PEDRO WESLEY VERTINO DE QUEIROZ

ESSAYS ON AGRICULTURAL TECHNOLOGY, RESOURCE ALLOCATION AND THE VALUE OF INFORMATION

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Ph.D.) in Agricultural Economics at the Graduate College of the University of Nebraska-Lincoln, USA, in co-tutelage with the Graduate Program in Applied Economics at Federal University of Viçosa, Brazil.

Advisers: Lilyan Estela Fulginiti
Alexandre Bragança Coelho

VIÇOSA - MINAS GERAIS
2020

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

Q3e
2020

Queiroz, Pedro Wesley Vertino de, 1989-

Essays on agricultural technology, resource allocation and the value of information / Pedro Wesley Vertino de Queiroz. – Viçosa, MG, 2020.

117 f. : il. (algumas color.) ; 29 cm.

Inclui anexos.
Orientador: Alexandre Bragança Coelho.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Inovações agrícolas. 2. Implementos agrícolas. 3. Agricultura de precisão. 4. Sistemas de informação geográfica. 5. Sistemas de recuperação da informação - Agricultura. I. Universidade Federal de Viçosa. Departamento de Economia Rural. Doutorado em Economia Aplicada. II. Título.

CDD 22. ed. 631.3

PEDRO WESLEY VERTINO DE QUEIROZ

ESSAYS ON AGRICULTURAL TECHNOLOGY, RESOURCE ALLOCATION AND THE VALUE OF INFORMATION

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Ph.D.) in Agricultural Economics at the Graduate College of the University of Nebraska-Lincoln, USA, in co-tutelage with the Graduate Program in Applied Economics at Federal University of Viçosa, Brazil.

APPROVED: April 13, 2020.

Assent:

Pedro Wesley Vertino de Queiroz
Author

Alexandre B. Coelho

Alexandre Braganqa Coelho
Adviser

To my father, Vandir Mendes de Queiroz (*in memoriam*), the best and the kindest human being I will ever know. Your love, generosity, and wisdom stay with us.

ACKNOWLEDGEMENTS

I had the opportunity to study and work at two great institutions during this Ph.D. program, at the Federal University of Viçosa (UFV), Brazil, and at the University of Nebraska-Lincoln (UNL), USA. I thank specially my advisors Dr. Lilyan Fulginiti and Dr. Richard Perrin who gave me the opportunity to come to the United States. The Ph.D. was a challenging experience but none of this process could have been accomplished without great mentors that really love what they do and care about their students. I will value you and appreciate you for the rest of my life.

Many thanks to my advisor at UFV, Dr. Alexandre B. Coelho and my committee members, Dr. Karina Schoengold and Dr. Mary McGarvey. To my professors at the Agricultural Economics and Economics departments at UNL. To the staff at UNL and UFV. To my friends at UFV.

Finally, I want to thank my family, friends, and colleagues who made life in Nebraska easier. Special thanks to my cohort Diya Ganguly, Ahmed Chennak, Brian Liston-Clark, Anubrata Deka, Susweta Ray, Saeedeh Ravan and Angela Graham. To Felipe Silva, Yasmin Lopes, Leonardo Lopes, Amandha Lopes, Brent Kumor, Charlie Brown, Thomas Miorini, Peter Schuster, Wilman Iglesias, Atanaele Bernardo, Luana Machado, Denis Mariano, Gabriel Morais, Ivo Zution, Renata Trefiglio, Vanessa de Oliveira, and my roommates Alexandre Miszura, Ândrea Plotski, Kelly Massarolo and Regina Stacke.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

ABSTRACT

QUEIROZ, Pedro Wesley Vertino de, D.Sc., Universidade Federal de Viçosa, April 2020. **Essays on agricultural technology, resource allocation and the value of information**. Advisers: Lilyan Estela Fulginiti and Alexandre Bragança Coelho.

Understanding agricultural technology adoption and resource allocation is crucial to achieve agricultural development and to adapt to changes in current institutional, environmental, and climatic conditions. In addition, decision-making can be improved because of new information and timely data available from precision agriculture (PA) and satellite imagery technologies. The objective of this dissertation was to, first, study agricultural technology adoption and resource allocation due to changes in the relative supply of land (land-and-labor ratio) in South American agriculture. Second, the payoff of a precision agriculture (PA) technology was assessed for U.S. farms. Third, I used remote sensing information to obtain a better measurement of the effect of droughts in Brazilian municipalities. Chapter 1 studied traditional (e.g., land and labor) and commercial (e.g., machinery and fertilizers) inputs in South American agriculture. Acemoglu's directed technical change framework was used to estimate the process of induced innovation in agriculture using deforestation patterns as source of exogenous variation for the agricultural land supply. The results indicated that deforestation was important to provide agricultural land in South America and because of a larger availability of land in intensive deforestation countries, more land-complementary inputs (machinery) were used relative to labor-complementary inputs (fertilizers). The induced innovation in South America was mainly driven by a larger "market size effect" indicating that technical change was biased towards land. Chapter 2 studied nitrogen fertilizer application in U.S. agriculture. Soil information (signal) obtained from the PA technology allowed the adoption of variable rate (VR) applications of nitrogen specific to the different plots (cells) of the field, with potential to increase the farmers' profitability and decrease the environmental damages due to excessive use. I provided a Bayesian structural model, based on the Expected Value of Sample Information (EVSI) approach, with a direct application using data from the Data-Intensive Farm Management (DIFM) project from University of Illinois to evaluate the expected returns of the VR technology. The results from the studied U.S. farms showed that the information from soil electroconductivity (EC) provided low expected returns. The insights from the model revealed that the low returns can be explained by EC being "poorly" correlated with the true soil conditions and/or the quality of the soil may be uniform across the fields, hence, not supporting the VR technology adoption. Chapter 3 used satellite remotely sensed information to estimate the effects of droughts on agriculture for Brazilian municipalities. First, the effect of droughts for all the corn- and soybeans-producing Brazilian municipalities was estimated, then a model adding remote sensing data was estimated for the municipalities from a soybeans-producing region of Southern Brazil, both for the 2002-2016 period. The results implied that the lack of biophysical variables in the model, reflecting the interaction among the soil, the plant, and the atmosphere, would bias the drought effects. This is an important result because economic decisions are made based on the effects of climate conditions in agriculture and remote sensing information can provide more reliable estimates of the true climatic effects.

Keywords: Agricultural technology. Innovation. Agricultural input use. Precision agriculture. Remote sensing information. Value of information.

RESUMO

QUEIROZ, Pedro Wesley Vertino de, D.Sc., Universidade Federal de Viçosa, abril de 2020. **Ensaios sobre tecnologia agrícola, alocação de recurso e o valor da informação**. Orientadores: Lilyan Estela Fulginiti e Alexandre Bragança Coelho.

Entender a adoção de tecnologia agrícola e a alocação de recursos é crucial para se atingir desenvolvimento agrícola e para se adaptar às mudanças nas atuais condições institucionais, ambientais e climáticas. Além disso, a tomada de decisões pode ser melhorada com novas informações e dados oportunos provenientes de tecnologias da agricultura de precisão e de imagens de satélite. O objetivo desta tese foi, primeiro, estudar a adoção de tecnologia agrícola e a alocação de recursos devido às mudanças da oferta relativa de terra (razão entre terra e trabalho) na agricultura sul-americana. Em seguida, o benefício de uma tecnologia da agricultura de precisão foi avaliado para fazendas dos EUA. Por fim, foram utilizados dados de sensoriamento remoto para obter uma melhor medida do efeito das secas em municípios brasileiros. O capítulo 1 estudou insumos tradicionais (terra e trabalho) e comerciais (maquinários e fertilizantes) na agricultura sul-americana. A teoria da mudança tecnológica direcionada (em inglês, *directed technical change*) por Acemoglu foi adaptada para estimar o processo de inovação induzida na agricultura utilizando desmatamento como fonte exógena de variação para a oferta relativa de terra. Os resultados indicaram que o desmatamento foi importante para prover terra agrícola na América do Sul e devido a maior disponibilidade de terra em países com desmatamento intensivo, mais insumos complementares à terra (maquinários) foram usados em relação aos insumos complementares ao trabalho (fertilizantes). A inovação induzida na América do Sul foi principalmente movida por um maior efeito de "tamanho de mercado" indicando que a mudança tecnológica aumentou o uso de terra. O capítulo 2 estudou a aplicação do fertilizante de nitrogênio na agricultura dos EUA. Informação sobre o solo (sinal) obtida da tecnologia de agricultura de precisão permitiu a adoção de aplicações variáveis (VR) de nitrogênio específicas às diferentes partes do campo, com potencial de aumentar a rentabilidade dos fazendeiros e diminuir os danos ambientais devido ao uso excessivo. Foi desenvolvido um modelo estrutural bayesiano, baseado na abordagem do Valor Esperado da Informação Amostral (EVSI), com uma aplicação direta usando dados do projeto *Data-Intensive Farm Management* (DIFM) da Universidade de Illinois para avaliar os retornos esperados da tecnologia VR. Os resultados das fazendas dos EUA mostraram que a informação obtida da eletro-condutividade (EC) do solo gerou retornos esperados baixos. As conclusões do modelo revelaram que os baixos retornos podem ser explicados por EC ser “fracamente” correlacionada com as verdadeiras condições do solo e/ou a qualidade do solo ser uniforme nos campos, assim, não dando suporte à adoção da tecnologia VR. O capítulo 3 utilizou informações de sensoriamento remoto para estimar os efeitos das secas na agricultura para municípios brasileiros. Primeiro, o efeito das secas foi estimado para todos os municípios brasileiros produtores de milho e soja, e então um modelo adicionando dados de sensoriamento remoto foi estimado para municípios de uma região produtora de soja no sul do Brasil, ambos para o período 2002-2016. Os resultados implicaram que a falta de variáveis biofísicas no modelo, refletindo a interação entre o solo, a planta, e a atmosfera, gerariam viés nos efeitos das secas. Esse resultado é importante porque decisões econômicas são feitas baseadas nos efeitos das mudanças climáticas na agricultura e informações de sensoriamento remoto podem fornecer estimativas mais confiáveis dos verdadeiros efeitos climáticos.

Palavras-chave: Tecnologia agrícola. Inovação. Uso de insumos agrícolas. Agricultura de precisão. Informação de sensoriamento remoto. Valor da informação.

TABLE OF CONTENTS

INTRODUCTION

Innovations in agriculture allowed for historical worldwide agricultural productivity gains over the past years. Understanding past technology adoption and resource allocation is essential to adapt to the changes in the institutional, environmental, and climatic conditions affecting agriculture. According to the “induced innovation hypothesis” (IIH) by Hicks, prices are usually the main factors determining innovations, but recent theories on development economics and growth associated the creation and the adoption of technologies to the resource endowments in the regions, deriving new forces behind the process of the induced innovation. The “directed technical change” model developed by the MIT Professor Daron Acemoglu provides an appealing setting to study innovations in agriculture, without the need for information on input prices. Chapter 1 of this dissertation uses this framework to study innovations in South American agriculture for the 1961-2015 period, estimating the process of the induced innovation and providing new insights on the relationship between deforestation and agricultural land supply.

New agricultural practices have relied more and more on soil and climatic information allowing for the adoption of variable rate technology (VRT) for the application of inputs and for better measurements of the effects of climate on agriculture. Chapter 2 assesses the expected payoff of VRT for the application of nitrogen fertilizer, characterized as one of the precision agriculture (PA) technologies, compared to a uniform rate technology (URT) application. In the United States, the adoption of precision agriculture (PA) has increased tremendously in recent years. Despite a large literature on the returns of PA, very few studies provided a structural model to assess the expected payoff and its determining parameters. The value of information (VOI) theory provided the theoretical environment to calculate the Expected Value of Sample Information (EVSI). VRT can only be adopted through acquiring some soil information (often referred as soil *signal*) that can affect the allocation of resources in agriculture. The EVSI is then the difference between expected payoffs with and without information, hence, providing a measurement of the expected payoff of VRT.

Lastly, remote sensing technologies provide information that allows improving the use of resources as an adaptation strategy for the effects of climate change. Specifically, satellite imagery can deliver important biophysical indicators to improve the estimation of weather fluctuations impacts in agriculture. Remote sensing satellite data sources have the advantage of containing detailed spatial and temporal records on variables affecting agricultural regions. Different stakeholders across the agricultural industry supply chain are particularly interested on the outcomes from dry seasons. From a public policy perspective, knowing the impacts of droughts on agriculture through remote sensing information can help policy makers define eligibility criteria in government assistance programs supporting farmers that are affected by severe droughts. This would be then one of the main motivations for public investment on national drought monitor centers such as the United States Drought Monitor (USDM), in the United States, and the recent Northeast Drought Monitor (NDM), in Northern Brazil. For example, in the United States, the United States Drought Monitor (USDM) index, which is constructed as a combination of satellite data, experts' inputs, and weather station data, is the main drought indicator used by the government to provide drought-alleviating financial aid to farmers. Chapter 3 of this dissertation uses weather and biophysical variables from weather stations and satellite remotely sensed data to estimate the effect of droughts on agriculture for two samples of Brazilian municipalities, one consisting of all the corn- and soybean-producing municipalities in Brazil, and the other of the municipalities from a soybean-producing region in three Southern Brazilian states (Paraná, Santa Catarina and Rio Grande do Sul).

Considering these approaches of studying agricultural technology, resource allocation, and the value of information in agricultural decision-making, this dissertation aimed to answer three main research questions, each addressed in one chapter:

Chapter 1: How did the availability of land and labor affect the use of commercial inputs that incorporate land-complementary technologies (machinery) and labor-complementary technologies (fertilizers) in South American agriculture?

Chapter 2: What is the theoretical expected value of sample information (EVSI) for the adoption of a precision agriculture (PA) technology? What is the estimated EVSI from soil electroconductivity information used for the adoption of a variable rate (VR) nitrogen fertilizer application in U.S. farms?

Chapter 3: Does remote sensing information improve the estimation of drought effects on Southern Brazilian agriculture?

The contribution of this research is mainly related to the literature on agricultural development, agricultural productivity, and production economics. Chapter 1 follows a different approach than the usual productivity analysis obtained from parametric and non-parametric estimations of production and distance functions. The process of induced innovation was expressed in one equation from the directed technical change model allowing to estimate the elasticities of substitution of the inputs and outputs in agriculture with publicly available country-level panel data and econometric models with fixed effects. Furthermore, the results allowed to identify the mechanism of the induced innovation through an endogenously determined technical change that would either increase the use of inputs that are more abundant or scarcer, depending on which forces are driving the decisions of the innovators. It also used an instrumental variable (IV) approach as an attempt to treat the endogeneity bias in the equation provided by the model. This allowed a straightforward link between deforestation and the directed technical change. Agricultural expansion through deforestation is common in South American countries and future research can benefit from Acemoglu's model to study the effects of deforestation control policies on agricultural productivity.

Chapter 2 contributes to the literature on the returns of precision agriculture by providing a structural approach to insert the soil information into the yield function. It provided an equation that incorporated the producer's Bayesian-updated belief about the true soil conditions, following a non-ad hoc way to study the economic value of information and the adoption of new agricultural technologies.

Chapter 3 contributes on the potential use of satellite information about soil biophysical conditions to study the effect of droughts on agriculture. On the same fashion of the USDM index, the drought measurements were constructed to reflect months or weeks of droughts in Brazilian municipalities, allowing the estimation of the effect of droughts using panel data models with fixed effects and assessing the extent to which remote sensing variables can improve the estimation. The last section of the dissertation presents the final remarks.

CHAPTER 1

INDUCED INNOVATION IN SOUTH AMERICAN AGRICULTURE: DEFORESTATION AND DIRECTED TECHNICAL CHANGE

1.1 Introduction

The development of new technologies has allowed for historical worldwide agricultural productivity gains over the past years. Yet the creation and the adoption of innovations vary significantly by country and depend on many factors such as the country's resources, institutions, and climate. These factors impact how innovations can increase/decrease specific inputs in agriculture. The induced innovation hypothesis (IIH) (Hicks, 1932) is a classic economic concept that states that prices are the main drivers of innovations, i.e., technology is adopted to reduce costs in production by decreasing inputs that are more expensive. The classic studies of Hayami and Ruttan (1970, 1971) contributed to the development of many empirical tests for the IIH in agriculture (Hayami and Ruttan, 1970, 1971; Binswanger, 1974; Kawagoe, Otsuka and Hayami, 1986; Thirtle and Ruttan, 1987; Fulginiti and Perrin, 1993; Thirtle et al., 2002; Liu and Shumway, 2009; Cowan, Lee, and Shumway, 2015)[1]. Most of these studies used information on input prices as the main determinants of innovations because, in addition to reflecting scarcity, they represent the mechanism through which economic agents respond to changes in the markets caused by shifts in the supply and/or the demand curves. These shifts are key in the theories of induced innovation (Hicks, 1932; Binswanger et al., 1978; Thirtle and Ruttan, 1987; Acemoglu, 1998, 2002) [2]. In recent theoretical developments, Acemoglu (2002) formally introduced his "directed technical change" structural model where the creation and adoption of innovations are directly affected by changes in the availability of resources (i.e., changes in the factor endowments ratio).

The objective of this paper is to provide an application of Acemoglu's directed technical change to agriculture by estimating the effect of changes in agricultural land relative to labor on the direction of technological adoption in South American agriculture. We use the

---

[1] Binswanger (1974) is also a classic study on the IIH. The author used the duality theory to create a measurement of the bias of technical change in terms of changes in cost shares which allows testing with a multi-factor production function. He found that during the period of his analysis a strong fertilizer-using technical change was accompanied by a decline in fertilizer price in the United States agriculture.

[2] In particular, the induced innovation theories focused on the demand shifts induced by technical change and its effect on prices. Acemoglu (1998) studied the shifts in the supply of skilled workers in the U.S. since 1970s which have increased the skill premium due to skill-biased technical change, i.e., the abundance of skilled workers shifted the demand for skills and caused an increase in wages.

framework developed in Acemoglu (2002) to establish a relationship between the ratios of two traditional inputs (land-and-labor) and two commercial inputs (machinery-and-fertilizers). The commercial inputs embody technologies that allow for saving one traditional input while complementing the other in agricultural production. The hypothesis is that shifts in the relative land supply (land-and-labor ratio) induced innovations to expand land use and to save labor by increasing the use of land-complementary inputs (e.g., machinery) more than the labor-complementary inputs (e.g., fertilizers)[3]. We explore changes in deforestation patterns as source of exogenous variation in the relative land supply. This is the first study to provide an application of Acemoglu's directed technical change model to agriculture linking deforestation to this framework.

Previous studies have tested the IIH in multiple ways. Most studies performed econometric tests to explain the changes in input ratios in terms of the changes in lagged own input price ratios finding a negative relationship (Hayami and Ruttan, 1970, Thirtle et al., 2002, Liu and Shumway, 2009). Lagged prices help isolate, from the usual substitution effect, the development of technologies that reduce the expensive inputs; this means that not only the quantity demanded decreases due to an increase in prices but also innovations help reduce costs in the long run. Thirtle et al. (2002) found support for the IIH by developing a time series test to separate the factor substitution effect from the technological change effect for the U.S. agriculture in the 1880-1990 period. Liu and Shumway (2009) tested the induced innovation by different methods – time series, direct econometric and nonparametric – and found little support for the IIH for the U.S. agriculture over the period 1960-1999. The authors concluded that their study, as well as previous studies, have focused only on the industry that incorporates innovations (demand side), while the changes in the industry that creates innovations[4] (supply side), should be considered. More recent tests included variables that captured the changes in the marginal costs of innovations (Popp, 2002; Cowan, Lee, and Shumway, 2015). These tests focused on the innovations themselves by using patent data, R&D expenditures, or data on the technologies. The specification must then consider demand-side variables (e.g., input prices), and supply-side variables, such as measures of scientific advancements and changes in the innovation marginal costs. Popp (2002) found support for the IIH in the energy sector using

---

[3] Alternatively, land-complementary can be read as land-using. Also, for a production function with two inputs, e.g., land and labor, land-complementary means labor-saving.

[4] Olmstead and Rhode (1993) were the main critics of the Hayami and Ruttan findings and the IIH hypothesis in U.S. agriculture. By using different data, they rejected the hypothesis which caused a major influence on the more recent literature to include not only demand-side factors but also supply-side factors on the process of innovation.

patent data on energy-efficient technologies, energy prices and “knowledge stock” variables. Cowan, Lee, and Shumway (2015), using U.S. public agricultural research expenditures, found support for the IIH for U.S. agriculture over the period 1927-2009. They estimated the ratio of public research expenditures on input-saving technologies as a function of the ratio of the own input prices.

The modern theories of induced innovation combined both the supply and demand factors into one equilibrium model where innovations are developed by technology monopolists. These theories reveal several other forces in the process of the induced innovation such as the “market size effect” (Acemoglu, 1998; 2002). The “market size effect” is when the technology monopolists find it more profitable to produce innovations to be used with the inputs that are more abundant. Acemoglu (2002) developed a directed technical change structural model that relates the optimal levels of innovations used in the production of a final good to the factor endowments. The model explicitly indicates what is behind the process of developing and adopting technologies that are complementary to the inputs in production by incorporating the technology monopolist’s market and the supply of innovations. The conclusion of the model is that irrespective of the elasticity of substitution, σ, between the inputs (e.g., land and labor), the IIH holds. Furthermore, the theory provides the process of the induced innovation in terms of the shifts in the factor supplies, without the need for information on factor prices. This study estimates it empirically for South American agriculture. To our knowledge, this is the first paper to provide an estimation of Acemoglu’s directed technical change structural parameters for the agricultural sector with an attempt to correct for the endogeneity bias[5].

We study ten South American countries for the periods of 1961-2015 and 1990-2015[6]. The agricultural sector in these countries is important for economic growth and incentives for industrialization have motivated a shift of resources from agriculture to the industry impacting the relative input supplies in these sectors. Given industrialization growth and agricultural expansion accompanied by increases in deforestation, the adoption of technologies can be related to the expansion of land in South American agriculture. Additionally, in countries with intensive deforestation, such as Brazil, and with large availability of land, there can be a larger

---

[5] Klump and Cabrera (2008) revisited the Hayami-and-Ruttan IIH for agriculture by estimating a version of the directed technical change model in Acemoglu (2002). However, they do not explicitly estimate the parameters of the model or attempt to correct for the endogeneity bias due to omitted variables.

[6] We separate the analysis between these two periods because our proxy for deforestation is only available from 1990.

market for the land-complementary commercial inputs[7]. The identification strategy relies on the assumption that deforestation affects agricultural productivity by providing area for the expansion of agriculture.

This study is related to the previous literature on the IIH in agriculture and directed technical change (Hayami and Ruttan, 1970; Acemoglu, 2002; Klump and Cabrera, 2008; Hanlon, 2015), country-productivity differences (Acemoglu and Zilliboti; 2001), deforestation policies and agricultural expansion in South America (Faria and Almeida, 2016; Nolte et al, 2017; le Polain de Waroux et al., 2019; Silva, Perrin and Fulginiti, 2019; Koch et al., 2019), and the more recent literature on environmental policies and directed technical change (Acemoglu et al., 2012; Calel and Dechezleprêtre, 2016; Aghion et al., 2016). Klump and Cabrera (2008) used the directed technical change framework to revisit the classic Hayami-Ruttan IIH empirically using panel data for ninety-three countries. They found that more relative land increased more machinery than fertilizers in agriculture. Thus, concluding that the IIH holds because, as machinery complements land in production, the demand for land increases. We differ from their analysis by directly estimating the directed technical change parameters from the theory[8] and by using deforestation patterns as source of exogenous variation for the relative land supply (land-and-labor ratio)[9]. Hanlon (2015) used the directed technical change framework to test the induced innovation in the British cotton industry during the U.S. civil war (1861-1865). The author used the timing of the war as an exogenous shock in the supply of U.S. cotton finding that more technologies were used to process Indian cotton since it became the main imports source to the British cotton industry during the period. The empirical strategy consisted of comparing Indian-cotton-processing technologies with the technologies used to process cotton from the U.S. and other countries before, during and after the period of the war, in a difference-in-difference setting. This study follows two empirical strategies, we estimate ordinary least squares models with fixed effects (OLS-FE) for the 1961-2015 period and instrumental variable techniques with fixed effects (IV-FE) for the 1990-2015

---

[7] Bustos et al. (2016) tested how agricultural productivity has affected industrialization in Brazil. They found that labor-saving technologies in agriculture increased labor supply in the industry between 1995 and 2006.

[8] There is an important difference in how we interpret Acemoglu's model and their interpretation. Although the empirical specification in this study is also in terms of machinery and fertilizers, we do not provide a test for the IIH per se but estimate the path through which it should hold. In addition, we provide estimates of the structural directed technical change parameters such as the elasticities of substitution between the agricultural commodities (livestock and crops) and between the traditional inputs (land and labor).

[9] This is an effort to treat the endogeneity between the commercial and traditional inputs in agriculture since it is hard to obtain exogenous shocks in the factor endowments ratio from the available data.

period. The sample is restricted to the 1990-2015 period in the IV estimations so we can use a proxy of deforestation from FAO as instrument for the land-and-labor ratio.

We find that an increase in the availability of agricultural land relative to labor increased machinery more than fertilizers in South American agriculture. Because machinery is complementary to land, we suggest that the IIH holds through a stronger market size effect, which also implies substitutability between land and labor (i.e., $\sigma > 1$) in the production of livestock and crops. Similarly, from the structural parameters of the model, we find that both outputs can be considered substitutes in the composition of the total agricultural output, but with a certain degree of complementarity. We test the different proxies for deforestation available from FAO and find that the results are robust in terms of capturing the effect of the forest on agricultural land and how that would impact machinery and fertilizers. Intensive deforestation countries have a stronger response in the machinery-and-fertilizers ratio as the land-and-labor ratio increases indicating that forest conversion is more correlated with the land endowment in these countries. We also find that the land-and-labor parameter would be overestimated in models without year fixed effects and that the inclusion of other controls could not recover the parameter from the OLS and IV models with country and year fixed effects. Finally, the inclusion of control variables in the models enabled a brief discussion on the potential determinants of technology adoption incorporated in the commercial inputs.

The paper is organized as follows. Section 2 summarizes the main results from Acemoglu's directed technical change and the application to agriculture. Section 3 provides some background on agriculture in South America and their relation to deforestation. Section 4 presents the empirical implementation. Section 5 presents the results. Finally, the last section concludes. The Appendix contains further theoretical discussions, data descriptions, and other estimation results.

1.2 Acemoglu's directed technical change applied to agriculture

Model structure and equilibrium solution

This section summarizes the directed technical change model presented in Acemoglu (2002) and provides the motivating framework to study innovations in agriculture. The framework considers an industry with two traditional inputs and one final agricultural output[10].

---

[10] Consumption in this model is given by representative consumer who maximizes utility over time subject to the CES production function of the final agricultural output.

We consider land (Z) and labor (L) as the two traditional inputs, and total agricultural production (Y) as the final output. Agricultural production is composed by two aggregate commodities, livestock (Y_Z) and crops (Y_L). Livestock is intensive in land while crops are intensive in labor. The agricultural output is represented by a CES index of the two commodities such as $Y \equiv \left[\lambda Y_L^{\frac{\varepsilon-1}{\varepsilon}} + (1-\lambda) Y_Z^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}$, where ε is the elasticity of substitution between Y_Z and Y_L and $\lambda \in (0,1)$ is the share of each commodity in the composition of the final agricultural output. The price of the agricultural output is set to 1.

Livestock and crops are each produced in perfectly competitive markets with Cobb-Douglas-type production functions represented by $Y_i = \frac{1}{1-\beta}\left(\int_0^{N_i} x_i(j)^{1-\beta}\, dj\right) \cdot i^{\beta}$ where $i \in (Z, L)$ and $\beta \in (0,1)$. The integral sums all the type-j commercial inputs ($x_i(j)$) used to produce each commodity and N_i is the total of the available commercial inputs that are complementary to input i[11]. Therefore, N_Z is the total of the land-complementary commercial inputs, represented by the quantity of machinery, and N_L is the total of the labor-complementary commercial inputs, represented by the quantity of fertilizers, all used in agriculture.

From the commodity markets, profit-maximizers producers choose the optimal levels of the traditional and commercial inputs. It is possible to obtain the derived demands for the aggregates N_Z and N_L. Likewise, the derived demands for each traditional input, Z and L, can be obtained from the following maximization problem:

$$\underset{i, x_i(j)}{\text{Max}}\ P_i Y_i - w_i i - \int_0^{N_i} q_i(j) x_i(j)\, dj, \qquad s.t.\ Y_i = f\big(i, x_i(j)\big) \tag{1}$$

where $i \in (Z, L)$, w_i is the price of input i, $q_i(j)$ is the price of the commercial input $x_i(j)$ and P_i is the price of commodity Y_i.

In the application of this study, the technology markets from Acemoglu's model are the markets of machinery and fertilizers, the commercial inputs that embody the land- and labor-complementary technologies, respectively. These markets are characterized by profit-maximizer monopolists who produce based on the prospect of greater profits. The monopoly profits are equal to $\pi_i = (q_i(j) - \psi) \cdot x_i(j)$ for $i \in (Z, L)$, where the profit-maximizing

---

[11] The commercial inputs represent inputs that embody specific technologies. Acemoglu describes these inputs as "machines".

commercial input prices are $q_Z(j) = q_L(j) = \frac{\psi}{1-\beta}$, and ψ is the monopolists' marginal cost. Acemoglu normalizes $\psi = 1 - \beta$ so that $q_Z(j) = q_L(j) = 1$.

First, we can characterize the equilibrium solution of this model for given levels of N_Z and N_L. In equilibrium, $q_Z(j) = q_L(j)$ are the commercial input prices that maximize the technology monopolists' profits. The optimal levels of Z, L, N_Z, and N_L maximize profits in the commodities markets while w_Z, w_L, P_Z, and P_L are the market clearing prices.

From the commodity producers' maximization problem in equation 1, we can obtain the aggregate commercial inputs derived demands, N_Z and N_L, in terms of Z and L. Plugging these demands into the production functions of the commodities, and inserting the results in the competitive market clearing price condition, expressed by the first equality below, gives[12]:

$$\frac{P_Z}{P_L} = \left(\frac{1-\lambda}{\lambda}\right)\left(\frac{Y_Z}{Y_L}\right)^{-\frac{1}{\varepsilon}} = \left(\frac{1-\lambda}{\lambda}\right)^{\frac{\beta\varepsilon}{\sigma}} \cdot \left(\frac{N_z Z}{N_L L}\right)^{-\frac{\beta}{\sigma}} \tag{2}$$

Now, from the inverse derived demands of the traditional inputs and the result in equation 2, we obtain the following relative inverse derived demand for land:

$$\frac{w_z}{w_L} = \left(\frac{P_Z}{P_L}\right)^{\frac{1}{\beta}}\left(\frac{N_z}{N_L}\right) = \left(\frac{1-\lambda}{\lambda}\right)^{\frac{\varepsilon}{\sigma}}\left(\frac{N_z}{N_L}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{Z}{L}\right)^{-\frac{1}{\sigma}} \tag{3}$$

where σ is the derived elasticity of substitution between land (Z) and labor (L)[13].

Figure 1 summarizes Acemoglu's directed technical change framework in the agricultural land market relative to labor. Note that equation 3 represents the inverse derived demand for land relative to labor which is a function of the quantities of land, labor, machinery, and fertilizers. One the one hand, we observe how a shift in the relative supply of land from S to S' causes a movement along the demand curve from A to B. This movement represents the case where technical change is exogenous ($\frac{N_z}{N_L}$ is constant). An increase in the relative supply of land decreases its relative price and quantity demanded increases at point B. This is the usual substitution effect reflected by the inverse relationship between $\frac{w_z}{w_L}$ and $\frac{Z}{L}$ in equation 3.

---

[12] We provide a summary of the results of Acemoglu's model as an application to agriculture. For more details and the results of all algebraic expressions, see Acemoglu (2002).

[13] This elasticity of substitution is different than the elasticity of substitution between the two commodities ε in the final output CES production function. Acemoglu derives the elasticity of substitution between Z and L to be: $\sigma \equiv \varepsilon - (\varepsilon - 1)(1 - \beta)$.

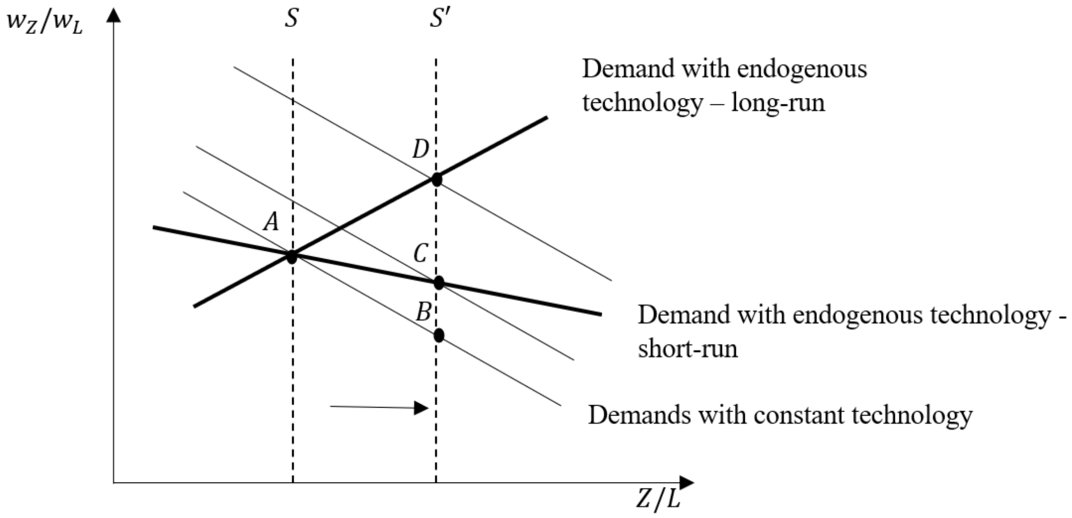

Figure 1. Acemoglu's directed technical change graph representation. Constant technology: A→B. Endogenous technology (short-run): A→ C. Endogenous technology (long-run): A→D (adapted from Acemoglu (2002, p. 784)).

On the other hand, when technical change is endogenous, the shift in $\frac{Z}{L}$ also affects the optimal $\frac{N_z}{N_L}$. One possibility is that an increase in the relative land supply creates a lot of incentive (in terms of profits) for the technology monopolists to produce more of the land-complementary commercial inputs (machinery) than the labor-complementary commercial inputs (fertilizers), and that ends up increasing the optimal use levels of $\frac{N_z}{N_L}$. This is called the "market size effect". As shown below, this is not the only effect at play in the model, there is also the "price effect". Therefore, depending on the size of the "market size effect" or the "price effect", the shift in the relative supply of land can increase or decrease the optimal levels of $\frac{N_z}{N_L}$.

From equation 3, the effect of changing the ratio of machinery and fertilizers, $\frac{N_z}{N_L}$, on the relative price of land, $\frac{w_z}{w_L}$, depends on whether $\sigma > 1$ or $\sigma < 1$. Nevertheless, when land becomes more abundant, the theory provides two main results in which there is a shift in the relative demand for land (i.e., technical change is biased towards land) irrespective of the elasticity of substitution. These are[14]:

---

[14] Hanlon (2015) was the first to test econometrically the predictions of this approach in the historical context of innovations in British cotton textile industry due to substitution of U.S. cotton by Indian cotton during the U.S. civil war. The author found that the increase in the relative supply of Indian cotton increased the relative supply of technologies to process it, and therefore, found empirical support for result 1. He also found that the relative price of Indian cotton rebounded to the higher pre-war levels despite the increase in its relative supply; this empirical evidence gives support to result 2.

1) The *weak induced-bias hypothesis*: irrespective of the elasticity of substitution between Z and L (it must be different than 1), technical change is biased towards the input that becomes more abundant. (This is shown in Figure 1 by the demand shift from B to C).
2) The *strong induced-bias hypothesis*: if $\sigma > 2$, technical change is biased towards the more abundant input, but the relative input price increases in the long-run (This is shown in Figure 1 by the demand shift from B to D).

Technology monopolists' incentives and supply of innovations

We have defined the equilibrium solution for given $\frac{N_Z}{N_L}$. However, the production of the commercial inputs depends on the technology monopolists' incentives to make them available. Monopolists, holding a patent to produce and sell their technologies, consider the discounted present value of the inventions in the decision to invest in them. The discounted present value of the commercial inputs that are used with input $i \in (Z, L)$ is given by $V_i = \frac{\pi_i}{r} = \frac{\beta[P_i]^{\frac{1}{\beta}}.i}{r}$, where π_i is the monopoly profits, P_i is the price of the commodity and r is the interest rate. This expression shows that the returns to type-i innovations (N_i) is increasing in the quantity of input i. For example, if i is land (Z) and Y_Z is livestock, then we have $V_Z = \frac{\beta[P_Z]^{\frac{1}{\beta}}.Z}{r}$. This expression shows that there are greater returns to the production of commercial inputs that are complementary to land (machinery) as more land is available ($\frac{\partial V_Z}{\partial Z} > 0$). This is the *market size effect* which is one of the two opposing effects in the model. The *price effect* is the other force at play because the returns to the production of machinery are also increasing in the price of the livestock commodity ($\frac{\partial V_Z}{\partial P_Z} > 0$). These effects are opposing because more land used in the production of livestock causes the supply of livestock to shift and the equilibrium price of the commodity decreases. Because of the direct relationship with the commodity price, the returns to innovations in machinery decreases. What determines which effect is predominant is the elasticity of substitution between land and labor, σ. Thus, a stronger market size effect causes the direction of technical change to favor inputs that are more abundant (i.e., $\frac{N_Z}{N_L}$ increases with $\frac{Z}{L}$) while a stronger price effect would favor the inputs that are scarcer (i.e., $\frac{N_Z}{N_L}$ decreases with $\frac{Z}{L}$).

The supply side of innovations is also considered in the model. The changes in the innovation possibility frontier (i.e., which reflects changes in the relative marginal costs of innovation) may affect the direction of technical change[15].

The production functions of new commercial inputs are $\Delta N_Z = c_Z R_Z$ and $\Delta N_L = c_L R_L$ where R_Z and R_L are R&D expenditures to produce commercial inputs for livestock (Y_z) and crops (Y_L), respectively. The parameters c_Z and c_L reflect the marginal opportunity costs of these innovations because there is a fixed R&D budget to allocate in both. Notice that the opportunity cost to produce N_Z is c_L. Notice also that $\frac{\partial \Delta N_Z}{\partial R_Z} = c_Z$ and $\frac{\partial \Delta N_L}{\partial R_L} = c_L$ does not depend on the levels of N_Z and N_L, so c_Z/c_L is constant[16].

In the balanced growth path (steady-state equilibrium), prices are constant (P_Z and P_L) and N_Z and N_L grow at the same rate. This means that the monopolists' profit ratio $\frac{V_Z}{V_L}$ is constant. Acemoglu defines one last condition for the technology monopolists to invest in both types of commercial inputs: the relative profitability should equate the relative opportunity costs: $\frac{V_Z}{V_L} = \frac{\pi_Z}{\pi_L} = \frac{c_L}{c_Z}$ or $c_Z \pi_Z = c_L \pi_L$. This condition is required so that there is incentive to innovate in both sectors in the steady-state equilibrium.

By solving this last condition in terms of the market clearing price condition, the equilibrium solution in the commodity markets, and the monopoly profits, the following result is obtained:

$$\frac{N_Z}{N_L} = \left(\frac{c_Z}{c_L}\right)^{\sigma} \left(\frac{1-\lambda}{\lambda}\right)^{\varepsilon} \left(\frac{Z}{L}\right)^{\sigma-1} \tag{4}$$

Equation 4 shows that a change in the factor endowments $\left(\frac{Z}{L}\right)$ affects the optimal quantities of commercial inputs $\left(\frac{N_Z}{N_L}\right)$ used. These quantities are also affected by the opportunity costs of innovations $\left(\frac{c_Z}{c_L}\right)$ and the share of the commodities in the composition of the agricultural output, captured in λ and $(1-\lambda)$. Therefore, the model generates the following testable prediction for the agricultural sector:

---

[15] The innovation possibility frontier is like a production possibility frontier (PPF) for innovations. It was introduced by Kennedy (1964) to capture the trade-off between different innovations. In his model of induced innovation, firms choose the technologies to maximize "cost reduction" for given factor proportions.

[16] This is what Acemoglu defined as "no state dependence" where the current levels of technologies do not affect future costs of innovations. In most models of endogenous growth, such as the models of "learning by doing", this assumption is relaxed and spillovers from past R&D change future costs.

PREDICTION 1. *Let the total agricultural output be composed of livestock and crops commodities and let these commodities be produced using traditional (land and labor) and commercial inputs (machinery and fertilizers). When the competitive commodity markets are in equilibrium, and when the producers of machinery and fertilizers are profit-maximizers technology monopolists, an increase in the land-and-labor ratio* (Z/L) *must increase (decrease) the machinery-and-fertilizers ratio* $\left(\frac{N_Z}{N_L}\right)$ if $\sigma > 1$ $(\sigma < 1)$.

This solution links the markets of land and labor with the markets of machinery and fertilizers through the production of livestock and crops commodities. Thus, the model provides the process of the induced innovation where the direction of the technological progress is affected by the factor endowments instead of the factor prices. The *weak-induced bias* can be shown by combining PREDICTION 1 with equation 3, where the term $\left(\frac{N_Z}{N_L}\right)^{\frac{\sigma-1}{\sigma}}$ always increases, irrespective of σ (although it must be different than 1). The *strong-induced bias* can be seen by plugging equation 4 into equation 3:

$$\frac{w_Z}{w_L} = \left(\frac{c_Z}{c_L}\right)^{\sigma-1} \left(\frac{1-\lambda}{\lambda}\right)^{\varepsilon} \left(\frac{Z}{L}\right)^{\sigma-2} \tag{5}$$

Now we can see clearly that the shift in the relative demand for land induced by the change in the relative supply can increase the relative price of land if $\sigma > 2$.

1.3 Background

Agriculture in South America

This section presents the evolution of the agricultural production and resources in the ten South American countries for the 1961-2015 period. The data used in this section is from FAO and Fuglie (2015), described in detail below. The descriptive analysis focuses on two groups of five countries, named "deforesters" and "non-deforesters" according to the Hosonuma et al. (2012) and Lablois et al. (2017) studies[17]. The deforesters group are the countries that were classified as having intensive deforestation around the period of their study. The non-deforesters were countries that were classified in the other forest transition phases

---

[17] Hosonuma et al. (2012) created a forest transition indicator to identify whether countries are moving from deforestation to reforestation. Lablois et al. (2017) defined these phases as, **Phase 1: "undisturbed forests", Phase 2: "intensive deforestation"**, **Phase 3: "transition is occurring"**, and **Phase 4: "net forest cover is increasing**". They used this classification to control for the forest transition phases in their study of the determinants of deforestation in developing countries.

other than “intensive deforestation”, hence they may still have some significant deforestation rates (Hosonuma et al., 2012; Lablois et al., 2017). The deforesters are Brazil, Bolivia, Ecuador, Paraguay, and Venezuela, and the non-deforesters are Argentina, Chile, Colombia, Peru, and Uruguay[18].

Alternatively, in the empirical analysis, we define “FAO deforesters” according to the FAO’s report “State of the World’s Forests” (FAO, 2016). This report indicated the changes in forest area and agricultural area for the countries in the world for the 2000-2010 period. Brazil, Paraguay, Argentina, and Peru presented net gain in agricultural area and net loss in forest area. These are the “FAO deforesters” group used in some estimations of the empirical model shown in the Results section. According to the same report, Bolivia and Venezuela presented small change while Ecuador and Colombia had loss in both land areas. Uruguay had net loss in agricultural area and net gain in forest area. Chile was the only country with net gain in both.

Figure 2 shows the evolution of the average agricultural output composed by the values of livestock and crops commodities, in million USD, for the deforesters and non-deforesters groups (according to the Hosonuma-Lablois classification). This figure illustrates the discrepancy between the agricultural sector in both groups, indicating that despite starting at similar output levels, growth in the deforesters was much more accentuated, reaching an average of almost $40 billion USD in 2015, while the average of the other group was below $20 billion. Despite the differences in size and agricultural potential, this difference can indicate a positive correlation between deforestation and agricultural expansion, also shown in several studies for South America (Schmitz et al., 2015; Faria and Almeida, 2016; FAO, 2016, Koch et al., 2019).

---

[18] The countries in Phase 1 are Colombia and Peru. Phase 2: Brazil, Bolivia, Venezuela, Paraguay, and Ecuador. Phase 3: Argentina. Phase 4: Uruguay and Chile. For this section, we define the deforesters as those countries in Phase 2, and the non-deforesters, the ones in the other phases.

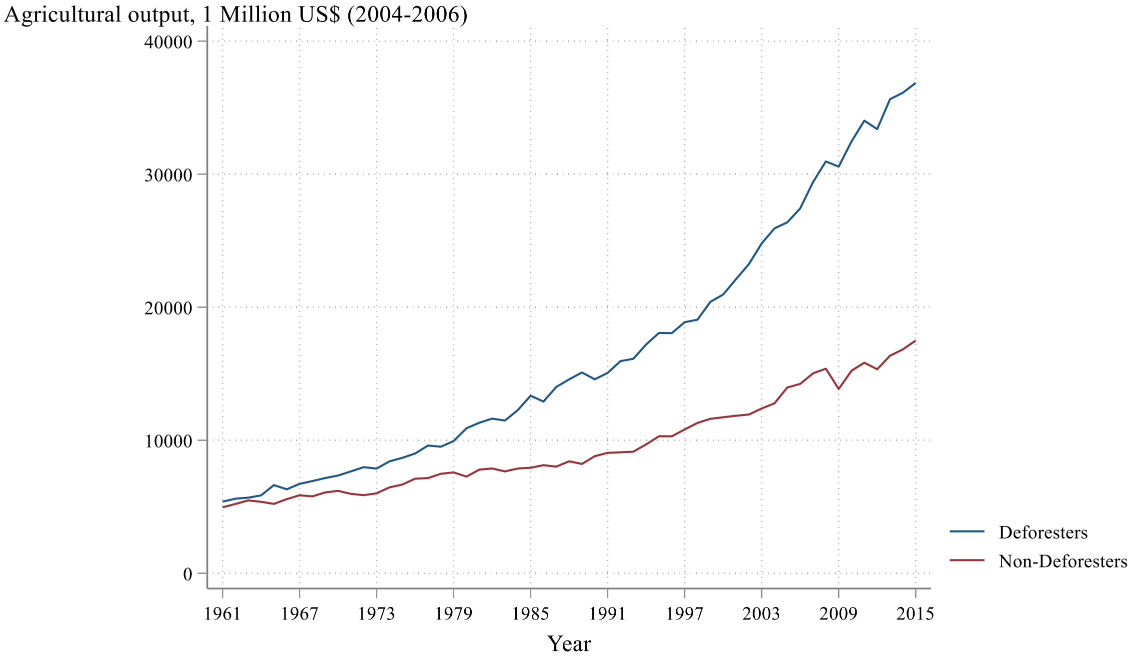

Figure 2. Evolution of Agricultural Output (1 million US$ Constant 2004-2006) for Deforesters and Non-Deforesters (average), 1961-2015.

The expansion of agricultural production can arise from productivity gains and/or the use of more resources to agriculture. Trindade and Fulginiti (2015) investigated whether there was a slowdown in productivity in the same set of South American countries. They found that productivity accounted for half of the output growth for the 1969-2009 period and most productivity gains were related to innovations. We study agricultural productivity by following a different approach than theirs. The focus is on the use of resources and the relationship between the traditional and commercial inputs. Figure 3 shows the evolution of the inputs in agricultural production (measured in logarithm) for the average of the deforesters and non-deforesters groups.

We observe constant use of land for the non-deforesters but increasing use of land for the deforesters. Although labor seems to be increasing for both groups, the average levels of the deforesters are higher, which indicates that agriculture is more labor-intensive, although with slower growth from mid-1990s. Interestingly, the logarithm of fertilizers and machinery had higher levels over the whole period for the non-deforesters. Given the greater availability of resources in the deforesters[19] and the substitutability relationship between the traditional and commercial inputs, it is no surprise that agriculture should be more intensive in fertilizers and machinery in countries with less endowments of land and labor. However, the use of machinery

---

[19] Brazil has the largest representation of the group in the size of the agricultural sector, resources, and the potential for agricultural expansion.

increased substantially for the deforesters while remaining relatively constant for the other group. Fertilizers use had growth for the average of all countries.

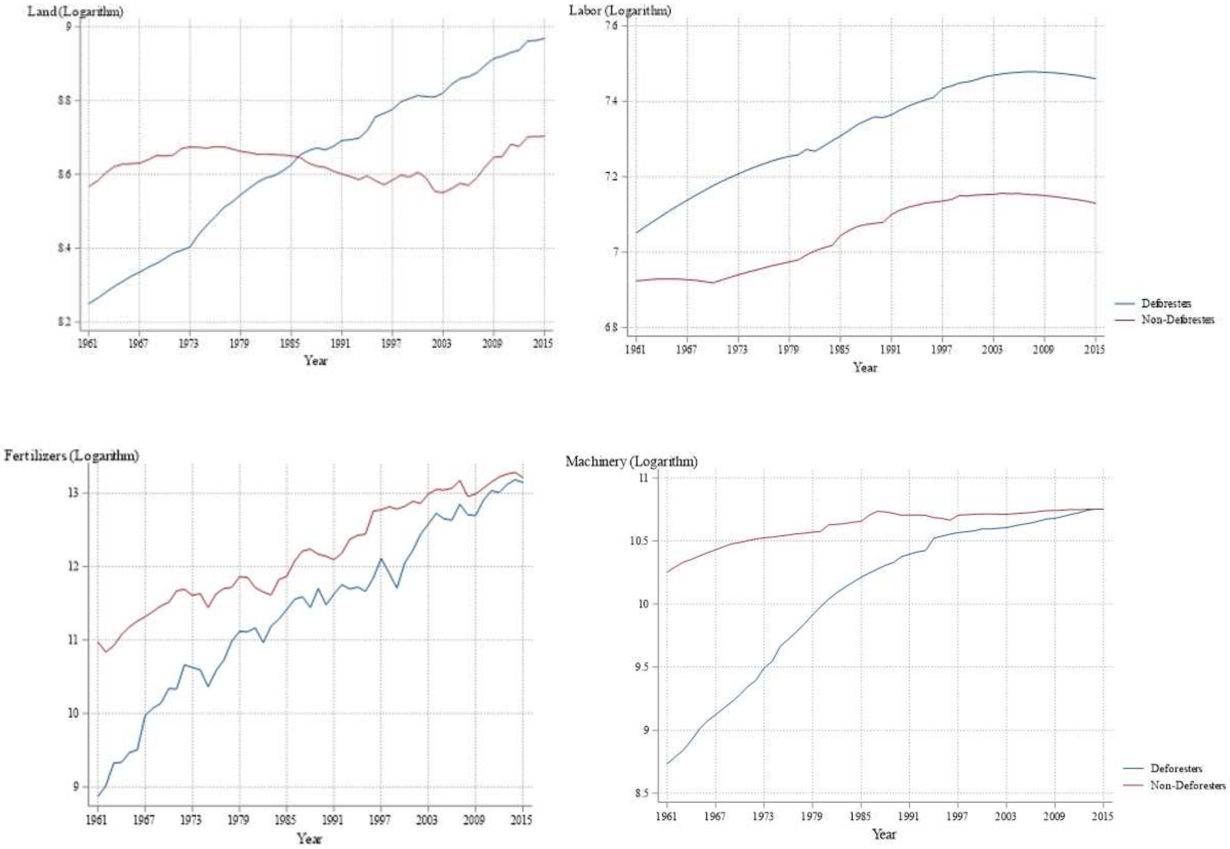

Figure 3. Evolution of Land (top-left), Labor (top-right), Fertilizers (bottom-left) and Machinery (bottom-right) (in logarithm) for Deforesters and Non-Deforesters (average), 1961-2015.

Deforestation and agricultural expansion

One fundamental question for the discussion presented in this paper is whether agriculture affected deforestation or deforestation affected agriculture in South America. Most studies focused on the first direction of causation through agricultural international trade (Schmitz et al, 2015; Faria and Almeida, 2016; Leblois et al., 2017). One example of how agriculture and deforestation are related is the slash-and-burn method. It is the clearing of vegetation and forestland to produce crops. This is a method that increases soil fertility in the short run, however, because these new cleared areas become degraded with time, deforestation tends to increase. Even though, deforestation may be caused by agriculture in this context, the only way it should affect new agricultural enterprises, and hence agricultural expansion/productivity, is through increasing land availability. Deforestation can be related to the land endowment and, therefore, the supply of land for agriculture. This is not an unrealistic assumption because of the explicit trade-off between preserving the forest and expanding agriculture, especially in South America where the agricultural frontiers significantly reduced the areas from the Amazon forest in Brazil (Silva, Perrin, and Fulginiti, 2019) and the regions from the Gran Chaco and Chiquitano in Argentina, Bolivia, and Paraguay (le Polain de Waroux et al., 2016).

In the empirical part, we use deforestation as an instrument for relative land supply (land-and-labor ratio) to explain the use of inputs that embody complementary technologies (machinery-and-fertilizers ratio). The direction of causation in this case is from deforestation to agriculture. The papers of le Poulain de Waroux et al. (2019) and Koch et al. (2019) studied the impact of deforestation policies on agricultural expansion in South America. The former found that control on deforestation decreased the production of soy and beef which was attenuated by shifts of production to less regulated areas. The latter studied a specific deforestation policy in the Brazilian Amazon region, the Priority list, where municipalities with high deforestation were targeted by increased enforcement of the environmental laws, increased monitoring, restricted credit, and reputation damage. They found that controlling deforestation induced productivity gains in cattle production (cattle per hectare) and had no effect on dairy or crop production. These studies are examples on how deforestation patterns can relate to investments in the farms and agricultural productivity.

1.4 Empirical implementation

Data

We use panel data from 10 South American countries for the 1961-2015 period. We combine data from FAO, Fuglie (2015) and other sources. We use agricultural land from Fuglie (2015) which is a quality-adjusted measurement of land expressed in hectares of rainfed cropland[20]. Labor is from FAO and is measured in number of agricultural workers (male and female), fifteen years old or more. Machinery and fertilizers are both from Fuglie (2015); these variables were constructed as to allow for international comparisons. Machinery is measured in 40-cv tractor units and fertilizer in metric tons[21].

We also use data from the World Development Indicators (WDI) and institutional variables data from Polity IV project (2017) and The Freedom House (2018) indicators.

---

[20] Agricultural land is an aggregation of cropland (rainfed and irrigated) and pastureland weighted by relative quality. This variable is expressed in rainfed cropland equivalents for international comparisons.

[21] Fuglie (2012, 2015) describes the methodologies for the construction of these variables in more detail.

Variable descriptions, sources and descriptive statistics are in Tables B.1 and B.2 in the Appendix.

Empirical model

The application of directed technical change model to agriculture established, in equation 4, a relationship between the ratios of the traditional inputs and the commercial inputs. PREDICTION 1 summarizes this relationship in terms of the elasticity of substitution, σ, between land and labor. Taking the logarithm of equation 4 yields:

$$\ln\left(\frac{N_Z}{N_L}\right) = \sigma \ln\left(\left(\frac{c_Z}{c_L}\right)\right) + \varepsilon \ln\left(\left(\frac{1-\lambda}{\lambda}\right)\right) + (\sigma - 1)\ln\left(\frac{Z}{L}\right) \tag{6}$$

where c_Z and c_L are the marginal opportunity costs of innovating in the land-complementary (N_Z) and the labor-complementary commercial inputs (N_L), respectively. λ and $(1-\lambda)$ are the shares of the labor-intensive commodity (crops) and the land-intensive commodity (livestock) in the final agricultural output. ε is the elasticity of substitution between crops and livestock in the CES production index, and the traditional inputs elasticity of substitution is given by σ.

Equation 6 can be estimated if we have information on the marginal opportunity costs of innovations, shares of crops and livestock in the agricultural output, and the factor endowments. Given the panel structure of the data, the effects of the marginal costs can be estimated through country fixed effects (country-FE), year fixed effects (year-FE), and the inclusion of other control variables. Adding a random error term, ϵ_{ct}, to equation 6, the following estimating equation can be obtained:

$$ln\left(\frac{Machinery}{Fertilizers}\right)_{ct} = \alpha + \varepsilon\, ln\left(\frac{Livestock\%}{Crops\%}\right)_{ct} + \beta ln\left(\frac{Land}{Labor}\right)_{ct} + X_{ct}\theta + \mu_c + \varphi_t + \epsilon_{ct} \tag{7}$$

where c indexes the country and t indexes time, $Machinery_{ct}$ represents the commercial inputs that embody the land-complementary technologies, $Fertilizers_{ct}$ represents the commercial inputs that embody the labor-complementary technologies[22], $Livestock\%_{ct}$ is the share of the value of livestock commodities and $Crops\%_{ct}$ is the share of the value of crops in the total value of agricultural GDP. $Land_{ct}$ is the sum of cropland and pastureland, $Labor_{ct}$ is the number of people (above fifteen) working in the agricultural sector, X_{ct} represents other control variables that differ across countries and time, μ_c are the country-specific fixed effects (FE) and

---

[22] Land-complementary technologies can also be interpreted as land-using (labor-saving) for a production function with only land and labor as traditional inputs.

λ_t are the year-specific FE. The parameters to be estimated are α, ε, β, μ_c, φ_t and a vector of parameters θ.

The response to the factor endowments ratio is measured in terms of the parameter β; if it is positive and statistically significant then, according to PREDICTION 1, as land becomes more abundant relative to labor, there is an increase in the commercial inputs that embody technologies that are complementary to land (machinery) relative to the inputs that embody technologies that are complementary to labor (fertilizers). This implies that the elasticity of substitution between land and labor is greater than 1 ($\sigma > 1$) and, thus, they are considered substitutes in production. An elasticity of substitution greater than 1 would also imply a stronger "market size effect" in the process of induced innovation, i.e., the development and the adoption of commercial inputs are greater for the inputs that are complementary to the more abundant resource[23].

Identification

Equation 7 presents several identification challenges if we want to find a causal relationship. One is the simultaneity bias: the decision to use commercial inputs can also affect the use of land and labor in agriculture. The other source of endogeneity arises when unobservable variables affect both the use of commercial inputs and traditional inputs. These could be, for example, economic conditions, institutional setting, environmental regulations, among other factors. Using country and year fixed effects (FE) helps capture the effects of the marginal opportunity costs of innovations from the theory but also decreases the endogeneity problem due to simultaneity and/or omitted variables. However, the endogeneity problem can persist if there are omitted variables that cannot be captured in the country- and year-FE. In this case, one must find suitable instrumental variables (IV) for the factor endowment ratio and obtain an IV estimate of β.

The assumption we make to identify the changes in the land-and-labor ratio is through exogenous variables that must affect the endowments and hence the supplies of these resources. Historically, agricultural land expansion in the set of South American countries studied in this paper is related to the conversion of forests (le Poulain de Waroux et al., 2016; Silva, Perrin, and Fulginiti, 2019). Therefore, we use "net forest conversion" as an instrument for the land-

---

[23] This is the interpretation from Acemoglu's directed technical change theory. The counterpart interpretation from the Hayami-Ruttan classic IIH is that as the factor becomes more available, then technical change is factor-using. In addition, the induced innovation process would be captured through the changes in the factor prices.

and-labor ratio to capture the effect of deforestation patterns on the agricultural land supply[24]. Since deforestation is mainly correlated with land, we need to include an instrument(s) that captures the dynamics of labor. We include as an instrument for labor the number of people in the rural areas of these countries (rural population). The use of instruments is conditional on them being good predictors of the land-and-labor ratio (i.e., instruments are valid). The second condition is that the instruments must not be correlated with the error term of the main equation of interest (i.e., the exclusion restriction). The proxy for deforestation is the area of forests that were converted to other land uses and, according to FAO (2016), most of these areas were converted to agriculture in South America. Therefore, this should represent an exogenous shock to the land endowment. Likewise, the dynamics of the rural population should also represent changes in the labor endowment. One must be aware that all the IV results are conditional on these assumptions to hold. Using instrumental variables is an attempt to treat the endogeneity problem in the process of the induced innovation obtained from Acemoglu's directed technical change structural model. Although the scope of this paper is not to study the instruments themselves, we provide an advance in the discussion on how deforestation and/or deforestation control policies can be sources of variation for agricultural land and their potential effects on agricultural productivity. In addition, the IV approach is a simple way to establish a link between deforestation and the directed technical change. It can play a fundamental role in process of innovations in the agricultural sector, especially in the South American countries with intensive deforestation.

1.5 Results

OLS main estimation results

The OLS estimates of equation 7 with and without fixed effects (FE) are presented in this section. Table 1 shows the results without FE and no controls (column 1), with fixed effects (OLS-FE) and basic controls (column 2), and the OLS-FE estimations weighted by average agricultural GDP (column 3) and by the average population (column 4). In addition to factor endowment ratio and the ratio of the share of livestock and crops, all FE models contain three

---

[24] This variable is one of three variables available from FAOSTAT to measure the effects of deforestation, reforestation, and afforestation within the countries. The other two are net CO_2 emissions and forestland. We use net forest conversion as our main instrument but also estimate with the other definitions as a robustness test.

basic control variables[25]: share of arable land[26], share of the credit provided by financial institutions, and a measurement of the average annual temperature change. The temperature change variable is calculated by the FAO and measures the average meteorological year anomaly in Celsius degrees compared to a baseline climatology period. The share of arable land is to account for the changes in land area that are specific to temporary crops and land that is temporarily fallow. Controlling for arable land can isolate the decision of farmers to expand land use through land that has already been used in agriculture. Leaving this variable out of the estimation could significantly bias the OLS estimate of β, because we want to capture the effect of changes in the agricultural land that are caused by exogenous land supply shocks (i.e., changes in the factor endowment). The other controls are also an attempt to identify omitted variables that could affect the decisions of using the traditional and the commercial inputs in agriculture. We include them as proxies for available credit and the effects of weather.

In columns 1-4, the coefficient of the logarithm of the factor endowment ratio varies from 0.68-0.74 and are all statistically significant at the conventional levels of significance. It implies that a one-percent increase in the land-and-labor ratio increases the machinery-and-fertilizers ratio in 0.7 percent in agriculture. The ratio of the shares of livestock and crops has statistically significant coefficients at the conventional levels in the OLS-FE models (columns 2-4) ranging from 0.52-0.96. The share of arable land coefficient is consistently estimated to be negative and significant at 5-percent, while the coefficients for the other controls were not significant at any conventional level. Using the R^2-adjusted and the statistical information criteria (AIC and BIC), the preferred model is given in column 3. This gives us an estimate of the directed technical change coefficients of 0.68 (standard error = 0.265) for β and 0.96 (standard error = 0.267) for ε.

From the structural model, recall that ε is the elasticity of substitution between livestock and crops in the CES production index, and σ is the elasticity of substitution between land and labor. The obtained estimates are $\hat{\sigma} = \left(1 + \hat{\beta}\right) = 1.68$ and $\hat{\varepsilon} = 0.96$. The former indicates that land and labor are substitutes in the production of the agricultural commodities, and the latter

---

[25] We only have 550 observations and adding too many controls in the models with country-FE and year-FE can raise concerns to the degrees of freedom in the estimation.

[26] Arable land is obtained from the WDI, which is calculated based on the FAO definition as: land under temporary crops (double-cropped land is counted once), temporary meadows for mowing or for pasture, land under market or kitchen gardens, and land temporarily fallow (WDI metadata glossary).

indicates that the livestock and crops commodities are also substitutes in the composition of the final agricultural output, although with a small degree of complementarity.

Table 1: OLS and OLS-FE results, 1961-2015

Dependent variable: **Ln (Machinery/Fertilizers)**				
	(1)	(2)	(3)	(4)
	OLS w/o FE and controls	OLS-FE w/ basic controls	OLS-FE weighted by Ave. Ag. GDP	OLS-FE weighted by Ave. Population
ln (Land/Labor)	0.705***	0.736**	0.676**	0.689*
	(0.211)	(0.297)	(0.265)	(0.305)
ln (Livestock%/Crop%)	0.235	0.519**	0.958***	0.778*
	(0.240)	(0.215)	(0.267)	(0.345)
Arable land (%)		-0.203**	-0.163**	-0.166**
		(0.0631)	(0.0504)	(0.0604)
Financial credit (%)		0.000187	0.000297	0.000351
		(0.00210)	(0.000936)	(0.00114)
Temperature change		-0.0168	0.0242	0.0409
		(0.0944)	(0.0639)	(0.0853)
		(0.195)	(0.212)	(0.261)
Constant	-2.411***	-0.442	-0.341	-0.667
	(0.509)	(0.305)	(0.340)	(0.375)
Country-FE	No	Yes	Yes	Yes
Year-FE	No	Yes	Yes	Yes
Obs.	550	550	550	550
R^2 within		0.758	0.877	0.822
R^2 between		0.265	0.275	0.296
R^2 overall		0.454	0.447	0.464
R^2 adjusted	0.284	0.728	0.863	0.800
AIC	1563.9	457.6	69.21	152.9
BIC	1576.9	496.4	108.0	191.7

Note: Model 1 is a simple OLS regression without fixed effects and controls. Models 2-4 include country-FE, year-FE, and includes the basic controls (share of arable land, share of financial credit and average annual temperature change), which unless specified are in all models. In Model 3, the observations are weighted by the average agricultural GDP and, in Model 4, by average total population, for the 1961-2015 period. Robust standard errors in parentheses. *, **, and ***, denote significance at the 10, 5 and 1 percent level.

Without year-FE and including other controls[27]

Table 2 presents the OLS results without the year-FE but with the country-FE, the basic controls, and a new set of control variables, to observe the extent that the year-FE and potential omitted variables would affect the estimates of the main parameters of the model. We present these results to investigate whether these new explanatory variables have a role in determining the decisions of using machinery and fertilizers in agriculture. Columns 1-5 vary according to

---

[27] This section is inspired by Mundlak, Butzer and Larson (2012) on estimating heterogeneous agricultural production functions using country-level panel data. They focued on capturing the economic environment and on the replacement of country- and year-FE by observed state variables.

the inclusion of the explanatory variables while all models include the basic controls and country-FE.

Column 1 includes the proxies that capture the overall effect of past agricultural prices and agricultural trade (lagged agricultural GDP and lagged value of agricultural exports). Column 2 includes the set of institutional variables[28]. Column 3 includes the lagged agricultural GDP, lagged export values, and the institutional variables in the same model. Column 4 adds to the model in column 1 the logarithm of the value of forest exports. Finally, column 5 includes the whole set of controls.

The land-and-labor ratio parameter is estimated on the range of 1.12-1.20, and the parameter of the ratio of the shares of livestock and crops, on the range of 0.6-1.17. The coefficient of the share of arable land has a robust estimate around -0.2 but is only significant at the conventional levels in columns 1 and 4. The coefficient of the share of financial credit has positive and statistically significant effect, except in column 2, and the coefficient of the temperature change is not statistically significant at any conventional level. Because the variables are in logarithm, we estimate a negative and statistically significant elasticity for the lagged values of agricultural GDP, ranging from -1 to -1.2, and positive and significant elasticity (0.14) for the lagged value of agricultural exports in columns 3 and 5.

In column 4, we observe that an increase in 1-percent in the value of forest exports decreases the ratio of machinery and fertilizers by -0.08 percent, which indicates a statistically significant tradeoff between the forest and agriculture. In column 5, this effect is not statistically significant but continues to be negative while the lagged value of agricultural exports is statistically significant and positive. This indicates that the exports of agricultural products have a stronger effect on the allocation of resources to agriculture. The institutional variables do not have overall statistical significance in the models with country-FE and the other controls (columns 3 and 5). However, in column 2, the coefficients of Polity 2 and Durable are both negative and statistically significant at the 5-percent level. The only institutional variable that is significant in the full model (column 5) is the measure of civil liberties, indicating that countries with less institutional freedom would use less machinery relative to fertilizers.

---

[28] The Political rights and Civil liberties indicators, obtained from The Freedom House (2018), are measures of institutional quality, they vary from 1-7, where 1 indicates the highest level of freedom and 7 indicates the lowest. The other two institutional variables, Polity 2 and Durable, from the Polity IV project (2017), are indicators of democracy, from most democratic to the most autocratic regimes. Polity 2 varies from -10 to 10, where the lowest score indicates the most autocratic regime, and the highest score indicates the most democratic regime. Durable indicates the number of years since the last regime change and is a measure of political stability.

Table 2: OLS results with country-FE and other controls, 1961-2015

Dependent variable: **Ln (Machinery/Fertilizers)**					
	(1)	(2)	(3)	(4)	(5)
	OLS-FE w/ ag. GDP & export value	OLS-FE w/ institutional variables	OLS-FE w/ ag. GDP, export value and institutional variables	OLS-FE w/ ag. GDP, export value and forest exports	OLS-FE w/ all controls
ln (Land/Labor)	1.116**	1.237*	1.169**	1.205***	1.117**
	(0.356)	(0.657)	(0.469)	(0.299)	(0.481)
ln (Livestock%/Crop%)	1.153**	0.623	1.117**	0.952**	1.167**
	(0.421)	(0.507)	(0.487)	(0.325)	(0.377)
Arable land (%)	-0.223**	-0.288	-0.212	-0.251***	-0.197
	(0.0767)	(0.165)	(0.134)	(0.0620)	(0.121)
Financial credit (%)	0.00163***	0.00121	0.00198**	0.00198***	0.00226***
	(0.000415)	(0.000663)	(0.000735)	(0.000496)	(0.000649)
Temperature change	-0.0446	-0.0813	0.0854	-0.0153	0.0840
	(0.0945)	(0.160)	(0.111)	(0.0934)	(0.100)
$\ln(\text{Ag. GDP})_{-1}$	-1.181***		-1.174***	-1.002**	-1.218***
	(0.337)		(0.209)	(0.327)	(0.260)
$\ln(\text{Exports value})_{-1}$	0.0399		0.147*	0.0978	0.149**
	(0.0839)		(0.0704)	(0.0987)	(0.0633)
Civil Liberties		-0.0533	-0.0803		-0.0971**
		(0.0450)	(0.0468)		(0.0382)
Political Rights		-0.0410	-0.0162		-0.0291
		(0.0342)	(0.0322)		(0.0269)
Polity 2		-0.0256**	-0.00715		-0.00505
		(0.00850)	(0.00805)		(0.00587)
Durable		-0.0350**	-0.0208*		-0.0171
		(0.0116)	(0.00989)		(0.00963)
ln (Forest Exports)				-0.0851***	-0.0507
				(0.0240)	(0.0848)
Constant	17.84***	-0.861*	16.38***	14.94***	17.83***
	(4.518)	(0.436)	(2.933)	(4.128)	(3.438)
Country-FE	Yes	Yes	Yes	Yes	Yes
Year-FE	No	No	No	No	No
Obs.	540	430	430	515	414
R^2 within	0.758	0.704	0.748	0.777	0.766
R^2 between	0.186	0.347	0.249	0.209	0.240
R^2 overall	0.276	0.409	0.304	0.326	0.305

Note: All models contain country-FE, basic controls, but no year-FE, and the observations were weighted by the average agricultural GDP for the 1961-2015 period. Model 1 adds to the basic control variables (share of arable land, share of financial credit and temperature change) the lagged values of agricultural GDP and exports in logarithm. Model 2 contains the basic controls and institutional variables. Models 3 combines the variables in Models 1 and 2. Model 4 adds to Model 1 the logarithm of the value of forest exports. Model 5 includes all control variables. The observations differ because the institutional variables are available from 1973, and the zero values in forest exports. Robust standard errors in parentheses. *, **, and ***, denote significance at the 10, 5 and 1 percent level.

Not including the year-FE would overestimate the coefficient of the land -and-labor ratio, and the inclusion of new controls is not sufficient to capture these effects. This means that the year-FE not only capture the effects of these omitted variables but also of other unknown variables in the model. Because we cannot measure the changes in the marginal opportunity costs of innovation, country-FE and year-FE are necessary to obtain unbiased estimates of the main parameters of this framework. Overall, we observe that the lagged agricultural GDP has a negative effect on the machinery-and-fertilizers ratio and that lagged value of agricultural exports has a positive effect. An increase in the value of forest exports seems to reduce the ratio of machinery and fertilizers in agriculture and the institutional variables have small or no effect in the use of machinery and fertilizer in agriculture.

IV main estimation results

This section presents the IV estimations for the 1990-2015 period. Because we are unable to establish exogenous shocks in the relative land supply from the data, the IV estimation is more appealing if the endogeneity between the commercial and traditional inputs persists after controlling for the country- and year-FE. We restrict the period of the sample because of the availability of the deforestation proxies in the FAO dataset. As the main deforestation instrumental variable, we use the area of net forest conversion which is a measurement in hectares of the change in land use from forest to agriculture and other practices. We also include rural population as an instrumental variable to explain the land-and-labor ratio, considering it as a strong determinant of the agricultural labor.

Table 3 presents in column 1 a simple OLS estimation without controls or fixed effects for the 1990-2015 period. Column 2 presents the result of the IV-FE with basic controls, and column 3, the IV-FE with basic controls, lagged agricultural GDP and lagged exports value. Column 4 presents the same IV-FE model as in column 3 but with the observations weighted by the average agricultural GDP. The OLS model in column 1 allows to compare the extent which the parameters are biased in the presence of endogeneity in the model. We see that the coefficient of the ratio of land and labor is downward biased as compared to the IV-FE model in column 2 (0.53 vs. 0.67) while the standard error reduces a little in column 2.

The inclusion of the lagged agricultural GDP and the lagged exports value indicates a larger downward bias in column 3 (0.53 vs. 0.77) and a larger reduction in the standard error. The R^2 adjusted and the information criteria (AIC and BIC) indicate a better specification of the model with the inclusion of these controls. Because of the heterogeneity in the agricultural sector in these countries, we weight the observations by the average agricultural GDP for the period in column 4. The R^2 adjusted increases to 0.83 and the AIC and BIC are the lowest as compared to the other models. This is our preferred IV model (later, 'baseline IV-FE' model) and we are mainly interested in interpreting its parameters for the 1990-2015 period.

Table 3: OLS and IV-FE results, 1990-2015

Dependent variable: **Ln (Machinery/Fertilizers)**

	(1) OLS w/o FE and controls	(2) IV-FE w/ basic controls	(3) IV-FE w/ ag. GDP and export value	(4) IV-FE weighted by Ave. ag. GDP and w/ ag. GDP and export value
ln (Land/Labor)	0.528^{**}	0.669^{***}	0.772^{***}	0.400^{**}
	(0.195)	(0.190)	(0.184)	(0.170)
ln (Livestock%/Crop%)	0.274	0.286	0.375^{**}	0.751^{***}
	(0.172)	(0.175)	(0.171)	(0.188)
Arable land (%)		-0.174^{***}	-0.174^{***}	-0.106^{***}
		(0.0383)	(0.0371)	(0.0300)
Financial credit (%)		0.00272^{**}	0.00211^{*}	-0.000473
		(0.00130)	(0.00125)	(0.000958)
Temperature change		-0.0670	-0.0600	-0.0842
		(0.0937)	(0.0895)	(0.0814)
$\ln(\text{Ag. GDP})_{-1}$			-0.849^{***}	-0.709^{***}
			(0.239)	(0.256)
$\ln(\text{Exports value})_{-1}$			0.107^{***}	0.0972^{**}
			(0.0329)	(0.0391)
Constant	-2.637^{***}			
	(0.456)			
Country-FE	No	Yes	Yes	Yes
Year-FE	No	Yes	Yes	Yes
Obs.	260	260	260	260
R^2 adjusted	0.297	0.646	0.658	0.828
AIC	606.4	128.7	121.2	-112.8
BIC	617.1	235.5	235.1	1.130

Note: Model 1 is a simple OLS regression without FE and controls. Model 2 is the IV-FE model with the basic controls. Model 3 adds to the basic controls the lagged values of agricultural GDP and exports, in logarithm. Model 4 is the IV-FE model, as in Model 3, with observations weighted by the average agricultural GDP for the 1990-2015 period. Robust standard errors in parentheses. *, **, and ***, denote significance at the 10, 5 and 1 percent level.

The directed technical change parameters indicate a statistically significant coefficient for the land-and-labor ratio of 0.4, which means an elasticity of substitution between land and labor, σ, of 1.4. The elasticity of substitution between livestock and crops, ε, is estimated at 0.75, also statistically significant at the standard levels. This result is consistent with the estimation of the preferred OLS-FE (Table 1, column 3) for the longer period in the previous section. Both elasticities of substitution indicated substitution between land and labor and a certain degree of complementarity between livestock and crops, although the IV-FE model shows stronger complementarity (0.75).

The other results show a negative and statistically significant effect for the share of arable land (-0.11), positive and statistically significant effect for the share of financial credit in columns 2 and 3 but not significant at any conventional level in column 4. Like in the previous OLS estimations, the temperature change coefficient is not significant in any of the IV estimations. The lagged agricultural GDP has a negative and statistically significant parameter (-0.71), and the lagged value of agricultural exports has a positive and significant parameter at the conventional levels (0.10). Since the variables are in logarithm, it implies that, on average, a 1-percent increase in the agricultural GDP in the previous year reduces the current use of machinery per ton of fertilizers in 0.71-percent, while an increase in 1-percent in the value of agricultural exports would cause an increase of 0.10-percent in this ratio.

Table 4 compares the estimated directed technical change parameters of the baseline IV-FE model for all countries (column 1) and for the two groups of deforesters. The first group is composed by Brazil, Bolivia, Ecuador, Paraguay, and Venezuela, which are the countries with intensive deforestation by the Hosonuma et al. (2012) and Lablois et al. (2017) classification. The second group is composed by Brazil, Argentina, Paraguay, and Peru. These are the countries with forest area reduction and agricultural area expansion as reported in FAO (2016).

The land-and-labor ratio parameter is larger in magnitude for both subsamples. The standard errors are lower relative to the magnitude of the parameters, which indicates more efficient IV estimators for the groups separately. A larger effect is found for the group in column 2 (1 vs. 0.4) and a less strong effect in column 3 (0.7 vs. 0.4). Note that the elasticity of substitution between livestock and crops is poorly estimated in column 2; the standard error is larger, and the magnitude is close to zero. However, in column 3, this elasticity is statistically significant at the standard levels and very similar in magnitude to the estimated elasticity in column 1.

The IV-FE models provide more reliable coefficients under the usual assumptions that the instruments must hold for the land-and-labor ratio. From the first stage estimation of the models in Table 4, shown in Table B.3 in the Appendix, the F-statistic values of 369.68, 283.39, and 273.67 for columns 1, 2 and 3, respectively, indicate that forest conversion and rural population are strongly correlated with the logarithm of the land-and-labor ratio. In addition, the overidentifying Hansen test could not reject the null that the instruments can be excluded from the machinery-and-fertilizers equation for columns 1 and 3. This is an indication that these instruments are valid. The IV approach does not seem to be ideal for the deforesters group in column 2, which is an indication that deforestation is a more appropriate instrument for the countries that had both a reduction in the forest area and an increase in agricultural area[29].

Table 4: Baseline IV-FE results for different sample of countries, 1990-2015

Dependent variable: **Ln (Machinery/Fertilizers)**			
	(1)	(2)	(3)
	All countries	Intensive Deforestation (Hosonuma et al., 2012)	Forest area reduction + agricultural area expansion (FAO, 2016)
ln (Land/Labor)	0.400^{**}	1.030^{***}	0.718^{***}
	(0.170)	(0.161)	(0.252)
ln (Livestock%/Crop%)	0.751^{***}	0.0105	0.740^{***}
	(0.188)	(0.306)	(0.215)
Controls	Yes	Yes	Yes
Country-FE	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes
Obs.	260	130	104

Note: First column presents the baseline IV-FE model (Table 3, column 4) for all countries. Second column restricts the sample to the intensive deforestation countries (Brazil, Bolivia, Ecuador, Paraguay, and Venezuela) by the Hosonuma et al. (2012) classification. Third column restricts the sample to the countries that showed reduction in forest area and expansion in the agricultural area (Brazil, Argentina, Paraguay, and Peru) as reported in FAO (2016). Observations were weighted by average ag. GDP for the 1990-2015 period. Robust standard errors in parentheses. *, **, and ***, denote significance at the 10, 5 and 1 percent level.

This exercise shows that the machinery-and-fertilizers ratio has a larger response to an increase in the availability of land relative to labor in the countries with intensive deforestation and the countries that had expansion in agricultural land as well as a reduction in forest land. It also indicates that the induced innovation in these countries arises from a stronger market size effect and that the deforestation is more correlated to the agricultural land supply in these countries, reinforcing the relevance of using it as an instrumental variable within the directed technical change approach.

Forest exports and technological bias

This section compares in Table 5 the baseline IV-FE model (column 1) with models including the value of forest exports as explanatory variable. Forests provide economic benefits to these countries; the inclusion of this variable controls for the opportunity costs of exploiting the forests in other enterprises instead of using it for the expansion of agriculture. Column 2 adds the logarithm of forest exports to the baseline model.

---

[29] Table A.3. shows that the exclusion restriction is not met for the estimation using the subsample of deforesters classified by Hosonuma et al. (2012) and Lablois et al. (2017). In addition, the endogeneity test indicated that the land-and-labor ratio is exogenous, so the IV approach would not be the best.

Table 5: Baseline IV-FE results with forest exports (all countries and deforesters), 1990-2015

Dependent variable: **Ln (Machinery/Fertilizers)**	(1)	(2)	(3)	(4)
	Baseline IV-FE	With Forest Exports	With Forest Exports for Deforesters (Hosonuma, 2012)	With Forest Exports for Deforesters (FAO, 2016)
ln (Land/Labor)	0.400^{**}	0.416^{**}	0.986^{***}	0.855^{***}
	(0.170)	(0.174)	(0.157)	(0.270)
ln (Livestock%/Crop%)	0.751^{***}	0.783^{***}	0.0332	0.712^{***}
	(0.188)	(0.180)	(0.269)	(0.188)
Arable land (%)	-0.106^{***}	-0.0973^{***}	-0.284^{***}	-0.138^{***}
	(0.0300)	(0.0302)	(0.0435)	(0.0408)
Financial credit (%)	-0.000473	-0.000615	0.00324^{**}	-0.00467^{***}
	(0.000958)	(0.000974)	(0.00149)	(0.00111)
Temperature change	-0.0842	-0.0866	-0.0526	-0.0345
	(0.0814)	(0.0833)	(0.126)	(0.0972)
$\ln(\text{Ag. GDP})_{-1}$	-0.709^{***}	-0.667^{***}	0.590	-0.0891
	(0.256)	(0.248)	(0.437)	(0.509)
$\ln(\text{Exports value})_{-1}$	0.0972^{**}	0.0655^{*}	-0.0557	-0.485^{*}
	(0.0391)	(0.0395)	(0.0481)	(0.261)
ln (Forest exports)		0.0816^{**}	0.0762^{***}	0.175^{***}
		(0.0327)	(0.0248)	(0.0547)
Country-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs.	260	260	130	104
R^2 adjusted	0.828	0.833	0.900	0.927
AIC	-112.8	-119.9	-145.7	-109.4
BIC	1.130	-2.413	-51.05	-22.17

Note: Column 1 is the baseline IV-FE model (Table 3, column 4). Column 2 adds to the baseline IV-FE model the value of forest products exports, in logarithm. Column 3 restricts the model in Column 2 to the deforesters group classified by Hosonuma et al. (2012). Column 4 restricts the sample to the deforesters group classified by FAO (2016). Observations were weighted by average ag. GDP for the 1990-2015 period. Robust standard errors in parentheses. *, **, and ***, denote significance at the 10, 5 and 1 percent level.

Columns 3 and 4 restrict the sample to the two groups of deforesters and add the logarithm of forest exports. No substantial changes in the magnitude and significance of the parameters are observed from column 1 to column 2, but a better statistical fit in column 2 is shown by the R^2-adjusted and the information criteria. This indicates that forest exports can be an important determinant of the technological bias, represented by the allocation of machinery per ton of fertilizers. Note that there is a change in sign in the forest exports coefficient compared to the OLS-FE estimate (Table 2, column 4) for the 1961-2015 period. It is now positive and, since the variable is in logarithm, we interpret it as an elasticity. An increase in 1-percent in forest exports would increase machinery per ton of fertilizers in 0.08-percent (column 2). The forest exports coefficient is positive and statistically significant, at the conventional levels, for both groups of deforesters. The magnitude is much larger for the FAO group, indicating that an increase in 1-percent in the value of forest exports would increase machinery per ton of fertilizers in 0.17-percent. Also, we observe a negative and significant coefficient for the value of agricultural exports in this group, which indicates that an increase in 1-percent in agricultural exports in the previous year, would reduce the machinery-and-fertilizers ratio in 0.5-percent. This effect is the opposite of the estimate for the panel of all ten South American countries.

It is hard to determine the potential determinants of technological bias in agriculture and their mechanisms, so one must rely on the data and the estimation to tell if there is a significant effect and whether it should be positive or negative. The same logic is true for the other control variables other than the variables that can be explicitly obtained from the theory. The parameters are then reduced forms, providing only suggestive evidence of the determinants of the technological bias, keeping in mind that, from the theoretical standpoint, they should reflect the changes on the marginal opportunity costs of innovations, and, from the econometric standpoint, the omitted variables that could be causing the endogeneity problem, even after controlling for the country- and year-FE.

Robustness

This section tests the robustness of the IV-FE model as shown in Table 6. First, the other definitions for the deforestation instrument from the FAO data (columns 2 and 3) are used, then the institutional variables (column 4), and finally, a model with institutional variables but without year-FE is estimated in column 5. Column 1 presents the baseline IV-FE results where the area of net forest conversion is used as instrument. Column 2 presents the results of the model using net CO_2 emissions from forestland normalized by the forest area as instrument. Column 3 uses the annual change in forestland as instrument. Columns 4 and 5 add the institutional variables to the baseline IV-FE model; however, in column 5, we exclude the year-FE.

Because we are capturing the differences in forest area in the three first models, we can see that irrespective of the definition of deforestation from the FAO data, and hence the instrument used in the IV-FE model, we obtain robust directed technical change parameters that are statistically significant at the conventional levels. The land-and-labor ratio is only larger in magnitude (0.49) when we use the net CO_2 emissions as instrument. The coefficient of the ratio of the shares of livestock and crops is estimated in the range of 0.69-0.75.

From columns 1-4, the results for the other control variables indicate robust negative and statistically significant coefficient of the share of arable land, with a magnitude around -

0.10, and no significant parameters for the share of financial credit and temperature change. The elasticity of the lagged value of the agricultural GDP is robust and negative, varying from -0.62 to -0.79. While the elasticity of the lagged value of agricultural exports is not statistically significant when we include the institutional variables in column 4, although it continues to be positive.

Table 6: Baseline IV-FE robustness check results, 1990-2015

Dependent variable: **Ln (Machinery/Fertilizers)**					
	(1) Baseline IV-FE	(2) Net CO2 emissions per forest area as instrument	(3) Annual change in forest area as instrument	(4) With institutional variables	(5) With institutional variables and no year-FE
ln (Land/Labor)	0.400**	0.495***	0.401**	0.417*	0.848***
	(0.170)	(0.181)	(0.168)	(0.242)	(0.241)
ln (Livestock%/Crop%)	0.751***	0.731***	0.703***	0.691***	0.659***
	(0.188)	(0.188)	(0.175)	(0.165)	(0.216)
Arable land (%)	-0.106***	-0.118***	-0.0954***	-0.108***	-0.125***
	(0.0300)	(0.0314)	(0.0286)	(0.0311)	(0.0370)
Financial credit (%)	-0.000473	-0.000528	-0.000389	-0.000356	0.00191**
	(0.000958)	(0.000955)	(0.000932)	(0.00104)	(0.000754)
Temperature change	-0.0842	-0.0824	-0.108	-0.0895	-0.0124
	(0.0814)	(0.0817)	(0.0794)	(0.0839)	(0.0751)
$\ln(\text{Ag. GDP})_{-1}$	-0.709***	-0.790***	-0.670***	-0.623*	-1.866***
	(0.256)	(0.258)	(0.213)	(0.327)	(0.177)
$\ln(\text{Exports value})_{-1}$	0.0972**	0.103***	0.0973***	0.0406	0.326***
	(0.0391)	(0.0394)	(0.0376)	(0.0705)	(0.0748)
Civil Liberties				0.0403	0.0314
				(0.0490)	(0.0379)
Political Rights				-0.0779*	-0.130***
				(0.0444)	(0.0487)
Polity 2				-0.0178	-0.0262*
				(0.0134)	(0.0146)
Durable				0.00539	-0.0124***
				(0.00381)	(0.00439)
Country-FE	Yes	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes	No
Obs.	260	260	250	260	260
R^2 adjusted	0.828	0.828	0.822	0.830	0.777
AIC	-112.8	-112.4	-137.2	-111.7	-62.44
BIC	1.130	1.521	-28.02	16.44	-23.27

Note: Column 1 is the baseline IV-FE model (Table 3, column 4). Columns 2-5 are variations of the baseline model in Column 1. Column 2 uses the net CO_2 emissions per 1,000 ha of forest area as instrument in the IV estimation. Column 3 uses the negative of the annual absolute change in forest area as instrument. Column 4 adds the institutional variables as controls to the baseline model. Column 5 maintains the institutional variables but excludes the year-FE. Observations were weighted by average ag. GDP for the 1990-2015 period. Robust standard errors in parentheses. *, **, and ***, denote significance at the 10, 5 and 1 percent level.

Comparing columns 4 and 5, we conclude that the year-FE seem to be capturing the effects of most institutional variables. However, column 4 indicates that the Political Rights measure of freedom is the only institutional variable that has a negative statistically significant coefficient in the IV-FE model with country- and year-FE. Since an increase in this indicator means less political rights, we interpret that less freedom would induce a reduction in the use of machinery per ton of fertilizers in agriculture. Like in the OLS-FE models, the exclusion of the year-FE and the inclusion of institutional variables would provide overestimated elasticity for the factor endowments ratio (column 5), which highlights the importance of the fixed effects to control for the omitted variables in the model and the changes in the opportunity costs of innovations that come from the theory.

1.6 Conclusion

This study provided an application of Acemoglu’s directed technical change model to South American agriculture. First, it was presented a summary of the model with an application related to the production of livestock and crops commodities using traditional and commercial inputs. The traditional inputs considered were land and labor, and the commercial inputs were machinery and fertilizers. The time- and cross-sectional-variation in a panel of ten South American countries were explored using OLS-FE and IV-FE estimation techniques. This is the first study to estimate the parameters of the directed technical change model applied to agriculture with an attempt to correct for the endogeneity bias in the process of the induced innovation. We used deforestation as an instrument for the factor endowments ratio. The results indicated substitutability between land and labor with an elasticity of substitution estimated at 1.7 and 1.4 for the 1961-2015 and the 1990-2015 periods, respectively. The obtained elasticities of substitution between the livestock and crops commodities were 0.96 (1961-2015) and 0.75 (1990-2015), indicating more substitutability than complementarity in the total agricultural output. We provided new estimations of the process of the induced innovation in terms of factor quantities ratios, instead of factor prices, highlighting the potential of the directed technical framework to study innovations and their determinants in agriculture. A strong market size effect seemed to prevail in South American agriculture where more land-complementary commercial inputs were used as more land became available relative to labor. Finally, the direction of causation from deforestation to agricultural productivity was explored, suggesting that, as a determinant of agricultural expansion, deforestation patterns can be linked to the directed technical change framework in South America, although more research using disaggregated data is required to fully understand the effects of deforestation on agricultural land supply.

REFERENCES

ACEMOGLU, D. Directed technical change. **Review of Economic Studies**, 69, 2002, p. 781–809.

ACEMOGLU, D. Why Do New Technologies Complement Skills? Directed technical change and wage inequality. **The Quarterly Journal of Economics**, 113 (4), 1998, p. 1055-1089.

ACEMOGLU, D.; AGHION, P.; BURSZTYN, L.; HEMOUS, D. The Environment and Directed Technical Change. **American Economic Review**, 102, 2012, p. 131–166.

ACEMOGLU, D.; ZILIBOTTI, F. Productivity Differences. **The Quarterly Journal of Economics**, 116 (2), 2001, p. 563-606.

AGHION, P.; DECHEZLEPRÊTRE, A.; HEMOUS, D.; MARTIN, R.; REENEN, J. V. Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry. **Journal of Political Economy**, 124 (1), 2016, p. 1-51.

BINSWANGER, H.; RUTTAN, V. W. et al. **Induced innovation: Technology, Institutions and Development**. The John Hopkins University Press, Baltimore. 1978.

BINSWANGER, H.P. The measurement of technical change biases with many factors of production. **American Economic Review**. 64, 1974, p. 964–976.

BUSTOS, P.; CAPRETTINI, B.; PONTICELLI, J. Agricultural Productivity and Structural Transformation: Evidence from Brazil. **American Economic Review**, 106 (6), 2016, p. 1320–65.

CALEL, R., DECHEZLEPRÊTRE; A. Environmental Policy and Directed Technological Change: Evidence from the European Carbon Market. **Review of Economics and Statistics**, 98(1), 2016, p. 173-191.

COWAN, B. W.; LEE, D.; SHUMWAY, C. R. The induced innovation hypothesis and U.S. public agricultural research. **American Journal of Agricultural Economics**, 97(3), 2015, p. 727–742.

FAO. 2016. **State of the World's Forests 2016**. Forests and agriculture: land-use challenges and opportunities. Rome.

FARIA, W. R.; ALMEIDA, A. N. Relationship between openness to trade and deforestation: Empirical evidence from the Brazilian amazon. **Ecological Economics**, 121, 2016, p. 85–97.

FUGLIE, K. O. Productivity Growth and Technology Capital in the Global Agricultural Economy. In: Fuglie K., Wang, S.L., and Ball, V.E. (eds.) **Productivity Growth in Agriculture: An International Perspective**. CAB International, Wallingford, UK, 2012, p. 335-368.

FUGLIE, KEITH O. Accounting for Growth in Global Agriculture. **Bio-based and Applied Economics**, 4(3), 2015, p. 221-54.

FULGINITI, L. E.; PERRIN, R. Price and Productivity in Agriculture. **Review of Economics and Statistics,** 75 (3), 1993, p. 471-482.

HANLON, W. W. Necessity Is the Mother of Invention: Input Supplies and Directed Technical Change. **Econometrica**, 83 (1), 2015, p. 67–100.

HAYAMI, Y.; RUTTAN, V. W. **Agricultural development: an international perspective**. The John Hopkins University Press, Baltimore, 1971.

HAYAMI, Y.; RUTTAN, V. W. Factor prices and technical change in agricultural development: The United States and Japan, 1880-1960. **Journal of Political Economy**, 78, 1970, p. 1115-1141.

HICKS, J. **The Theory of Wages**. London: Macmillan, 1932.

HOSONUMA, N.; HEROLD, M.; SY, V. D.; FRIES, R. S. D.; BROCKHAUS, M.; VERCHOT, L. et al. An assessment of deforestation and forest degradation drivers in developing countries. **Environmental Research Letters**, 7(4), 044009, 2012.

KAWAGOE, T.; OTSUKA, K; HAYAMI, Y. Induced Bias of Technical Change in Agriculture: The United States and Japan, 1880–1980. **Journal of Political Economy**, 94, 1986, p. 523–44.

KENNEDY, C. Induced Bias in Innovation and the Theory of Distribution. **The Economic Journal**, LXXIV, 1964, p. 541-547.

KLUMP, R.; CABRERA, C. M. Biased Technological Change in Agriculture: The Hayami-Ruttan Hypothesis Revisited. **DEGIT Conference Papers c013_016, DEGIT, Dynamics, Economic Growth, and International Trade**, 2008.

KOCH, N.; ZU ERMGASSEN, E. K. H. J.; WEHKAMP, J.; OLIVEIRA FILHO, F. J. B.; SCHWERHOFF, G. Agricultural Productivity and Forest Conservation: Evidence from the Brazilian Amazon. **American Journal of Agricultural Economics**, 101(3), 2019, p. 919–940.

LE POLAIN DE WAROUX, Y.; GARRETT, R. D.; GRAESSER, J.; NOLTE, C.; WHITE, C.; LAMBIN, E. F. The Restructuring of South American Soy and Beef Production and Trade Under Changing Environmental Regulations. **World Development**, 121, 2019, p. 188–202.

LE POLAIN DE WAROUX, Y.; GARRETT, R. D.; HEILMAYR, R.; LAMBIN, E. F. Land-use policies and corporate investments in agriculture in the Gran Chaco and Chiquitano. **Proceedings of the National Academy of Sciences (PNAS),** 113(15), 2016, p. 4021-4026.

LEBLOIS, A.; DAMETTE, O.; WOLFERSBERGER, J. What has Driven Deforestation in Developing Countries Since the 2000s? Evidence from New Remote-Sensing Data. **World Development**, 92, 2017, p. 82–102.

LIU, Y.; SHUMWAY, C. R. Induced innovation in U.S. agriculture: time-series, direct econometric, and nonparametric tests. **American Journal of Agricultural Economics**, 91(1), 2009, p. 224-236.

MUNDLAK, Y; BUTZER, R; LARSON, D. F. Heterogeneous technology and panel data: The case of the agricultural production function. **Journal of Development Economics**, 99, 2012, p. 139-149.

NOLTE, C.; LE POLAIN DE WAROUX, Y.; MUNGER, J.; REIS, T. N. P.; LAMBIN, E. F. Conditions influencing the adoption of effective anti-deforestation policies in South America’s commodity frontiers. **Global Environmental Change**, 43, 2017, p. 1–14.

OLMSTEAD, A. L.; RHODE, P. W. Induced Innovation in American Agriculture: A Reconsideration. **Journal of Political Economy**, 101(1), 1993, p. 100–118.

POLITY IV. **Political Regime Characteristics and Transitions, 1800-2017**, a dataset, 2017.

POPP, D. Induced Innovation and Energy Prices. **American Economic Review**, 92 (1), 2002, p. 160-180.

SCHMITZ, C.; KREIDENWEIS, U.; LOTZE-CAMPEN, H.; POPP, A.; KRAUSE, M.; DIETRICH, J. P.; MULLER, C. Agricultural trade and tropical deforestation: interactions and related policy options. **Regional Environmental Change**, 15(8), 2015, p. 1757–1772.

SILVA, F.; PERRIN, R. K.; FULGINITI, L. E. The opportunity cost of preserving the Brazilian Amazon forest. **Agricultural Economics**, 50, 2019, p. 219-227.

THE FREEDOM HOUSE. **Freedom in the World 2018: The annual survey of political rights and civil liberties**. Rowman & Littlefield, 2018.

THIRTLE, C.; RUTTAN, V. The Role of Demand and Supply in the Generation and Diffusion of Technical Change. **Fundamentals of Pure and Applied Economics**, vol. 21, London: Harwood Academic Publishers, 1987.

THIRTLE, C.; SCHIMMELPFENNIG, D.; TOWSEND, R. F. Induced innovation in United States agriculture, 1880–1990: time series tests and an error correction model. **American Journal of Agricultural Economics**, 84 (3), 2002, p. 598-614.

TRINDADE, F. J.; FULGINITI, L. E. Is there a slowdown in agricultural productivity growth in South America? **Agricultural Economics**, 46, 2015, p. 69–81.

APPENDIX

A. Theory appendix

A.1. Two-factor test of induced innovation in terms of factor prices.

The traditional tests of induced innovation in agriculture, starting from the classic study of Hayami and Ruttan (1970), involved decomposing the changes in the land-labor (Z/L) ratio over time or across different countries into the *price substitution effect* and the *biases of technical change*. For the two-factor case, differentiating Z/L with respect to factor prices yields:

$$d\left(\frac{Z}{L}\right) = \left(\frac{1}{Z}.\frac{\partial Z}{\partial P_z} - \frac{Z}{L^2}\frac{\partial L}{\partial P_z}\right) dP_z + \left(\frac{1}{Z}.\frac{\partial Z}{\partial P_L} - \frac{Z}{L^2}\frac{\partial L}{\partial P_L}\right) dP_L \qquad (A.1)$$

where Z=land and L=labor, and P_z and P_L are the price of land and labor, respectively. We can express (A.1) in terms of the derived demand elasticities (η):

$$dln\left(\frac{Z}{L}\right) = (\eta_{ZZ} - \eta_{LZ})dlnP_z - (\eta_{LL} - \eta_{ZL})dlnP_L \qquad (A.2)$$

The change measured in (A.2) involves only the change in the land-labor ratio along the same isoquant, and therefore, is only the price substitution effect. The total change must include the effect of technical change. Let $dln\left(\frac{Z}{L}\right)^*$ be the total land-labor percentage change, then:

$$dln\left(\frac{Z}{L}\right)^* = dln\left(\frac{Z}{L}\right) + B \qquad (A.3)$$

where B represents the bias of technical change. If $B > 0$, then technical change is land-using (labor-saving), if $B < 0$, technical change is land-saving (labor-using). The bias of technical change can be measured as:

$$B = dln\left(\frac{Z}{L}\right)^{*} - \; dln\left(\frac{Z}{L}\right) \qquad (A.4)$$

Dividing both sides of (A.4) with respect to the change in logarithm of the factor prices ratio $dln\left(\frac{P_L}{P_Z}\right)$ yields:

$$\frac{B}{dln\left(\frac{P_L}{P_Z}\right)} = \sigma_{ZL}{}^{*} - \sigma_{ZL} \qquad (A.5)$$

where $\sigma_{ZL}{}^{*}$ and σ_{ZL} are the elasticities of substitution between land and labor given by $\frac{dln\left(\frac{Z}{L}\right)}{dln\left(\frac{P_L}{P_Z}\right)}$ for output Y held constant. Mundlak (1968)[30] expresses the pairwise elasticity of substitution (σ_{ZL}) in terms of the elasticities of factor demand:

$$\sigma_{ZL} = \frac{dln\left(\frac{Z}{L}\right)}{dln\left(\frac{P_Z}{P_L}\right)} = \frac{(\eta_{ZZ} - \eta_{LZ})dlnP_z - (\eta_{LL} - \eta_{ZL})dlnP_L}{dlnP_L - dlnP_z} \qquad (A.6)$$

Expression (A.5) provides a way to test the induced innovation due to changes in relative prices based on the elasticity of substitution and, given $(A.6)$, this test is equivalent to test the induced innovation based on the factor demand elasticities. If $\sigma_{ZL}{}^{*} = \sigma_{ZL}$, then $B = 0$ and technical change is neutral. Figure A.1, extracted from Hayami and Ruttan (1971, p. 126), shows how the induced innovation process happens in agriculture based on factor prices.

---

[30] Mundlak, Y. (1968). “Elasticities of substitution and the theory of derived demand”, Review of Economic Studies, 35, p: 225-236.

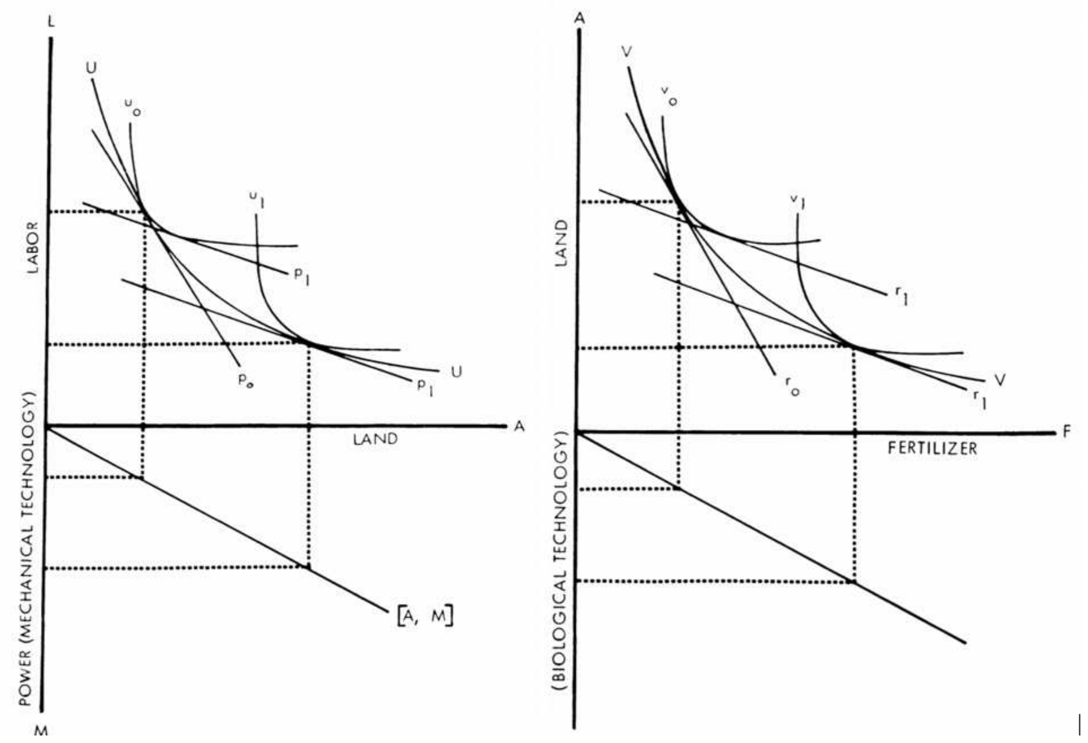

Figure A.1. Factor prices and induced technical change. Extracted from Hayami and Ruttan (1971, p. 126).

The change in relative prices from P_0 to P_1 causes a movement on the optimal input levels from isoquant u_0 to u_1 on the left. Similarly, a change in relative prices from r_0 to r_1 shifts the optimal levels from isoquant v_0 to v_1 on the right. This effect shows a labor-saving (land-using) technical change on the left and a land-saving (fertilizer-using) technical change on the right.

Notice that we have an isoquant UU on the left that is tangent to u_0 at relative prices P_0 and tangent to u_1 at relative prices P_1. The same for isoquant VV on the right. These isoquants would provide the necessary elasticity of substitution to explain factor changes only by price changes given by ${\sigma_{ZL}}^*$ in $(A.5)$. Obtaining an estimate of this critical elasticity of substitution would allow to determine whether the technology is equal between two periods/two countries (i.e., technical change is neutral) or not and how prices affect the direction of technical change.

Figure A.1 also shows how agricultural mechanical innovations and biological innovations are related to the inputs. Mechanical innovations are shown to be land-complementary and labor-saving while biological innovations are fertilizer-complementary and land-saving. Previous studies measured the induced innovation in agriculture through commercial inputs based on the extent that they incorporate the innovations in this sector. Given Hayami and Ruttan's (1970) complementarity and substitutability patterns between commercial and traditional inputs and how they are related to innovations, we use fertilizer and machinery as proxies for the technology inputs described in Acemoglu's model in our empirical test.

A.2. *Biases of technical change.*

There are three main definitions of bias of technological change. Although they all measure the change in input intensity due to technological change, they differ in magnitude. With factor prices held constant, the bias of technical change can be measured based on:

1) Factor shares, with many factors:

$$\mathrm{B_i} = \frac{\partial S_i}{\partial t}\frac{1}{S_i} \lesseqgtr 0: \begin{cases} i - saving \\ i - neutral \\ i - using \end{cases} \qquad (A.7)$$

This definition is from Binswanger (1974) and is useful for a production function with many inputs. It measures the changes in cost share (S_i) due to technical change.

2) Factor ratios, for two factors:

$$\mathrm{B}_{LK} = \frac{\partial\left(\frac{K}{L}\right)}{\partial t}\frac{1}{\frac{K}{L}} \gtreqless 0: \begin{cases} L - saving \\ neutral \\ L - using \end{cases} \qquad (A.8)$$

This measures the changes in the capital-labor ratio if there is a change in the technology set holding output constant. In the third definition, the factor ratio is held constant, and is based on:

3) Marginal products (f_K and f_L), for two factors:

$$\mathrm{B}_{LK} = \frac{\partial\left(\frac{f_K}{f_L}\right)}{\partial t}\frac{1}{\frac{f_K}{f_L}} \gtreqless 0: \begin{cases} L - saving \\ neutral \\ L - using \end{cases} \qquad (A.9)$$

This is the classic Hicksian pairwise bias of technical change and indicates that technical change is biased toward the input whose marginal product is increasing more relative to the other.

A.3. *Note on factor-augmenting technical change and bias of technical change.*

A more general CES production function, $y = F(L, Z, A)$ where L and Z are the inputs and A is the technology index, can illustrate the relationship between factor augmentation and the biases of technical change. Acemoglu (2002) gives the following CES production function which, by construction, has two factor-augmentation technology terms, A_L and A_Z:

$$Y = \left[q(A_L L)^{\frac{\sigma-1}{\sigma}} + (1-q)(A_Z Z)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \qquad (A.10)$$

where $\sigma \in (0, \infty)$ is the elasticity of substitution between L and Z, $q \in (0,1)$ and A_L and A_Z can represent, as defined by Acemoglu (2002), the labor-complementary and land-complementary technologies, respectively. Factor-augmentation means that as more technologies are developed, more “effective” use of the factor is observed in the production process. The bias of technical change is given by the change in the marginal products ratio given the development of new technologies. For the production function in (A.10), this ratio is:

$$\frac{f_Z}{f_L} = \frac{1-q}{q}\left(\frac{A_Z}{A_L}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{Z}{L}\right)^{\frac{-1}{\sigma}} \qquad (A.11)$$

Equation (A.11) shows that if we observe a Z-augmenting technical change (increase in A_Z), technical change will also be Z-biased if $\sigma > 1$, i.e., the marginal product ratio increases. However, technical change will be L-biased if $\sigma < 1$. This result shows that the way that factor-augmenting technical change affects the marginal products, and hence the bias of technical change, depends on the elasticity of substitution between the inputs for the general CES production function.

B. Data Appendix

B.1. Variable descriptions and sources.

Table B.1: Variable descriptions and sources

Variable	Definition	Source
Agricultural labor	Economically active adults 15+ in the ag. sector (1,000 people)	FAO
Temp. change	Average annual change in temperature compared to the 1951-1980 baseline climatology (Celsius degree)	
Livestock share	Share of the value of livestock production on the Agricultural GDP	
Crops share	Share of the value of crops production on the Agricultural GDP	
Forest exports	Value of all exported forest products (1,000 US dollars)	
Agricultural GDP	Value of production (189 crop and livestock commodities) (1,000 US$ (2004-2006 average international prices))	
Ag. Exports	Value of exported crop and livestock commodities (1,000 US dollars)	
Net Forest Conversion	(Two-year change) Forest area converted to other land uses (1,000 ha)	
Agricultural land	Quality-adjusted agricultural land (1,000 ha) measured in rainfed cropland equivalents. Weights: 1.00 for cropland, 0.0298 for permanent pastureland and 1.0094 for irrigated cropland	Fuglie (2015)
Machinery	Number of 40 cv tractor equivalent	
Fertilizer	Metric tons of N, P2O5, K2O consumption	
Rural population	Number of people in the rural population	WDI
Arable land	Percentage of arable land under temporary crops (double-cropped area counted once), temporary meadows, land under market and kitchen gardens, and fallow land in total land area	
Financial credit	Percentage of domestic credit by the financial sector in the GDP	
Polity 2	Annual score that ranges from -10 (strongly autocratic regimes) to 10 (strongly democratic regimes)	Polity IV project (2017)
Durable	Number of years since the last political regime change	
Political rights	Measured on a one-to-seven scale, 1 indicating the highest degree of freedom and 7 the worst.	The Freedom House (2018)
Civil liberties	Measured on a one-to-seven scale, 1 indicating the highest degree of freedom and 7 the worst.	

B.2. Descriptive statistics.

Table B.2.: Descriptive statistics.

Variable	Units	Obs	Mean	Std. Dev.	Min	Max
Agricultural Labor	1,000 (people)	550	2614.0	3924.4	182.0	16345.0
Machinery	Units	550	103280.2	203985.9	1466.0	1095925.0
Fertilizer	Metric tons	550	689755.0	1885574.0	600.0	14000000
Agricultural Land	1,000 (ha)	550	12401.0	19648.8	1097.2	92476.4
Livestock share	[0,1]	550	.4385978	.1348234	.1671593	.8364666
Crops share	[0,1]	550	.8538722	.073989	.6285031	.9732273
Ag. GDP	1 million USD	550	13100.0	24600.0	608	161000.0
Arable land	% in land area	550	5.2	3.1	1.2	14.5
Financial Credit	% in GDP	540	37.6	25.3	5.7	212.9
Temperature change	Celsius	550	0.3	0.4	-0.8	1.4
Rural Population	1,000 (people)	550	7380.986	10474.08	170.019	42210.9
Net Forest conversion	1,000 (ha)	260	386.9678	665.8519	0.0	2996.8
Civil Liberties	1 to 7	430	3.2	1.2	1.0	6.0
Political Rights	1 to 7	430	3.0	1.6	1.0	7.0
Polity 2	-10 to 10	550	3.9	6.4	-9.0	10.0
Durable	Years	550	12.9	11.8	0.0	58.0
Agricultural exports	1,000 (USD)	550	4044620.0	10300000	2910.0	83900000
Forest exports	1,000 (USD)	526	465316.8	1278037.0	0.0	8723114.0

Note: ha denotes hectares and USD denotes U.S. dollars.

B.3. First-stage estimation results of the baseline IV-FE models in Table 4 and tests statistics.

Table B.3: First-stage estimation results for all countries and the deforesters groups, 1990-2015

Dependent variable: ln (Land/Labor)

	(1) All countries	(2) Intensive Deforestation (Hosonuma et al., 2012)	(3) Forest area reduction + agricultural area expansion (FAO, 2016)
Net Forest Conversion	-0.000083***	-0.0000863***	-0.000074***
	(0.0000096)	(0.0000147)	(0.000013)
Rural Population	-0.000043***	-0.0000514***	-0.000043***
	(0.0000030)	(0.0000037)	(0.000005)
Obs.	260.00	130.00	104.00
F-statistic	F (2, 217)	F (2, 92)	F (2, 67)
	369.68	283.39	273.67
Hansen J statistic (p-value)	1.421 (0.233)	11.025 (0.0009)	0.439 (0.5078)
Endogeneity test (p-value)	0.402 (0.526)	2.824 (0.0929)	0.472 (0.492)

Note: The rule-of-thumb to reject that the instruments are weak is an F-statistic above 10. The Hansen J statistic shows the result from the exclusion restriction test based on a Chi-square distribution with 1 degree of freedom ($\chi^2(1)$). The null hypothesis is that the instruments are valid (i.e., not correlated with the error term of the main equation), so they can be excluded from the second stage. The reported endogeneity test statistics also follows a $\chi^2(1)$. The null hypothesis is that the endogenous variable is exogenous. Failure to reject the null implies that an IV approach would be preferred. Robust standard errors in parentheses. *, **, and ***, denote significance at the 10, 5 and 1 percent level.

CHAPTER 2

EX-ANTE EXPECTED PAYOFF FROM A VARIABLE NITROGEN RATE APPLICATION: AN EXPECTED VALUE OF SAMPLE INFORMATION (EVSI) APPROACH

2.1 Introduction

This paper studies the use of soil sample information as a decision-making tool for the adoption of a precision agriculture (PA) technology[31,32]. The benefits of PA are given by the increase in the value of the agricultural inputs' applications (e.g., fertilizers, pesticides, water, etc.) over costs (profits) because it can help decrease the uncertainty about soil conditions affecting the performance of inputs in agriculture[33]. Variable rate technology (VRT) for agricultural inputs, characterized as one of the PA technologies, refers to a technology that adjusts the application rate for cells within a field, based on information that is unique for each cell. The technology requires an observable signal at each cell that conditions input response in that cell, and a combination of software and hardware capable of changing the application rate across subunits.

Computer-controlled VRT technologies have been commercially available in the U.S. since the late 1980s. Economic analyses of VRT beginning in the early 1990s have shown that VRT for fertilizer on grain crops is seldom profitable. Despite these adverse profitability findings, by 2013, VR technologies were used on about a third of U.S. corn-soybean cropland (Schimmelpfennig, 2016), and reported studies suggested that by 2017, across states in the U.S. from 43% to 73% of farmers had adopted VRT for fertilizer application (Lowenberg-DeBoer and Erickson, 2019). Adoption rates this high challenge the results of economic studies showing the practice to have little if any economic benefit. This suggests that further economic analysis is warranted.

---

[31] Soil sample information, soil test information and soil signal are used interchangeably throughout the paper.

[32] Precision agriculture is defined as toolkit of several technologies from which farmers can choose to adopt to improve input application efficiency. For example, it can involve yield monitors, variable rate (VR) applicators, remote and local sensors, and GPS-guided automated farm operations.

[33] Additionally, some environmental benefits can be attributed to the adoption of precision agriculture when inputs' applications are site-specific, tailored according to characteristics of the field, requiring less need to overapply inputs, and avoiding soil and water pollution, to ensure a given level of productivity.

VRT has been used for seed, fertilizer, and irrigation, and from the studies on the assessment of this technology, such as in Perrin (1976), Swinton and Lowenberg-DeBoer (1998), Bullock and Bullock (2000), there seems to be little guidance to identify ex-ante whether VRT might be profitable on a specific field. The main limitation for the evaluation of VRT is the availability of experimental data. Farmer experimentation faces difficulties: the costs are high for using the technology on a single field, and the year-to-year variability in outcomes would require several years of experimentation to reach a reliable conclusion whether it pays off or not. The present study offers an analysis for a new dataset from the Data-Intensive Farm Management (DIFM) Project (DIFM, 2018; Bullock, et al., 2019) at the University of Illinois, which has been working since 2016 with commercial farmers to run agronomic field trials on entire fields and several years. DIFM’s methods employ precision technology, including variable rate input applicators and equipment-mounted yield monitors to gather large amounts of data, at very low costs.

The modelling approach of this paper employs the concept of the Expected Value of Sample Information (EVSI) from information theory to examine the expected payoff of VRT. We assume that the response to the input is related to an unobservable soil characteristic distributed with some prior density across the field. We further assume that there is a characteristic that we can observe, correlated with the unobservable characteristic, which we refer to as a signal. Observation of this signal at a given point changes expectations about the unobservable characteristic at that point, and thus changes the optimal application rate. We apply this approach to examine the expected payoff to nitrogen (N) fertilizer application, using soil electroconductivity (EC) as the signal, which we compare to the expected payoff to a uniform rate technology (URT). The DIFM experimental data in this paper are from randomized applications in Illinois, Ohio, and Nebraska farms in 2016 and 2017. Figure 1 shows an example of a DIFM random trial of variable nitrogen application in a 32-ha farm in Illinois in 2018. We can observe from the figure the different application rates across the cells (grids) of the field. Precision application allows to vary the application rates and at the same time to measure the level of soil EC in each plot.

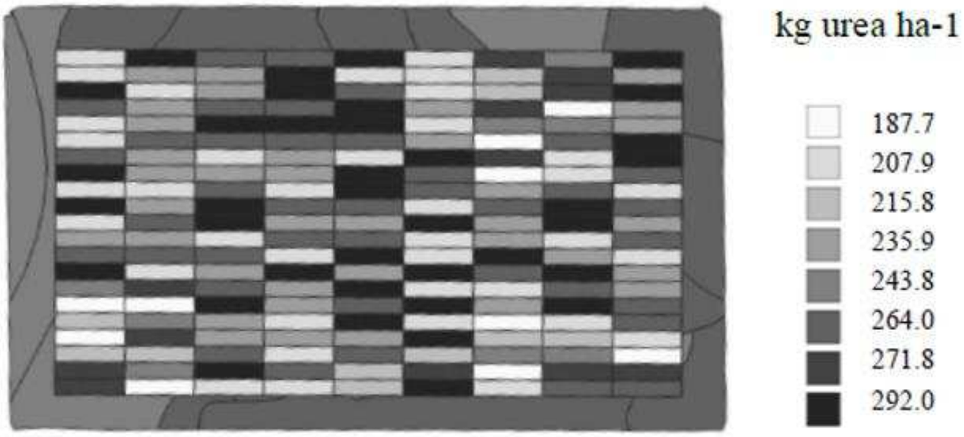

Figure 1. A 2018 on-farm N fertilizer trial conducted on a 32-ha field in Central Illinois. Source: DIFM data.

Related literature

It is intuitively clear that the economic benefits of VRT for nitrogen application depend on many circumstances: the variability of the soil across the field, the yield response to nitrogen application and how it is affected by the soil characteristics, economic variables, such as the prices of the crop and fertilizer, and the cost of obtaining measures of the soil information and implementing the technology. Several economic studies have examined the benefits of adopting VRT nitrogen application (Babcock, Carriquiry and Stern, 1996; Bullock and Bullock, 2000; Hurley, Malzer and Kilian, 2004; Hurley, Oishi and Malzer, 2005; Bullock et al., 2009). Very few though provided an ex-ante analysis and/or studied the role of the soil signal distribution as determinant of the expected returns.

An ex-ante optimal N rate, for a risk-neutral producer, is the rate that maximizes expected profits. Therefore, it is a decision that is made before observing the draw of a given random variable that affects yields (Bullock et al., 2009). Most studies insert the randomness in the producer's decision-problem through weather variables and/or soil nitrogen levels (Babcock, 1992, Ruffo et al., 2006, Liu et al., 2006, Bullock et al., 2009). We follow this definition of ex-ante, but instead of the uncertainties regarding the "weather", we define two correlated random variables that can affect the optimal rates: an unobserved specific soil characteristic and an observed soil sample information (signal). Therefore, we estimate the ex-ante expected value of VRT application with respect to the distributions of the "true" soil characteristic and the distribution of the measurement of soil electroconductivity (EC). Ex-post evaluations of VRT on specific fields have shown widely-varying results[34,35] (Hurley, Malzer and Kilian, 2004; Anselin, Bongiovanni and Lowenberg-DeBoer, 2004; Hurley, Oishi and Malzer, 2005; Lambert, Lowenberg-DeBoer and Malzer, 2006); while ex-ante evaluations such as those in Babcock, Carriquiry and Stern (1996), Liu et al. (2006), Ruffo et al. (2006), and Bullock et al. (2009) provided different ways of performing an ex-ante analysis depending on how the researchers are interpreting the randomness in the producer's decisions. In addition, different methods have been used to calculate the ex-ante returns to VRT. For example, Liu et al. (2006) defined ex-ante optimal rate based on "weather variables"[36]. They used a random coefficient model to estimate the ex-ante optimal nitrogen fertilizer application by assuming that the parameters of the yield function were multivariate normally distributed; then they performed a Monte Carlo experiment that generated 3,000 yield functions, based on the distribution of the parameters, and bootstrapped 80% confidence intervals for the ex-ante expected VRT payoff. Bullock et al. (2009) used weather variables distributed over time, so the ex-ante expected VRT payoff is the average of all the payoffs of the applications over 54 years of weather data. These studies did not compute the ex-ante expected VRT payoff based on the distributions of soil characteristics and/or soil signals.

Instead of considering the soil information as given, which would yield an ex-post analysis with respect to the soil information, Babcock, Carriquiry and Stern (hereafter, BCS) (1996) used experimental data to calculate the expected N rates using the posterior distribution of soil nitrate levels (i.e., their unknown soil characteristic) obtained from the Bayes' rule after the soil test information was made available. The unknown soil characteristic is then the actual soil nitrate levels, from which farmers have a prior belief, and the known signal is the soil test measurement. The ex-ante expected profits is obtained by integrating expected profits over the distribution of the signal itself. Their procedure is based on calculating three functions estimated from the data: the crop production function, the prior density function of nitrate levels (producer's prior belief) and the sampling distribution of soil tests[37]. We differ from BCS

---

[34] Bullock et al. (2009) defined ex-post optimal rates as those used in years which the weather is the same as in the year of the experiment.

[35] Most studies found little payoff from this type of PA technology. For a detailed literature review on the values of VR technology and information, see Bullock and Lowenberg-DeBoer (2007).

[36] Although they called them "weather variables", they did not actually use weather data in the analysis. Instead, they used a random coefficient model to simulate parameters normally distributed to represent that the yield function parameters are functions of a set of random variables.

[37] The sampling distribution of soil tests is represented below in our theoretical model as the following density function $v(s|\gamma)$, which is the probability of obtaining a signal s given the true soil characteristic represented by γ. The sampling distribution of soil tests are used to calculate the marginal density function of the signal, $\phi(s)$, which is then used to calculate the ex-ante expected payoffs.

(1996) in how we approach the yield function (we use a quadratic while they used a plateau yield function) and how we perform the ex-ante analysis. Following Kihlstrom (1976) and Lawrence (1999), we develop an empirical structural model that provided a simple way to insert the signal into the yield function without having to calculate from the data the prior and the posterior distributions of the soil characteristic, and the sampling distribution of the signal. The advantage of this model is that it does not require to compute analytically (or numerically) the posterior distribution of the soil characteristic and the sampling distribution of the signal which could be quite complicated depending on the assumed distributions.

Small vs. large-plot agronomic trials

The general approach of the studies on VRT adoption is to use input and output data from randomized agronomic field trials to estimate yield response functions, and to use those estimates to examine how much value can be derived from site-specific input management. However, most studies used small-plot experiments. Scientists have been conducting small-plot agronomic trials for 175 years (Odell, et al., 1984), and data from small-plot trials has always been the principal source of data from which yield response functions have been estimated. We benefit from the large number of trials on whole fields in different years available from the DIFM dataset, providing a more precise estimation of the yield response to nitrogen application and capturing the spatial and temporal heterogeneities in the estimate of the ex-ante returns.

As is widely recognized, small-plot trials have small inference spaces; empirical results gained from experiments on very small plots of land in one geographic location may provide very little information about the same yield response relationships in the rest of the farm field, let alone in distant fields. BCS (1996), for example, used data from a small-plot (nitrogen fertilizer (N), corn) trial in Iowa, with thirty experimental plots, conducted over six years. Ten N-rates were applied each year, with three replications of each, and then the mean yield of each rate was obtained, providing sixty observations. Hurley, Malzer, and Killian (2004) and Hurley, Oishi, and Malzer (2005) used data from two 1995 small-plot trials in Minnesota. Each experiment was conducted on a 164 m x 274 m plot of land (1.21 hectares), each divided into six “regions,” with each region containing three replications of six N rates. Therefore, inference has been based on few N rate replications, small-plots, and few observations. The main advantage of the data used in this paper is that it comes from randomized trials in entire fields. Figure 2 illustrates a corn seed rate trial run on a 30-ha cornfield in central Illinois in 2017 from the DIFM project. A computer program was written to design the trial, in the form of an ESRI shapefile, which "instructed" a variable rate planter to randomize among four seed rates as the farmer drove the planter through the field, in the usual manner. At harvest, a yield monitor was used to record yield data, with each record being assigned a (longitude, latitude) coordinate.

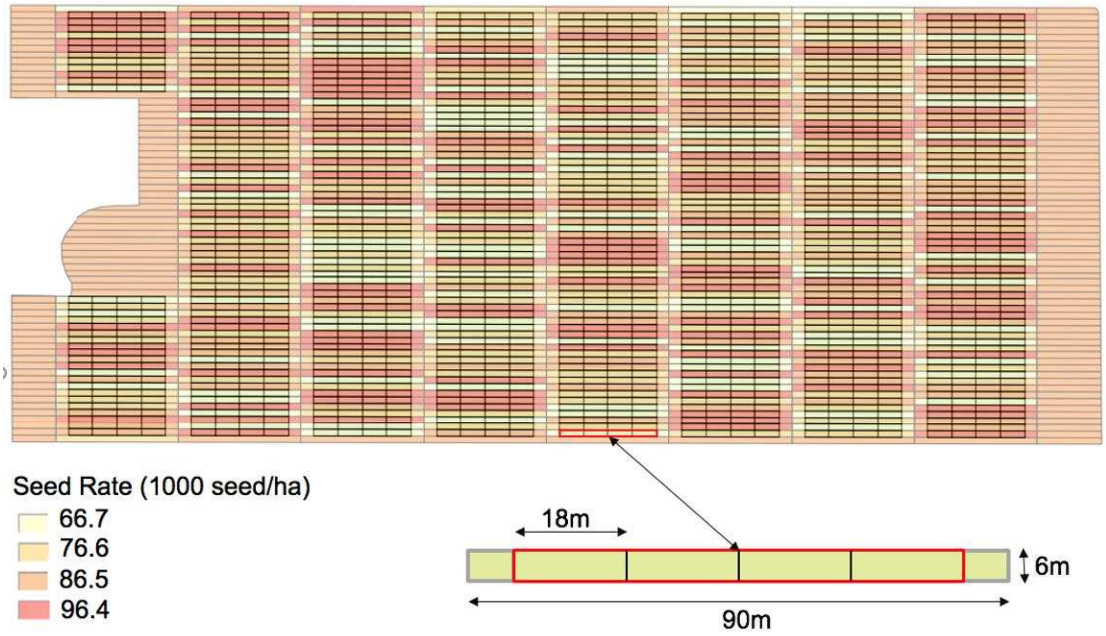

Figure 2. Seed rate trial design conducted on a 30-ha field in Central Illinois, 2017, and magnification of a single plot to show dimensions. Source: DIFM data.

Figure 3 shows the "as-applied" nitrogen rates and yield rates, which were the median observations for each 18m x 6 m "subplot" in the field, of farms from the 2017 trial. We use a subsample of the DIFM data on nitrogen applications, their respective corn yields, and the soil electroconductivity measure for each plot, calculated from entire U.S. farms in 2016 and 2017.

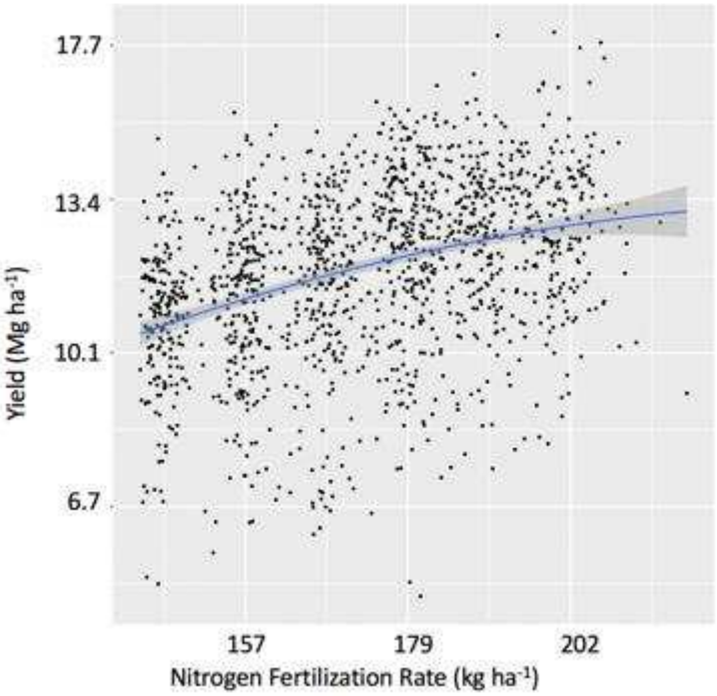

Figure 3. Scatter plot of (N, Corn Yield) data from 2017 trial. Source: DIFM.

This paper contributes to the literature by using a unique dataset on large agronomic random trials on nitrogen fertilizer application and by presenting a theoretical structural model to incorporate the signal into the producer's decision problem to calculate the ex-ante returns of VRT based on the EVSI approach. The paper is organized as follows. Next section presents the theoretical model and application. The third section presents the empirical strategy. The fourth section presents the results. The fifth section concludes.

2.2 Theoretical model and application

A simple Bayesian decision making framework

The general framework is based on Kihlstrom (1976) where farm profits are given as a function of a vector of known inputs, represented by $x = (x_1, \dots, x_n)$, and an unknown parameter represented by $\gamma \in G$. The set G represents all the possible values of the unknown parameter which in the model reflects one specific soil characteristic. Farm profit is described as:

$$\pi(x, \gamma) = p.f(x, \gamma) - w.x \tag{1}$$

where $f(x, \gamma)$ is the production or yield function, p is the price of the farm output, the vector of input prices is given in w. A profit-maximizing farmer chooses x to maximize expected profit as:

$$\max_x E_\gamma[\pi] = \int_G \pi(x,\gamma)\; g(\gamma)d\gamma = E_\gamma[p.f(x,\gamma) - w.x] \tag{2}$$

where $g(\gamma)$ is the density function of the prior probability distribution of γ. We denote as x' the level of input that maximizes expected profits in equation (2). If the signal is not observed at any point, x' is optimal for the entire field and is thus the optimal uniform rate under URT.

Following Kihlstrom (1976), we introduce the possibility of the farmer to obtain some soil sample information (signal) $s \in S$, that is correlated with the true unknown soil characteristic γ. Having observed s, the expected profit maximization problem becomes:

$$\max_x E_{\gamma|s}[\pi] == \int_G \pi(x,\gamma)\; h(\gamma|s)d\gamma = E_{\gamma|s}[p.f(x,\gamma) - w.x] \tag{3}$$

where $h(\gamma|s)$ is the density function of the posterior probability distribution of γ, given s, obtained by Bayes's rule[38]. We denote $x''(\boldsymbol{s})$ as the application rate that maximizes expected profits when s is observed. It can be interpreted as a contingency plan describing the input decision in response to any message s that might be observed.

We assume that prior to adopting VRT on a given field, the prior distribution of γ and the response function are known. When a decision maker decides whether to obtain the signal, she does not know what message the signal will provide. The decision is thus based on an expected profit maximization in which the profit associated with each possible message provided by the signal is weighted by the probability of receiving that message. EVSI is the extra expected profit from observing s, with expectations taken with respect to the density functions of both γ and s:

$$\text{EVSI} = \int_S \left(\int_G \pi(x''(\boldsymbol{s}),\gamma)\; h\big(\gamma|s = s_{ij}\big)d\gamma - \int_G \pi(x',\gamma)\; g(\gamma)d\gamma \right) \phi(s)ds \tag{4}$$

where $\phi(s)$ is the marginal probability density function of the signal. The first expression on the right-hand side identifies the expected payoff from first observing the signal s and then applying the rate that maximizes expected profit given that signal, *evaluated prior to observing the signal*[39]. The expectation is taken with respect to the density of s across the field, $\phi(s)$, and

---

[38] Bayes's rule shows that the posterior density function can be calculated as:

$$h(\gamma|s) = \frac{v(s|\gamma) \cdot g(\gamma)}{\phi(s)} \text{ where } \phi(s) = \int_G v(s|\gamma)\; g(\gamma)d\gamma$$

where $v(s|\gamma)$ represents the sampling distribution of the signal and $\phi(s)$ is the marginal density function of the signal.

[39] This is known as a preposterior analysis, see for example Raiffa and Schlaifer (1961).

since the profit is scaled to the level of one acre, it is the expected profit per acre using VRT across the field. Similarly, the second expression is the comparable expected profit per acre if the optimal uniform rate is applied across the field. Thus equation (4) is the expected extra profit per acre from observing and using the signal compared to a uniform rate. Note that this is the *gross value* of observing the signal, from which cost of adopting the VRT package must be deducted to determine the net benefit of VRT relative to URT.

A specification with a quadratic response function and bivariate Normal distributions [40]

The result presented in equation (4) requires an analytical solution to EVSI. In the same Bayesian decision-making setting, Lawrence (1999) proposed a simple quadratic yield function providing a neat solution where the probability distribution parameters that determine EVSI can be examined. The model can be applied to different decision problems, but we are mainly interested in a producer that chooses nitrogen application (x) to maximize profits under uncertainty. Let the quadratic yield function $f(x, \gamma)$ have the following general specification[41]:

$$y = \alpha + \beta_1 x + \beta_2 x^2 + \beta_3 \gamma + \beta_4 \gamma . x \qquad (5)$$

where $\beta_2 < 0$. Given the yield function in (5), the solutions from the maximization problems in (2) and (3) are, respectively, $x' = x(w, p, \mu_\gamma) = \frac{r - \beta_1 - \beta_4 \mu_\gamma}{2\beta_2}$ and $x''(s) = x(w, p, \mu_{\gamma|s}) = \frac{r - \beta_1 - \beta_4\, \mu_{\gamma|s}}{2\beta_2}$ where $r = \frac{w}{p}$ and μ_γ and $\mu_{\gamma|s}$ are the prior and the posterior mean of γ, respectively. Plugging back these two optimal levels into the profit generates two maximum profit functions. We define the maximum profit obtained based on the prior distribution of γ as $\mathrm{V}(\mu_\gamma)$, and hence a function of the prior mean (μ_γ). The other maximum profit function is $\mathrm{V}(\mu_{\gamma|s})$, where the optimal decision is based on the posterior distribution after observing $s \in S$, and hence a function of the posterior mean $(\mu_{\gamma|s})$. The theoretical EVSI in (4), now in terms of expectations, is:

$$\mathrm{EVSI} = \mathrm{E}_s\left[\mathrm{V}(\mu_{\gamma|s})\right] - \mathrm{V}(\mu_\gamma) \qquad (6)$$

---

[40] The quadratic yield function was commonly used in previous studies (Anselin, Bongiovanni and Lowenberg-DeBoer, 2004; Liu et al., 2006). Liu et al. (2006) performed nested F-tests comparing the quadratic with the linear response and the plateau and could not reject the quadratic in 5 of the 8 fields they tested. We use the quadratic because of its tractability and because it is commonly accepted in the literature. It also allows the application proposed in Lawrence (1999) to determine analytically the ex-ante expected value of information.

[41] We omit subscripts for cells and fields for simplification.

where $\mathrm{E_s}[.]$ indicates the expectation over the distribution of the signal. For the quadratic yield function, these maximum profit functions become:

$$\mathrm{V}(\mu_\gamma) = c_1\mu_\gamma^2 + c_2\mu_\gamma + c_3$$

$$\mathrm{V}(\mu_{\gamma|s}) = c_1\mu_{\gamma|s}^2 + c_2\mu_{\gamma|s} + c_3 \tag{7}$$

where $c_1 = -\frac{p\beta_4^2}{4\beta_2}$, $c_2 = p\beta_3 + \frac{p\beta_4}{2\beta_2}\left(\frac{w}{p} - \beta_1\right)$, and $c_3 = p\alpha + \frac{1}{2\beta_2}\left(w\beta_1 - \frac{w^2}{p} - \frac{p\beta_1^2}{2}\right)$ are combinations of the output and input prices and the parameters of the quadratic yield function. Plugging (7) into (6) yields:

$$\mathrm{EVSI} = c_1\left[\mathrm{E_s}\left(\mu_{\gamma|s}^2\right) - \mu_\gamma^2\right] \tag{8}$$

Lawrence (1999, pp. 118-119, equation 5.7), shows that using the law of iterated expectations, equation (8) can also be presented in terms of the variances as:

$$\mathrm{EVSI} = c_1\left[\sigma_\gamma^2 - \mathrm{E_s}\left(\sigma_{\gamma|s}^2\right)\right] \tag{9}$$

where σ_γ^2 and $\sigma_{\gamma|s}^2$ are the prior and the posterior variances of γ, respectively. The expected value of obtaining information s about the unknown γ is proportional to the reduction in uncertainty about γ (ignoring the cost of the VRT technology package).

When γ and s are bivariate normally distributed with correlation ρ (i.e., $(\gamma, s) \sim Bivariate\ Normal\ (\mu_\gamma,\ \sigma_\gamma^2;\ \mu_s,\ \sigma_s^2; \rho)$), the posterior variance is $\sigma_{\gamma|s}^2 = \sigma_\gamma^2(1-\rho^2)$, Lawrence then shows (his equation 5.8) that equation (9) above can in this case be expressed as:

$$\mathrm{EVSI} = c_1\left[\rho^2\sigma_\gamma^2\right] = -\frac{p\beta_4^2}{4\beta_2}\left[\rho^2\sigma_\gamma^2\right] \tag{10}$$

Equation (10) is a fundamental contribution of this analysis. It expresses the value of VRT as an explicit function of parameters representing the underlying factors we intuitively identified as affecting the value of VRT: the variance of the state of nature across the field, σ_γ^2; the correlation between the signal and the state of nature, ρ ; the curvature of the response function[42] β_2; the effect of the state of nature on the response, β_4; and the price of output, *p*.

---

[42] The value of VRT varies inversely with the curvature of the response function, β_2, i.e., the flatter the response curve, the smaller is the potential loss from a suboptimal application rate and therefore the less is the value in knowing the optimum rate.

2.3 Empirical strategy

The sensitivity of the EVSI to the underlying parameters

While we cannot observe the underlying parameters ρ and σ_γ, the assumed bivariate normal distribution of γ and s allows us to derive an approximation of equation (10). By Bayes' rule, the posterior mean of the distribution of γ, given s , is given by:

$$E(\gamma|s) = \mu_{\gamma|s} = \mu_\gamma + \rho\frac{\sigma_\gamma}{\sigma_s}(s - \mu_s) \tag{11}$$

Taking the expectation of (6) with respect to the prior distribution of γ, we obtain the following expression, for given values of x:

$$E_\gamma(y) = \alpha + \beta_1 x + \beta_2 x^2 + \beta_3\mu_\gamma + \beta_4\mu_\gamma x \tag{12}$$

Similarly, we take the expectation of (4) with respect to the posterior distribution of γ to obtain:

$$E_{\gamma|s}(y) = \alpha + \beta_1 x + \beta_2 x^2 + \beta_3\left(\mu_\gamma + \rho\frac{\sigma_\gamma}{\sigma_s}(s - \mu_s)\right) + \beta_4\left(\mu_\gamma + \rho\frac{\sigma_\gamma}{\sigma_s}(s - \mu_s)\right)x \tag{13}$$

Substituting $\tilde{s} = \frac{(s-\mu_s)}{\sigma_s}$, so that $\tilde{s} \sim Normal\ (0, 1)$, re-arranging terms and adding a random error term ϵ , we obtain the following reduced form estimating equation:

$$E_{\gamma|s}(y) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3\tilde{s} + \theta_4\tilde{s}\,x + \epsilon\,, \tag{14}$$

where $\theta_0 = \alpha + \beta_3\mu_\gamma$, $\theta_1 = \beta_1 + \beta_4\mu_\gamma$, $\theta_2 = \beta_2$, $\theta_3 = \beta_3\rho\sigma_\gamma$ and $\theta_4 = \beta_4\rho\sigma_\gamma$ are the reduced form parameters.

Given that the expected value of $\tilde{s}$ is zero, equivalent to the first order condition for the expected profit maximizing application rate without observing s, the optimal uniform rate applied to all cells is:

$$x'(\mu_\gamma) = \frac{w}{2\theta_2 p} - \frac{\theta_1}{2\theta_2}\,, \tag{15}$$

and the first order condition for the expected profit maximizing application rate conditional on having observed s (the variable rate for a given cell) yields:

$$x''(\tilde{s}) = \frac{w}{2\theta_2 p} - \frac{\theta_1}{2\theta_2} - \frac{\theta_4\tilde{s}}{2\theta_2} \tag{16}$$

Notably, EVSI measure of the value of the VRT technology in equation (10) becomes:

$$EVSI = -\frac{p\beta_4^2}{4\beta_2}\left[\rho^2\sigma_Y^2\right] = -\frac{p\theta_4^2}{4\theta_2} \tag{17}$$

Sensitivity

Equation (17) reveals some of the determinants of the payoff from VRT technology. In brackets, we see that EVSI increases with the correlation between the signal and the state of nature, and with the variability of the state of nature across the field. There is no benefit to variable rate application if the field is perfectly uniform, or if the correlation between the signal and the state of nature is zero. From equation (17), the elasticity of EVSI with respect to the correlation ρ, is:

$$\frac{\partial lnEVSI}{\partial ln\rho} = 2 \tag{18}$$

Similarly, elasticity of EVSI with respect to the variance of the state of nature, σ_Y^2 is 1.0, and the elasticity with respect to the curvature β_2, is -1.0, while the elasticity with respect to the interaction coefficient, β_4, is 2.0.

Data

We use a rich set of experimental data from the Data-Intensive Farm Management (DIFM, 2018) project at University of Illinois[43]. To estimate the response function (14), we pooled data from 10 farm fields, 4 in Illinois in 2016, 5 in Illinois and 1 in Ohio in 2017, with a total of 7,294 cells[44]. We used 1,431 cells from the 3 farms (two in Illinois and one in Nebraska in 2017) to generate new distributions of the signal for the EVSI robustness analysis. The data consists of corn yields, nitrogen application (N) and soil electroconductivity (EC), which is the observed signal for each cell i in field j. EC is usually associated with the availability of nitrate in the soil in which high levels are expected to increase yields (Johnson et al., 2003, Liu et al., 2006). It is a soil signal that correlates with some soil properties as: texture, drainage, cation-exchange capacity, and subsoil characteristics (Grisso et al., 2005). Data on EC can be obtained in a shorter period and is more cost-efficient than traditional grid-based soil testing. Figure 4 illustrates the location of the 2017 trials.

---

[43] See https://publish.illinois.edu/data-intensive-farm-managment/2016/02/23/hello-world/

[44] We refer to sub-units within fields as “cells”, but in the precision agriculture literature they are also referred to variously as “management zones”, "sites", “plots”, or “grids”, depending somewhat on how the subunits are identified.

Figure 4. Location of the six 2017 trials used in the study. Source: DIFM.

Table 1 presents the descriptive statistics for the 10 fields used in the estimation of the yield function.

Table 1: Descriptive statistics for variables.

Variable	Observations	Units	Mean	Std. Dev.	Min	Max
Corn dry yield	7,294	bushels/acre	232.65	31.45	0.0	315.67
Applied nitrogen (N)	7,294	pounds/acre	174.15	35.89	29.88	315.94
Soil EC	7,294	Veris EC scale	37.32	9.36	7.6	80.2
Standardized EC ($\widetilde{EC}$)	7,294	Veris EC scale	0.0	1.0	-3.17	4.58

2.4 Results

Estimated response function

To estimate equation (14) the signal $\tilde{s}$ is represented as standardized observed ECs, referred to as $\widetilde{EC}$, and the nitrogen application *x*, represented by N, is measured per acre. Table 2 presents the estimates of the pooled yield response function for all 10 fields, 7,294 observations. The nitrogen and the nitrogen squared coefficients are both statistically significant at 10%. We included dummy variables for fields 2 to 10 so coefficients of these dummies are intercept changes relative to field 1. From the variable *N* response equation (14), we conclude that, because estimates of θ_2 and θ_4 are both negative, there is an inverse relationship between $\widetilde{EC}$ and optimal *N* rate. In other words, $\widetilde{EC}$ is a substitute for *N*.

Table 2: Pooled corn yield response estimation results.

Variable	Coefficient	Estimate (Standard Error)
N_{ij}	θ_1	0.455*
		(0.225)
N_{ij}^2	θ_2	-0.00101*
		(0.000547)
$\widetilde{EC}_{ij}$	θ_3	9.540***
		(2.182)
$\widetilde{EC}_{ij}.N_{ij}$	θ_4	-0.0494***
		(0.0113)
Fixed effect, field 2	d2	77.75***
		(0.455)
“	d3	87.53***
		(0.791)
“	d4	86.84***
		(0.831)
“	d5	38.82***
		(0.932)
“	d6	99.53***
		(1.435)
“	d7	86.84***
		(1.889)
“	d8	119.7***
		(1.281)
“	d9	109.9***
		(0.767)
“	d10	78.23***
		(2.408)
Constant		90.80***
		(21.99)

Note: D2-D10 are the dummy parameters for fields 2 to 10. Standard errors in parentheses are clustered at the farm level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3 presents for each field the mean, standard deviation, minimum and maximum of EC and optimal VRT rates across the cells within each field, calculated using estimated equations (12) and (14), evaluated at the values of N and EC for each cell in each field. Ordering from lowest to highest average EC, we observe the clear inverse relationship between optimal VRT rate and EC, as implied by the negative estimate of the interaction coefficient for N and EC, θ_4. For example, fields 7 and 10, which have the lowest average EC, have the largest average optimal VRT applications of 180.80 and 183.23 lbs/ac. Fields 1 and 4, with the highest average EC, have the lowest average optimal VRT application of 143.62 and 130.77 lbs/ac. The relationship between the standard deviation of EC and the standard deviation of optimal VRT application is similarly monotonic.

Table 3: Soil electroconductivity (EC) and estimated optimal VRT Nitrogen application per field, Illinois (4 fields in 2016, 5 in 2017), and Ohio (1 in 2017).

		EC				Optimal VR Nitrogen Application			
Field	Obs.	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
1	127	42.25	6.66	32.68	66.97	143.62	17.46	78.76	168.71
2	160	38.42	3.24	31.40	44.26	153.66	8.51	138.33	172.07
3	160	35.58	4.05	24.73	42.48	161.11	10.63	143.01	189.58
4	256	47.15	6.31	34.58	60.41	130.77	16.56	95.99	163.74
5	581	35.47	8.86	20.12	79.57	161.41	23.24	45.71	201.68
6	1,548	40.44	6.22	22.81	59.22	148.37	16.32	99.10	194.61
7	682	28.08	11.48	7.60	78.85	180.80	30.11	47.60	234.52
8	819	33.06	9.07	14.57	59.45	167.72	23.79	98.48	216.23
9	2,347	41.28	6.98	22.68	80.22	146.17	18.31	44.01	194.97
10	614	27.15	6.17	15.92	48.32	183.23	16.20	127.69	212.69
Pooled cells	7,294	37.32	9.36	7.6	80.2			44.01	234.52

[a]In lbs/a, using corn price=$3/bu and nitrogen price=$0.42/lb.

Using equation (17) with the parameter estimates for the pooled sample of 10 fields, we estimate EVSI to be 1.81/ac:

$$\widehat{\text{EVSI}} = -\frac{p\hat{\theta}_4^2}{4\hat{\theta}_2} = -\frac{3*(-.0494)^2}{4(-.00101)} = \$1.81/ac \qquad (19)$$

We also use equation (14) to estimate EVSI for each individual cell, using the observed EC. Table 4 presents the average and range of cell-level EVSI estimates by field, using observed EC and estimated optimal VRT by cell from Table 3. These estimates range from $0 to $38.21/ac, but average $1.82/ac; essentially the same as the estimate in equation (19).

The $1.81/ac from equation (19) is our best ex-ante estimate of gross return to VRT for fields drawn from a distribution of fields like those in our sample. It is the expected return from using the signal from each cell to generate an optimal rate, ignoring the costs of the VRT technology package. While it is below the estimates in the literature of those costs, this result is like the ones in other empirical studies. BCS (1996) estimated higher comparable payoffs at $2.93-10.03 per acre. Bullock, et al (2009) found comparable payoffs to be a dollar per acre, or less, and concluded that prospects for VRT “are generally dim”.

Table 4: Estimated EVSI by cell, average and range per field in Illinois and Ohio, 2016-2017. (US$/a).

Farm ID	Obs.	Mean (Std. Dev)	Min	Max
1	127	1.42 (3.29)	0	18.25
2	160	0.24 (0.23)	0	0.99
3	160	0.40 (0.72)	0	3.29
4	256	2.83 (2.59)	0	11.06
5	581	1.70 (3.50)	0	37.07
6	1,548	1.00 (1.17)	0	9.95
7	682	4.50 (3.97)	0	35.81
8	819	2.08 (2.13)	0	10.75
9	2,347	1.33 (2.31)	0	38.21
10	614	2.94 (2.29)	0	9.52
Pooled Cells	7,294	1.82 (2.64)	0	38.21

Figure 7 shows the estimates of Kernel regression between the cell-level EVSI and EC values. This shows that individual cell EVSI will be higher as we move away from the average EC of all fields (37.32).

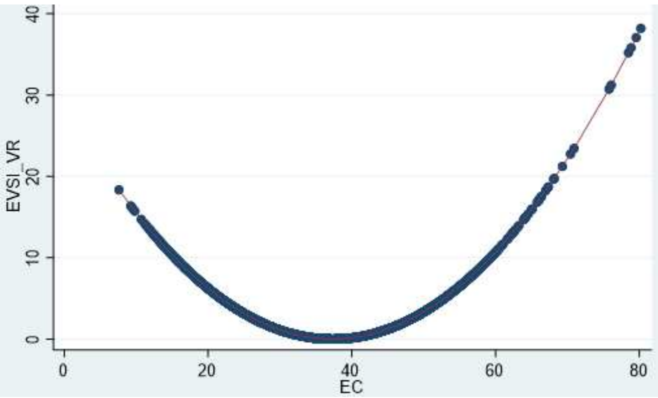

Figure 7. Kernel regression of cell-level EVSI on EC for 7,294 observations.

Sensitivity results

The estimated EVSI of $1.81/ac is too low to warrant adoption. We examine here changes in various parameters that would be sufficient to achieve an EVSI of (arbitrarily) $10/ac, which is about 5.6 times higher than the estimated level. We solve equation (17) for the various parameters to derive estimates of the changes in individual parameters that would be sufficient to increase the EVSI by about 5.5 times, to $10/ac. Solving (17) for dlnρ yields:

$$dln\rho = \frac{5.52}{2} = 2.76. \qquad (20)$$

We estimate that a ρ 2.76 times larger than the ρ of these fields would be sufficient to raise EVSI to $10/ac. But of course, we do not have an estimate of ρ. Judging from the generally low VRT payoff measured here, this correlation must be low – perhaps ρ=0.1 or as little as ρ=0.01. If ρ=*0.1*, then from the equation above, ρ would need to increase from 0.1 to 0.376. If ρ=*0.01*, ρ would need to increase from 0.01 to 0.038. However, if $\rho \geq 0.362$, apparently there is no increase that would yield an EVSI of $10/ac or more. In any case, if a 2.8-fold increase in ρ is required, it seems that electrical conductivity is not sufficiently correlated with N response on these fields to be a profitable signal.

Considering now the variance of the state of nature over the field, σ_γ^2, from equation (11) and given that the elasticity of EVSI with respect to σ_γ^2 is 1.0, the necessary percentage increase in variance to achieve an EVSI of $10/ac is 10/1.81= 5.52. If the distribution of γ is like that of EC, this would imply an increase of σ_γ^2 from about 87.6 (the variance of EC across all fields) to 484, which is much higher than the EC variance of 132 in field 7, the most variable of any of the fields. Similarly, the curvature coefficient, β_2, would need to increase from an absolute value of 0.001 to 0.0065, which is an indication that profits in our sample are not highly sensitive to the level of N applied. The interaction coefficient, β_4, would need to change from -0.0494 to -0.186, a further indication that N response in this sample is not greatly affected by the level of EC.

EVSI robustness analysis: the role of the EC distribution

This section performs a robustness exercise for the estimated EVSI to study the role of the signal (EC) distribution in our ex-ante analysis. This exercise is useful to test whether the estimated $1.82/acre changes with the characteristics of the EC distribution (mean and variance) for all fields, which would go against the analytical expression (equation 10) obtained from the theoretical model. The empirical model assumed that EC is normally distributed across the cells for the pooled fields and allowed the inclusion of the standardized EC values in the pooled regression estimation. As we can observe from Figure 5, estimating a regression for each field separately would not allow the assumption of normality of the EC distribution (i.e., EC is not normally distributed for the individual fields). Different distribution assumptions would be necessary because of the heterogenous variability of EC within each field, requiring different approaches, such as simulation methods, to compute the ex-ante expected payoff for the VRT[45].

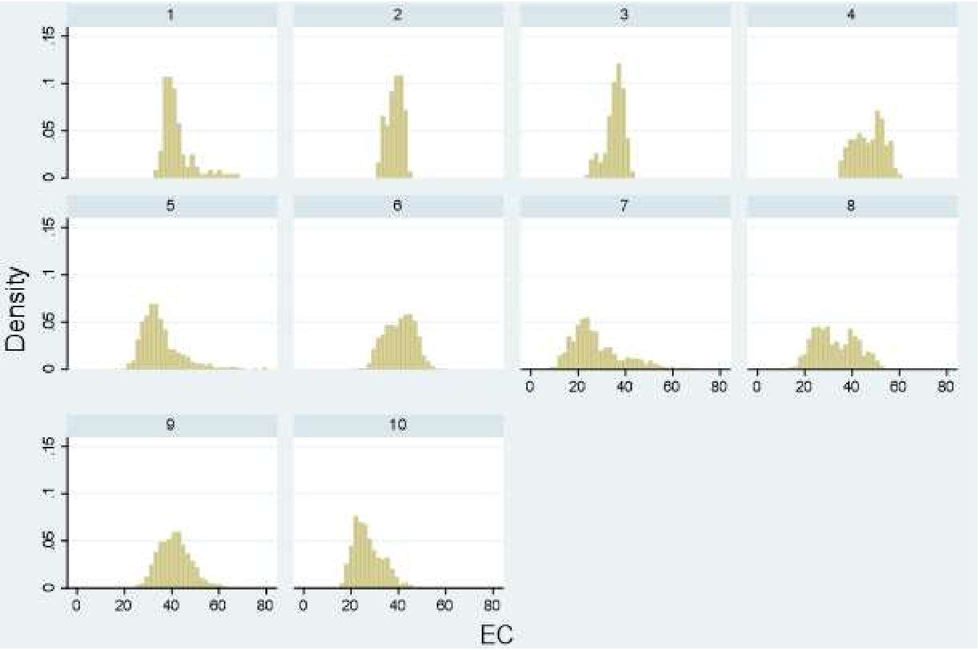

Figure 5. Histogram of Electroconductivity (EC) for each field.

Figure 6 shows the histogram of EC for the 10 pooled fields used in the estimation. The standardized EC was obtained based on the mean and the variance of this distribution, where EC is the outcome of random draws from the same (Normal) population distribution across all fields. Similarly, the average EVSI for the pooled cells in Table 4 was calculated based on the average of the cell-level EVSI's calculated using the EC values from this distribution.

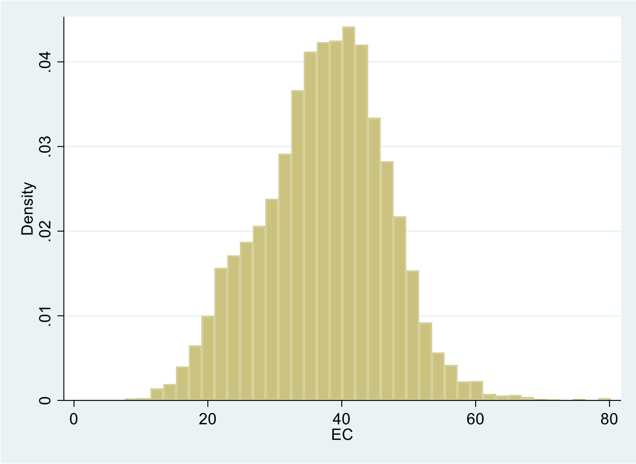

Figure 6. Histogram of Electroconductivity (EC) for 7,294 cells in 10 fields in 2016 and 2017.

---

[45] Therefore, the bivariate normal case was crucial to estimate the analytical and empirical EVSI results of this paper.

The robustness exercises tests if the estimated (pooled cells) EVSI is sensitive to other EC distributions. New EC values for the 7,294 observations were generated from normal distributions based on the mean and the variance of each of the 3 fields not included in the estimation. Then, the EVSI for each distribution was computed using the estimated parameters from the quadratic yield function. Let EC11, EC12 and EC13 correspond to the generated normal distributions based on the means and variances of fields 11, 12 and 13, respectively. Table 5 shows the estimated EVSI's.

Table 5: EVSI based on different EC distributions.

Distributions		Ex-ante EVSI		
EC11		Mean	Min	Max
$\mu_{s_{11}}$	$\sigma_{s_{11}}$			
92.34	23.49	1.82	0	27.68
EC12				
$\mu_{s_{12}}$	$\sigma_{s_{12}}$			
30.98	7.38	1.82	0	29.89
EC13				
$\mu_{s_{13}}$	$\sigma_{s_{13}}$			
58.92	11.19	1.82	0	20.78

Note: μ_{s_i} and σ_{s_i} are the mean and standard deviations of field $i \in \{11, 12, 13\}$.

As expected, the EC distribution does not affect the expected return from VRT and the soil electroconductivity information (EVSI). However, the individual EVSI does change with the distribution of EC as shown by the changes in the maximum values. The result in Table 5 confirms the theoretical implication that the distribution of signal has no role in determining the EVSI for the VRT on fields where the true soil characteristic and the signal are bivariate normally distributed. The low expected value of VR application may be solely determined by the other parameters of the model, such as a potential low correlation between the true and observed variables, and the variability of the soil condition, which are both unobserved to the researcher.

2.5 Conclusion

In this paper we have adapted insights from the decision theory literature on the value of information to provide an economic model of the value of VRT (variable rate technology) as the expected value of sample information (EVSI). The sample information in our case was the observed electroconductivity (EC) of the soil as a signal for an unobservable soil characteristic affecting nitrogen response. Our theoretical results provided an estimate of the expected value of VRT as an explicit function of parameters representing five underlying factors: variability of soil characteristic across the field; curvature of the response function; effect of the soil characteristic on input response; correlation of the signal with the soil characteristic; and the ratio of crop price to input price. The expected value of VRT is taken with respect to the frequency distribution of the state of nature and the distribution of the signal obtained across all cells in all ten fields/years observed. This expected value can therefore be taken as the ex-ante expected payoff from adopting VRT on any field drawn from the same population of field/years as those observed.

To obtain tractable analytical results for this approach, we assumed that the yield response to fertilizer is quadratic in applied nitrogen (N) and the state of nature, γ. The state of nature cannot be observed, but a signal s can be observed for each cell. A second assumption critical to our results is that γ and s are distributed bivariate normally across cells in these fields. Given this underlying structure, individual cell application rates can be adjusted to the level that maximizes the expected payoff conditional on the signal. The difference between this optimal expected payoff and the expected payoff from an optimal uniform application rate (UAR) provides the expected gross payoff from the adoption of VRT. In the decision theory literature, this is known as the expected value of sample information (EVSI).

We applied this approach to estimate the ex-ante expected payoff to VRT using data from field-level experimental trials with nitrogen on 10 farmers' corn fields in Illinois and Ohio in 2017 and 2018, consisting of 7,294 gridded cells. The signal used to adjust the fertilizer rate for each cell is electrical conductivity. Our EVSI estimate of the ex-ante expected payoff of VRT is \$1.81/ac (prior to subtracting VRT implementation costs). This is insufficient to warrant VRT implementation costs, which we believe to be in the range of \$10/ac. Our analysis suggests that for VRT benefits to reach this level, the correlation between state of nature and signal would need to increase by roughly 2.8 times, though we are not able to estimate the level of that correlation. Alternatively, the same improvement in VRT value could be attained by an increase of similar size for either the curvature coefficient or the N times EC interaction coefficient, or a five-fold increase in the variance of the state of nature could reach the same outcome. Clearly, some of these changes could occur if we had a more robust measure than electrical conductivity (EC) of the state of nature affecting N response.

Our approach has obtained some results regarding the determinants of VRT payoff that were previously understood intuitively, but not analytically or quantitatively. Our claim for our results to be a measure of the expected payoff of VRT on a similar field, is that the $1.81/ac is a plausible estimate of the expected gross benefit of VRT across a field with cells drawn from the same distribution as the 7,294 cells in our sample. Perhaps such variables as soil classification, remote sensing data, etc., may provide coarser but cheaper signals for calibrating application levels.

REFERENCES

ANSELIN, L.; BONGIOVANNI, R.; LOWENBERG-DEBOER, J. A spatial econometric approach to the economics of site-specific nitrogen management in corn production. **American Journal of Agricultural Economics**, 86(3), 2004, p. 675–687.

BABCOCK, B. A. The Effects of Uncertainty on Optimal Nitrogen Applications. **Review of Agricultural Economics**, 14, 1992, p. 271-280.

BABCOCK, B.; CARRIQUIRY, A.; STERN, H. S. Evaluation of Soil Test Information in Agricultural Decision-Making. **Applied Statistics**, 45(4), 1996, p. 447-461.

BULLOCK, D. G.; BULLOCK, D. S. Quadratic and Quadratic-Plateau Models for Predicting Optimal Nitrogen Rate of Corn: A Comparison. **Agronomy Journal**, 86, 1994, p. 191-194.

BULLOCK, D. S.; LOWENBERG-DEBOER, J. Using Spatial Analysis to Study the Value of Variable Rate Technology and Information. **Journal of Agricultural Economics**, 53, 2007, p. 517-35.

BULLOCK, D. S.; MAXWELL, B. A.; TAO, H.; BOERNGEN, M. A.; LUCK, J. D.; SHIRATSUCHI, L. S.; PUNTEL, L.; MARTÍN, AND N. F. **The Data-Intensive Farm Management Project: Changing Agronomic Research Through On-farm Precision Experimentation**. Working Paper. University of Illinois Department of Agricultural and Consumer Economics, 2019.

BULLOCK, D. S.; RUFFO, M. L.; BULLOCK, D. G.; BOLLERO, G. A. The Value of a Variable Rate Technology: An Information-Theoretic Approach. **American Journal of Agricultural Economics**, 91(1), 2009, p. 209-223.

BULLOCK, D.S.; BULLOCK, D. G. From Agronomic Research to Farm Management Guidelines: A Primer on the Economics of Information and Precision Technology. **Precision Agriculture**, 2, 2000, p. 71-101.

DATA INTENSIVE FARM MANAGEMENT (DIFM) Advisory Board Report. 2018. Available at: http://publish.illinois.edu/data-intensive-farm-managment/files/2018/02/final_DIFM-Advisory-Board-Report.pdf.

GRISSO, R. D.; ALLEY, M. M.; HOLSHOUSER, D. L.; THOMASON, W. E. **Precision Farming Tools: Soil Electrical Conductivity**. Virginia Cooperative Extension, Publication 422-508. 2005. URI: http://hdl.handle.net/10919/51377.

HURLEY, T. M.; OISHI, K.; MALZER, G. L. Estimating the potential value of variable rate nitrogen applications: A comparison of spatial econometric and geostatistical models. **Journal of Agricultural and Resource Economics**, 30(2), 2005, p. 231–249.

HURLEY, T.M.; MALZER, G.L.; KILIAN, B. Estimating Site-Specific Nitrogen Crop Response Functions: A Conceptual Framework and Geostatistical Model. **Agronomy Journal**, 96, 2004, p. 1331- 43.

JOHNSON, C. K.; MORTENSEN, D. A.; WIENHOLD, B. J.; SHANAHAN, J. F.; DORAN, J.W. Site-specific Management Zones Based on Soil Electrical Conductivity in Semiarid Cropping System. **Agronomy journal**, 95(2), 2003, p. 303-315.

KIHLSTROM, R. Firm Demand for Information about Price and Technology. **Journal of Political Economy**, 84(6), 1976, p. 335-341.

LAMBERT, D.; LOWENBERG-DEBOER, J.; MALZER, G. L. Economic analysis of spatial-temporal patterns in corn and soybean response to nitrogen and phosphorous. **Agronomy Journal**, 98, 2006, p. 43–54.

LAWRENCE, D. **The Economic Value of Information**. Springer-Verlag, 1999, 393 pp.

LIU, Y.; SWINTON, S. M.; MILLER, N. R. Is site-specific yield response consistent over time? Does it pay? **American Journal of Agricultural Economics**, 88(2), 2006, p. 471–483.

LOWENBERG-DEBOER, J; ERICKSON, B. Setting the Record Straight on Precision Agriculture Adoption. **Agronomy Journal**, 111, 2019, p. 1552-1569.

ODELL, R.T.; WALKER, W.M.; BOONE, L.V.; OLDHAM, M.G. **The Morrow Plots: A Century of Learning**. Bulletin 775. Urbana: University of Illinois Agricultural Experiment Station, 1984.

PERRIN, R. K. The Value of Information and the Value of Theoretical Models in Crop Response Research. **American Journal of Agricultural Economics**, 58(1), 1976, p. 54-61.

RAIFFA, H; SCHLAIFER, R. **Applied Statistical Decision Theory**. Boston, Massachussets: Division of Research, Harvard Business School, 1961.

RUFFO, M.; BOLLERO, G.; BULLOCK, D. S.; BULLOCK, D. G. Site-specific production functions for variable rate corn nitrogen fertilization, **Precision Agriculture**, 7, 2006, p. 327– 342.

SCHIMMELPFENNIG, D. **Farm Profits and Adoption of Precision Agriculture**. ERR-217, U.S. Department of Agriculture, Economic Research Service, 2016.

SWINTON, S.M.; LOWENBERG-DEBOER, J. Evaluating the Profitability of Site-Specific Farming, **Journal of Production Agriculture**, 11(4), 1998, p. 439-446.

CHAPTER 3

ESTIMATING THE EFFECT OF DROUGHT ON BRAZILIAN AGRICULTURE USING REMOTE SENSING INFORMATION

3.1 Introduction

Drought incidences have been reported more frequently for agricultural regions[1], especially in the developing world. Together, it is observed greater economic interest on estimating the agricultural effects of droughts to assist government agricultural policies such as the financial drought-alleviating assistance programs oriented to farmers exposed to severe droughts. These policies aim to mitigate the adverse effects of droughts (ex-post) because the agricultural sector is the most affected by these events (Kuwayama et al., 2019). Measurements of past and current drought events can now be easily constructed with data from meteorological stations and/or satellite imagery technologies and expressed in periods such as days, weeks, or months. Moreover, remote sensing satellite imagery and computational advancements have allowed to recover high resolution climatic and biophysical data (i.e., remote sensing information) for different regions over time. Therefore, the marginal effect of an extra period of drought, and the resulting agricultural losses[2], can be more precisely estimated.

The objective of this study is to estimate the effect of drought in agriculture for two samples of Brazilian municipalities, one consisting of all the corn- and soybeans-producing Brazilian municipalities (hereafter, Brazil sample), and the other for the municipalities of a soybeans-producing region in Southern Brazil (hereafter, Southern Brazil sample). First, we present a motivating theoretical framework to discuss the use of drought indicators for a government irrigation subsidy plan to mitigate the effect of droughts in agriculture (ex-ante). This framework allows to identify the determinants of the theoretical economic value of information, highlighting the role of the correlation between actual droughts and the drought indicators to provide better informed agricultural decision-making. Then, we estimate the effect of drought on agricultural outcomes (yield and value of production) for all Brazilians municipalities that produced corn and soybeans over the 2002-2016 period, measuring droughts

---

[1] These reports are supported by the increasing global trend for drought events shown by climatic data and climate prediction models (Dai et al., 2013).

[2] Droughts are usually characterized by abnormally low levels of precipitation and abnormally high levels of temperature, decreasing the level of soil moisture, and are expected to negatively affect agricultural yields (Wilhite, 2000; Behrangi et al., 2016; Battisti et al., 2017).

in months of the growing season. Next, we restrict the sample to a soybean production area in the Southern Brazilian states of Paraná, Santa Catarina and Rio Grande do Sul for the same period. The sample of the Southern Brazilian municipalities was used to match the agricultural indicators with the remote sensing drought indicators, one constructed from precipitation data, measured in weeks, and the others constructed from biophysical variables capturing the characteristics of the agricultural land, the plant conditions, and the atmosphere. The effect of drought is given by the average response in agricultural outcomes due to one more month (for the Brazil sample) or week (for the Southern Brazil sample) of severe and/or extreme drought during the plant growing season in a year.

Several indicators based on remote sensing data are usually combined with the common meteorological variables (i.e., precipitation and temperature) to measure the incidence of droughts. The United States Drought Monitor (USDM)[3] index is a map released weekly to show which parts of the United States had one of five drought classifications during a week[4]. A commonly used drought indicator, also considered in the USDM, is the Standardized Precipitation Index (SPI). The SPI is a meteorological indicator that uses only observed precipitation, but studies from agronomic and remote sensing sciences have reported that it alone poorly captures the relationship among the soil, the plant, and the atmosphere. Other biophysical conditions influence the performance of the plant and that should be considered when assessing the effects of droughts in agriculture (Anderson et al., 2016, Mariano, 2019). However, few studies in agricultural economics have quantified the marginal effect of an extra period of drought in agriculture[5] and included biophysical indicators other than the usual climatic variables.

Previous economic studies have estimated the impact of climate change on agriculture (Schenkler, Hanneman and Fisher, 2005, 2006; Deschênes and Greenstone, 2007; Fisher et al., 2012; Cui, 2020). Others focused on the adaptation strategies to droughts (Ding, Schoengold and Tadesse, 2009; Hornbeck and Keskin, 2014), the effects of drought insurance on

---

[3] The USDM index is jointly developed by National Drought Mitigation Center (NDMC) at the University of Nebraska-Lincoln, the National Oceanic and Atmospheric Administration (NOAA), and the U.S. Department of Agriculture (USDA).

[4] The USDM then combines data from satellite imagery with other climatological station data to indicate drought intensity levels across the country. The five drought classifications are: abnormally dry (D0), moderate (D1), severe (D2), extreme (D3) and exceptional (D4). More information is available at: https://droughtmonitor.unl.edu/.

[5] The main reason for that is that there is no one universal definition of drought (Kuwayama et al., 2019). Despite varying across regions and time, the definition of drought can depend on the socioeconomic characteristics. Regions that depend heavily on the climatic conditions, such as agricultural producing areas, can consider one mild drought level as strong depending on the impact on the local economy.

agricultural conservation practices (Schoengold, Ding and Headlee, 2015), the effects of drought and irrigation on natural resources and agriculture using hydroeconomic models (Maneta et al., 2009; Silva, Fulginiti, Perrin, and Schoengold, 2019), and the value of water for irrigation in terms of agricultural production controlling for the effect of climatic variables (degree days and precipitation) (García-Suárez, Fulginiti and Perrin, 2019). Meanwhile, few studies have quantified the marginal effects of droughts, except for Bernknopf et al. (2018) and Kuwayama et al. (2019). Both papers estimated the relationship between the USDM drought severity indicators and agriculture using county-level data for the United States. Because the USDM is a weekly classification of drought at the county level, it allows computing the number of weeks that the county was under drought for a given year.

Kuwayama et al. (2019) found that one more week of the most severe drought can reduce yields by up to 1.2% in dryland counties and up to 8% in dryland counties in the U.S. Midwest. However, they found no significant effect of the drought indicators on farm income. Bernknopf et al. (2018) added remote sensing variables in their empirical model with the USDM drought indicators to predict yields and farm income in the U.S. counties. They found significant differences among the models with only the USDM indicators and the models that included the USDM indicators and the remote sensing data, concluding that having more information on groundwater storage and soil moisture from remotely sensed satellite imagery decreased the uncertainty about the drought measurements. They concluded that adding more variables to the model increased the overall correlation between the drought indicators and the actual droughts, providing more robust estimates of the drought effects. This study follows closely the papers of Kuwayama et al. (2019) and Bernknopf et al. (2018) by estimating the marginal effect of an additional month/week of drought and including remote sensing biophysical variables in the econometric models.

We differentiate from their studies in mainly two ways. First, by comprehensively assessing the effect of drought for Brazilian agriculture. Brazil is a country that relies heavily on rainfed agriculture and had severe drought episodes in the previous years. Although these events are mostly common in the Northeast semiarid area, droughts happen all over the country. For example, a massive drought hit the Southeast region and parts of the South in 2014[6] (De Nys, Engle and Magalhães, 2017), and a severe drought was observed in the Amazon region in

---

[6] This drought caused one of the largest water shortages experienced in the State of São Paulo, in which the Cantareira System, which is comprised of reservoirs that supply water for around 70% of the São Paulo city metropolitan area, reached historically low levels (De Nys, Engle and Magalhães, 2017).

2015 (Mariano, 2019). We provide estimations for the pooled soybeans- and corn-producing municipalities and for the Brazilian geographic regions separately[7]. Second, we define a specific soybeans-producing region in the South of Brazil to identify the extent the effect of drought is biased if biophysical variables, constructed from remote sensing satellite imagery, are omitted from the models. Few studies analyzed the effect of droughts on Southern states, although both the Northeast and the South regions are the most affected by the El Niño Southern Oscillation (ENSO) (Mariano et al., 2018, Mariano, 2019). Southern Brazil has a chance of an extreme drought event once in every ten years (Awange et al., 2016). Also, it comprises the states with the second and the third largest quantities of soybeans production in Brazil (Rio Grande do Sul and Paraná, respectively), accounting for approximately 30% of total production in 2017 (IBGE, Censo Agropecuário 2017).

There is no study, to our knowledge, that quantified the effect of drought on Brazilian agriculture using municipality level data and drought indicators in panel data regressions with fixed effects (FE)[8]. The main advantage of panel data models is that we can control for fixed unobservable variables for each municipality and year that may be correlated with the drought measurement, providing more reliable estimates of the drought effect. Furthermore, including biophysical variables can improve the model specification. We use different periods of the growing season to identify these indicators. The importance of remote sensing information in these models is highlighted mainly because it increases the correlation between the drought variables and the actual droughts. The insights from the motivating theoretical framework, based on the value of information (VOI) theory, indicate that this correlation is a strong determinant of the value of information obtained from the meteorological drought indicators (e.g., from observed precipitation). Therefore, remote sensing information can improve decision making when these indicators are considered in the decision-making process ex-ante. The VOI theory and applications have evolved significantly in the earth and energy sciences as new technologies, such as remote sensing, allowed to obtain dynamic and spatial information on the landscapes. Eidsvik, Mukerji, and Bhattacharjya (2015) present a complete framework on the VOI for decision analysis with applications in the earth sciences. Although commonly

---

[7] Brazil has five geographic regions. We present in the Appendix a map of the Brazilian geographic region division.

[8] Most studies on droughts in Brazil are for the Northeast region. Historically, the semi-arid region of the Brazilian Northeast has suffered from severe droughts with extremely high temperatures and low levels of precipitation in the Summer (Cunha et al., 2018, Alvalá et al., 2019). The most recent extended drought in the Northeast was from 2010-2015, where, starting from 2012, most reservoirs have dried up and large losses in agricultural output were observed (De Nys, Engle and Magalhães, 2017).

used in economics and finance[9], there is still need for more investigation related to the economic value of remote sensing information, especially for agricultural decisions. It is not the scope of this study to insert remote sensing information directly into the decision-making process, but we highlight its importance mainly as a tool to evaluate the expansion of drought monitoring to different agricultural regions. VOI is relevant in the context of Brazil because there have been efforts to expand the Brazilian Northeast Drought Monitor (NDM) [10], created in 2014, to other regions of the country. The creation of a drought monitor in Brazil is a movement that started from the prolonged recent droughts in the country and is an ongoing debate among experts and stakeholders (Gutiérrez et al., 2014; De Nys, Engle, and Magalhães, 2017). The construction of countrywide drought indicators such as the USDM can provide better guidance for agricultural decision-makers and agricultural public policies as an effort to shift away from reactionary response to long-term solutions to droughts. Moreover, to improve monitoring and management in Brazil, government investments in satellite technologies can be justified if the cost of droughts has a large share of the total value of agricultural production[11].

The paper is organized as follows. Section 2 presents our motivating theoretical framework to indicate the importance of the correlation between the drought indicators and the actual droughts, and the role of remote sensing biophysical variables. Section 3 presents the description of the data and the construction of the variables for the Brazil and the Southern Brazil samples. Section 4 contains the analysis for all the Brazilian municipalities that produced corn and soybeans for the period studied. Section 5 shows the results for the soybeans production region in Southern Brazil. Section 6 concludes. The Appendix brings further results.

3.2 Motivating theoretical model

Actual droughts vs. drought indicators from public sources

Remote sensing information can help improve the prediction about the intensity of a drought event expressed in a drought indicator. A Bayesian approach is useful to evaluate how decision-makers change the expectation about an uncertain outcome of interest when they can obtain some related information (Lawrence, 1999). Following Bernknopf et al. (2018), let T_{mt}

---

[9] VOI is also commonly used in medicine.

[10] The NDM is a joint effort of state, federal and private agents to create maps of drought incidences using satellite imagery, specialists' inputs, and other meteorological indicators in the Northeast region. It is an initiative like the United States Drought Monitor (USDM) that aims to provide prompt information from public agents to individual decision makers.

[11] For examples of studies that use the value of information theory to justify investments in geological maps and remote sensing technology, see Bernknopf et al. (1997) and Macauley (2006).

be a random variable that represents the “true” drought intensity in municipality *m*, year *t*, and let D_{mt} be the respective drought indicator that can be obtained through public sources such as meteorological stations, providing observed precipitation and temperature, or the USDM.

The meteorological drought indicator contains information related to the uncertain drought intensity realization. The distributions of T_{mt} and D_{mt} are assumed to be normal (N) where $T_{mt} \sim N(\mu_{T_{mt}}, \sigma^2_{T_{mt}})$ and $D_{mt} \sim N(\mu_{D_{mt}}, \sigma^2_{D_{mt}})$. We assume that both random variables are correlated and that they jointly follow a bivariate normal (BN) distribution such that $(T_{mt}, D_{mt}) \sim BN(\mu_{T_{mt}}, \sigma^2_{T_{mt}};\ \mu_{D_{mt}}, \sigma^2_{D_{mt}}; \rho)$, where ρ is the correlation coefficient between the actual drought and the meteorological drought indicator. The posterior distribution of T_{mt} given that we can observe $D_{mt} = d_{mt}$ is:

$$(T_{mt} | D_{mt} = d_{mt}) \sim N \left\{ \mu_{T_{mt}} + \rho \frac{\sigma^2_{T_{mt}}}{\sigma^2_{D_{mt}}} \left(d_{mt} - \mu_{D_{mt}} \right), (1 - \rho^2) \sigma^2_{T_{mt}} \right\} \tag{1}$$

where d_{mt} represents the observed drought indicator in municipality m and year t. Notably, the correlation between the actual drought and the meteorological indicator represented by $\boldsymbol{\rho}$ is what can reduce the uncertainty faced by a decision-maker. If more information, such as biophysical remote sensing data, can increase ρ, we observe that the posterior variance, $(1 - \rho^2)\sigma^2_{T_{mt}}$, decreases, which allows the decision-maker to better estimate the true intensity of the drought.

The correlation coefficient and the economic value of information obtained from the meteorological drought indicator

Suppose the government wants to subsidize farmers for the adoption of an irrigation plan as a (ex-ante) strategy to mitigate the effect of droughts. A quadratic payoff function can be used to obtain the optimal irrigation subsidy plan. The expected payoff function represents *the value of the agricultural losses that can be avoided by adopting the irrigation plan with the government subsidy* and is a function of the true drought intensity T_{mt}, and the amount of the irrigation subsidy, I_{it}. We use expectation to express the uncertainty over the true drought intensity.

$$E[V(T_{mt}, I_{mt}))] = E[\alpha_1 + \alpha_2 T_{mt} + \alpha_3 I_{mt} + \alpha_4 T_{mt} I_{mt} + \alpha_5 T^2_{mt} + \alpha_6 I^2_{mt}] \tag{2}$$

where $\alpha_6 < 0$ for concavity of the payoff function. Obtaining the first-order conditions, we can easily calculate the optimal irrigation subsidy plan using the prior distribution of T_{mt} as $I'_{mt} =$

$\frac{\alpha_3+\alpha_4\mu_{T_{mt}}}{2\alpha_6}$, where $\mu_{T_{mt}}$ is the prior mean. The government can use a public available meteorological drought indicator (either past or current), $D_{mt} = d_{mt}$, to decrease the uncertainty over the true drought events on a municipality at a given year and then decide the optimal subsidy level. By using the posterior density function in equation (1), the optimal irrigation subsidy is given by $I''_{mt} = \frac{\alpha_3+\alpha_4\mu_{T_{mt}|D_{mt}=d_{mt}}}{2\alpha_6}$, where $\mu_{T_{mt}|D_{mt}=d_{mt}}$ is the posterior mean equal to $\mu_{T_{mt}} + \rho\frac{\sigma^2_{T_{mt}}}{\sigma^2_{D_{mt}}}\left(d_{mt} - \mu_{D_{mt}}\right)$.

The value of information (VOI) by using a meteorological drought indicator ($D_{mt} = d_{mt}$) in the decision concerning the subsidy level farmers should receive is obtained from the difference between the expected payoff function using the optimal allocation I''_{mt}, represented by $E[V''_{mt}]$, and the expected payoff function using the optimal allocation I'_{mt}, represented by $E[V'_{mt}]$. Given that the distributions of T_{mt} and D_{mt} are known and that they are bivariate normally distributed with correlation ρ, Lawrence (1999), pp118-119, shows that the VOI can be calculated as:

$$VOI_{mt} = E[V''_{mt}] - E[V'_{mt}] = r.\{\rho^2\sigma^2_{T_{mt}}\} \tag{3}$$

where r is a constant consisting of a combination of the parameters of the expected payoff function. From this result, we observe that the VOI is increasing in ρ. Identifying the size of this correlation is difficult because the actual drought intensity is not something that can be directly measured. Bernknopf et al. (2018) used agricultural variables as proxies for the "true state of droughts" and compared different models by adding remote sensing data to analyze whether they improve the correlation between the publicly available drought indicators and the actual drought intensity, *proxied* by the agricultural variables.

In this paper, we follow these authors and proxy drought realizations with agricultural outcomes (yield and value of production) and assess whether the correlation coefficient, measured by the overall statistical fit of the econometric regressions, improve with the addition of the biophysical remote sensing data. Note that in the irrigation subsidy decision above, using a meteorological drought indicator provides some economic benefit (i.e., VOI) to the decision-maker by decreasing the uncertainty over the actual drought intensity, and a key determinant is the correlation coefficient ρ. Thus, the role of the remote sensing information in our context is to improve the estimation of this correlation allowing better estimates of the effect of droughts on agriculture captured in drought indicators constructed from observed precipitation.

3.3 Data and variables construction

Brazilian municipality samples and period of study

Two different samples of Brazilian municipalities are used in this study for the 2002-2016 period. One is for all Brazilian municipalities (Brazil sample) that produced corn and soybeans, and the other is for 281 municipalities that produced soybeans in the Southern region of Brazil (Southern Brazil sample), in the states of Paraná, Santa Catarina and Rio Grande do Sul. The period of the study is defined as 2002-2016 because from the 57 municipalities created between the agricultural censuses of 1996 and 2006, 53 were installed in 2001 and four in 2005. In addition, five municipalities were installed after 2010, reaching a total of 5,570 municipalities in the country. The nine municipalities created after 2002 were aggregated back to their municipalities of origin[12], hence they are considered minimum comparable areas (or AMCs, acronym in Portuguese). The AMCs are commonly used in studies involving Brazilian municipalities because of the historical changes in the geographic division within the country.

Weather and agricultural production data

The weather and agricultural production data sources are:

Precipitation data – for the Brazil sample, precipitation is the monthly accumulated precipitation, in millimeters (mm), obtained from the Global Meteorological Forcing Dataset for land surface modeling (Sheffield et al., 2006) by the Terrestrial Hydrology Research Group at Princeton University[13]. The data was aggregated at the municipality level per month.

For the Southern Brazil sample, precipitation is the weekly accumulated precipitation, in millimeters (mm). Daily precipitation was obtained from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS-v.2) delivered at 5-km resolution (Funk et al., 2015) [14], aggregated to a 7-day (week) period at the municipality level.

---

[12] The municipalities created in 2005, their respective municipality of origin (in parentheses), and their states were: Aroeiras do Itaim (Picos) – Piauí; Figueirão (Camapuã) - Mato Grosso do Sul; Ipiranga do Norte (Tapurah) – Mato Grosso; Itanhangá (Tapurah) – Mato Grosso. The municipalities created after 2010, their respective municipality of origin (in parentheses), and their states were: Pescaria Brava (Laguna) – Santa Catarina; Balneário Rincão (Içara) – Santa Catarina; Mojuí dos Campos (Santarém) – Pará; Pinto Bandeira (Bento Gonçalves) – Rio Grande do Sul; Paraíso das Águas (Costa Rica) – Mato Grosso do Sul.

[13] The weather data is a combination of near-surface meteorological data and other terrestrial modeling systems and is available at the 1, 0.5- and 0.25-degree spatial resolution and at 3-hourly, daily, and monthly temporal resolution. This dataset was compiled at Federal University of Viçosa, Brazil. We thank Dr. Marcelo Braga and the *Instituto de Políticas Publicas e Desenvolvimento Sustentável* (IPPDS) for providing the monthly weather data for the Brazilian municipalities. The original source is: https://hydrology.princeton.edu/data.pgf.php.

[14] The database website is: http://legacy.chg.ucsb.edu/data/chirps/.

Temperature – temperature is the average monthly temperature, in Celsius degrees (°C), from the Global Meteorological Forcing Dataset for land surface modeling at Princeton University.

Agricultural production – municipality level data was recovered for soybeans and corn from the SIDRA website of the Brazilian Institute of Geography and Statistics (or IBGE, abbreviation in Portuguese)[15], measured in yield (kilograms/hectare) and value of production (one thousand 2017 Brazilian *reais* (BRL))[16].

Precipitation drought variables in months

Two meteorological drought indicators were constructed based on monthly accumulated precipitation data for all the corn- and soybeans-producing Brazilian municipalities over the 2002-2016 period. The growing season for soybeans and corn was defined using the agricultural year calendars from the public agency CONAB (*Companhia Nacional de Abastecimento*) available at the state level[17]. The agricultural year is defined as starting at the end of the Winter season of the previous year (t-1) up to the end of the Winter season in the current year (t) (from September(t-1) to September(t)). These calendars revealed that although there is variability across the five Brazilian geographic regions, the most important months for soy and corn production are during the Spring and the beginning of the Summer, from October to January. However, some states have a second "short" season of corn production (*milho safrinha*) during the Summer and beginning of the Fall, implying that corn is also grown from January to May[18]. The main drought variable for the Brazil sample, denoted as DM1, is defined over a growing season from October to January (4 months), but we also calculated a drought indicator for the January-May period (5 months), denoted as DM2, and included it in estimations for corn to account for the second short season. The construction of DM1 and DM2 is based on the Standardized Precipitation (SPrec) considering different averaging periods based on the reference month within the growing season. Standardized Precipitation follows the concept of the Standardized Precipitation Index (SPI) (McKee et al., 1993) and calculates the monthly precipitation deviations from the average of the meteorological period (i.e., 2002-2016). Following Mariano (2019), SPrec is calculated directly

---

[15] IBGE's main data platform from the agricultural censuses and other censuses/surveys is the SIDRA website at https://sidra.ibge.gov.br/home/pnadcm.

[16] The value of production was deflated by the *Índice Geral de Preços - Disponibilidade Interna* (IGP-DI) elaborated by Fundação Getúlio Vargas (FGV).

[17] https://www.conab.gov.br.

[18] The seasons in Brazil are defined as: Spring, from October to December, Summer, from January to March, Fall, from April to June, and Winter, from July to September.

from the data as the monthly z-scores obtained for each month based on a 3-month average including the reference month (1 month before + reference month + 1 month after = 3 months), then we calculate the mean and the standard deviations of these 3-month averages over the years to compute the z-scores as shown below. Considering a 4-month growing season (from October of the previous year to January of the current agricultural year), we first obtained the 3-month average for reference month *i*, municipality *m* and year *t*. Then, for each month *i*, and municipality *m*, we calculated the 3-month average and the standard deviation for 2012-2016 period, denoted as $\mu_{Prec_{3,i,m}}$ and $\sigma_{Prec_{3,i,m}}$, respectively. The $\mathrm{SPrec}_{i,m,t}$ for month i, municipality *m*, and year *t*, was calculated as:

$$\mathrm{SPrec}_{i,m,t} = \frac{Prec_{i,m,t} - \mu_{Prec_{3,i,m}}}{\sigma_{Prec_{3,i,m}}} \tag{4}$$

where $Prec_{i,m,t}$ is month *i* accumulated precipitation for municipality *m* in year *t* and the subscript 3 indicates the 3-month averaging period around the reference month.

Figures A.2, A.3, A.4, A.5 in the Appendix show the average 4-month accumulated z-scores for each Brazilian geographic region for corn- and soybeans-producing municipalities. These figures illustrate the dryer and wetter years related to the average of the 2002-2016 period. We can observe a lot of dryer years for the Northeast region for both groups compared to the other regions. For corn, we see 2005, 2009 and 2013 with the lowest average 4-month accumulated z-cores, while for soybeans 2006 was the lowest. The South region had similar patterns as the Northeast, but 2012 and 2014 were also dry years for the corn- and soybeans-municipalities. The Southeast and Midwest had a dryer year in 2008 and a stronger indication of drought is seen for 2015. Meanwhile, the North region had 2015 and 2016 as the driest years for the corn and soybeans areas. The monthly z-scores were used to classify the drought events at each municipality and year within the categories in Table 1. This classification is from the SPI seminal paper (McKee et al., 1993) and is frequently used in the environmental and climate sciences literature to define droughts.

Table 1: Drought categories.

SPrec values	Drought Category
0 to -0.99	mild drought
-1 to -1.49	moderate drought
-1.5 to -1.99	severe drought
≤-2	extreme drought

Source: McKee et al. (1993, p. 2).

DM1 and DM2 were then constructed by counting the number of months that SPrec was below -1.5 in a year within the growing season, characterizing droughts as severe and/or extreme in the given months. Therefore, for the October-January growing season, DM1 can vary from 0 to 4, where 0 indicates no months with severe droughts and 4 indicates that severe and/or extreme droughts happened throughout the four months of the growing season for a given municipality and year.

Precipitation drought variables in weeks

For the Southern Brazil sample, the drought variables were based on weekly accumulated precipitation data for each municipality over the 2002-2016 period. We also use the Standardized precipitation (SPrec) indicator using different averaging periods for each reference week[19]. The z-scores are obtained for each week based on a 5-week average considering the reference week (2 weeks before + reference week + 2 weeks after = 5 weeks), then we calculated the mean and the standard deviations of these 5-week averages over the years to compute the z-scores. Let $\mu_{Prec_{5,w,m}}$ be the mean and $\sigma_{Prec_{5,w,m}}$ the standard deviation of the 5-week averages for reference week w and municipality *m*, over the 2002-2016 period. Then we compute the $\text{SPrec}_{w,m,t}$ for week *w*, municipality *m* and year *t*, as:

$$\text{SPrec}_{w,m,t} = \frac{Prec_{w,m,t} - \mu_{Prec_{5,w,m}}}{\sigma_{Prec_{5,w,m}}} \tag{5}$$

where $Prec_{w,m,t}$ is week *w* accumulated precipitation for municipality *m* in year *t*. Similarly, we classify droughts as in Table 1 based on the weekly standard values of precipitation.

The drought indicator, denoted as DW, is measured by the number of weeks that SPrec was below -1.5 in a year for the growing season which characterizes severe and/or extreme droughts. For an approximately 18-week growing season, counting from the beginning of October to the end of January, DW can vary from 0 to 18.

Remote sensing sources

The biophysical remote sensing data come from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensors onboard on Terra and Aqua satellites[20]. This data was

---

[19] Standardized precipitation is a simpler version of the Standardized Precipitation Index (SPI) (McKee et al., 1993). The SPI is derived from a standardized normal probability distribution generated from the observed precipitation data (McKee et al., 1993).

[20] https://modis.gsfc.nasa.gov/data/dataprod/index.php.

initially constructed by Mariano (2019). However, in this study the data was aggregated in a panel format dataset where the cross-sectional variation comes from 281 soy-producing municipalities in Southern Brazil and the temporal variation is from the 2002-2016 period. The obtained MODIS products provide a measure of vegetation index, evapotranspiration (ET), and land surface temperature dynamics (LSTD) since these variables are related to crop response to drought. Vegetation indices are common proxies for biomass and plant health, whereas ET and LSTD are related to hydrothermal and thermal stress (Anderson et al., 2016). For vegetation index, the chosen variable was the Normalized Difference Vegetation Index (NDVI) from the product MOD13A1 and MYD13A1 (Terra and Aqua, respectively), which have 500 m spatial resolution and, when combined, achieve a temporal resolution of 8 days. For ET data, which are the input to calculate the Evaporative Stress Index (ESI, Anderson et al., 2007), the MOD16A2 product was used, also delivered in 500 m every eight days. For daytime LSTD, the used product was MOD11A2, delivered in 1000 m spatial resolution every eight days. All the datasets were submitted to time-series smoothening to remove outliers due to cloud contamination and limitation of the sensors.

Biophysical remote sensing data extraction and aggregation

For the construction and aggregation of the remote sensing drought variables, we limit the analysis for the soybean production area in the 281 municipalities in the Southern region of Brazil (Paraná, Santa Catarina and Rio Grande do Sul). Crop maps from the study area were used to extract pixel values from these areas; then, pixel values were averaged for the crop areas within a municipality. The period of reference for the extraction of the remote sensing variables is the plant growing season. The NDVI time-series were used to obtain crop phenology metrics determining the start and peak of the growing season (SOS and POS), which varies between municipalities and years since the sowing window and management are not uniform across the region (Mariano, 2019).

The period between SOS and POS for each year and each municipality was equally divided into three. The first period (stage 1) is related to crop emergence and early stages. The second period (stage 2) is related to biomass accumulation and is characterized by an increase in NDVI. The third period (stage 3) is the reproductive and grain-filling stages when the plant uses water to fill-up grains, but no longer accumulates green biomass (as in the second period). Since the studied agriculture happens in the Summer of the southern hemisphere, the soybean is sowed in one year and harvested next year. Therefore, we refer to years as the harvest (agricultural) year (2016 is the season that started in 2015 and ended in 2016). Soybean crops have different sensitivity to drought during the plant cycle, and the phenology metrics extraction is an approach to deal with the fact that drought has different impacts on soybeans depending on when it occurs (AghaKouchak et al., 2015; Bolton, 2013).

The growing season based on the NDVI goes from mid-October to mid-January, varying from 100 days in Paraná to 86 days in Rio Grande do Sul. Mariano (2019) explored mainly the temporal variation of the remote sensing indicators and used time series analysis to test for how many weeks before the POS there is highest correlation between the indicator and yields for an early detection of drought. The spatial and temporal variation using panel data models is explored in this study, so the data is aggregated at the municipality level for each year during each phenological stage of the plant. This also allows capturing the effects of biophysical conditions at the earlier stages of the plant.

The three biophysical drought indicators used were: Evaporative Stress Index (ESI), Normalized Difference Vegetation Index (NDVI) and Land Surface Temperature Dynamics (LSTD). ESI measures anomalies in evapotranspiration (ET) which indicates water use levels over the land surface. It is a water stress index that compares crop water requirements versus water use and is calculated as $1 - \frac{ET_a}{ET_p}$. ET_a and ET_p are actual and potential evapotranspiration levels, respectively. As an indicator, ESI can provide information on whether higher or lower than normal rates of water use in plant growth is observed in one region for the reference period. Therefore, anomalies are measured in standardized variations of the 8-day ESI from a long-term average condition for a given period. First, the yearly average ESI was calculated based on five 8-day periods, two before and two after the reference 8-day period. So, yearly averages are based on 40-day intervals that comprise the reference 8-day period. Then, the ESI z-score for reference period j, $ESI_{j,m,t}$, is computed based on the historical mean and standard deviation of these averages such as:

$$ESI_{j,m,t} = \frac{ESI'_{j,m,t} - \mu_{ESI\,j,m}}{\sigma_{ESI\,j,m}} \tag{6}$$

where $ESI'_{j,m,t}$ is the 8-day ESI for reference period j in municipality m and year t, $\mu_{ESI\,j,m}$ and $\sigma_{ESI\,j,m}$ are the mean and the standard deviation of the 40-day yearly average ESI for all years.

NDVI is a measurement of plant health which varies from -1 to 1. It is related to how the plant can absorb sunlight radiation which affects the ability of the plant to do photosynthesis. The more the plant is absorbing sunlight during the growing season, the better it will be the plant condition. So, higher levels of NDVI indicate good condition of plant growth. LSTD measures land heat in Celsius degrees. It is the remotely sensed indicator of temperature. We average all biophysical indicators ESI, NDVI and LSTD to reflect each stage of the plant in the econometric estimation.

3.4 The effect of drought on Brazilian agriculture

Panel regression model with fixed effects

Due to the yearly and cross-sectional (municipality-level) variations in the data, the effect of droughts on agriculture can be estimated using panel data models with fixed effects (FE). The meteorological drought indicators are expressed in months and the agricultural outcomes in logarithm of yield (kilograms/hectare) or value of production (2017, 1,000 BRL) for soybeans and corn. The model provides the average effect of one additional month of severe and/or extreme drought[21] for the Brazilian soybeans- and corn-producing municipalities over the 2002-2016 period. The panel data regression model with FE for municipality m, state s and year t, is given by:

$$\mathrm{Y_{mst}} = \gamma_0 + \gamma_1' \mathrm{DM_{mst}} + \gamma_3' \mathrm{T_{mst}} + \lambda_\mathrm{t} + \phi_\mathrm{m} + \psi_\mathrm{s}(\mathrm{t}) + \epsilon_\mathrm{mst} \qquad (7)$$

where $\mathrm{Y_{mst}} \in$ {ln (soybeans yield), ln (corn yield), ln (value of soybeans production), ln (value of corn production)}. $\mathrm{DM_{mst}}$ represents the drought variables, DM1 or DM2, and $\mathrm{T_{mst}}$ is a vector of the monthly average temperatures starting from September of the previous year (t-1) to September of the current agricultural year (t). λ_t represents the year fixed effects (FE) and ϕ_m represents the municipality FE. Following Kuwayama et al. (2019), state-specific trends, $\psi_\mathrm{s}(\mathrm{t})$, were included to account for the positive trends in soybeans and corn agricultural production within each state over the years. $\psi_\mathrm{s}(\mathrm{t})$ is obtained by interacting the time variable with the dummy variables for states. ϵ_it is the random error term. We obtained the standard errors clustered at the state level to account for the potential spatial autocorrelation among the municipalities within the states. The drought effect is measured by the parameter γ_1, indicating the percentage change in the agricultural outcome given one additional month of severe and/or extreme drought.

Brazil Sample descriptive statistics

---

[21] For which the value of the precipitation z-score is below -1.5.

The final Brazil sample contains fifteen years of data from 5,538 municipalities plus 8 AMCs (i.e., 5,546 cross-sectional units and 83,190 observations)[22]. Table 2 shows the descriptive statistics for the variables in the analysis.

Table 2: Descriptive statistics for the Brazil sample (2002-2016).

Variable	Obs.	Mean	Std. Dev.	Min	Max
Soy yield	28,256	2559.359	685.9117	0	12000
Corn yield	77,479	2860.319	2160.158	0	16725
Value of Soy Production	28,257	35168.76	96108.58	0	2114906
Value of Corn Production	77,988	5364.28	19199.78	0	888300.1
DM1 (Oct-Jan)	83,190	.630821	.7790253	0	8
DM2 (Jan-May)	83,190	.8251833	.8709176	0	9
Temp. (Sept(t-1))	83,190	23.90	4.05	11.74	34.81
Temp. (Oct(t-1))	83,190	25.07	3.20	12.37	33.77
Temp. (Nov(t-1))	83,190	25.18	2.69	12.82	33.17
Temp. (Dec(t-1))	83,190	25.53	2.25	12.96	33.21
Temp. (Jan(t))	83,190	25.85	1.75	13.02	32.45
Temp. (Feb(t))	83,190	25.94	1.69	13.49	30.67
Temp. (Mar(t))	83,190	25.47	1.80	13.69	30.53
Temp. (Apr(t))	83,190	24.59	2.54	13.54	30.44
Temp. (May(t))	83,190	22.88	3.68	12.00	31.14
Temp. (Jun(t))	83,190	22.00	4.13	10.54	30.44
Temp. (Jul(t))	83,190	21.82	4.30	9.79	31.84
Temp. (Aug(t))	83,190	22.82	4.15	11.00	33.10
Temp. (Sept(t))	83,190	23.84	4.07	11.74	34.81
Prec. (Sept(t-1))	83,190	66.16	71.55	0	796.91
Prec. (Oct(t-1))	83,190	105.66	83.67	0	735.80
Prec. (Nov(t-1))	83,190	127.15	90.58	0	1181.55
Prec. (Dec(t-1))	83,190	171.46	116.82	0	1380.60
Prec. (Jan(t))	83,190	204.47	125.48	0	1660.03
Prec. (Feb(t))	83,190	160.67	99.56	0	1835.59
Prec. (Mar(t))	83,190	173.09	109.75	0	2691.26
Prec. (Apr(t))	83,190	134.28	95.75	0	826.17
Prec. (May(t))	83,190	100.71	93.18	0	852.21
Prec. (Jun(t))	83,190	68.54	80.88	0	622.06
Prec. (Jul(t))	83,190	64.45	75.86	0	877.61
Prec. (Aug(t))	83,190	46.68	58.88	0	769.94
Prec. (Sept(t))	83,190	63.44	70.25	0	796.91

Note: The ‘Brazil Sample’ is all Brazilian municipalities with non-zero or non-missing values for soy and corn production. Value of production is in $1,000 2017 Brazilian reais (BRL). Yield is in kilograms per hectare. DM1 and DM2 are expressed in months of severe and/or extreme drought calculated from the monthly standardized precipitation for the reference period. Temp. is the monthly average temperature measure in Celsius (°C) and Prec. is the accumulated monthly precipitation in millimeters (mm).

---

[22] The number of observations varies according to each agricultural outcome due to either unavailable data for some municipalities or because some municipalities did not have any soybeans and/or corn production during the period. Unavailable data can be due to misreported values or for confidentiality reasons in municipalities with too few farms.

Notably, corn is grown in almost all Brazilian municipalities while soybeans are concentrated in the Midwest, South, and, more recently, it has been expanding to the North and Northeast regions[23]. Interestingly, the monthly average temperatures did not vary much across the months of the agricultural year, but precipitation had greater levels for January, February, and March. The means of the drought variables show approximately one month in severe and/or extreme droughts, with a one-month standard deviation[24].

Brazil sample results

This section provides the estimation of the effect of droughts for all soybeans- and corn-producing municipalities for the 2002-2016. Table 3 presents the results of the panel data regression FE-estimations. The effects of one additional month of severe and/or extreme drought are found to be negative and statistically at the conventional levels for all agricultural outcomes. Because the agricultural outcomes are in logarithm, we interpret the coefficients as the percent change in yields and value of production due to one additional month of severe and/or extreme drought. On average, for the Brazilian municipalities, soybeans yield decreases by 5.59%, while corn yield decreases by 6.0%, with each additional month of drought.

The same pattern is observed for the value of production with a larger decrease for corn (-7.9%) than soybeans (-3.7%). We can also interpret the effects of the monthly average temperatures as the percent change in the agricultural outcomes but due to a 1-degree Celsius increase in the monthly average temperature. From the months around the growing season (October-January), we observe a large negative effect on the value of soybeans production for the month of October, where an increase in 1-degree Celsius in the average temperature would decrease the value of soybeans production in 8.37%. January has negative and statistically significant coefficients at the standard levels for the value of soybeans production (-9.7%), corn yield (-7.69%) and the value of corn production (-12.8%). Meanwhile, higher average temperatures in November and December increase both agricultural outcomes for corn and soybeans, respectively.

The second short seasons of corn (*milho safrinha*) are considered in the estimations presented in Table 4. We include the drought variable comprising the months from January to

---

[23] The Brazilian agricultural frontier is expanding mainly across the Northeast region. The region named MATOPIBA comprises states from the North (Tocantins (TO)) and Northeast (Maranhão (MA), Piauí (PI) and Bahia (BA)) and is considered the new agricultural frontier.

[24] The maximum values of the drought variables exceed four months for DM1 and five months for DM2 because of the aggregation of some municipalities into AMCs. Thus, one month of severe and/or extreme drought in two municipalities belonging to an AMC means two months of drought in the AMC.

May (DM2) in the models of columns 2 and 4 and compare them with the previous results for corn yield and value of corn production in columns 1 and 3.

Table 3: Effect of droughts on Soybeans and Corn for Brazil, 2002-2016.

	Soybeans (Log)		Corn (Log)	
Variables	Yield	Value of Production	Yield	Value of Production
Drought in months:				
DM1 (Oct-Jan)	-0.0559***	-0.0365	-0.0599***	-0.0787***
	(0.0147)	(0.0213)	(0.0202)	(0.0197)
Mean Temperature:				
Sept(t-1)	-0.0101	0.0393	0.0274*	0.0630**
	(0.0179)	(0.0332)	(0.0154)	(0.0285)
Oct(t-1)	-0.00278	-0.0837**	-0.0253	-0.0215
	(0.0182)	(0.0393)	(0.0171)	(0.0268)
Nov(t-1)	-0.0169	-0.0420	0.104**	0.149***
	(0.0106)	(0.0361)	(0.0393)	(0.0413)
Dec(t-1)	0.0416*	0.151***	0.0281	0.0252
	(0.0203)	(0.0340)	(0.0308)	(0.0392)
Jan(t)	-0.0650	-0.0977**	-0.0769***	-0.128***
	(0.0382)	(0.0463)	(0.0244)	(0.0306)
Feb(t)	-0.0103	-0.0405	-0.0665***	-0.0872***
	(0.0269)	(0.0388)	(0.0226)	(0.0302)
Mar(t)	0.0231	0.226***	-0.0118	0.0158
	(0.0148)	(0.0428)	(0.0187)	(0.0344)
Apr(t)	0.0835***	0.178***	-0.0695	-0.125*
	(0.0202)	(0.0290)	(0.0454)	(0.0649)
May(t)	0.0360	-0.00116	0.0421	0.0590*
	(0.0258)	(0.0417)	(0.0264)	(0.0325)
Jun(t)	-0.0605***	-0.109***	-0.0533**	-0.0584
	(0.0141)	(0.0328)	(0.0234)	(0.0387)
Jul(t)	-0.00262	-0.0770**	-0.0235	-0.0140
	(0.00933)	(0.0309)	(0.0164)	(0.0248)
Aug(t)	-0.0669***	-0.0761***	-0.00151	-0.0179
	(0.0186)	(0.0204)	(0.0142)	(0.0211)
Sept(t)	0.0564**	0.0856**	0.0952***	0.142***
	(0.0239)	(0.0305)	(0.0262)	(0.0331)
Constant	7.347***	3.076*	8.316***	6.922***
	(0.509)	(1.490)	(1.562)	(1.775)
Obs.	28,200	28,189	77,478	77,413
R^2 w	0.376	0.333	0.105	0.109
R^2 b	0.00310	0.000788	0.130	0.0309
R^2 o	0.193	0.0272	0.0424	0.00878
R^2 adj.	0.375	0.332	0.104	0.108

Note: Standard errors in parentheses are clustered at the state level. All models include municipality-FE, year-FE, and state-specific time trends. Significance levels are denoted as: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

On average, an increase in one month of severe and/or extreme drought would reduce corn yields by 4.7% during the second growing season of corn. The DM2 coefficient for the value of corn production is not statistically significant at any conventional level. One should note that the reduction in corn yield by an increase in one month of drought during the first growing season goes from 6% to 4.4% with the inclusion of the second drought variable. The statistical fit of both outcomes for corn improves when accounting for the second season drought effect in the estimations, as shown by the R^2-adjusted and the information criteria AIC and BIC (bold), therefore, the meaningful drought effect on the value of corn production must be the -6.7 percent (Table 4, column 4).

Table 4: Effects of droughts on Corn with and without considering a second season in the measurement of drought in months, Brazil, 2002-2016

	Corn Yield		Value of Corn Production	
Variables	(1)	(2)	(3)	(4)
Drought in months:				
DM1 (Oct-Jan)	-0.0599***	-0.0444***	-0.0787***	-0.0677***
	(0.0202)	(0.0157)	(0.0197)	(0.0175)
DM2 (Jan-May)		-0.0468***		-0.0333
		(0.0131)		(0.0202)
Constant	8.316***	8.302***	6.922***	6.912***
	(1.562)	(1.560)	(1.775)	(1.793)
Obs.	77,478	77,478	77,413	77,413
R^2 w	0.105	0.110	0.109	0.110
R^2 b	0.130	0.150	0.0309	0.0329
R^2 o	0.0424	0.0381	0.00878	0.00860
R^2 adj.	0.104	**0.109**	0.108	**0.109**
AIC	96499.3	**96077.2**	186392.1	**186325.7**
BIC	96740.0	**96317.9**	186632.8	**186566.4**

Note: Standard errors in parentheses are clustered at the state level. All models include municipality-FE, year-FE, state-specific time trends, and the monthly average temperature variables from September(t-1) to September(t). Significance levels are denoted as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Regional Analysis

The results of the panel data regression FE-estimations for the five Brazilian regions (i.e., North, Northeast, Southeast, South and Midwest) and the four agricultural outcomes are shown in the Appendix. The heterogeneous effects and significance levels indicate the importance of the drought variables constructed from the monthly observed precipitation data for each region and crop separately. For corn yield, negative and statistically significant drought effects are found for most regions, except for the North and Midwest. For the Northeast, an increase in one month of severe and/or extreme drought would decrease corn yield by 2.68%. The reduction for the Southeast region is of 3.25%, and the largest effect was observed for the South (-6%). The only statistically significant effect at the conventional levels for the value of corn production was for the Southeast, with an estimate of -8%. The estimation for corn also included the DM2 variable to account for the second growing season. The results indicate statistically significant effects for the South (-2.59%) and Midwest (-1.73%) on corn yield, and for the Northeast (-8%) and South (-4.4%) on the value of production.

For Soybeans yield, statistically significant coefficients at the standard levels were found for the Northeast (-7.6%) and Midwest (-2.75%), while the soybeans value of production was significant only for the Midwest (-5.67%). The statistically significant soybeans effects on the Midwest are not surprising because the states of Mato Grosso, Goiás, and Mato Grosso do Sul are the first, fourth, and fifth largest producers of soybeans in Brazil, respectively (Censo Agrícola, 2017). Meanwhile, the non-statistically significant effects for soybeans yield and value of production for the South were not expected because the second and third largest producers of soybeans are the states of Rio Grande do Sul and Paraná, both in this region. Although the estimated effect on soybeans yield is found to be -3.42% for the South, it was not statistically significant at any standard level. The coefficient on the value of production is positive but also not statistically significant[25].

3.5 The effect of drought on soybeans production in Southern Brazil

Despite having the advantage of the FE-estimations due to the panel format of the data, it is important to consider that, so far, we have not claimed for causality regarding the drought effects because we are unsure of the variables that would potentially affect the agricultural outcomes but were not included in the previous models. This section goes one step further on this discussion by assessing a specifically defined soybeans production region in Southern municipalities and including biophysical variables to estimate the effect of droughts on yield and value of production. This analysis complements the previous analysis in the sense that to estimate the causal effect of drought, it would require more information on the relationship among the plant, the soil, and the atmosphere. Furthermore, it highlights the importance of using satellite imagery and remote sensing technologies to obtain high-resolution temporal and spatial data that capture biophysical characteristics at different stages within the growing season, allowing for more precise estimation of the effect of droughts in agriculture.

Panel regression model with fixed effects and remote sensing information

For the 281 soybeans-producing municipalities in the Southern Brazil sample, the following panel data regression FE-model was estimated:

$$\mathrm{Y_{mst}} = \theta_0 + \theta_1' \mathrm{DW_{mst}} + \theta_2' \mathrm{Temp_{mst}} + \theta_3' \mathrm{RS_{mst}} + \lambda_t + \phi_m + u_{mst} \qquad (8)$$

where $\mathrm{Y_{mst}}$ is the log of soybeans yield or the log of the value of soybeans production in municipality m, state s and year t. $\mathrm{DW_{mst}}$ is the drought indicator expressed in weeks, $\mathrm{Temp_{mst}}$

---

[25] We re-estimated the models without including the state-specific trends and the results did not change.

is the monthly average temperatures covering the agricultural year (from September(t-1) to September(t))[26]. RS_{mst} includes the remote sensing (RS) variables (ESI, NDVI and LTSD) constructed to reflect each of the three stages of the phenological growing season. λ_t are year-FE and ϕ_m are municipality-FE. u_{mst} is the random error. We are mainly interested in the drought effect measured by the parameter θ_1 in the model, which is interpreted as the percentage response on yields due to one additional week of severe or extreme drought in a year. We estimate different models by iteratively including the different RS variables. Finally, we compare the statistical properties of the models and the magnitudes of the drought effect parameters to decide whether the inclusion of these new variables improve the agricultural drought effect estimation.

Southern Brazil sample descriptive statistics

Table 5 shows the descriptive statistics for the main variables in the Southern Brazil sample. We observe great variability from the mean in the ESI values and the same is observed for NDVI. As expected, NDVI has no negative values because it is related to the plant vegetation. The closer to the end of the growing season the greater the NDVI values.

Table 5: Summary statistics – 281 municipalities in Southern Brazil (2002-2016)

Variable	Obs	Mean	Std. Dev.(s.d.)	Min	Max
Soy yield	4,212	2,684.55	707.21	146	6,988
Value of Soy Production	4,212	57,838.22	57,932.59	625.49	533,542.7
DW	4,212	4.87	2.17	0	11
ESI1	4,212	0.01	0.87	-2.1	2.83
ESI2	4,212	0.07	0.83	-2.38	2.68
ESI3	4,212	0.08	0.79	-1.71	2.81
NDVI1	4,212	0.46	0.08	0.26	0.67
NDVI2	4,212	0.49	0.08	0.3	0.77
NDVI3	4,212	0.66	0.08	0.35	0.87
LSTD1	4,212	34.71	4.09	24.66	46
LSTD2	4,212	36.36	3.20	27.41	46.89
LSTD3	4,212	32.73	3.12	26.24	44.78

Note: The Southern Brazil municipalities comprise the soy production area in the states of Paraná, Santa Catarina and Rio Grande do Sul. Soy yield is measured in kilograms per hectare and Value of Soy Production in \$1,000 2017 *Brazilian reais* (BRL). DW is the drought indicator based on standardized precipitation measured in weeks. ESI1 is the average *Evaporative Stress Index* (z-score) based on 8-day periods during stage 1 of the plant. ESI2 and ESI3 are the respective average z-scores for stages 2 and 3. NDVI1 is the average *Normalized Difference Vegetation Index* (this measure varies between -1 to 1) for stage 1, while NDVI2 and NDVI3 are the stage 2 and 3 respective averages. LSTD1, LSTD2 and LSTD2 are the average *Land Surface Temperature Dynamics* (measured in °C) for stage 1, 2 and 3, respectively.

---

[26] We include monthly average temperature for the agricultural year, defined from September of the previous year to September of the current year, to capture long-term effects of weather on yields that happen before and after the phenological growing season.

Land surface temperature (LSTD) has significant large variability over the growing season. On average we observe 5 of the 18 weeks of the growing season with severe or extreme drought and a maximum of 11 weeks. Figure A.6 in the Appendix indicate the potential wet and dry years measured by the 18-week accumulated precipitation z-scores for the three states in Southern Brazil. In Paraná, the state with the largest level of soybeans production in the region, we can see dry years in 2004, 2006, 2008, 2012 and 2014. Similar patterns are seen for the other two states, where 2012 and 2014 were the driest years for the whole region. Figure A.7 indicate where the 281 municipalities are in Southern Brazil and the total number of weeks that each municipality was in drought for the whole period. It shows that around forty-five municipalities were in severe and/or extreme drought for a period equivalent to five growing seasons (approximately ninety weeks) during the 2012-2016 period. Figures A.8 and A.9 illustrate the negative correlation between soybeans yield and the weeks in drought comparing a dry (2012) and a wet year (2016), respectively. We can see that only 17 municipalities achieved yields equal or above 3,000 kilograms per hectare during the dry year as opposed to 188 municipalities in the wet year.

Southern Brazil sample results

This section presents the panel FE-regression results of different models obtained from the estimation of equation 8. Table 6 shows eight models where we iteratively add and remove the RS variables. The baseline model (column 1) includes only DW and the monthly average temperatures as explanatory variables. Models (2) to (4) include each biophysical indicator separately, but all models contain the drought indicator constructed from observed precipitation (DW) and the temperature variables. From models (5) to (7), we combine two RS indicators at a time. Finally, model (8) includes all the meteorological and the RS variables in the same estimation. We can compare the models statistically using the adjusted R^2 (R^2 adj.), the Akaike information criterion (AIC) and Bayesian information criterion (BIC). Models (6) and (8) are the best according to the R^2 adj while the AIC favors model (8), and the BIC, model (1)[27]. Given the significance of the parameters in model (8) (7 significant parameters out of 11) and the statistical support of the R^2 adj and AIC, we choose the full specification for interpretation.

Each indicator is an important determinant of yields, but the response is different in each phenological stage. For example, ESI has a positive significant effect at 1% on yields for stage 1 but no significant effect in the other stages. The direction of the effect of each RS

---

[27] For each criterion, the chosen model is in bold in Table 3.

indicator is hard to be established given the endogeneity of the biophysical indicators. Each indicator depends on the values of the other, hence their effects on yields must be determined empirically for different regions and different periods. Unfortunately, it is still hard to claim for causality using these models because the biophysical variables are by nature endogenous in determining agricultural outcomes. Although studying the specific effects of the RS indicators is beyond the scope of this paper, this is a common research agenda for the natural and remote sensing sciences (Anderson et al., 2016). On the other hand, there are still few studies that explore these variables to quantify the overall effect of droughts in the agricultural economic literature.

Interestingly, ESI anomalies are found to be the main indicator that captures early detection of droughts in agriculture in the remote sensing studies of Anderson et al. (2016) and Mariano (2019). This aligns with our results regarding the statistically significance of the ESI indicator in the first stage of the plant. NDVI is significant at 10% with a negative effect on yields for the second stage and a positive effect significant at 5% on the third stage. NDVI is a vegetation index so we would expect the third stage to be predominant in affecting yields which is what the results show; however, it is not clear why it affects negatively yields on the second stage. We observe that land surface temperature (LTSD) affects yields positively in stages 1 and 3. It is not surprising given that these stages require more solar radiation for photosynthesis (stage 1) and plant filling (stage 3).

Comparing the models in Table 6 shows that the drought coefficient based on DW seems to be smaller in the complete model as compared to the model with only the meteorological variables. With the inclusion of the biophysical information, we observe a decrease in absolute value from -1% to -0.7%. Since we chose to interpret the complete model, we conclude that the average effect of one additional week of severe and/or extreme drought in these 281 Southern municipalities is a 0.7% reduction in soy yields. Kuwayama et al. (2019) found values ranging from -0.1% to -1.2% for corn and soy yields for the U.S. Also, we can observe that each RS variable affects yields differently in each stage of the plant. This is relevant for studies that try to measure the causal effect of droughts because omitting these variables may bias the estimates. From the regional analysis presented in the previous section, the one-month drought coefficient for the Southern region was much larger (-3.42%) but not statistically significant. Although this coefficient is not directly comparable to the one-week drought coefficient of -0.7 percent for the 281 municipalities in Southern Brazil, we can observe that defining a specific and more homogeneous soybeans production region and using other determinants of agricultural yields such as the biophysical variables provides a lower more precise average drought effect (also from comparing columns 1-8 in Table 6).

Table 6: Effect of drought on log of soybean yield using remote sensing data – Southern Brazil.

	Soybean yield (Log)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Var.	DW + Temp	ESI	NDVI	LTSD	ESI + NDVI	ESI + LSTD	NDVI +LTSD	All
DW	-.009***	-.008***	-.009***	-.009***	-.008***	-.008***	-.008***	-.007***
	(.002)	(.002)	(.002)	(.002)	(.0027)	(.0028)	(.0027)	(.002)
ESI1		.03***			.03***	.04***		.04***
		(.008)			(.008)	(.008)		(.008)
ESI2		-.0043			-.0007	-.00030		.00012
		(.007)			(.007)	(.0077)		(.007)
ESI3		-.0117			-.00848	-.00383		-.0031
		(.007)			(.0074)	(.0080)		(.008)
NDVI1			.136		.127		.169	0193
			(.132)		(.134)		(.136)	(.140)
NDVI2			-.2***		-.24***		-.22	-.25*
			(.087)		(.0896)		(.135)	(.134)
NDVI3			.0725		.0695		.236**	.227**
			(.082)		(.0841)		(.0949)	(.096)
LSTD1				.0062*		.01***	.00655*	.01***
				(.003)		(.0035)	(.00360)	(.003)
LSTD2				.00273		.00196	-.000279	-.0012
				(.001)		(.0019)	(.00307)	(.003)
LSTD3				.007***		.008***	.0108***	.011***
				(.002)		(.0025)	(.00261)	(.002)
Const.	5.96***	5.9***	6.1***	5.7***	5.97***	5.31***	5.66***	5.3***
	(.600)	(.580)	(.596)	(.608)	(.582)	(.569)	(.599)	(.569)
Obs.	4,212	4,212	4,212	4,212	4,212	4,212	4,212	4,212
R^2 w	.661	.663	.662	.663	.663	.665	.663	.665
R^2 b	.0026	.00019	.0003	.00012	.0003	.0007	.0023	.0041
R^2 o	.403	.401	.398	.365	.394	.349	.369	.353
R^2 adj.	.659	.660	.659	.660	.661	**.662**	.660	**.662**
AIC	-1666.9	-1682	-1668	-1679	-1682	-1701	-1680	**-1702**
BIC	**-1489**	-1485	-1471	-1482	-1466	-1485	-1464	-1467

Note: The first row indicates the dependent variable, and the third row indicates the drought indicators used in each model. All models were estimated with the monthly average temperature (Temp) from September(t-1) to September(t). Robust standard errors are in parentheses. All models include municipality-FE and year-FE. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7 shows the results for the log of the value of soybeans production. One more week of drought during the growing season has an average effect of -1.6% on the value of production. This result is contrary to the evidence in Kuwayama et al. (2019) where no effects on farm income was found and, in our case, the effect on the value of production is even larger than the effect on yields.

Table 7: Effect of drought on log of value of soy production using remote sensing data – Southern sample.

	Value of soybeans production (Log)	
	(9)	(10)
Variables	DW+Temp	All
DW	-0.0194***	-0.0159***
	(0.00405)	(0.00418)
ESI1		0.0470***
		(0.0119)
ESI2		-0.0151
		(0.0114)
ESI3		0.0334***
		(0.0122)
NDVI1		-0.536**
		(0.211)
NDVI2		-0.726***
		(0.214)
NDVI3		0.144
		(0.132)
LSTD1		0.0173***
		(0.00552)
LSTD2		-0.00292
		(0.00414)
LSTD3		0.0166***
		(0.00426)
Constant	9.080***	8.821***
	(0.954)	(0.884)
Obs.	4212	4212
R^2 w	0.667	0.677
R^2 b	0.00118	0.00931
R^2 o	0.148	0.153
R^2 adj.	0.664	0.674
AIC	1038.0	922.0
BIC	1215.6	1156.8

Note: The first row indicates the dependent variable, and the row indicates the drought indicators used in each model. All models were estimated with the monthly average temperature (Temp) from September(t-1) to September(t). Robust standard errors are in parentheses. All models include municipality-FE and year-FE. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.6 Conclusion

This study estimated the effect of droughts using remote sensing information with Brazilian data. First, we used all Brazilian municipalities that produced corn and soybeans for the 2002-2016 period to estimate a one-month effect of drought measured by the standardized precipitation. In general, corn was more affected to drought events than soybeans. One additional month of drought reduced agricultural yields, on average, by 6 percent for corn and 5.5 percent for soybeans. The effects on value of production were also greater for corn. Second, we estimated models adding remote sensing information from a soybeans production area in Southern Brazil. The effect of droughts is better estimated with the inclusion of these variables considering the statistical properties of the models. One additional week of drought reduced soybeans yield, on average, by 0.7 percent. The one-week effect on the value of soybeans production was greater, approximately -1.6 percent.

From the motivating theoretical framework, we observed that a key parameter determining the economic value of remote sensing information is the correlation between the drought indicator and agricultural yields. In this sense, by improving the statistical properties of the yield response to the meteorological drought variables and hence the estimation of the drought coefficient, remote sensing information can improve decision-making when the publicly available drought indicator obtained from observed precipitation is used to determine the allocation of government assistance to farms facing severe drought events or the adoption of ex-ante strategies such as irrigation to decrease the dependence on the climate and to alleviate the agricultural impacts from the exposure to droughts.

Furthermore, we found that when considering a specific region such as Southern Brazil where remote sensing information allowed to identify the exact production area and to obtain biophysical variables affecting agriculture, we can estimate more reliable drought effects. For this region, remote sensing information also allowed to construct a drought measurement in weeks instead of months; remember that the drought effect for the Southern region was not statistically significant when the meteorological drought variable was measured in months. This shows that more information can help obtain unbiased drought effects and to predict agricultural losses more precisely. Although we use panel data models with fixed effects, which help with the endogeneity issues in the regressions, note that because the biophysical variables may be endogenously determined, we are still careful to interpret the drought coefficient as a causal effect for the soybeans production region in Southern Brazil. On the other hand, these variables are important determinants of yields and should still be considered, hence a statistically significant effect for the meteorological indicator in a panel data model with fixed effects and with remote sensing information can be a good approximation of the true marginal effect of one additional period of severe and/or extreme drought in agriculture as provided in this study.

Finally, we have moved forward on the discussion to expand the Northeast Drought Monitor (NDM) in Brazil by studying the agricultural effect of droughts in Southern Brazil, which is also a region severely affected by these events. Government, public and private agents' efforts to combat the effects of droughts ex-ante have increased the interest to invest in monitoring and long-term solutions such as the adoption of subsidized irrigation plans in the country. Investment in satellite imagery technologies and in the remote sensing sciences could significantly change the country's strategies to adapt to the more frequent droughts in agricultural regions all over Brazil.

As a research agenda, studying the effect of remote sensing information on the decisions themselves, instead of using it to improve the correlation between the meteorological drought indicator and the true drought event (proxied by the agricultural outcomes), can provide the economic value of remote sensing information, which could then be compared with the cost of investing on (and adopting) the remote sensing technologies. One should first consider how the decisions would change when remote sensing information is available and compare with the counterfactual case (i.e., no information); this is an established exercise/research in the earth and energy sciences (see, for example, Eidsvik, Mukerji, and Bhattacharjya, 2015); however, there is a large opportunity to expand the use of the value of information (VOI) theory in agricultural economics studies as the access to these technologies increases and as the data obtained can be combined with decisions regarding the allocation of resources.

REFERENCES

AGHAKOUCHAK, A.; FARAHMAND, A.; MELTON, F. S. S.; TEIXEIRA, J.; ANDERSON, M. C. C.; WARDLOW, B. D. D.; HAIN, C. R. R. Remote sensing of drought: Progress, challenges, and opportunities. **Reviews of Geophysics**, 53(2), 2015, p. 452–480.

ALVALÁ, R.; CUNHA, A. P.; BRITO, S.; SELUCHI, M. E.; MARENGO, J. A.; MORAES, O.; CARVALHO, M. Drought monitoring in the Brazilian Semiarid region. **Annals of the Brazilian Academy of Sciences**, 91 (Supp 1): e20170209, 2019.

ANDERSON, M. C.; NORMAN, J. M.; MECIKALSKI, J. R.; OTKIN, J.; KUSTAS, W. P. A climatological study of evapotranspiration and moisture stress across the continental

United States based on thermal remote sensing: 1. Model formulation. **Journal of Geophysical Research**, 112(D10), D10117, 2007.

ANDERSON, M. C.; ZOLIN, C. A.; SENTELHAS, P. C.; HAIN, C. R.; SEMMENS, K.; TUGRUL YILMAZ, M.; …TETRAULT, R. The Evaporative Stress Index as an indicator of agricultural drought in Brazil: An assessment based on crop yield impacts. **Remote Sensing of Environment**, 174, 2016, p. 82–99.

AWANGE, J. L.; MPELASOKA, F.; GONCALVES, R. M. When every drop counts: Analysis of Droughts in Brazil for the 1901-2013 period. **Science of the Total Environment**, 566-567:1472–88, 2016.

BATTISTI, R.; SENTELHAS, P. C.; BOOTE, K. J.; GIL, G. M.; FARIAS, J. R.; BASSO, J. Assessment of soybean yield with altered water-related genetic improvement traits under climate change in Southern Brazil. **European Journal of Agronomy**, 83, 2017, p. 1–14.

BEHRANGI A; FETZER, E. J.; GRANGER, S. L. Early detection of drought onset using near surface temperature and humidity observed from space. **International Journal of Remote Sensing**, 37(16), 2016, p. 3911–23.

BERNKNOPF, R. L.; BROOKSHIRE, D. S.; MCKEE, M.; SOLLER, D. R. Estimating the Social Value of Geologic Map Information: A Regulatory Application. **Journal of Environmental Economics and Management**, 1997, 32: 204-218.

BERNKNOPF, R.; BROOKSHIRE, D.; KUWAYAMA, Y.; MACAULEY, M.; RODELL, M; THOMPSON, A.; VAIL, P.; ZAITCHIK, B. The Value of Remotely Sensed Information: The Case of a GRACE Enhanced Drought Severity Index. **Weather, Climate, and Society**, 10 (1), 2018, p. 187–203.

BOLTON, D. K.; FRIEDL, M. A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. **Agricultural and Forest Meteorology**, 173, 2013, p. 74–84.

CUI, X. Climate change and adaptation in agriculture: Evidence from US cropping patterns. **Journal of Environmental Economics and Management**, 2020, v. 101, 102306. doi.org/10.1016/j.jeem.2020.102306

CUNHA, A. P.; TOMASELLA, J.; RIBEIRO-NETO, G. G.; BROWN, M.; GARCIA, S. R.; BRITO, S. B.; CARVALHO, M. A. Changes in the spatial–temporal patterns of droughts in the Brazilian Northeast. **Atmospheric Science Letters**, 19: e855, 2018.

DAI, A. G. Increasing drought under global warming in observations and models. **Nature Climate Change**, 2013, 3(1): 52–58. https://doi.org/10.1038/nclimate1633

DE NYS, E.; ENGLE, N. L.; AND MAGALHÃES, A. R. (eds.). **Drought in Brazil: Proactive Management and Policy**. CRC Press, Boca Raton, FL, 2017.

DESCHÊNES, O.; GREENSTONE, M. The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather. **American Economic Review,** 2007, 97 (1): 354–85.

DING, Y.; SCHOENGOLD, K; TADESSE, T. The Impact of Weather Extremes on Agricultural Production Methods: Does Drought Increase Adoption of Conservation Tillage Practices? **Journal of Agricultural and Resource Economics**, 2009, 34 (3): 395–411.

EIDSVIK, J.; MUKERJI, T.; BHATTACHARJYA; D. **Value of Information in the Earth Sciences: Integrating Spatial Modeling and Decision Analysis**. Cambridge: Cambridge University Press. 2015. doi:10.1017/CBO9781139628785

FISHER, A. C.; HANEMANN, W. M.; ROBERTS, M. J.; SCHLENKER, W. The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather: Comment. **American Economic Review**, 2012, 102 (7): 3749-60.

FUNK, C.; PETERSON, P; LANDSFELD, M.; PEDREROS, D.; VERDIN, J.; SHUKLA, S.; …MICHAELSEN, J. The climate hazards infrared precipitation with stations - a new environmental record for monitoring extremes. **Scientific Data**, 2015, 2, 150066. https://doi.org/10.1038/sdata.2015.66

GARCÍA-SUÁREZ, F.; FULGINITI, L. E.; PERRIN, R. K. What is the use value of irrigation water from the High Plains aquifer? **American Journal of Agricultural Economics**, 2019, 101 (2): 455-466.

GUTIÉRREZ, A. P. A.; ENGLE, N. L.; DE NYS, E.; MOLEJÓN, C.; MARTINS, E. S. Drought preparedness in Brazil. **Weather and Climate Extremes**, 2014, 3: 95–106. https://doi.org/10.1016/j.wace.2013.12.001

HORNBECK, R.; KESKIN, P. The Historically Evolving Impact of the Ogallala Aquifer: Agricultural Adaptation to Groundwater and Drought. **American Economic Journal: Applied Economics**, 2014, 6, (1): 190–219.

IBGE, Censo Agropecuário 2017. Accessed in March 2020. Available at: https://censoagro2017.ibge.gov.br/

KUWAYAMA, Y.; THOMPSON, A.; BERNKNOPF, R.; ZAITCHIK, B.; VAIL, P. Estimating the Impact of Drought on Agriculture Using the U.S. Drought Monitor, **American Journal of Agricultural Economics**, 101(1), 2019, p. 193–210.

LAWRENCE, D. **The Economic Value of Information**. Springer-Verlag, 1999, 393 pp.

MACAULEY, M. K. The value of information: Measuring the contribution of space-derived earth science data to resource management. **Space Policy**, 2006, 22: 274–282.

MANETA, M. P.; TORRES, M. O.; WALLENDER, W. W.; VOSTI, S.; HOWITT, R; RODRIGUES, L; BASSOI, L. H.; PANDAY, S. A spatially distributed hydroeconomic model to assess the effects of drought on land use, farm profits, and agricultural employment, **Water Resources Research**, 2009, 45, W11412. doi:10.1029/2008WR007534

MARIANO, D. **Drought impacts assessment in Brazil - a remote sensing approach**. A DISSERTATION. *Dissertations & Theses in Natural Resources*. 291. University of Nebraska-Lincoln, 2019. Available at: https://digitalcommons.unl.edu/natresdiss/291/

MARIANO, D.; SANTOS, C. A. C.; WARDLOW, B. D.; ANDERSON, M. C.; SCHILTMEYER, A. V.; TADESSE, T.; SVOBODA, M. D. Use of remote sensing indicators to assess effects of drought and human-induced land degradation on ecosystem health in Northeastern Brazil. **Remote Sensing of Environment**, 2018, 213: 129–143.

MCKEE, T.B.; DOESKEN, N. J.; KLEIST, J. **The relationship of drought frequency and duration to time scales.** In: Eighth Conference on Applied Climatology. Anaheim, California: American Meteorological Society, 1993, p. 17–22.

SCHLENKER, W.; HANEMANN, W.M.; FISHER, A. C. Will U.S. Agriculture Really Benefit from Global Warming? Accounting for Irrigation in the Hedonic Approach. **American Economic Review**, 2005, 95 (1): 395–406.

SCHOENGOLD, K.; DING, Y.; HEADLEE, R. The Impact of Ad hoc Disaster and Crop Insurance Programs on the Use of Risk Reducing Conservation Tillage Practices, **American Journal of Agricultural Economics**, 2015, 97(3): 897-919.

SHEFFIELD, J.; GOTETI, G.; WOOD, E. F. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. **Journal of Climate**, 19 (13), 2006, p. 3088–3111.

SILVA, F.; FULGINITI, L.; PERRIN, R. K.; SCHOENGOLD, K. The Effects of Irrigation and Climate on the High Plains Aquifer: A County-Level Econometric Analysis. **Journal of the American Water Resources Association,** 2019, 55 (5): 1– 17.

WILHITE, D. A. **Drought as a natural hazard: Concepts and definitions**. In Drought: A Global Assessment (pp. 3–18), 2000.

APPENDIX

A.1. Brazilian geographic region division.

Figure A.1. Brazilian geographic region division.

A.2. Figures of the 4-month accumulated z-scores for the corn- and soybeans municipalities per geographic region.

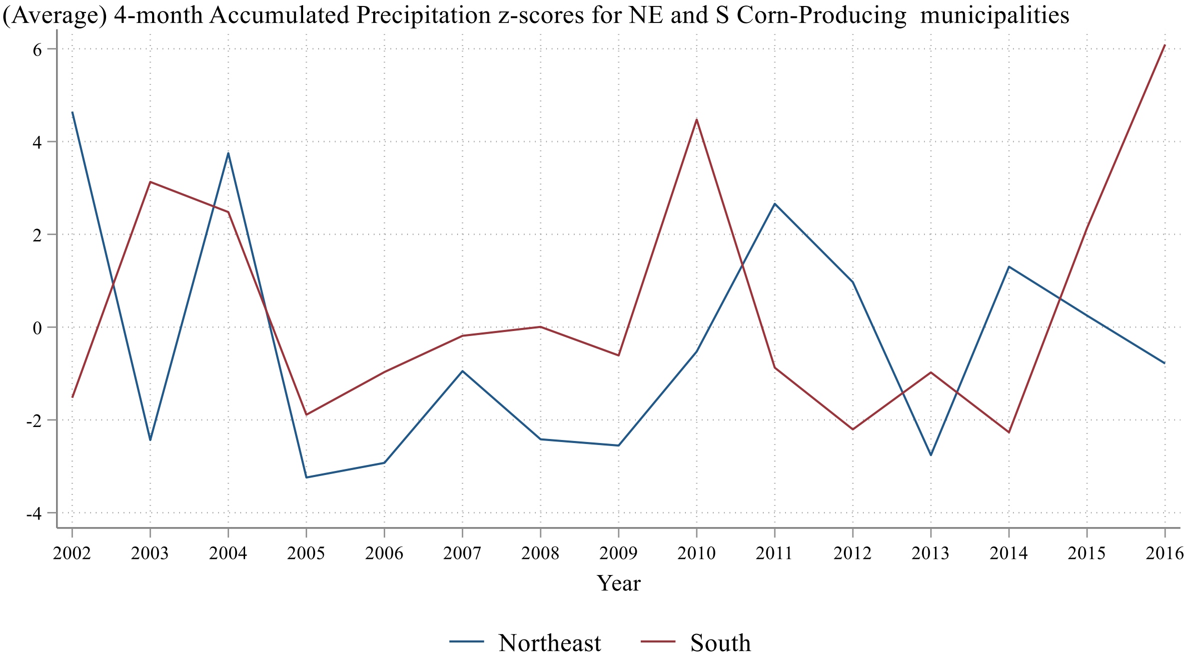

Figure A.2. (Average) 4-month accumulated precipitation z-scores for Northeast (NE) and South (S) Corn-producing municipalities.

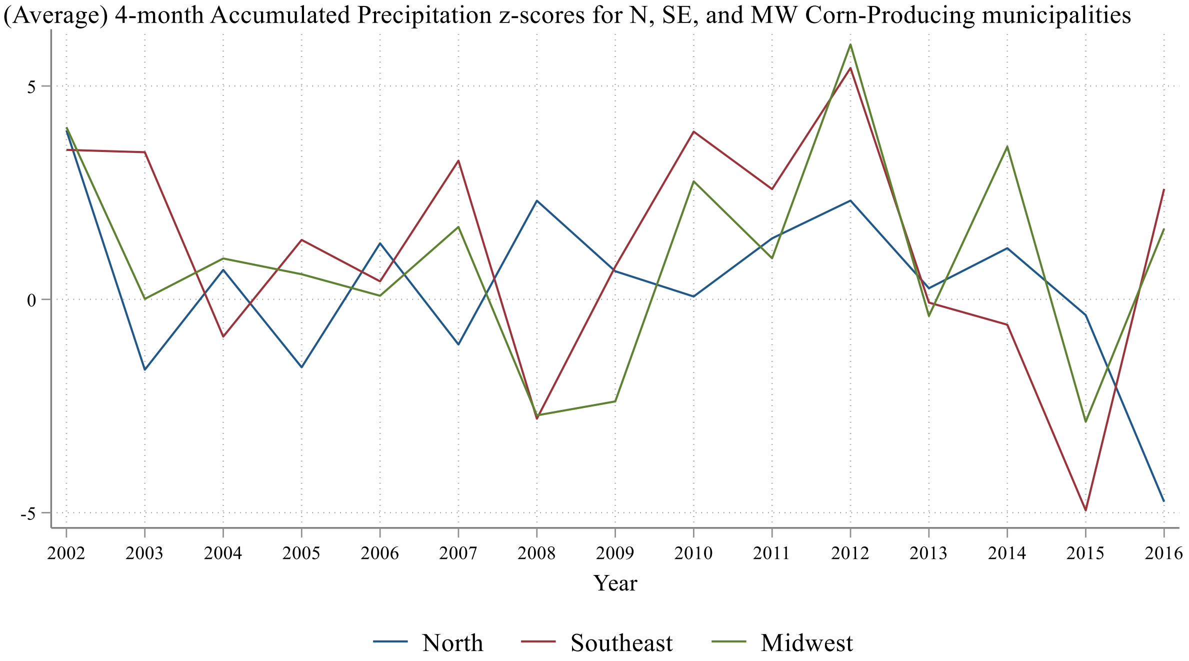

Figure A.3. (Average) 4-month accumulated precipitation z-scores for North (N), Southeast (SE), and Midwest (MW) Corn-producing municipalities.

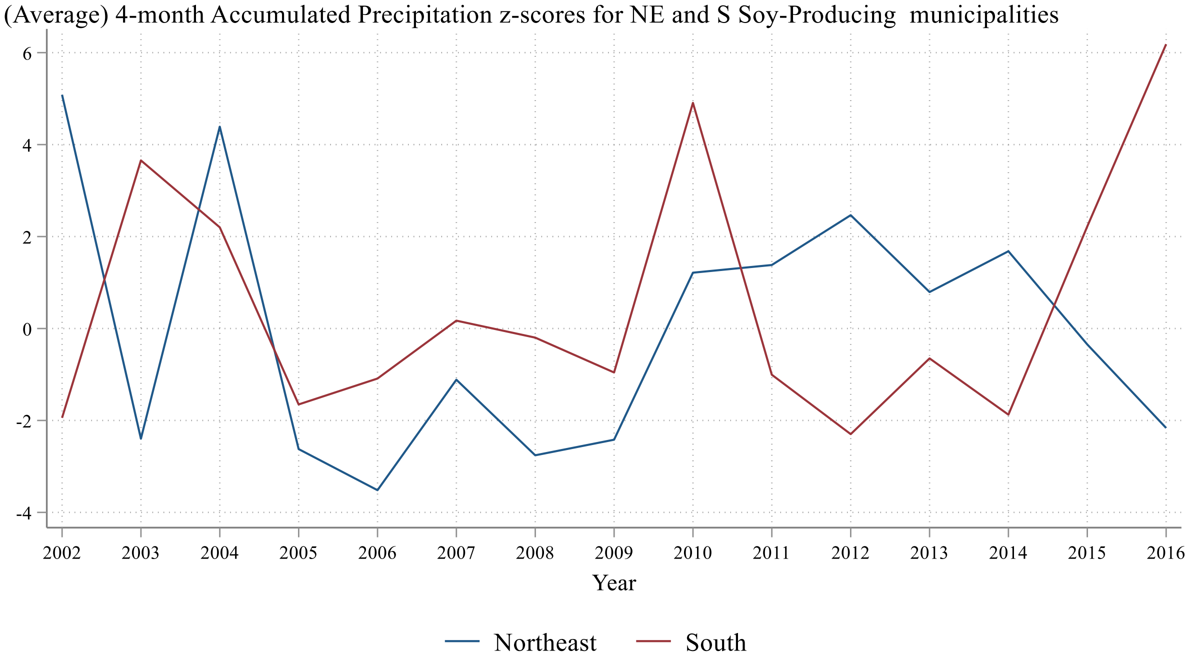

Figure A.4. (Average) 4-month accumulated precipitation z-scores for Northeast (NE) and South (S) Soybeans-producing municipalities.

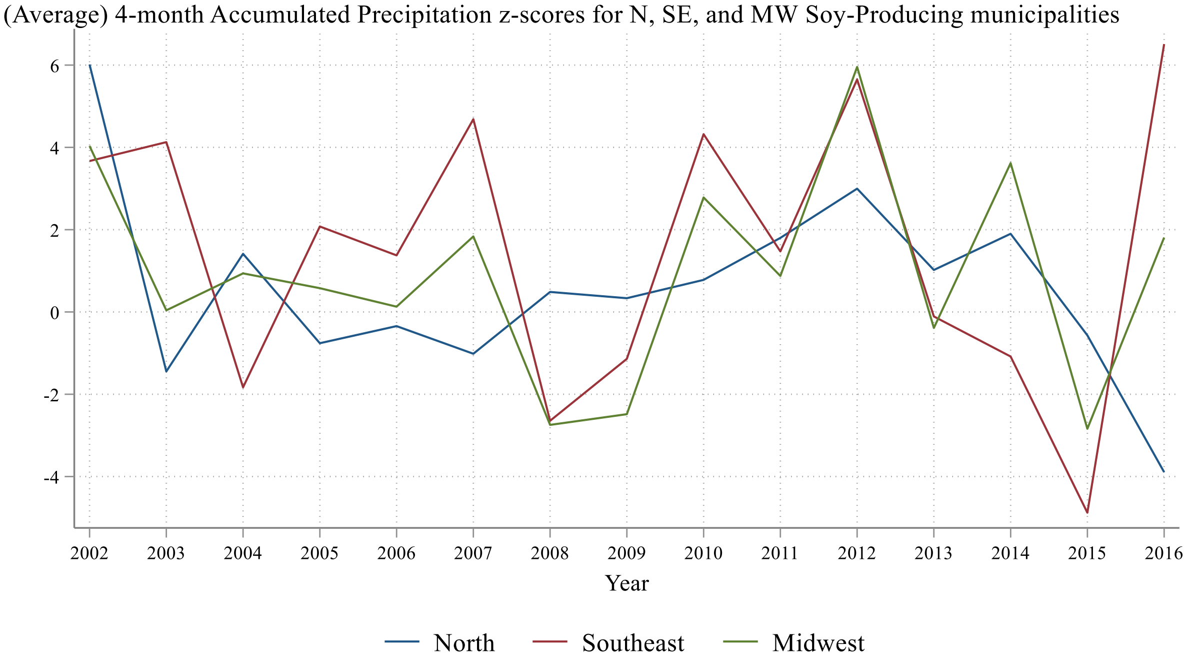

Figure A.5. (Average) 4-month accumulated precipitation z-scores for North (N), Southeast (SE), and Midwest (MW) Soybeans-producing municipalities.

A.3. Regional estimates of effect of droughts on Corn and Soybeans for the 2002-2016 period.

Table A.1. Effect of droughts on soybeans yield for the Brazilian regions, 2002-2016.

	Soybeans Yield (Log)				
Variables	North	Northeast	Southeast	South	Midwest
Drought in months:					
DM1 (Oct-Jan)	-0.0111	-0.0772***	0.00868	-0.0342	-0.0248
	(0.00749)	(0.00994)	(0.00164)	(0.0176)	(0.0122)
Mean Temperature:					
Sept(t-1)	0.0279***	0.0165	-0.0131	0.0365	-0.00942
	(0.00559)	(0.0687)	(0.00454)	(0.0342)	(0.00520)
Oct(t-1)	0.0202**	-0.0734	0.000805	-0.0566	0.0423**
	(0.00813)	(0.0723)	(0.00782)	(0.0272)	(0.0133)
Nov(t-1)	-0.0115	0.0166	0.0107	-0.0304	-0.0283
	(0.0113)	(0.0568)	(0.0146)	(0.0257)	(0.0165)
Dec(t-1)	-0.0372	-0.0877*	-0.0303	0.0868***	-0.0284
	(0.0252)	(0.0387)	(0.0278)	(0.00874)	(0.0259)
Jan(t)	-0.0201	0.113	0.0242*	-0.199*	-0.0342
	(0.0168)	(0.122)	(0.00205)	(0.0551)	(0.0404)
Feb(t)	-0.00672	-0.0714	-0.0536	0.0747	-0.0182
	(0.0127)	(0.0433)	(0.0178)	(0.0479)	(0.0139)
Mar(t)	0.00319	-0.0581**	0.0435	-0.0792	0.0307
	(0.00902)	(0.0153)	(0.0122)	(0.0295)	(0.0218)
Apr(t)	0.0325*	-0.0137	0.0293	0.115**	0.0155
	(0.0157)	(0.0889)	(0.0132)	(0.0210)	(0.00862)
May(t)	-0.0122	-0.0782	-0.0195	0.119*	0.0377
	(0.0208)	(0.118)	(0.00999)	(0.0318)	(0.0240)
Jun(t)	0.0128	0.0309	-0.0434	-0.0621***	-0.0327**
	(0.0239)	(0.0766)	(0.00970)	(0.000974)	(0.00730)
Jul(t)	-0.0105	-0.0250	0.00641	-0.0403	0.0348
	(0.00681)	(0.0582)	(0.00665)	(0.0239)	(0.0320)
Aug(t)	0.00584	0.0180	0.00171	-0.0412	-0.0438*
	(0.00686)	(0.0762)	(0.00589)	(0.0218)	(0.0179)
Sept(t)	0.0129	-0.0236	0.0200	0.0642	0.0151
	(0.00943)	(0.0243)	(0.00704)	(0.0546)	(0.0165)
Constant	7.125***	14.00***	8.150*	8.359*	8.369***
	(0.595)	(1.763)	(0.738)	(2.484)	(1.027)
Obs.	1,600	982	6,647	13,842	5,129
R^2 w	0.301	0.319	0.161	0.590	0.257
R^2 b	0.331	0.0187	0.0779	0.230	0.109
R^2 o	0.162	0.114	0.108	0.487	0.199
R^2 adj.	0.286	0.296	0.158	0.589	0.252

Note: Standard errors in parentheses are clustered at the state level. All models include municipality-FE, year-FE, and state-specific time trends. Significance levels are denoted as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2. Effect of droughts on value of soybeans production for the Brazilian regions, 2002-2016.

	Value of Soybeans Production (Log)				
Variables	North	Northeast	Southeast	South	Midwest
Drought in months:					
DM1 (Oct-Jan)	0.0298	-0.0480	0.0146	0.0333	-0.0533*
	(0.0575)	(0.0341)	(0.0483)	(0.0457)	(0.0212)
Mean Temperature:					
Sept(t-1)	0.158	-0.184*	0.109	0.0280	0.0692
	(0.0811)	(0.0773)	(0.0541)	(0.0579)	(0.0334)
Oct(t-1)	-0.00920	-0.0167	-0.180	-0.134**	-0.0616
	(0.0520)	(0.0872)	(0.0531)	(0.0187)	(0.0413)
Nov(t-1)	0.102*	0.0388	-0.0302	-0.105	-0.176
	(0.0457)	(0.109)	(0.0475)	(0.0993)	(0.0801)
Dec(t-1)	-0.150***	0.0166	0.0792	0.289***	0.0700*
	(0.0316)	(0.139)	(0.0718)	(0.0163)	(0.0273)
Jan(t)	-0.0519	-0.0569	0.0273	-0.227	-0.0612
	(0.0467)	(0.220)	(0.0471)	(0.0839)	(0.0457)
Feb(t)	-0.140*	0.0831	-0.0155	0.155	0.00101
	(0.0606)	(0.268)	(0.0286)	(0.0833)	(0.0589)
Mar(t)	-0.0601	-0.315	0.321	0.0217	0.128*
	(0.168)	(0.290)	(0.103)	(0.0507)	(0.0477)
Apr(t)	0.153*	0.164	0.0902	0.239**	0.0703**
	(0.0683)	(0.220)	(0.0492)	(0.0341)	(0.0120)
May(t)	-0.219	-0.290	-0.150	0.0436	-0.0270
	(0.114)	(0.144)	(0.163)	(0.100)	(0.0258)
Jun(t)	0.131	0.150**	-0.0337	-0.0983	0.0465**
	(0.148)	(0.0497)	(0.119)	(0.0466)	(0.00981)
Jul(t)	-0.117**	0.198	-0.186	-0.0536	-0.0830
	(0.0405)	(0.103)	(0.0604)	(0.0226)	(0.0468)
Aug(t)	0.111**	0.0825	-0.0931	0.0470	-0.0589***
	(0.0423)	(0.0874)	(0.0523)	(0.0320)	(0.00791)
Sept(t)	0.0380	-0.141*	0.0741	0.0108	0.0343
	(0.0707)	(0.0551)	(0.0263)	(0.0469)	(0.0162)
Constant	7.449	15.39*	5.164	2.725	9.576**
	(5.842)	(7.613)	(4.079)	(3.966)	(2.767)
Obs.	1,596	982	6,647	13,835	5,129
R^2 w	0.460	0.405	0.180	0.492	0.382
R^2 b	0.00246	0.0803	0.0875	0.0000435	0.0372
R^2 o	0.0786	0.00512	0.00943	0.0554	0.0164
R^2 adj.	0.448	0.384	0.176	0.491	0.378

Note: Standard errors in parentheses are clustered at the state level. All models include municipality-FE, year-FE, and state-specific time trends. Significance levels are denoted as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3. Effect of droughts on corn yield for the Brazilian regions, 2002-2016.

	Corn Yield (Log)				
Variables	North	Northeast	Southeast	South	Midwest
Drought in months:					
DM1 (Oct-Jan)	-0.00649	-0.0265*	-0.0326*	-0.0600*	-0.00602
	(0.00447)	(0.0128)	(0.0130)	(0.0169)	(0.00376)
DM2 (Jan-May)	-0.000160	-0.0207	-0.000307	-0.0255*	-0.0142
	(0.00615)	(0.0239)	(0.00678)	(0.00776)	(0.00876)
Mean Temperature:					
Sept(t-1)	0.00941	-0.107***	-0.00417	0.0443	0.0307
	(0.0131)	(0.0201)	(0.0127)	(0.0216)	(0.0205)
Oct(t-1)	-0.0145	0.0273	-0.0259	-0.0421	-0.0198
	(0.00766)	(0.0532)	(0.0177)	(0.0352)	(0.0423)
Nov(t-1)	-0.0129	0.215**	0.0859*	-0.149	-0.00360
	(0.0120)	(0.0929)	(0.0334)	(0.0599)	(0.0341)
Dec(t-1)	0.00739	0.0642	0.0406***	0.0980	-0.0103
	(0.00648)	(0.0808)	(0.00592)	(0.0558)	(0.0341)
Jan(t)	0.00137	-0.0543	0.0461***	-0.167*	-0.0126
	(0.0130)	(0.0509)	(0.00643)	(0.0524)	(0.0319)
Feb(t)	-0.0158	-0.293***	-0.0153**	-0.0141	-0.00599
	(0.0135)	(0.0442)	(0.00473)	(0.0436)	(0.0245)
Mar(t)	-0.00447	-0.00531	-0.0962*	-0.0258	0.0412***
	(0.0268)	(0.112)	(0.0370)	(0.0716)	(0.00209)
Apr(t)	0.0168	-0.213*	-0.0619	0.140**	0.0271*
	(0.0150)	(0.105)	(0.0315)	(0.0160)	(0.00959)
May(t)	-0.00731	0.0501	-0.0300	0.0467	-0.0130
	(0.0101)	(0.0491)	(0.0134)	(0.0473)	(0.0239)
Jun(t)	-0.0108	-0.145**	-0.00295	-0.0208	-0.0158
	(0.00922)	(0.0613)	(0.0176)	(0.00928)	(0.0121)
Jul(t)	0.00784	0.0878	0.0396	-0.0816	-0.00850
	(0.0191)	(0.0576)	(0.0252)	(0.0334)	(0.0369)
Aug(t)	-0.00732	0.000610	-0.00603	0.0231	0.00352
	(0.00489)	(0.0695)	(0.0125)	(0.0138)	(0.00970)
Sept(t)	0.000425	0.0871	0.0278	0.0311	0.0177
	(0.0135)	(0.0624)	(0.0165)	(0.0386)	(0.0186)
Constant	8.027***	13.57***	8.406***	11.06**	7.124***
	(0.758)	(3.396)	(0.193)	(2.345)	(0.524)
Obs.	6,418	23,761	23,031	17,500	6,768
R^2 w	0.218	0.186	0.122	0.511	0.354
R^2 b	0.0747	0.0272	0.181	0.0334	0.00711
R^2 o	0.0759	0.0208	0.102	0.231	0.124
R^2 adj.	0.214	0.185	0.121	0.510	0.351

Note: Standard errors in parentheses are clustered at the state level. All models include municipality-FE, year-FE, and state-specific time trends. Significance levels are denoted as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4. Effect of droughts on value of corn production for the Brazilian regions, 2002-2016.

	Value of Corn Production (Log)				
Variables	North	Northeast	Southeast	South	Midwest
Drought in months:					
DM1 (Oct-Jan)	-0.0327	0.0199	-0.0813**	-0.0691	-0.103*
	(0.0184)	(0.0277)	(0.0239)	(0.0287)	(0.0424)
DM2 (Jan-May)	0.0120	-0.0803*	0.0345	-0.0449*	-0.00470
	(0.0259)	(0.0398)	(0.0187)	(0.0142)	(0.0226)
Mean Temperature					
Sept(t-1)	0.0312	-0.147***	0.0439*	0.0641	0.0167
	(0.0702)	(0.0359)	(0.0153)	(0.0518)	(0.0289)
Oct(t-1)	0.0218	0.149*	-0.0529**	-0.0404	0.0411
	(0.0401)	(0.0778)	(0.0162)	(0.0191)	(0.0394)
Nov(t-1)	-0.0943	0.0507	0.122	-0.0602	0.0108
	(0.0605)	(0.122)	(0.100)	(0.0893)	(0.0715)
Dec(t-1)	0.157*	0.238**	0.0813	0.00763	-0.104
	(0.0756)	(0.0882)	(0.0520)	(0.0447)	(0.0884)
Jan(t)	0.0384	-0.136**	0.0757*	-0.134	-0.0898
	(0.0308)	(0.0579)	(0.0258)	(0.0625)	(0.0431)
Feb(t)	-0.0105	-0.486***	-0.0336**	-0.104	0.112
	(0.0560)	(0.109)	(0.0104)	(0.0813)	(0.108)
Mar(t)	-0.191**	0.171	-0.150	0.00186	0.169**
	(0.0748)	(0.155)	(0.104)	(0.0866)	(0.0472)
Apr(t)	0.0123	-0.289*	-0.116	0.124**	-0.000622
	(0.0465)	(0.132)	(0.0596)	(0.0265)	(0.0195)
May(t)	0.0953**	0.0562	-0.0792	0.0418	0.140
	(0.0373)	(0.0965)	(0.0423)	(0.0553)	(0.0623)
Jun(t)	0.0966	-0.226**	0.0943**	-0.108***	-0.0751
	(0.0610)	(0.0948)	(0.0232)	(0.0103)	(0.0440)
Jul(t)	-0.0471	0.168**	0.0315	-0.0519	-0.00957
	(0.0266)	(0.0677)	(0.0579)	(0.0226)	(0.0427)
Aug(t)	-0.0376	-0.0754	-0.0374	-0.00464	0.00499
	(0.0359)	(0.109)	(0.0171)	(0.0161)	(0.0521)
Sept(t)	0.0228	0.209***	0.0274	0.0777	0.0494
	(0.0755)	(0.0572)	(0.0268)	(0.0311)	(0.0444)
Constant	3.300	13.32***	7.465**	12.56***	0.280
	(2.860)	(3.050)	(1.761)	(0.806)	(3.094)
Obs.	6,416	23,704	23,027	17,498	6,768
R^2 w	0.0397	0.219	0.124	0.192	0.128
R^2 b	0.149	0.0949	0.111	0.00154	0.0880
R^2 o	0.00710	0.134	0.0524	0.0167	0.00714
R^2 adj.	0.0344	0.218	0.123	0.190	0.124

Note: Standard errors in parentheses are clustered at the state level. All models include municipality-FE, year-FE, and state-specific time trends. Significance levels are denoted as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.4. Drought incidence in Southern Brazil

Figure A.6. (Average) 18-month accumulated precipitation z-scores for soybeans-producing municipalities in Paraná, Santa Catarina, and Rio Grande do Sul.

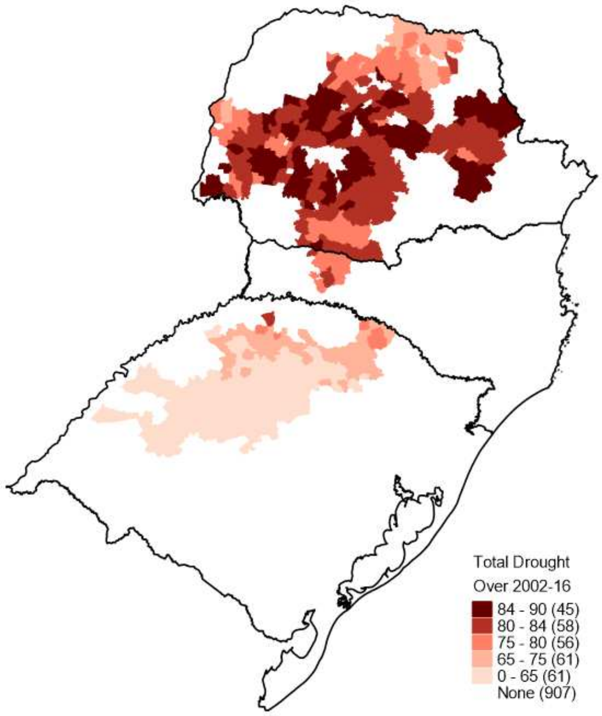

Figure A.7. Total number of weeks in severe and/or extreme drought in Southern Brazil for the 2002-2016 period.

A.4. Soybeans Production and drought incidence for selected dry and wet years

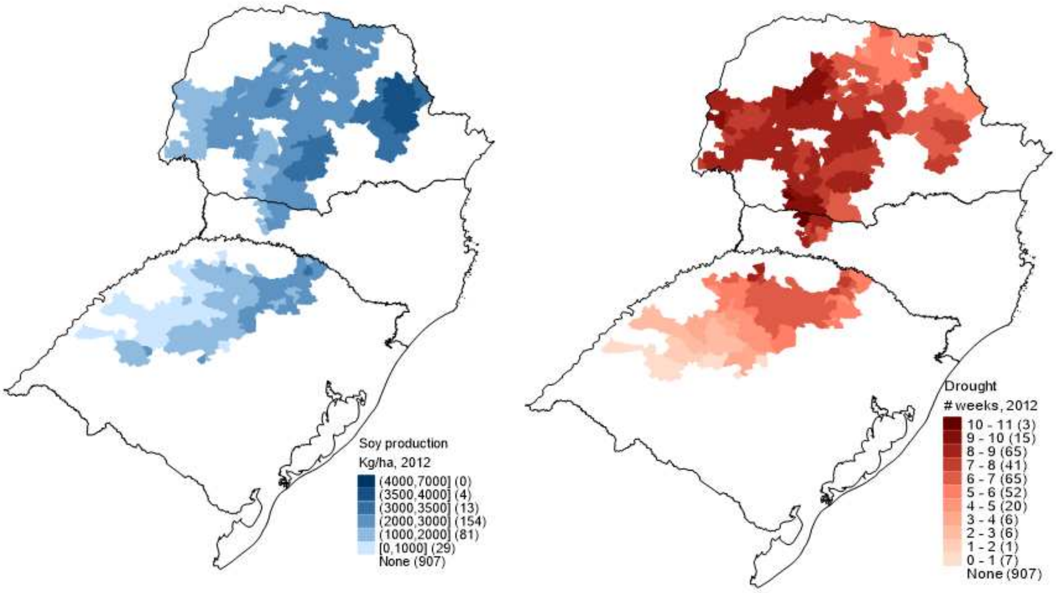

Figure A.8. Soybeans yield and drought incidence in weeks for Southern Brazil in a selected dry year, 2012.

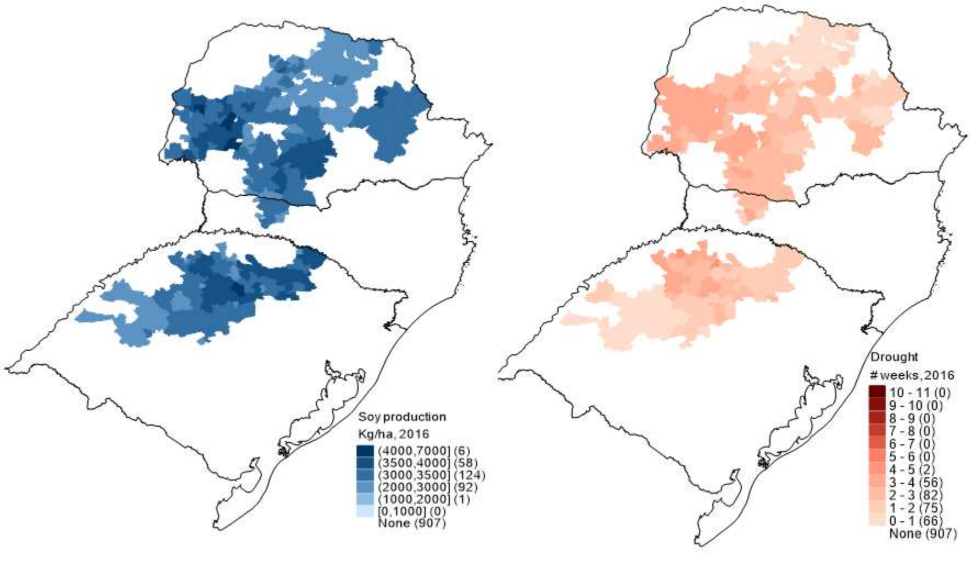

Figure A.9. Soybeans yield and drought incidence in weeks for Southern Brazil in a selected wet year, 2016.

FINAL REMARKS

This dissertation studied in three chapters the adoption of agricultural technologies, the allocation of resources, and the value of information for agricultural decisions. It brings a combination of theory and methodologies related to agricultural productivity, agricultural development and the decisions involving the allocation of resources for agricultural production. The first chapter studied the use of traditional (labor and land) and commercial (machinery and fertilizers) inputs used to produce crops and livestock commodities for a group of ten South American countries. I used Acemoglu's directed technical change theory to obtain an empirical model that allowed me to estimate elasticities of substitutions between the inputs and between the outputs with country-level panel data and to study the process of induced innovation in South American agriculture. Additionally, I provided instrumental variables (IV) estimations where deforestation was used as an instrument given its association with the supply of agricultural land in intensive deforestation South American countries. This chapter contributed mainly on the discussion related to the mechanization of agriculture in developing countries as more machinery was used relative to fertilizers motivated by the expansion of the agricultural frontiers and the displacement of agricultural workers to the industrial sectors. Also, it highlighted the potential of using the directed technical change theory to study the effect of deforestation control policies in South America and their impacts on agricultural productivity.

The second chapter provided a structural model that allowed to insert a soil signal (i.e., soil electroconductivity (EC)) into the empirical specification of the corn yield function used to estimate the impact of nitrogen fertilizer decisions on yields. The purpose of the chosen approach was to estimate the expected benefits of the adoption of variable rate applications (VRA) of nitrogen within the plots of a field as compared to a uniform rate application (URA). The chapter combined the tools from Bayesian decision theory and U.S. farms experimental data to estimate the Expected Value of Sample Information (EVSI) which is the expected return to adopting VRA technology in the studied farms. The EVSI also contains the value of using the soil EC information to guide the VRAs. The theoretical model and the application of the model allowed me to identify the determinants of EVSI, and to empirically estimate it with unique experimental data from ten U.S. farms in Illinois and Ohio for 2016 and 2017. The estimated EVSI was $1.81 U.S. dollars per acre, which is considered too low for the adoption of VRA technology if costs are estimated at $10 per acre. By identifying the parameters that determine EVSI in this chapter, the main conclusion was that either EC is a "poor" soil signal, because its correlation with the true soil conditions is low, or the field's soil quality is significantly uniform, which would not require variable rates in the first place.

Finally, the third chapter studied the effects of weather anomalies on Brazilian agriculture by estimating the marginal effect of one additional period of severe and/or extreme drought on corn and soybeans yields and values of production. First, I provided a comprehensive assessment of the effect of droughts in all corn- and soybeans-producing Brazilian municipalities for the 2002-2016 period, and the estimated effects by region. The meteorological drought variable used was obtained from observed precipitation and was measured in months of the growing season. Then, I used remote sensing information to identify a soybeans agricultural region in Southern Brazil and matched soybeans yield and value of production with a meteorological drought indicator, also obtained from observed precipitation but measured in weeks. For Southern Brazil, I added remotely sensed biophysical indicators related to the plant, the soil, and the atmosphere. The chapter estimated panel data regression models with fixed effects to obtain the marginal effect of an extra period of drought and showed that adding remote sensing biophysical variables improved the estimation of the drought effect for Southern Brazil. The motivating theoretical framework indicated that this result can be explained mainly by the increase in the correlation between the meteorological drought indicator and the actual droughts (proxied by the agricultural outcomes). It also highlighted the importance of the value of information (VOI) theory to study how information from the publicly available meteorological drought indicator can affect decisions regarding agricultural policies aimed at alleviating the agricultural impacts of droughts ex-ante such as the promotion and adoption of irrigation plans. I also briefly discussed the use of the VOI theory as a tool to calculate the expected benefits of the remote sensing data used directly in agricultural decisions, as opposed to only affecting the correlation between the meteorological indicator and the actual drought, to guide public policies concerning the expansion of drought monitoring to all regions of Brazil and investments in satellite imagery and remote sensing technologies.