each unit q of grassland conserved. The subsidy can be different between two time periods, denoted by y. Each farmer of type θ decides what quantity q of grassland conservation to provide, given its opportunity cost of provision. The farmers targeted by the policy are those that have some grassland on their farm, such that $q > 0$. θ represents the heterogeneity in cost related to unobserved farm characteristics. In reality, some of the variation in θ could be observed by the regulator, but he might choose not to use it. I refer to θ as farmers' type which is continuously distributed on $[\underline{\theta}, \bar{\theta}]$, with F_θ as cumulative distribution function (CDF) and f_θ as density function (PDF). $\underline{\theta}$ represents a low-cost farmer, in the sense that he has a low cost of conserving grassland because his unobserved characteristics make other agricultural activities more expensive on his land.

The opportunity cost function represents the difference between the profit a farmer can get without conservation and the profit he can get by conserving q hectares of grassland. It can be written as $C(q, y, p, x, \theta)$ [5], where q represents the quantity of grassland conserved, y is the year (2000 or 2005), p defines input and output prices, such as the price of fertilizers, of labor, of crops, etc., x is the municipality land quality and climatic factors and θ expresses all the unobserved farm characteristics.

To identify the opportunity cost function, some assumptions are needed in order to reduce the dimensionality of the problem.

Assumption 1 *The input and output prices and the municipality characteristics can be*

[5] Capital letters correspond to random variables, while lowercase letters correspond to the realizations of these variables.

grouped in a single category, m, that is common to all farmers living in the same municipality.

$$C(q, y, p, x, \theta) = C(q, y, m, \theta)$$

This assumption is made with the purpose of reducing the dimensionality of the problem and it states that farmers living in the same municipality are exposed to the same input and output prices and to the same geographical characteristics.

Assumption 2 *The opportunity cost function is separable between the farm-level characteristics and the common opportunity cost of grassland conservation, and is increasing in θ and increasing and convex in q (i.e. $C_\theta > 0$, $C_q > 0$, $C_{qq} > 0$).*

$$C(q, y, m, \theta) = \theta C(q, y, m)$$

Assumption 2 states that the only difference in the opportunity cost of grassland among farmers with the same characteristics is due to the unobserved heterogeneity term θ.

Assumption 3 *The common opportunity cost of grassland does not change over time between the two periods considered here (i.e. between 2000 and 2005).*

$$\theta C(q, y, m) = \theta C(q, m)$$

This assumption means that the change in the amount of subsidy that occurred with the introduction of PHAE is not related to a change in the cost of grassland conservation. In a report by the Ministry of Agriculture evaluating the effectiveness of PHAE, it is stated that the increase in payment in 2003 is actually a catch-up payment as the level of support did not change between 1995 and 2003. Moreover, according to Eurostat, agricultural input prices[6] were rather stable between 2000 and 2003 in real terms, while crop output prices[7] decreased by 6%. Therefore, it is safe to believe that the increase in payment in 2003 is not related to an increase in the opportunity cost of grassland conservation. For simplicity, in what follows y will be omitted in the notation of t and q, but the reader should keep in mind that both these variables depend on it. Moreover, without loss of generality, I will use $C(q)$ instead of $C(q, m)$.

A farmer of type θ chooses to conserve the quantity of grassland q that maximizes

[6] The agricultural inputs covered are intermediate consumption of goods and services (fertilisers, pesticides, feed, seed, energy and lubricants, maintenance and repairs, etc.).

[7] The output price indices cover crop output, excluding fruits and vegetables.

her profit:

$$\max_{q} \quad tq(\theta) - \theta C(q(\theta))$$

The optimal quantity $q^f(\theta)$ is defined by the first-order condition:

$$t = \theta C'(q^f(\theta)) \tag{3.1}$$

Using the assumption that C is convex in q, I can define the supply function $s = C'^{-1}(.)$. The solution becomes:

$$q^f(\theta) = s\left(\frac{t}{\theta}\right)$$

3.3.2 The regulator's program

In what follows, I assume that the environmental benefit of grassland conservation is the same overall France, and it captures the benefits of carbon sequestration in the soil, improved water quality, pollination, hunting and landscape amenities.[8] Thus, I assume that the regulator values a hectare of grassland conserved by a constant value B. As mentioned before, the regulator offers a payment t for each hectare of grassland conserved. Each monetary unit costs him $(1 + \lambda)$, where λ represents the opportunity cost of public funds. Therefore, when deciding what type of contract to propose to farmers, the regulator will take into account the environmental benefit of the grassland conservation, the farmers' profit and the taxpayers' surplus.

Complete information case

If the regulator had perfect information about the farmers' opportunity costs, he would implement the first-best policy by designing a contract tailored for each farmer. In this case, he would maximize the social welfare taking into account the participation constraint (PC) of farmers:

[8] There are also other environmental benefits of grassland conservation such as protection against erosion and flooding or increased biodiversity levels.

$$\max_{q,t} \quad Bq(\theta) + [tq(\theta) - \theta C(q(\theta))] - (1+\lambda)tq(\theta)$$
$$\text{s.t.} \quad tq(\theta) - \theta C(q(\theta)) \geq 0 \tag{PC}$$

As participation in the contract is voluntary, farmers will participate only if the subsidy covers their opportunity cost of providing grassland conservation. Since transfers are costly, the regulator will choose to offer the lowest possible subsidy, such that $tq(\theta) = \theta C(q(\theta))$. Using this equality, the regulator's maximization problem becomes:

$$\max_{q} \quad Bq(\theta) - (1+\lambda)\theta C(q(\theta))$$

From the first-order condition with respect to q, the quantity under complete information contracts, q^{CI}, solves the equality between the marginal environmental benefit and the marginal opportunity cost of grassland conservation:

$$\frac{B}{1+\lambda} = \theta C'(q^{CI}(\theta))$$

Therefore, the quantity and transfer under complete information are given by:

$$q^{CI}(\theta) = s\left(\frac{B}{(1+\lambda)\theta}\right)$$
$$t^{CI} = \frac{\theta C(q^{CI}(\theta))}{s\left(\frac{B}{(1+\lambda)\theta}\right)}$$

It follows that $q^{CI}(y,\theta)$ is decreasing in θ, meaning that the quantity of grassland conservation provided is higher for a low-cost farmer than for a high-cost farmer. Moreover, the transfer is such that the subsidy paid per unit of grassland covers exactly the unit opportunity cost, in both low or high cost situations.

Nonlinear contracts

As it is almost impossible for the regulator to have full knowledge of the opportunity cost of farmers, economists suggest the use of nonlinear contracts, in which the regulator proposes a menu of contracts (t,q) that should induce farmers to choose the contract intended for their type. In this case, the maximization problem should take into account both this incentive compatibility constraint (IC) and the participation constraint (PC) of farmers:

$$
\begin{aligned}
\max_{q,t} \quad & \int_{\underline{\theta}}^{\overline{\theta}} \Big[Bq(\theta) + \big(t(q(\theta)) - \theta C(q(\theta))\big) - (1+\lambda)t(q(\theta))\Big] f(\theta)\, d\theta \\
\text{s.t.} \quad & t(q(\theta)) - \theta C(q(\theta)) \geq 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (PC) \\
& t(q(\theta)) - \theta C(q(\theta)) \geq t(q(\tilde{\theta})) - \theta C(q(\tilde{\theta})), \forall \tilde{\theta} \neq \theta, \tilde{\theta} \in [\underline{\theta}, \bar{\theta}] \qquad\qquad (IC)
\end{aligned}
$$

In order to solve this problem, I define the farmers' utility as $U(\theta) = t(q(\theta)) - \theta C(q(\theta))$ and I substitute it in the maximization problem, that now depends on q and U:

$$
\max_{q,U} \quad \int_{\underline{\theta}}^{\overline{\theta}} \Big[Bq(\theta) + U(\theta) - (1+\lambda)\big(U(\theta) + \theta C(q(\theta))\big)\Big] f(\theta)\, d\theta
$$

I then use integration by parts and the incentive compatibility constraints to eliminate the terms in $U(\theta)$:

$$
\begin{aligned}
\int_{\underline{\theta}}^{\overline{\theta}} U(\theta) f(\theta)\, d\theta =& \big[U(\bar{\theta})F(\bar{\theta}) - U(\underline{\theta})F(\underline{\theta})\big] - \int_{\underline{\theta}}^{\overline{\theta}} U'(\theta)F(\theta)\, d\theta \\
=& - \int_{\underline{\theta}}^{\overline{\theta}} U'(\theta)F(\theta)\, d\theta \\
=& \int_{\underline{\theta}}^{\overline{\theta}} C(q(\theta))F(\theta)\, d\theta
\end{aligned}
$$

Finally, I replace the equality above in the maximization problem, that now depends on q only:

$$
\max_{q} \quad \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq(\theta) - \left((1+\lambda)\theta + \lambda \frac{F(\theta)}{f(\theta)}\right) C(q(\theta))\right] f(\theta)\, d\theta
$$

The first-order condition with respect to q is given by:

$$
\frac{B}{1+\lambda} = \left[\theta + \frac{\lambda}{1+\lambda} \frac{F(\theta)}{f(\theta)}\right] C'(q(\theta))
$$

where $\frac{\lambda}{1+\lambda} \frac{F(\theta)}{f(\theta)} C'(q(\theta))$ represents the information rent that the regulator has to give to farmers such that they reveal their cost type. The information rent increases with θ and leads to inefficiencies in the provision of grassland conservation compared to the complete information case.

The quantities under nonlinear contracts solve the first-order condition and are defined by:

$$q^{NL}(\theta) = s\left(\frac{B}{(1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}}\right)$$

Compared to the complete information case, for the most efficient type (i.e. the lowest-cost farmer), there is no distortion in the provision of grassland conservation as $F(\underline{\theta}) = 0$, but for all the other types there is a downward distortion in provision. Overall, this leads to fewer hectares of grassland conserved than under complete information.

The nonlinear payments are found from the combination of the two definitions of $U(\theta) = \int_{\theta}^{\bar{\theta}} C(q(\tau))\, d\tau$ and $U(\theta) = t(q(\theta)) - \theta C(q(\theta))$:

$$t^{NL}(q^{NL}(\theta)) = \int_{\theta}^{\bar{\theta}} C(q^{NL}(\tau))\, d\tau + \theta C(q^{NL}(\theta))$$

Then, the payment per hectare of grassland is given by:

$$t^{NL} = \frac{t^{NL}(q^{NL}(\theta))}{q^{NL}(\theta)} = \frac{\int_{\theta}^{\bar{\theta}} C(q^{NL}(\tau))\, d\tau + \theta C(q^{NL}(\theta))}{q^{NL}(\theta)}$$

Compared to the complete information case, nonlinear contracts have a higher cost for the regulator, as the lower-cost type farmers must be compensated at a level above their opportunity cost in order to induce them to reveal their type.

Despite the appeal of these types of contracts, to my knowledge they were never applied in practice in the context of Payments for Ecosystem Services.

Linear-uniform contracts

What regulators often propose in practice are linear-uniform contracts that offer the same payment no matter the cost type of the farmer. The main advantage of these contracts is the ease of implementation, but this comes at the cost of a loss of flexibility in allocations that are no longer type-dependent. In this case, the regulator chooses the payment and the farmers choose the optimal quantity of grassland conservation to produce, given the payment. Thus, the regulator maximizes:

$$\begin{aligned} \max_{t} \quad & \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq(\theta) + \big(tq(\theta) - \theta C(q(\theta))\big) - (1+\lambda)tq(\theta))\right] f(\theta)\, d\theta \\ \text{s.t.} \quad & t = \theta C'(q(\theta)) \qquad\qquad (Farmers'FOC) \end{aligned}$$

From the farmers' FOC, $q(\theta) = s\left(\frac{t}{\theta}\right)$. Replacing q in the maximization problem, this becomes:

$$\max_{t} \quad \int_{\underline{\theta}}^{\overline{\theta}} \left\{ Bs\left(\frac{t}{\theta}\right) + \left[ts\left(\frac{t}{\theta}\right) - \theta C\left(s\left(\frac{t}{\theta}\right)\right)\right] - (1+\lambda)ts\left(\frac{t}{\theta}\right)\right\} f(\theta)\, d\theta$$

The first-order condition with respect to t is given by:

$$\int_{\underline{\theta}}^{\overline{\theta}} \left[Bs_t\left(\frac{t}{\theta}\right) - \lambda s\left(\frac{t}{\theta}\right) - \lambda t s_t\left(\frac{t}{\theta}\right) - \theta \frac{\partial C}{\partial s} s_t\left(\frac{t}{\theta}\right)\right] f(\theta)\, d\theta = 0$$

It follows that the linear-uniform payment per hectare of grassland conserved is defined as:

$$t^{LU}(y) = \frac{B}{\lambda} - \frac{\int_{\underline{\theta}}^{\overline{\theta}} s\left(\frac{t}{\theta}\right) f(\theta)\, d\theta}{\int_{\underline{\theta}}^{\overline{\theta}} s_t\left(\frac{t}{\theta}\right) f(\theta)\, d\theta} - \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{\partial C}{\partial s} s_t\left(\frac{t}{\theta}\right) f(\theta)\, d\theta}{\lambda \int_{\underline{\theta}}^{\overline{\theta}} s_t\left(\frac{t}{\theta}\right) f(\theta)\, d\theta}$$

To induce all types of farmers to participate in the program, the regulator must offer a linear-uniform payment that covers the opportunity cost of grassland provision of the high-cost type farmers. However, this leads to potentially large excess payments to low-cost farmers, especially when there are substantial differences in the opportunity cost of grassland conservation among farmers.

Spatially-targeted contracts

A solution to the cost-efficiency - ease of implementation trade-off could be a contract that takes into account some information about the farmers' opportunity cost that the regulator has at his disposal at a low cost. For example, the regulator could still use linear-uniform payments, but differentiated on geographic characteristics. Such spatially-targeted contracts are modelled as the linear-uniform contracts above, but the integrals are taken on a specific sub-population of farmers. In the welfare analysis below, I consider a regional-targeted contract, in which farmers are paid differently depending on the region they live in.

3.4 Identification and Estimation

This section presents the nonparametric identification and estimation of the marginal opportunity cost of grassland, C', and the distribution of farmers' type , $F_{\theta|Y}$. I follow closely the method proposed by D'Haultfoeuille and Février (2020), that is based on agents' (here farmers') program only and it exploits an exogenous change in the payment structure. Here, I exploit the increase in the payment that followed the replacement of the PMSEE program with PHAE in 2003. I argued in Section 3.3 that this change in payment is not related to a change in the cost of grassland conservation.

3.4.1 Nonparametric identification

Using the farmers' first-order condition 3.1 together with the exogenous change in contracts, D'Haultfoeuille and Février (2020) show that C' and $F_{\theta|Y}$ are point identified on a sequence of points (q_k) determined by the CDF of quantities Q, $F_{Q|Y}$. In the data, I observe the payments per hectare of grassland conserved and the quantities of grassland conserved under both PMSEE (in 2000) and PHAE (in 2005) programs. Therefore, the conditional distribution of Q, $F_{Q|Y}$, is known here. Moreover, using monotonicity arguments, they show that C' and $F_{\theta|Y}$ functions can be bounded and the best nonparametric bounds can be derived (see Theorem 3.2 in D'Haultfoeuille and Février (2020)).

First, D'Haultfoeuille and Février (2020) show that if C' is identified at q, it is also identified at $H(q)$, where $H(q) = F^{-1}_{Q|Y=2005} \circ F_{Q|Y=2000}(q)$ is the quantile-quantile transform between the two CDFs of Q,

$$C'(H(q)) = \frac{t(2005)}{t(2000)} C'(q). \tag{3.2}$$

Then, they use this fact to point identify $C'(.)$ on a sequence $(q_k)_{k\in\mathbb{Z}}$ determined by the CDF $F_{Q|Y}$ as follows:

- choose a starting value for q_0;
- if $k > 0$, then $q_{k+1} = H(q_k)$;
- if $k < 0$, then $q_{k-1} = H^{-1}(q_k)$.

By induction based on 3.2, they show that:

$$C'(q_k) = \left(\frac{t(2005)}{t(2000)}\right)^k C'(q_0) = \left(\frac{t(2005)}{t(2000)}\right)^k c_0 \tag{3.3}$$

where $C'(q_0) = c_0$ is a normalization. Thus, given a starting value c_0, they point identify $C'(q_k)$ for all the k points determined by the conditional distribution of Q.

Second, using the first-order condition 3.1 and equation 3.3 above, they define farmers' type θ_k as the ratio between the initial payment and the marginal opportunity cost of grassland at each quantity q_k:

$$\theta_k = \frac{t(2000)}{C'(q_k)} \tag{3.4}$$

Finally, by the strict monotonicity of $\theta(1, q_k)$, where $\theta(1, q_k)$ is the inverse function of $q_k(\theta)$, D'Haultfoeuille and Février (2020) find that $F_{\theta|Y}$ is the complementary CDF of Q:

$$F_{\theta|Y}(\theta_k) = 1 - F_{Q|Y}(q_k) \tag{3.5}$$

In short, given the starting values q_0 and c_0, I can identify $C'(q_k)$, then θ_k and finally $F_{\theta|Y}$.

3.4.2 Nonparametric estimation

The nonparametric estimation follows the same steps as the nonparametric identification. First, the conditional distributions of the quantities Q are known and are shown in Figure 3.1. As expected, $F_{Q|2005}$ dominates stochastically $F_{Q|2000}$ on most part of the $(0, 1)$ interval. The two distributions are further used to build the $(q_k)_{k \in \mathbb{Z}}$ sequence of points as explained before, and the median of $F_{Q|2000}$ is taken as a starting value (i.e. $q_0 = 58$).[9] This leads to the estimation of 42 q_k points, 19 to the left and 23 to the right side of q_0.

Second, for the estimation of the marginal opportunity cost function C', the starting values $q_0 = 58$ (i.e. the median value of $F_{Q|2000}$) and $c_0 = 46$ (i.e. the amount of the payment in the first period, $t(2000)$) were chosen. Replacing $t(2005) = 76$€ , $t(2000) = 46$€ and $k = 42$, in formula 3.3 yields the values for which $C'(q_k)$ is point identified. Figure 3.2 shows the bounds on C'. Their intersection represents the point identified values of C'. The convex form of the marginal opportunity cost shows that the increase in the cost of conserving one more hectare of grassland is quite big for high quantities of grassland.

Finally, the estimated $C'(q_k)$ is introduced in formula 3.4 to compute θ_k and estimate its distribution function, $F_{\theta|Y}$. Figure 3.3 shows that, as expected, the distribution of types is the same between 2000 and 2005.

[9] Other starting values q_0 do not modify the choice of parametric specifications.

3.4.3 Parametric estimation

The purpose of the nonparametric estimation is to help selecting a parametric specification for the marginal opportunity cost function and the distribution of farmers' types that is data driven and that will be used to compute counterfactual policy scenarios. The nonparametric estimates of $C'(q_k)$ and $F_{\theta|Y}$ are therefore used to identify which parametric function fits best the data. I choose three possible parametric functions that are plotted against the linear approximation of the nonparametric estimates. The parametric function for which the points are aligned is retained.

I consider $C'(q) = \alpha\phi(q)^{\beta}$, with $\phi(x) = ln(x)$, $\phi(x) = x$ and $\phi(x) = exp(\sqrt{x})$. I then plot in Figure 3.4 $lnC'(q_k)$ against $ln(ln(q_k))$, $ln(q_k)$ and $ln(exp(\sqrt{q_k}))$ and their corresponding fit given by the R^2. With an R^2 of 0.985, the best fit for $C'(q_k)$ is given by $\phi(x) = x$.

Similarly, for $F_{\theta|Y}$ I consider the Weibull, the Frechet and the lognormal distributions, i.e. $F_{\theta|Y} = 1 - exp(-a\theta^{b})$, $F_{\theta|Y} = 1 - exp(-a\theta^{-b})$ and $F_{\theta|Y} = \Phi(\frac{ln\theta - a}{b})$. I then plot $ln(-ln(1 - F_{\theta|Y}))$, $ln(-ln(F_{\theta|Y}))$ and $\Phi^{-1}(F_{\theta|Y})$ against $ln\theta$. Figure 3.5 shows that with an $R^2 = 0.997$, farmers' types are lognormally distributed.

Therefore, the best parametric fit is $C'(q_k) = \alpha q_k^{\beta}$ and $F_{\theta|Y} = \Phi(\frac{ln\theta_k - a}{b})$. I then estimate by OLS the parameters (α, β, a, b) from the linear approximations below. Table 3.2 shows the estimated values.

$$lnC'(q_k) = ln\alpha + \beta ln(q_k)$$

$$\Phi^{-1}(F_{\theta|Y}) = -\frac{a}{b} + \frac{1}{b} ln\theta_k$$

3.5 Welfare Analysis

In this section, the functional form recovered for the marginal opportunity cost function is replaced in the general form of contract designs presented in Section 3.3 and the different welfare functions using data for the French Grassland Conservation Program are compared. Table 3.3 shows the formulas for the quantity, the payment and the welfare functions under the different contract designs studied using $C'(q(y,\theta)) = \alpha q(y,\theta)^{\beta}$. This section presents the main results, the detailed computation are found in Appendix 3.7.1.

To obtain the empirical results[10], I replace in the formulas from Table 3.3 the esti-

[10] As the integrals do not have closed form solutions, I approximate them using Monte Carlo Integration.

mated parameters as follows:

- $\alpha = 0.000012$ and $\beta = 3.75$: OLS estimates of the parameters of $C'(q)$ (see Table 3.2);
- $\lambda = 0.3$: estimate of the opportunity cost of public funds taken from the literature, $\lambda = [0.1, 0.5]$;[11]
- $S = 421$€/ha : estimate of the environmental benefits of grassland cover taken from the literature;[12]
- $\theta = [0; 14000]$: values for estimated types of farmers, distributed as a lognormal function with mean $a = 0.93$ and standard deviation $b = 3.57$ (see Table 3.2).

3.5.1 Results

First, as it can be seen in Figure 3.6 and as predicted by the theory, I find that the quantity of grassland conservation is higher for farmers that have a low opportunity cost of provision and decreases with the increase in the opportunity cost. Moreover, the quantities provided under linear-uniform contracts are lower than those provided under nonlinear contracts for the low-cost farmers, but higher for the high-cost farmers. Also, the overall quantities of grassland conservation are lower when there is asymmetric information.

Second, under the parameter values considered here, I find that when the regulator has complete information, the payment per hectare of grassland would be equal to 68 Euro. The actual payment per hectare of grassland after the reform was 76 Euro. The small difference in values might suggest a small asymmetry of information in the case of the French Grassland Conservation Program. Taking into account adverse selection, the payment under a linear-uniform contract would be 174 Euro per hectare of grassland, with regional disparities ranging from 104 Euro in Normandy to 250 Euro in Provence-Alpes-Cote d'Azur in case of a spatially-targeted contract.

Third, as shown in Table 3.4, I find that the welfare loss associated with the use of linear-uniform instead of nonlinear contracts is small, of only 2.6%. This result suggests that simple linear-uniform contracts are a good trade-off between efficiency and ease of implementation. This conclusion contrasts the general idea that simple contracts are inefficient, but is in line with other studies supporting the opposite claim. For example, D'Haultfoeuille and Février (2020) find a loss of 16% from using simple compensations,

[11]Estimated values for France range from 0.1 to 0.5 (Beaud, 2008)

[12]This estimate combines an estimate of the benefits of carbon sequestration in soil from Baudrier et al. (2015) and estimates of other environmental benefits from Puydarrieux and Devaux (2013).

while Pollinger (2021) find that the optimal subsidy for solar panels in Germany is close to linear. This results could explain why simple tariffs are so extensively used in practice. However, if the regulator had the possibility to offer a regional-targeted contract, this would increase the welfare by 1.9% with respect to the welfare under uniform payments. The loss compared to nonlinear contracts reduces to 0.7%.

Finally, my results show a rather small cost of asymmetric information. The welfare under nonlinear contracts represents 87% of the welfare under complete information. This loss of 13% is smaller than the loss of 22% that D'Haultfoeuille and Février (2020) find. Moreover, the welfare under linear-uniform contracts is 85% of what it could be under complete information.

3.5.2 Influence of λ

The results I presented so far are obtained using a benchmark value for the opportunity cost of public funds of 0.3. This means that each Euro of public spending costs the regulator one Euro and 30 cents. The estimated values for France range from 0.1, estimated by Bernard and Vielle (2003) to 0.5, suggested by *Commissariat Général du Plan* in 1985 (Beaud, 2008). The value of 0.3 was recommended in the report by Lebegue (2005). Using this value, I find that the difference in welfare between nonlinear and linear-uniform contracts is 2.6%, represented by the red dot in Figure 3.7. The higher the opportunity cost of public funds, the bigger the welfare loss associated with the use of linear-uniform contracts instead of the nonlinear ones, and the more interesting spatially-targeted contracts become.

3.6 Conclusion

In this paper, I use a principal-agent model to study the optimal design of the French Grassland Conservation Program. I exploit an exogenous change in the payment structure that happened in 2003 to identify and estimate nonparametrically the farmers' opportunity cost function and the distribution of their types. I select parametric specifications based on the nonparametric estimates to evaluate welfare under three contract designs: nonlinear, linear-uniform and spatially-targeted. I find that linear-uniform contracts are actually a good compromise between efficiency and ease of implementation, as the welfare loss compared to nonlinear contracts is only 2.6%. This finding can explain why the French Government opted for this type of contract design in the implementation of the Grassland Conservation Program.

The model in this paper takes into account only hidden information issues and looks only at farmers that participated in the French Grassland Conservation Program both before and after the reform studied here (i.e. *always-takers*). For a complete analysis, the model should also include hidden action and should consider also the farmers that enter the grassland program after the reform (i.e. *compliers*).

First, hidden action appears in situations when monitoring the compliance with the contract is costly for the regulator and the landowners take advantage of this by avoiding fulfilling the contract requirements. However, in the context of the French Grassland Conservation Program this is unlikely to be the case, as the control procedures were the same between the two five-years programs. More precise controls based on satellite data were introduced only after 2007.

Second, the number of farmers that enter the new program determine the participation margin. Pollinger (2021) finds that ignoring the participation margin when designing a policy biases the estimate of the intensive margin downwards. Therefore, the results I obtained in this paper could be seen as a lower bound. One way to include the participation margin in this model is by making the payments depend on the proportion of grassland in total agricultural area. Farmers receive a subsidy only if they have at least 75% of grassland on their farm, otherwise they are not eligible to the program. Thus, *compliers* pass from a subsidy equal to zero to a subsidy equal to 76 Euro per hectare of grassland, compared to the *always-takers* that move from 46 Euro to 76 Euro. The lack of information on the *compliers* before entering the program makes the identification of the marginal opportunity cost function and the distribution of their types more complicated. To my knowledge, this setting was not yet studied in the literature and I leave it for future research.

3.7 Appendix

3.7.1 Computational Details

Contract Designs

Nonlinear contracts

$$
\begin{aligned}
\max_{q,t} \quad & \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq(\theta) + \big(t(q(\theta)) - \theta C(q(\theta))\big) - (1+\lambda)t(q(\theta))\right] f(\theta)\, d\theta \\
\text{s.t.} \quad & t(q(\theta)) - \theta C(q(\theta)) \geq 0 && (PC) \\
& t(q(\theta)) - \theta C(q(\theta)) \geq t(q(\tilde{\theta})) - \theta C(q(\tilde{\theta})) && (IC)
\end{aligned}
$$

First, I define the farmers' utility as $U(\theta) = t(q(\theta)) - \theta C(q(\theta))$ and I substitute it in the maximization problem, that now depends on q and U:

$$
\begin{aligned}
\max_{q,U} \quad & \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq(\theta) + U(\theta) - (1+\lambda)\big(U(\theta) + \theta C(q(\theta))\big)\right] f(\theta)\, d\theta \\
\Rightarrow \max_{q,U} \quad & \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq(\theta) - (1+\lambda)\theta C(q(\theta)) - \lambda U(\theta)\right] f(\theta)\, d\theta
\end{aligned}
$$

Second, I use integration by parts to eliminate the term in $U(\theta)$:

$$
\begin{aligned}
\int_{\underline{\theta}}^{\overline{\theta}} U(\theta) f(\theta)\, d\theta = & U(\theta)F(\theta)\Big|_{\underline{\theta}}^{\bar{\theta}} - \int_{\underline{\theta}}^{\overline{\theta}} U'(\theta)F(\theta)\, d\theta \\
= & \left[U(\bar{\theta})F(\bar{\theta}) - U(\underline{\theta})F(\underline{\theta})\right] - \int_{\underline{\theta}}^{\overline{\theta}} U'(\theta)F(\theta)\, d\theta \\
= & - \int_{\underline{\theta}}^{\overline{\theta}} U'(\theta)F(\theta)\, d\theta
\end{aligned}
$$

as $U(\bar{\theta}) = 0$ since a high-cost farmer will participate in the contract if he receives a compensation that covers his opportunity cost (i.e. the PC of the high-cost type should be binding) and $F(\underline{\theta}) = 0$.

Third, I use the ICs to substitute for $U'(\theta)$:

- the IC can be written as: $U(\theta) \geq U(\tilde{\theta}) + \tilde{\theta}C(q(\tilde{\theta})) - \theta C(q(\tilde{\theta}))$, where $U(\tilde{\theta}) + \tilde{\theta}C(q(\tilde{\theta})) = t(q(\tilde{\theta}))$
- the maximization problem of a farmer that chooses to announce $\tilde{\theta}$ when his true cost

type is θ is:

$$\max_{\tilde{\theta}} \quad t(q(\tilde{\theta})) - \theta C(q(\tilde{\theta}))$$

and the first-order condition is given by:

$$\frac{\partial t}{\partial q}\frac{\partial q}{\partial \tilde{\theta}} - \theta \frac{\partial c}{\partial q}\frac{\partial q}{\partial \tilde{\theta}} = 0$$

- IC implies that this FOC is satisfied when $\tilde{\theta} = \theta$:

$$\frac{\partial t}{\partial q}\frac{\partial q}{\partial \theta} - \theta \frac{\partial c}{\partial q}\frac{\partial q}{\partial \theta} = 0$$

- differentiating $U(\theta) = t(q(\theta)) - \theta C(q(\theta))$ with respect to θ gives:

$$\begin{aligned} U'(\theta) =& \frac{\partial t}{\partial q}\frac{\partial q}{\partial \theta} - \theta \frac{\partial c}{\partial q}\frac{\partial q}{\partial \theta} - C(q(\theta)) \\ =& - C(q(\theta)) \end{aligned}$$

as the first two terms equal zero from the previous first-order condition.

- replacing $U'(\theta) = -C(q(\theta))$ in the integration by parts formula, I obtain:

$$\int_{\underline{\theta}}^{\overline{\theta}} U(\theta) f(\theta)\, d\theta = \int_{\underline{\theta}}^{\overline{\theta}} C(q(\theta)) F(\theta)\, d\theta$$

Finally, I replace the term in $U(\theta)$ in the maximization problem that now depends on q only:

$$\begin{aligned} &\max_{q} \quad \int_{\underline{\theta}}^{\overline{\theta}} \left[\left(Bq(\theta) - (1+\lambda)\theta C(q(\theta)) \right) f(\theta) - \lambda C(q(\theta)) F(\theta) \right] d\theta \\ \Rightarrow &\max_{q} \quad \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq(\theta) - (1+\lambda)\theta C(q(\theta)) - \lambda \frac{F(\theta)}{f(\theta)} C(q(\theta)) \right] f(\theta)\, d\theta \\ \Rightarrow &\max_{q} \quad \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq(\theta) - \left((1+\lambda)\theta + \lambda \frac{F(\theta)}{f(\theta)} \right) C(q(\theta)) \right] f(\theta)\, d\theta \end{aligned}$$

The first-order condition with respect to q is given by:

$$\begin{aligned}
&\int_{\underline{\theta}}^{\bar{\theta}} \left[B - \left((1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}\right) C'(q(\theta))\right] f(\theta)\, d\theta = 0 \\
\Rightarrow &\int_{\underline{\theta}}^{\bar{\theta}} Bf(\theta)\, d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \left[(1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}\right] C'(q(\theta)) f(\theta)\, d\theta \\
\Rightarrow &B = \left[(1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}\right] C'(q(\theta)) \\
\Rightarrow &\frac{B}{1+\lambda} = \left[\theta + \frac{\lambda}{1+\lambda}\frac{F(\theta)}{f(\theta)}\right] C'(q(\theta))
\end{aligned}$$

From the first-order condition, $q^{NL}(\theta)$ is defined as:

$$\boxed{q^{NL}(\theta) = C'^{-1}\left(\frac{B}{(1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}}\right) = s\left(\frac{B}{(1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}}\right)}$$

Then, integrating $U'(\theta) = -C(q(\theta))$ on both sides, I obtain:

$$\begin{aligned}
U(\bar{\theta}) - U(\theta) &= -\int_{\theta}^{\bar{\theta}} C(q(\tau))\, d\tau \\
\Rightarrow U(\theta) &= \int_{\theta}^{\bar{\theta}} C(q(\tau))\, d\tau
\end{aligned}$$

as $U(\bar{\theta}) = 0$ from the binding participation constraint of the high-cost farmer.

Also, by definition, $U(\theta) = t(q(\theta)) - \theta C(q(\theta))$. Then, combining the two definitions of $U(\theta)$ I obtain:

$$\begin{aligned}
\int_{\theta}^{\bar{\theta}} C(q(\tau))\, d\tau &= t(q(\theta)) - \theta C(q(\theta)) \\
\Rightarrow t^{NL}(q^{NL}(\theta)) &= \int_{\theta}^{\bar{\theta}} C(q^{NL}(\tau))\, d\tau + \theta C(q^{NL}(\theta))
\end{aligned}$$

$t^{NL}(q^{NL}(\theta))$ represents the total payment a farmer receives. Thus, the payment per hectare of grassland conserved is given by:

$$\boxed{t^{NL} = \frac{t^{NL}(q^{NL}(\theta))}{q^{NL}(\theta)} = \frac{\int_{\theta}^{\bar{\theta}} C(q^{NL}(\tau))\, d\tau + \theta C(q^{NL}(\theta))}{q^{NL}(\theta)}}$$

Linear-uniform contracts

$$\max_{t} \quad \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq(\theta) + \big(tq(\theta) - \theta C(q(\theta))\big) - (1+\lambda)tq(\theta)\right] f(\theta)\, d\theta$$
$$\text{s.t.} \quad t = \theta C'(q(\theta)) \qquad\qquad (Farmers' FOC)$$

From the farmers' FOC, $q^{LU}(\theta) = s(\frac{t^{LU}}{\theta})$. I replace it in the maximization problem to obtain:

$$\max_{t} \quad \int_{\underline{\theta}}^{\overline{\theta}} \left[Bs\left(\frac{t}{\theta}\right) + \left(ts\left(\frac{t}{\theta}\right) - \theta C\left(s\left(\frac{t}{\theta}\right)\right)\right) - (1+\lambda)ts\left(\frac{t}{\theta}\right)\right] f(\theta)\, d\theta$$
$$\Rightarrow \max_{t} \quad \int_{\underline{\theta}}^{\overline{\theta}} \left[Bs\left(\frac{t}{\theta}\right) - \lambda ts\left(\frac{t}{\theta}\right) - \theta C\left(s\left(\frac{t}{\theta}\right)\right)\right] f(\theta)\, d\theta$$

The first-order condition with respect to t is given by:

$$\int_{\underline{\theta}}^{\overline{\theta}} \left[Bs_t\left(\frac{t}{\theta}\right) - \lambda s\left(\frac{t}{\theta}\right) - \lambda ts_t\left(\frac{t}{\theta}\right) - \theta\frac{\partial C}{\partial s}s_t\left(\frac{t}{\theta}\right)\right] f(\theta)\, d\theta = 0$$
$$\Rightarrow B\int_{\underline{\theta}}^{\overline{\theta}} s_t\left(\frac{t}{\theta}\right) f(\theta)\, d\theta - \lambda\int_{\underline{\theta}}^{\overline{\theta}} s\left(\frac{t}{\theta}\right) f(\theta)\, d\theta - \int_{\underline{\theta}}^{\overline{\theta}} \theta\frac{\partial C}{\partial s}s_t\left(\frac{t}{\theta}\right) f(\theta)\, d\theta = \lambda t\int_{\underline{\theta}}^{\overline{\theta}} s_t\left(\frac{t}{\theta}\right) f(\theta)\, d\theta$$

From the first-order condition it follows that t^{LU} is defined as:

$$t^{LU} = \frac{B\int_{\underline{\theta}}^{\overline{\theta}} s_t(\frac{t}{\theta})f(\theta)\,d\theta}{\lambda\int_{\underline{\theta}}^{\overline{\theta}} s_t(\frac{t}{\theta})f(\theta)\,d\theta} - \frac{\lambda\int_{\underline{\theta}}^{\overline{\theta}} s(\frac{t}{\theta})f(\theta)\,d\theta}{\lambda\int_{\underline{\theta}}^{\overline{\theta}} s_t(\frac{t}{\theta})f(\theta)\,d\theta} - \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta\frac{\partial C}{\partial s}s_t(\frac{t}{\theta})f(\theta)\,d\theta}{\lambda\int_{\underline{\theta}}^{\overline{\theta}} s_t(\frac{t}{\theta})f(\theta)\,d\theta}$$
$$\Rightarrow \boxed{t^{LU} = \frac{B}{\lambda} - \frac{\int_{\underline{\theta}}^{\overline{\theta}} s(\frac{t}{\theta})f(\theta)\,d\theta}{\int_{\underline{\theta}}^{\overline{\theta}} s_t(\frac{t}{\theta})f(\theta)\,d\theta} - \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta\frac{\partial C}{\partial s}s_t(\frac{t}{\theta})f(\theta)\,d\theta}{\lambda\int_{\underline{\theta}}^{\overline{\theta}} s_t(\frac{t}{\theta})f(\theta)\,d\theta}}$$

Then, $\boxed{q^{LU}(\theta) = s(\frac{t^{LU}}{\theta})}$

Welfare Analysis

To perform the welfare analysis on the French Grassland Conservation Program, I compute in this section the specific welfare functions by replacing the parametric specification of the marginal opportunity cost in the general results derived in the previous section. Thus, I will extensively use the following functional forms:

$$\begin{aligned} C'(q(\theta)) &= \alpha q(\theta)^{\beta} \\ \Rightarrow C(q(\theta)) &= \alpha \frac{q(\theta)^{(\beta+1)}}{\beta+1} \end{aligned}$$

Complete information case

1. The quantity

$$\begin{aligned} \frac{B}{1+\lambda} &= \theta C'(q^{CI}(\theta)) \\ \Rightarrow \frac{B}{1+\lambda} &= \theta \alpha q^{CI}(\theta)^{\beta} \\ \Rightarrow & \boxed{q^{CI}(\theta) = \sqrt[\beta]{\frac{B}{(1+\lambda)\theta\alpha}}} \end{aligned}$$

2. The payment

$$\begin{aligned} t^{CI} &= \frac{\theta C(q^{CI}(\theta))}{q^{CI}(\theta)} \\ \Rightarrow t^{CI} &= \frac{\theta\alpha \frac{q^{CI}(\theta)^{\beta+1}}{\beta+1}}{q^{CI}(\theta)} \\ \Rightarrow t^{CI} &= \frac{\theta\alpha}{\beta+1} q^{CI}(\theta)^{\beta+1} \frac{1}{q^{CI}(\theta)} \\ \Rightarrow t^{CI} &= \frac{\theta\alpha}{\beta+1} q^{CI}(\theta)^{\beta} \\ \Rightarrow t^{CI} &= \frac{\theta\alpha}{\beta+1} \frac{B}{(1+\lambda)\theta\alpha} \\ \Rightarrow & \boxed{t^{CI} = \frac{B}{(1+\lambda)(1+\beta)}} \end{aligned}$$

3. The welfare function

$$W^{CI} = \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq^{CI}(\theta) - (1+\lambda)\theta C(q^{CI}(\theta))\right] f(\theta)\, d\theta$$

$$\Rightarrow W^{CI} = \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq^{CI}(\theta) - (1+\lambda)\theta\alpha \frac{q^{CI}(\theta)^{\beta+1}}{\beta+1}\right] f(\theta)\, d\theta$$

$$\Rightarrow W^{CI} = \int_{\underline{\theta}}^{\overline{\theta}} \left[B\left(\frac{B}{(1+\lambda)\theta\alpha}\right)^{\frac{1}{\beta}} - \frac{(1+\lambda)\theta\alpha}{\beta+1}\left(\frac{B}{(1+\lambda)\theta\alpha}\right)^{\frac{\beta+1}{\beta}}\right] f(\theta)\, d\theta$$

$$\Rightarrow W^{CI} = \int_{\underline{\theta}}^{\overline{\theta}} \left[\frac{B^{\frac{\beta+1}{\beta}}}{((1+\lambda)\theta\alpha)^{\frac{1}{\beta}}} - \frac{B^{\frac{\beta+1}{\beta}}}{(\beta+1)[(1+\lambda)\theta\alpha]^{\frac{1}{\beta}}}\right] f(\theta)\, d\theta$$

$$\Rightarrow W^{CI} = \int_{\underline{\theta}}^{\overline{\theta}} \frac{\beta B^{\frac{\beta+1}{\beta}}}{(\beta+1)[(1+\lambda)\theta\alpha]^{\frac{1}{\beta}}} f(\theta)\, d\theta$$

$$\Rightarrow \boxed{W^{CI} = \frac{\beta B^{\frac{\beta+1}{\beta}}}{(\beta+1)[(1+\lambda)\alpha]^{\frac{1}{\beta}}} \int_{\underline{\theta}}^{\overline{\theta}} \theta^{-\frac{1}{\beta}} f(\theta)\, d\theta}$$

Nonlinear contracts

1. The quantity

$$B = \left[(1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}\right] C'(q^{NL}(\theta))$$

$$\Rightarrow B = \left[(1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}\right] \alpha q^{NL}(\theta)^{\beta}$$

$$\Rightarrow \boxed{q^{NL}(\theta) = \sqrt[\beta]{\frac{B}{\alpha\left[(1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}\right]}}}$$

2. The payment

$$
\begin{aligned}
t^{NL} &= \frac{\int_{\theta}^{\bar{\theta}} C(q^{NL}(\tau))\,d\tau + \theta C(q^{NL}(\theta))}{q^{NL}(\theta)} \\
\Rightarrow t^{NL} &= \frac{\int_{\theta}^{\bar{\theta}} \alpha \frac{q^{NL}(\tau)^{\beta+1}}{\beta+1}\,d\tau + \theta\alpha\frac{q^{NL}(\theta)^{\beta+1}}{\beta+1}}{q^{NL}(\theta)} \\
\Rightarrow t^{NL} &= \frac{1}{\left(\frac{B}{\alpha\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]}\right)^{\frac{1}{\beta}}} \int_{\theta}^{\bar{\theta}} \frac{\alpha\left(\frac{B}{\alpha\left[(1+\lambda)\tau+\lambda\frac{F(\tau)}{f(\tau)}\right]}\right)^{\frac{\beta+1}{\beta}}}{\beta+1}\,d\tau + \frac{\theta\alpha}{\beta+1}\frac{B}{\alpha\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]} \\
\Rightarrow t^{NL} &= \frac{\left(\alpha\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]\right)^{\frac{1}{\beta}}}{B^{\frac{1}{\beta}}} \frac{B^{\frac{\beta+1}{\beta}}}{(\beta+1)\alpha^{\frac{1}{\beta}}} \int_{\theta}^{\bar{\theta}} \left(\frac{1}{(1+\lambda)\tau+\lambda\frac{F(\tau)}{f(\tau)}}\right)^{\frac{\beta+1}{\beta}} d\tau + \frac{\theta B}{(\beta+1)\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]} \\
\Rightarrow & \boxed{t^{NL} = \frac{B\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]^{\frac{1}{\beta}}}{\beta+1} \int_{\theta}^{\bar{\theta}} \left(\frac{1}{(1+\lambda)\tau+\lambda\frac{F(\tau)}{f(\tau)}}\right)^{\frac{\beta+1}{\beta}} d\tau + \frac{\theta B}{(\beta+1)\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]}}
\end{aligned}
$$

3. The welfare function

$$
\begin{aligned}
W^{NL} &= \int_{\underline{\theta}}^{\bar{\theta}} \left[Bq^{NL}(\theta) - \left((1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right)C(q^{NL}(\theta))\right] f(\theta)\,d\theta \\
\Rightarrow W^{NL} &= \int_{\underline{\theta}}^{\bar{\theta}} \left[Bq^{NL}(\theta) - \left((1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right)\alpha\frac{q^{NL}(\theta)^{\beta+1}}{\beta+1}\right] f(\theta)\,d\theta \\
\Rightarrow W^{NL} &= \int_{\underline{\theta}}^{\bar{\theta}} \left[B\left(\frac{B}{\alpha\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]}\right)^{\frac{1}{\beta}} - \left((1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right)\frac{\alpha}{\beta+1}\left(\frac{B}{\alpha\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]}\right)^{\frac{\beta+1}{\beta}}\right] f(\theta)\,d\theta \\
\Rightarrow W^{NL} &= \int_{\underline{\theta}}^{\bar{\theta}} \left[\frac{B^{\frac{\beta+1}{\beta}}}{\left[\alpha\left((1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right)\right]^{\frac{1}{\beta}}} - \frac{1}{\beta+1}\frac{B^{\frac{\beta+1}{\beta}}}{\left[\alpha\left((1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right)\right]^{\frac{1}{\beta}}}\right] f(\theta)\,d\theta \\
\Rightarrow W^{NL} &= \int_{\underline{\theta}}^{\bar{\theta}} \frac{\beta B^{\frac{\beta+1}{\beta}}}{(\beta+1)\left[\alpha\left((1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right)\right]^{\frac{1}{\beta}}} f(\theta)\,d\theta \\
\Rightarrow & \boxed{W^{NL} = \frac{\beta B^{\frac{\beta+1}{\beta}}}{(\beta+1)\alpha^{\frac{1}{\beta}}} \int_{\underline{\theta}}^{\bar{\theta}} \left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]^{-\frac{1}{\beta}} f(\theta)\,d\theta}
\end{aligned}
$$

Linear-uniform contracts

1. The payment

$$
\begin{aligned}
t^{LU} &= \theta C'(q^{LU}(\theta)) \\
\Rightarrow t^{LU} &= \theta\alpha q^{LU}(\theta)^{\beta} \\
\Rightarrow q^{LU}(\theta) &= s\left(\frac{t^{LU}}{\theta}\right) = \left(\frac{t^{LU}}{\theta\alpha}\right)^{\frac{1}{\beta}}
\end{aligned}
$$

Then,

$$
\begin{aligned}
t^{LU} &= \frac{B}{\lambda} - \frac{\int_{\underline{\theta}}^{\overline{\theta}} s(\frac{t}{\theta}) f(\theta)\, d\theta}{\int_{\underline{\theta}}^{\overline{\theta}} s_t(\frac{t}{\theta}) f(\theta)\, d\theta} - \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \frac{\partial C}{\partial s} s_t(\frac{t}{\theta}) f(\theta)\, d\theta}{\lambda \int_{\underline{\theta}}^{\overline{\theta}} s_t(\frac{t}{\theta}) f(\theta)\, d\theta} \\
\Rightarrow t^{LU} &= \frac{B}{\lambda} - \frac{\int_{\underline{\theta}}^{\overline{\theta}} \left(\frac{t^{LU}}{\theta\alpha}\right)^{\frac{1}{\beta}} f(\theta)\, d\theta}{\int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{\beta}\left(\frac{1}{\theta\alpha}\right)^{\frac{1}{\beta}} t^{LU^{\frac{1-\beta}{\beta}}} f(\theta)\, d\theta} - \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta\alpha \frac{1}{\beta}\left(\frac{1}{\theta\alpha}\right)^{\frac{\beta+1}{\beta}} t^{LU^{\frac{1}{\beta}}} f(\theta)\, d\theta}{\lambda \int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{\beta}\left(\frac{1}{\theta\alpha}\right)^{\frac{1}{\beta}} t^{LU^{\frac{1-\beta}{\beta}}} f(\theta)\, d\theta} \\
\Rightarrow t^{LU} &= \frac{B}{\lambda} - \frac{t^{LU^{\frac{1}{\beta}}} \int_{\underline{\theta}}^{\overline{\theta}} \left(\frac{1}{\theta\alpha}\right)^{\frac{1}{\beta}} f(\theta)\, d\theta}{\frac{1}{\beta} t^{LU^{\frac{1-\beta}{\beta}}} \int_{\underline{\theta}}^{\overline{\theta}} \left(\frac{1}{\theta\alpha}\right)^{\frac{1}{\beta}} f(\theta)\, d\theta} - \frac{\frac{1}{\beta} t^{LU^{\frac{1}{\beta}}} \int_{\underline{\theta}}^{\overline{\theta}} \left(\frac{1}{\theta\alpha}\right)^{\frac{1}{\beta}} f(\theta)\, d\theta}{\lambda \frac{1}{\beta} t^{LU^{\frac{1-\beta}{\beta}}} \int_{\underline{\theta}}^{\overline{\theta}} \left(\frac{1}{\theta\alpha}\right)^{\frac{1}{\beta}} f(\theta)\, d\theta} \\
\Rightarrow t^{LU} &= \frac{B}{\lambda} - \beta t^{LU} - \frac{1}{\lambda} t^{LU} \\
\Rightarrow t^{LU} + \beta t^{LU} + \frac{1}{\lambda} t^{LU} &= \frac{B}{\lambda} \\
\Rightarrow \left(\frac{\lambda + \lambda\beta + 1}{\lambda}\right) t^{LU} &= \frac{B}{\lambda} \\
\Rightarrow & \boxed{t^{LU} = \frac{B}{(\beta+1)\lambda + 1}}
\end{aligned}
$$

2. The quantity

$$
\begin{aligned}
q^{LU}(\theta) &= \left(\frac{t^{LU}}{\theta\alpha}\right)^{\frac{1}{\beta}} \\
\Rightarrow q^{LU}(\theta) &= \left(\frac{1}{\theta\alpha} \frac{B}{(\beta+1)\lambda + 1}\right)^{\frac{1}{\beta}} \\
\Rightarrow & \boxed{q^{LU}(\theta) = \sqrt[\beta]{\frac{B}{((\beta+1)\lambda + 1)\theta\alpha}}}
\end{aligned}
$$

3. The welfare function

$$
\begin{aligned}
W^{LU} &= \int_{\underline{\theta}}^{\overline{\theta}} \left[Bs\left(\frac{t}{\theta}\right) - \lambda ts\left(\frac{t}{\theta}\right) - \theta C\left(s\left(\frac{t}{\theta}\right)\right)\right] f(\theta)\, d\theta \\
\Rightarrow W^{LU} &= \int_{\underline{\theta}}^{\overline{\theta}} \left[Bq^{LU}(\theta) - \lambda t^{LU} q^{LU}(\theta) - \theta\alpha \frac{q^{LU}(\theta)^{\beta+1}}{\beta+1}\right] f(\theta)\, d\theta \\
\Rightarrow W^{LU} &= \int_{\underline{\theta}}^{\overline{\theta}} \left[B\left(\frac{B}{((\beta+1)\lambda+1)\theta\alpha}\right)^{\frac{1}{\beta}} - \lambda \frac{B}{(\beta+1)\lambda+1}\left(\frac{B}{((\beta+1)\lambda+1)\theta\alpha}\right)^{\frac{1}{\beta}} - \right. \\
&\quad \left. - \frac{\theta\alpha}{\beta+1}\left(\frac{B}{((\beta+1)\lambda+1)\theta\alpha}\right)^{\frac{\beta+1}{\beta}}\right] f(\theta)\, d\theta \\
\Rightarrow W^{LU} &= \int_{\underline{\theta}}^{\overline{\theta}} \left[\frac{B^{\frac{\beta+1}{\beta}}}{\left[\left((\beta+1)\lambda+1\right)\theta\alpha\right]^{\frac{1}{\beta}}} - \lambda \frac{B^{\frac{\beta+1}{\beta}}}{\left((\beta+1)\lambda+1\right)^{\frac{\beta+1}{\beta}}(\theta\alpha)^{\frac{1}{\beta}}} - \right. \\
&\quad \left. - \frac{B^{\frac{\beta+1}{\beta}}}{(\beta+1)\left((\beta+1)\lambda+1\right)^{\frac{\beta+1}{\beta}}(\theta\alpha)^{\frac{1}{\beta}}}\right] f(\theta)\, d\theta \\
\Rightarrow W^{LU} &= \int_{\underline{\theta}}^{\overline{\theta}} \left[\frac{B^{\frac{\beta+1}{\beta}}}{\left[\left((\beta+1)\lambda+1\right)\theta\alpha\right]^{\frac{1}{\beta}}} - \frac{(\beta+1)\lambda+1}{\beta+1} \frac{B^{\frac{\beta+1}{\beta}}}{\left((\beta+1)\lambda+1\right)^{\frac{\beta+1}{\beta}}(\theta\alpha)^{\frac{1}{\beta}}}\right] f(\theta)\, d\theta \\
\Rightarrow W^{LU} &= \int_{\underline{\theta}}^{\overline{\theta}} \left[\frac{B^{\frac{\beta+1}{\beta}}}{\left[\left((\beta+1)\lambda+1\right)\theta\alpha\right]^{\frac{1}{\beta}}} - \frac{B^{\frac{\beta+1}{\beta}}}{(\beta+1)\left[\left((\beta+1)\lambda+1\right)\theta\alpha\right]^{\frac{1}{\beta}}}\right] f(\theta)\, d\theta \\
\Rightarrow W^{LU} &= \int_{\underline{\theta}}^{\overline{\theta}} \frac{\beta B^{\frac{\beta+1}{\beta}}}{(\beta+1)\left[\left((\beta+1)\lambda+1\right)\theta\alpha\right]^{\frac{1}{\beta}}} f(\theta)\, d\theta \\
\Rightarrow & \boxed{W^{LU} = \frac{\beta B^{\frac{\beta+1}{\beta}}}{(\beta+1)\left[\left((\beta+1)\lambda+1\right)\alpha\right]^{\frac{1}{\beta}}} \int_{\underline{\theta}}^{\overline{\theta}} \theta^{-\frac{1}{\beta}} f(\theta)\, d\theta}
\end{aligned}
$$

3.7.2 Data

Table 3.1: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Median	Min	Max
payment_pmsee	725	2,570.910	1,547.568	2,368.587	140.187	13,720.420
payment_phae	725	4,882.506	3,473.245	4,416.500	306.000	23,794.020
quantity_pmsee	725	85.331	107.293	57.120	5.060	1,317.260
quantity_phae	725	75.860	76.792	57.040	4.200	1,321.900
total_grassland_00	725	94.265	133.079	58.000	2.150	1,409.000
total_grassland_05	725	109.188	158.051	66.510	1.080	1,600.000
share_grassland_contract_pmsee	725	107.990	109.297	97.097	9.246	1,726.142
share_grassland_contract_phae	725	92.602	75.324	88.911	7.888	1,420.833
agric_area_00	725	106.379	136.150	70.000	3.010	1,409.000
agric_area_05	725	120.396	160.790	75.0	1	1,600
share_grassland_00	725	86.717	13.964	89.091	9.849	100.000
share_grassland_05	725	88.706	11.970	90.826	13.614	100.000
share_crops_00	725	6.748	8.293	5.171	0.000	90.151
share_crops_05	725	5.827	7.241	3.746	0.000	61.828
ugb_00	725	69.382	51.093	57.800	0.000	369.200
ugb_05	725	71.651	55.161	59.400	0.000	359.600
loading_ratio_00	725	8.096	32.129	2.671	0.000	660.000
loading_ratio_05	722	9.254	71.351	2.295	0.000	1,860.000

3.7.3 Figures

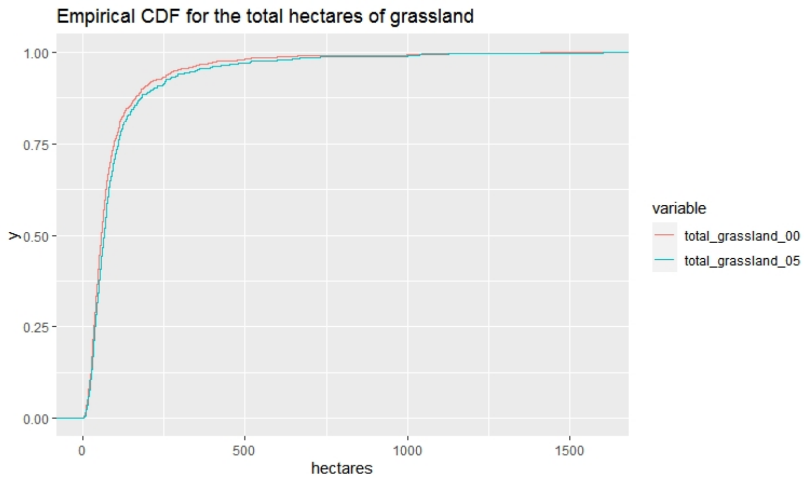


Figure 3.1: Empirical CDF of the total hectares of grassland in 2000 and 2005.

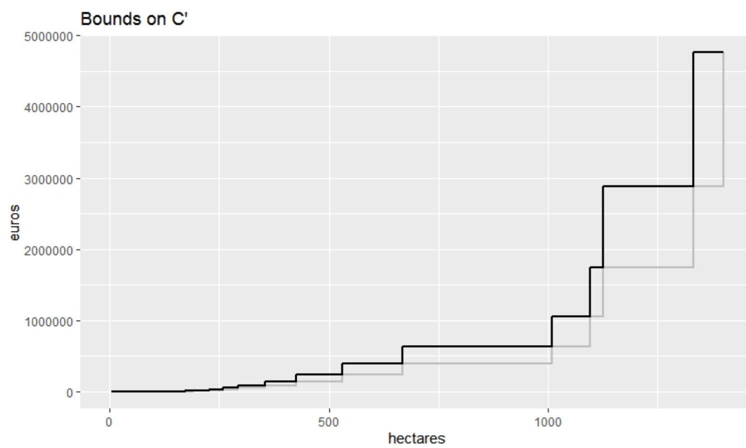
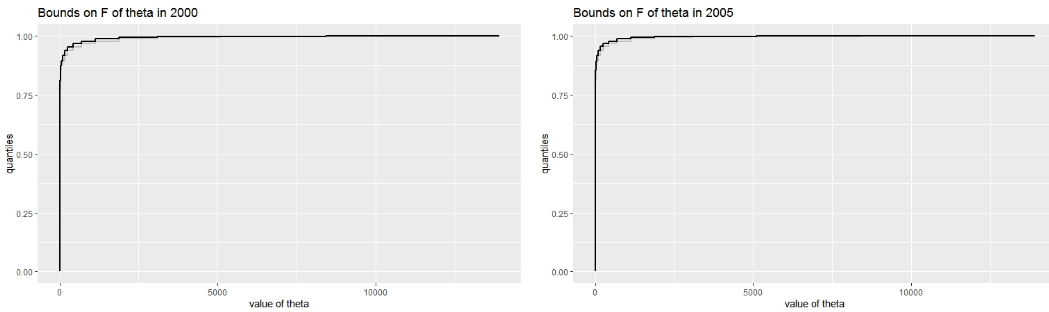


Figure 3.2: Nonparametric estimation of C'.



(a) Estimated bounds on $F_{\theta|2000}$ (b) Estimated bounds on $F_{\theta|2005}$

Figure 3.3: Nonparametric estimation of F_{θ}.

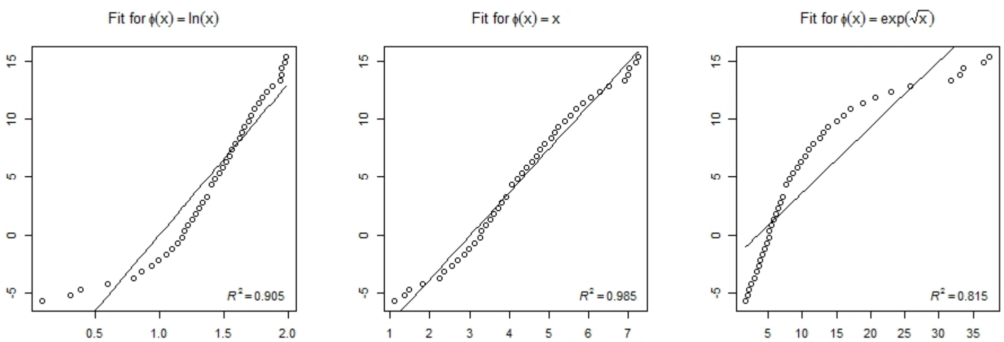


Figure 3.4: Choice of parametric function for C'.

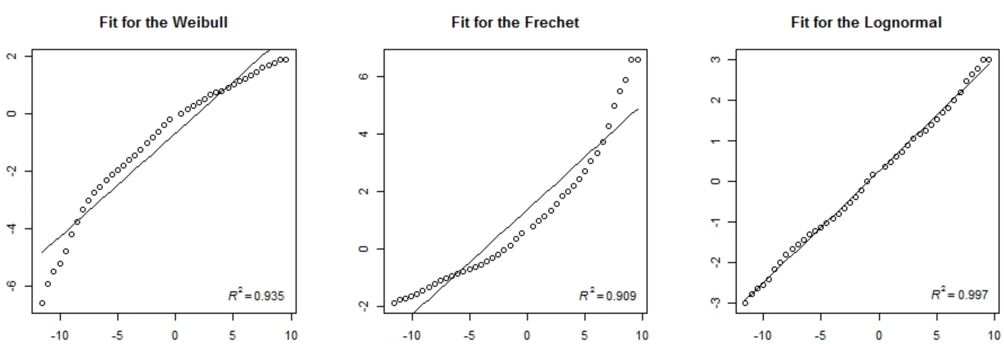


Figure 3.5: Choice of parametric function for $F_{\theta|Y}$.

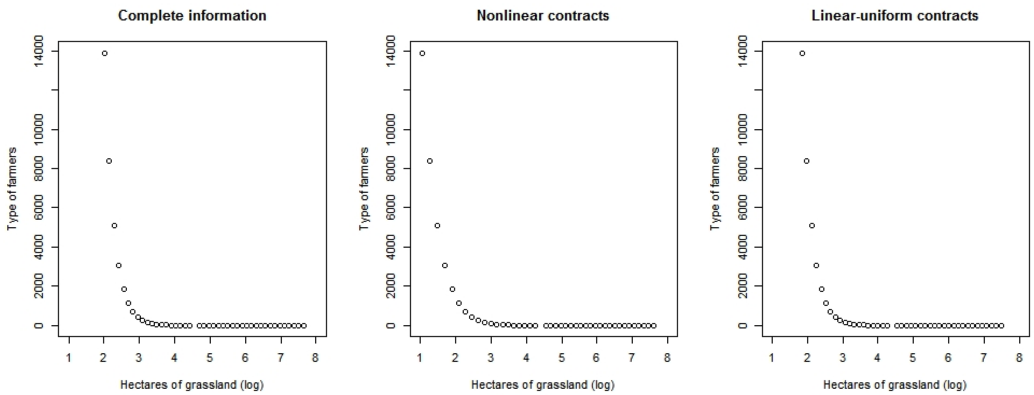


Figure 3.6: Estimated quantities (in logs) of grassland conservation per type of farmer, per contract design.

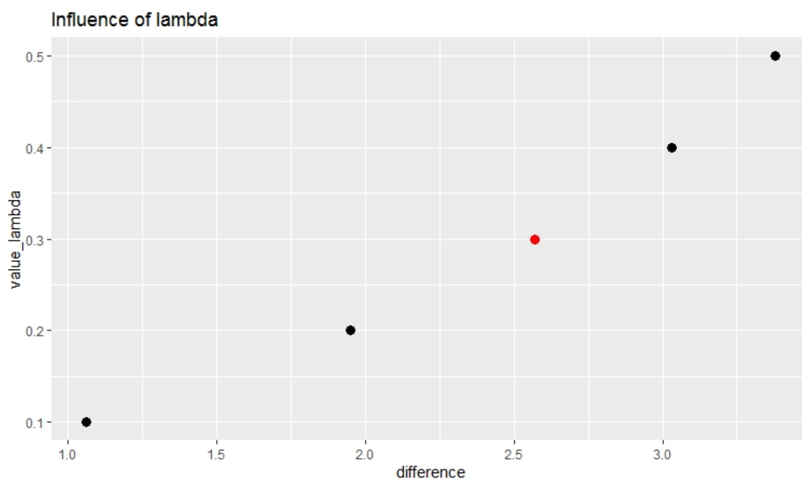


Figure 3.7: Influence of the value of λ on the difference (%) in welfare between nonlinear and linear-uniform contracts.

3.7.4 Tables

Table 3.2: OLS estimates of the parameters of $C'(q)$ and F_θ

Parameter	OLS estimate
α	0.000012
	(0.000004)
β	3.75
	(0.07)
a	-0.93
	(0.07)
b	3.57
	(0.03)
Observations	**725**

Note: OLS estimation. The standard errors in parenthesis are computed using the delta method.

Table 3.3: Contract design solutions using $C'(q(y,\theta)) = \alpha q(y,\theta)^{\beta}$

Contract design	**The quantity**	**The payment**	**The welfare function**
Complete information	$q^{CI}(y,\theta) = \sqrt[\beta]{\frac{B}{(1+\lambda)\theta\alpha}}$	$t^{CI}(y) = \frac{B}{(1+\lambda)(1+\beta)}$	$W^{CI} = \frac{\beta B^{\frac{\beta+1}{\beta}}}{(\beta+1)[(1+\lambda)\alpha]^{\frac{1}{\beta}}} \int_{\underline{\theta}}^{\overline{\theta}} \theta^{-\frac{1}{\beta}} f(\theta)\, d\theta$
Nonlinear	$q^{NL}(y,\theta) = \sqrt[\beta]{\frac{B}{\alpha\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]}}$	$t^{NL}(y) = \frac{B\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]^{\frac{1}{\beta}}}{\beta+1} \int_{\theta}^{\bar{\theta}} \left(\frac{1}{(1+\lambda)\tau+\lambda\frac{F(\tau)}{f(\tau)}}\right)^{\frac{\beta+1}{\beta}} d\tau + \frac{\theta B}{(\beta+1)\left[(1+\lambda)\theta+\lambda\frac{F(\theta)}{f(\theta)}\right]}$	$W^{NL} = \frac{\beta B^{\frac{\beta+1}{\beta}}}{(\beta+1)\alpha^{\frac{1}{\beta}}} \int_{\underline{\theta}}^{\overline{\theta}} \left[(1+\lambda)\theta + \lambda\frac{F(\theta)}{f(\theta)}\right]^{-\frac{1}{\beta}} f(\theta)\, d\theta$
Linear-uniform	$q^{LU}(y,\theta) = \sqrt[\beta]{\frac{B}{\left((\beta+1)\lambda+1\right)\theta\alpha}}$	$t^{LU}(y) = \frac{B}{(\beta+1)\lambda+1}$	$W^{LU} = \frac{\beta B^{\frac{\beta+1}{\beta}}}{(\beta+1)\left[\left((\beta+1)\lambda+1\right)\alpha\right]^{\frac{1}{\beta}}} \int_{\underline{\theta}}^{\overline{\theta}} \theta^{-\frac{1}{\beta}} f(\theta)\, d\theta$

Note: The formulas for the quantity, the payment and the welfare functions under the different contract designs studied, using the identified functional form, $C'(q(y,\theta)) = \alpha q(y,\theta)^{\beta}$. The formulas used for the spatially-targeted contract design are the same as for the linear-uniform contract design, applied to each region.

Table 3.4: Estimated welfare per type of contract design

Contract design	**Estimated welfare (in €)**	**95% Confidence interval**
Complete information	37,725	[35,336 , 40,386]
Nonlinear contract	32,789	[30,512 , 35,272]
Linear-uniform contract	31,946	[29,924 , 34,200]
Spatially-targeted contract	32,564	[30,843 , 34,757]

Note: The 95% confidence interval is computed by bootstrap using the Bias Corrected and Accelerated (BCa) method.

Conclusion

Payments for Ecosystem Services programs are being increasingly used in the context of development and environmental policies around the world. By targeting grassland or forest conservation, these programs could be an effective tool for climate change mitigation if the overall benefits exceed the costs of implementation. Most of the existing empirical literature focuses on evaluating the additional effect in terms of supplementary areas conserved thanks to Payments for Ecosystem Services programs. However, this measure doesn't say anything about environmental effectiveness, which is one of the main objectives of these programs. Therefore, the empirical literature on the environmental effectiveness of Payments for Ecosystem Services in terms of reduced emissions is rather sparse. My thesis tries to shed light on this question by estimating the benefit-cost ratios of the French Grassland Conservation Program and of 18 Forest Conservation Programs implemented in developing countries.

The results of the first two chapters of my thesis suggest that at current carbon prices, Payments for Ecosystem Programs are not cost-effective, in the sense that they cost much more than the environmental benefits they bring. The lack of cost-effectiveness could probably be due to insufficient additionality. At current additionality levels, these programs would become effective at higher carbon prices than the current estimate of around 42 Euro/tCO_2eq (77 Euro/tCO_2eq for forest conservation programs and 200 Euro/tCO_2eq for grassland conservation programs). The difference in required carbon prices comes from the fact that Forest Conservation Programs are implemented in developing countries, as they have the highest deforestation rates. In comparison, Grassland Conservation Programs are mainly implemented in the European Union and in the United States of America.

Given these findings, a natural question arises: what could we do to increase the additionality of this programs? In practice, Payments for Ecosystem Services are usually designed as linear-uniform, in the sense that all program participants are paid the same amount of money per hectare conserved. Despite of being easy to implement, this type of contract design could limit the cost-effectiveness of Payments for Ecosystem Programs by increasing the risk of windfall effects, that are especially high when participants have heterogeneous opportunity costs. A solution to this problem would be to offer nonlinear contracts that differentiate the payment by the quantity conserved or to offer payments that are spatially differentiated. In the last chapter I show that for the French Grassland Conservation Program, among the three possible contract designs, a linear uniform subsidy is actually the best trade-off between ease of implementation and efficiency. Of course, the efficiency can be slightly improved by offering spatially-targeted or nonlinear contracts. However, these solutions require additional costs of implementation that might undermine the gain in welfare. This result suggests that the low cost-effectiveness of Payments for Ecosystem Services cannot be fully explained by the asymmetric information linked to the linear-uniform design of such contracts.

The reason why Payments for Ecosystem Services are not cost-effective enough is still an open question. Maybe the answer is simply a lack in the elasticity of supply (as I show is the case for the French Grassland Conservation Program). But then the question that arises is whether Governments should continue to invest in them or rather use the money to finance other public policies aimed at reducing carbon emissions that are more cost-effective. The answer to this question requires a more complex analysis of the costs and benefits of Payments for Ecosystem Services that is beyond the scope of this thesis. Indeed, the values for programs' costs we used for analysis actually represent the payments that farmers receive per hectare of grassland or forest conserved. They represent a lower bound as real programs' costs include also administrative costs and the true costs of conservation for farmers that are not easily available. Similarly, the environmental benefits considered in this thesis are mostly related to avoided CO_2 emissions. Even if they represent the major part, there are also other environmental benefits brought about by Payments for Ecosystem Services schemes, such as increased biodiversity, landscape amenities and better water quality. In the case of the French Grassland Conservation Programs estimates of these benefits are available in the literature (except for biodiversity), but this is usually not the case. In addition to the environmental benefits analysed in this thesis, Payments for Ecosystem Services have also secondary social objectives, such as poverty alleviation (especially those programs implemented in developing countries), that are rather difficult to quantify. When deciding the fate of Payments for Ecosystem Services all these elements should be taken into account.

Bibliography

[1] Abadie A. and G.W. Imbens (2006). Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica 77(1)*, 235-267.

[2] Abito J. M. (2020). Measuring the welfare gains from optimal incentive regulation. *The Review of Economic Studies 87(5)*, 2019-2048.

[3] Agouridis C. T., S. R. Workman, R. C. Warner, G. D. Jennings (2005). Livestock grazing management impacts on stream water quality: a review. *Journal of the American Water Resources Association 41(3)*, 591-606.

[4] Alix-Garcia J., E. Shapiro, K. Sims (2012). Forest Conservation and Slippage: Evidence from Mexico's National Payments for Ecosystem Services Program. *Land Economics 88(4)*, 613-638.

[5] Alix-Garcia J., K. Sims, P. Yanez-Pagans (2015). Only One Tree From Each Seed? Environmental Effectiveness and Poverty Alleviation in Mexicos Payments for Ecosystem Services Program. *American Economic Journal: Economic Policy 7(4)*, 1-40.

[6] Alix-Garcia J. and H. Wolff (2014). Payment for Ecosystem Services from Forests. *IZA Discussion Paper*, No. 8179.

[7] Alston L. J., K. Andersson, S. M. Smith (2013). Payment for environmental services: hypothesis and evidence. *NBER Working Paper Series*, Working Paper 18740.

[8] Andrews I. and M. Kasy (2019). Identification of and Correction for Publication Bias. *The American Economic Review 109(8)*, 2766-2794.

[9] Arata L. and P. Sckokai (2016). The impact of agri-environmental schemes on farm performance in five E.U. member states: a DID-matching approach. *Land Economics 92(1)*, 167-186.

[10] Arguedas C. and D. P. van Soest (2011). Optimal conservation programs, asymmetric information and the role of fixed costs. *Environmental and Resource Economics 50*, 305-323.

[11] Arriagada R.A., E.O. Sills and S.K. Pattanayak (2011). Payments for Environmental Services and Their Impact on Forest Transition in Costa Rica. *Working paper*.

[12] Arriagada R.A., P.J. Ferraro, E.O. Sills, S.K. Pattanayak and S. Cordero-Sancho (2012). Do Payments for Environmental Services Affect Forest Cover? A Farm-Level Evaluation from Costa Rica. *Land Economics 88(2)*, 382-399.

[13] Arrouays D., J. Balesdent, J.C. Germon, P.A. Jayet, J.F. Soussana and P. Stengel. (2002). Contribution à la lutte contre l'effet de serre. Stocker du carbone dans les sols agricoles de France ? Expertise scientifique collective. INRA (France), 334 pp.

[14] Athey S. and G. W. Imbens (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica 74(2)*, 431-497.

[15] Baudrier M., V. Bellassen and C. Foucherot (2015). La precedente politique agricole commune (2003-2013) a reduit les emissions agricoles françaises. *Etude Climat*, Working Paper 49.

[16] Baylis K., J. Honey-Rosés and M.I. Ramírez. Conserving Forests: Mandates, Managament or Money? *Working paper*.

[17] Beaud M. (2008). Le coût social marginal des fonds publics en France. *Annales d'Economie et de Statistique 90*, 215-232.

[18] Bernard A. and M. Vielle (2003). Measuring the welfare cost of climate change policies: A comparative assessment based on the computable general equilibrium model GEMINI-E3. *Environmental Modeling and Assessment 8(3)*.

[19] Bertrand M., E. Duflo and S. Mullainathan (2003). How much should we trust Differences-In-Differences estimates?. *Quarterly Journal of Economics 119*, 249-275.

[20] Bontemps C. and D. Martimort (2014). Identification and estimation of incentive contracts under asymmetric information: An application to the French water sector. *Toulouse School of Economics*, Working Paper.

[21] Borenstein M., L.V. Hedges, J.P.T. Higgins and H.R. Rothstein (2009). *Introduction to Meta-Analysis*. United Kingdom, West Sussex: John Wiley & Sons Ltd.

[22] Bourgeon J-M., P-A. Jayet and P. Picard (1995). An incentive approach to land set-aside programs. *European Economic Review 39*, 1487-1509.

[23] Boyer M. and J-J. Laffont (1999). Toward a political theory of the emergence of environmental incentive regulation. *The RAND Journal of Economics 30(1)*, 137-157.

[24] Bretagnolle V., G. Balent, C. Thenail and E. Berthet (2012). Gestion de la biodiversité en milieu céréalier intensif: importance des prairies aux échelles locales et régionales. *Innovations Agronomiques 22*, 31-43.

[25] Börner J., K. Baylis, E. Corbera, D. Ezzine-de-Blas, J. Honey-Rosés, U.M. Persson and S. Wunder (2017). The Effectiveness of Payments for Environmental Services. *World Development 96*, 359-374.

[26] Canton J., S. de Cara and P-A. Jayet (2009). Agri-environmental schemes: Adverse selection, information structure and delegation. *Ecological Economics 68*, 2114-2121.

[27] Centre National pour l´ Aménagement des Structures des Exploitations Agricoles (CNASEA) (2008). Evaluation ex post du Plan de Développement Rural National (Marché CNASEA n°22-07). Paris, Ministère de l'agriculture et de la Pêche.

[28] Chabé-Ferret, S. and J. Subervie (2009). Estimation des effets propres des mesures agro-environnementales du PDRN 2000-2006 sur les pratiques des agriculteurs. *CEMAGREF*.

[29] Chabé-Ferret S. and J. Subervie (2012). Econometric methods for estimating the additional effects of agro-environmental schemes on farmers' practices. In *Evaluation of agri-environmental policies: selected methodological issues and case studies* (Ch. 10), Ed. by Alison Burrell, OECD.

[30] Chabé-Ferret S. and J. Subervie (2013). How much green for the buck? Estimating additional and windfall effects of French agro-environmental schemes by DID-matching. *Journal of Environmental Economics and Management 65*, 12-27.

[31] Chabé-Ferret S., E. Souza-Rodrigues, A. van Benthem and J. Subervie (2015). Governments do not know the best way to save the Amazon rainforest, and that needs to change. *The Economist*, Nov. 16^{th}.

[32] Chabé-Ferret S. and A. Voia (2021). Are Forest Conservation Programs a Cost-Effective Way to Fight Climate Change? A Meta-Analysis. *Toulouse School of Economics*, Working Paper.

[33] Chervier C. and S. Costedoat (2017). Heterogeneous Impact of a Collective Payment for Environmental Services Scheme on Reducing Deforestation in Cambodia. *World Development 98*, 148-159.

[34] Christensen P., P. Francisco, E. Myers, M. Souza (2021). Decomposing the Wedge between Projected and Realized Returns in Energy Efficiency Programs. *The Review of Economics and Statistics*, 146.

[35] Clements T. and E.J. Milner-Gulland (2015). Impact of Payments for Environmental Services and Protected Areas on Local Livelihoods and Forest Conservation in Northern Cambodia. *Conservation Biology 00(0)*, 1-10.

[36] Corine Land Cover France (2009). *Guide d'utilisation*.

[37] Costedoat S., E. Corbera, D. Ezzine-de-Blas, J. Honey-Rosés, K. Baylis, M.A. Castillo-Santiago (2015). How Effective are Biodiversity Conservation Payments in Mexico? *PLoS ONE 10(3)*.

[38] Crépin A-S. (2005). Incentives for wetland creation. *Journal of Environmental Economics and Management 50*, 598-616.

[39] DellaVigna S. and E. Linos (2021). RCTs to Scale: Comprehensive Evidence from Two Nudge Units. *Econometrica*, forthcoming.

[40] Delta Method (n.d.). Retrieved from http://www.montefiore.ulg.ac.be/kvansteen /MATH0008-2/ac20122013/Class2Oct/Supplementary.

[41] Dominici F., M. Greenstone and C. R. Sunstein (2014). Particulate matter matters. *Science 344(6181)*, 257-259.

[42] Ecosystem Markets Map (2019). Retrieved from https://www.forest-trends.org.

[43] FAO. 2015. *FAO Assessment of Forests and Carbon Stocks, 1990-2015*. Available at http://www.fao.org/3/a-i4470e.pdf.

[44] Ferraro P. J. (2008). Asymmetric information and contract design for payments for environmental services. *Ecological Economics 65*, 810-821.

[45] Fraser I.M. (1995). An analysis of management agreement bargaining under asymmetric information. *Journal of Agricultural Economics 46(1)*, 20-32.

[46] Gallic G. and V. Marcus (2019). La réforme 2015 de la Politique Agricole Commune: évaluation de l'impact de la réforme de l'Indemnité Compensatoire de Handicaps Naturels sur les surfaces herbagres. *Economie & prévision 215(1)*, 113-140.

[47] Gechert S., T. Havranek, Z. Irsova and D. Kolcunova (2021). Measuring capital-labor substitution: The importance of method choices and publication bias. *Review of Economic Dynamics*, In Press.

[48] Gillingham K. and J. H. Stock (2018). The Cost of Reducing Greenhouse Gas Emissions. *Journal of Economic Perspectives 32(4)*, 53-72.

[49] Global Forest Watch - Climate (2019). Available at http://climate.globalforestwatch.org/countries.

[50] Gren I-M. (2004). Uniform or discriminating payments for environmental production on arable land under asymmetric information. *European Review of Agricultural Economics 31(1)*, 61-76.

[51] Harrer M., P. Cuijpers, T.A. Furukawa and D.D. Ebert (2019). Doing Meta-Analysis in R: A Hand-on Guide. https://bookdown.org/MathiasHarrer/DoingMetaAnalysisinR.

[52] D'Haultfoeuille X. and P. Février (2020). The provision of wage incentives: A structural estimation using contracts variation. *Quantitative Economics 11*, 349-397.

[53] Higgins J.PT., S.G. Thompson, J.J. Deeks and D.G. Altman (2003). Measuring Inconsistency in Meta-Analyses. *British Medical Journal 327(7414)*.

[54] Honey-Rosés J., K. Baylis and M.I. Ramírez (2011). A Spatially Explicit Estimate of Avoided Forest Loss. *Conservation Biology 25(5)*, 1032-1043.

[55] Ioannidis J. P. A., T. D. Stanley and H. Doucouliagos (2017) The Power of Bias in Economics Research. *The Economic Journal 127(605)*, F236-F265.

[56] Institut de l´ Elevage (2007). La prairie: un enjeu économique et sociétal (Hors série. Spécial prairie). Paris, Département Economie de l´ Institut de l´ Elevage.

[57] Interagency Working Group on Social Cost of Greenhouse Gases, United States Government (2016). Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866. Retrieved from https://19january2017snapshot.epa.gov/sites/production/files/2016-12/documents.

[58] Jackson C. K. and C. L. Mackevicius (2021). The Distribution of School Spending Impacts, *mimeo*, Northwestern University.

[59] Jayachandran S., J. de Laat, E.F. Lambin, C.Y. Stanton, R. Audy and N.E. Thomas (2017). Cash for Carbon: A Randomized Trial of Payments for Ecosystem Services to Reduce Deforestation. *Science 357*, 267-273.

[60] Jones K.W., M.B. Holland, L. Naughton-Treves, M. Morales, L. Suarez and K. Keenan (2016). Forest Conservation Incentives and Deforestation in the Ecuadorian Amazon. *Environmental Conservation 44(1)*, 56-65.

[61] Kleven H. and M. Waseem (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *The Quarterly Journal of Economics 128(2)*, 669-723.

[62] Khufuss L. and J. Subervie (2018). Do European agri-environment measures help reduce herbicide use? Evidence from viticulture in France. *Ecological Economics 148(C)*, 202-211.

[63] Knapp G. and J. Hartung (2003). Improved Tests for a Random Effects Meta-Regression with a Single Covariate. *Statistics in Medicine 22(17)*, 2693-2710.

[64] Kvarven A., E. Stromland and M. Johannesson (2019). Comparing meta-analyses and preregistered multiple-laboratory replication projects. *Nature Human Behaviour 4*, 423434.

[65] Lebegue D. (2005). Le prix du temps et la décision publique: Révision du taux d'actualisation public. *La Documentation francaise*.

[66] Mason C.F. and A.J. Plantinga (2013). The additionality problem with offsets: Optimal contracts for carbon sequestration in forests. *Journal of Environmental Economics and Management 66*, 1-14.

[67] Ministry of Agriculture (2006). *Evaluation de la prime herbagère agroenvironnementale en France (2003-2006).* Rapport d'évaluation.

[68] Ministry of Agriculture (2005). *L'utilisation du territoire en 2004: nouvelle série 1992 à 2004*. Chiffres et Données - Série Agriculture - n^{o}169.

[69] Ministry of Agriculture (2008). *L'utilisation du territoire en 2007 Teruti-Lucas*. Chiffres et Données - Série Agriculture - n^{o}202.

[70] Mohebalian P.M. and F.X. Aguilar (2016). Additionality and Design of Forest Conservation Programs: Insights from Ecuador's Socio Bosque Program. *Forest Policy and Economics 71*, 103-114.

[71] Mohebalian P.M. and F.X. Aguilar (2018). Beneath the Canopy: Tropical Forests Enrolled in Conservation Payments Reveal Evidence of Less Degradation. *Ecological Economics 143*, 64-73.

[72] Motte E., J.-M. Salles and L. Thomas (2002). Incentive policies to farmers for conserving biodiversity in forested areas in developing countries. *LAMETA*, Working Paper.

[73] Moxey A., B. White and A. Ozanne (1999). Efficient contract design for agri-environmental policy. *Journal of Agricultural Economics 50(2)*, 187-202.

[74] Nemati M. and J. Penn (2020). The impact of information-based interventions on conservation behavior: A meta-analysis. *Resource and Energy Economics 62*, 101-201.

[75] Ozanne A., T. Hogan and D. Colman (2001). Moral hazard, risk aversion and compliance monitoring in agri-environmental policy. *European Review of Agricultural Economics 28*, 329-347.

[76] Pattanayak S.K., S. Wunder and P.J. Ferraro (2010). Show Me the Money: Do Payments Supply Environmental Services in Developing Countries? *Review of Environmental Economics and Policy 4(2)*, 254-274.

[77] Perrigne I. and Q. Vuong (2011). Nonparametric identification of a contract model with adverse selection and moral hazard. *Econometrica 79*, 1499-1539.

[78] Perrigne I. and Q. Vuong (2012). Nonlinear pricing in Yellow Pages. *Pennsylvania State University*, Working Paper.

[79] Pollinger S. (2021). Kinks know more: policy evaluation beyond bunching with an application to solar subsidies. *Toulouse School of Economics*, Working Paper.

[80] Pufahl A. and C.R. Weiss (2009). Evaluating the effects of farm programmes: results from propensity score matching. *European Review of Agricultural Economics 36(1)*, 79-101.

[81] Puydarrieux P. and J. Devaux (2013). *Quelle évaluation économique pour les services écosystémiques rendus par les prairies en France métropolitaine?* (Report n°92). Collection "Etudes et documents" du Service de l'Economie, de l'Evaluation et de l'Intégration

du Développement Durable (SEEIDD) du Commissariat Général au Développement Durable (CGDD).

[82] Robalino J., A. Pfaff, G.A. Sánchez-Azofeifa, F. Alpízar, C. León and C.M. Rodríguez (2008). Deforestation Impacts of Environmental Services Payments: Costa Rica's PSA Program 2000-2005. *Environment for Development Discussion Paper*, 08-24.

[83] Robalino J. and A. Pfaff (2013). Ecopayments and Deforestation in Costa Rica: A Nationwide Analysis of PSA's Initial Years. *Land Economics 89(3)*, 432-448.

[84] Robalino J., C. Sandoval, D.N. Barton, A. Chacon and A. Pfaff (2015). Evaluating Interactions of Forest Conservation Policies on Avoided Deforestation. *PLoS ONE 10(4)*.

[85] Samii C., M. Lisiecki, P. Kulkarni, L. Paler and L. Chavis (2014). Effects of Payment for Environmental Services (PES) on Deforestation and Poverty in Low and Middle Income Countries: A Systematic Review. *Campbell Systematic Reviews*, 11.

[86] Sheriff G. (2009). Implementing second-best environmental policy under adverse selection. *Journal of Environmental Economics and Management 57*, 253-268.

[87] Simonet G., A. Agrawal, F. Bénédet, C. de Perthuis, D. Haggard, N. Jansen, A. Karsenty, W. Liang, P. Newton, A.-M. Sales, B. Schaap and C. Seyller (2016). ID-RECCO, International Database on REDD+ projects, linking Economic, Carbon and Communities data. version 2.0. http://www.reddprojectsdatabase.org.

[88] Simonet G., J. Subervie, D. Ezzine-de-Blas, M. Cromberg and A.E. Duchelle (2018). Effectiveness of a REDD+ Project in Reducing Deforestation in the Brazilian Amazon. *American Journal of Agricultural Economics 101(1)*, 211-229.

[89] Simula M. (2010). *Analysis of REDD+ Financing Gaps and Overlaps*. REDD+ Partnerchip. Available at http://reddpluspartnership.org.

[90] Snilsveit B., J. Stevenson, L. Langer, N. Tannous, Z. Ravat, P. Nduku, J. Polanin, I. Shemilt, J. Eyers and P. J. Ferraro (2019). Incentives for climate mitigation in the land use sector —effects of payment for environmental services on environmental and socioeconomic outcomes in low– and middle–income countries: A mixed-methods systematic review. *Campbell Systematic Reviews 15(3)*, e1045.

[91] Soussana J.-F., P. Loiseau, N. Vuichard, E. Ceschia, J. Balesdent, T. Chevallier and D. Arrouays (2004). Carbon cycling and sequestration opportunities in temperate grasslands. *Soil Use and Management 20*, 219-230.

[92] Stanley, T.D. and H. Doucouliagos (2012). Meta-Regression Analysis in Economics and Business. Routledge.

[93] Stanley, T.D. and H. Doucouliagos (2015). Neither fixed nor random: weighted least squares metaanalysis. *Statistics in Medecine*, 34(13), 2116-2127.

[94] United States Department of Agriculture - Natural Resources Conservation Service (USDA-NRCS) (2010). Final Benefit-Cost Analysis for the Grassland Reserve Program (Food, Conservation, and Energy Act of 2008; Title II Conservation; Subtitle E Farmland Protection and Grassland Reserve; Section 1238N Grassland Reserve Program). Washington D.C., United States Department of Agriculture.

[95] Viechtbauer W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics 30(3)*, 261-293.

[96] Wertz-Kanounnikoff S., B. Locatelli, S. Wunder, M. Brockhaus (2011). Ecosystembased adaptation to climate change: what scope for payments for environmental services? *Climate and Development 3(2)*, 143-152.

[97] Wiik E., R. d'Annunzio, E. Pynegar, D. Crespo, N. Asquith and J.P.G. Jones (2019). Experimental Evaluation of the Impact of a Payment for Environmental Services Program on Deforestation. *Conservation Science and Practice 1(2)*, e8.

[98] Wolosin M. and N. Harris (2018). Tropical Forests and Climate Change: The Latest Science. *World Resources Institute Working Paper*.

[99] Wu J. and B.A. Babcock (1996). Contract design for the purchase of environmental goods from agriculture. *American Journal of Agricultural Economics 78*, 935-945.

[100] Wunder S. (2005). Payments for Environmental Serivces: Some Nuts and Bolts. *CIFOR Occasional Paper*, No. 42.

[101] Wunder S. (2007). The Efficiency of Payments for Environmental Services in Tropical Conservation. *Conservation Biology 21(1)*, 48-58.

[102] Wunder S. (2008). Payments for Environmental Services and the Poor: Concepts and Preliminary Evidence. *Environment and Development Economics 13(3)*, 279-297.