Global Trade and Margins of Productivity in Agriculture

Farid Farrokhi
Purdue

Heitor S. Pellegrina
NYU Abu Dhabi*

May 2020

Abstract

We study the effects of globalization on agricultural productivity across countries. We develop a multi-country general equilibrium model that incorporates choices of crops and technologies in agricultural production at the micro-level of fields covering the surface of the earth. We estimate our model using field-level data on potential yields of crops under different technologies characterized by factor and input intensity. We evaluate the welfare and productivity gains from reductions in trade costs of agricultural outputs and inputs across countries between 1980 and 2015. In addition to gains from international crop specialization, we find notable gains from access to foreign agricultural inputs. This mechanism operates through a shift from traditional (labor-intensive) technologies to modern (input-intensive) ones.

*We thank conversations with Jonathan Eaton, Tom Hertel, David Hummels, Samreen Malik, Lucas Scottini, Sebastian Sotelo, Farzad Taheripour, Chong Xiang, and participants in seminars at Purdue, NYU Abu Dhabi, NOITS and NBER Agricultural Markets and Trade Policy. We would like to thank excellent research assistance from Karolina Wilckzynska. We would like to thank the help from the ITaP team of Purdue with our high-performance computing. All errors are our own. Email: ffarrokh@purdue.edu and heitor.pellegrina@nyu.edu.

1 Introduction

In recent decades, agricultural productivity has grown remarkably around the world. This growth has been crucial to sustain the rising global demand for food, and to increase the economy-wide efficiency in many countries, particularly in developing ones where the share of employment in agriculture is high. Hence, understanding drivers of agricultural productivity has been central to discussions about food consumption and welfare across countries. In this paper, we study how the dramatic fall in the international barriers to trade in recent decades, often referred to as globalization, shaped agricultural productivity worldwide.

Globalization can affect agricultural productivity through the output or the input side of agricultural markets. On the output side, it can bring productivity gains by allowing countries to specialize in crops in which they have a comparative advantage. On the input side, it facilitates the procurement of inputs such as machinery, fertilizers, and pesticides that are crucial for the use of modern, input-intensive technologies. Cross-country variations in these two margins of productivity emerge from the enormous spatial heterogeneity in land suitability for different crops and from large differences in countries' access to foreign agricultural inputs.

In this paper, we develop a quantifiable, multi-country general equilibrium model of trade with an extremely rich spatial structure that incorporates these two margins of productivity gains in agriculture. We bring our model to micro-level data produced by agroclimate scientists on agricultural productivity of crops under different technologies for more than 1.3 million fields covering the surface of the earth. We construct, for each field, a production possibility frontier (PPF) that governs choices of crops and technologies. Using our framework, we address a few key questions. What were the consequences of globalization for land and labor productivity of crops across the world geography? How important was the access to internationally supplied inputs for technology adoption and efficiency gains in agriculture? How important were these input-side mechanisms for productivity gains in agriculture as opposed to output-side mechanisms that operate through comparative advantage?

In addition to our field-level data, which comes from the Food and Agriculture Organization's Global Agro-Ecological Zones (FAO-GAEZ) project, we assemble a country-level dataset that integrates information on bilateral trade and production of crops and agricultural inputs over more than three decades. We start our analysis by documenting three empirical patterns that emerge from these data: (i) as a global trend, agricultural productivity and the import content of agricultural input have largely grown since 1980; (ii) across countries, agricultural productivity is strongly and positively associated with measures of agricultural input intensity; and (iii), at the level of fields, there are large premia in yields of modern, input-intensive technologies over the traditional, labor-intensive ones. Such patterns indicate that productivity gains in agriculture are associated with the use of input-intensive production technologies, and that many countries by and large procure these inputs through imports.

Guided by these empirical observations, we develop our theoretical framework. Agriculture producers in each field choose which crops to grow and with which technology to grow them. A technology is characterized by factor and input intensities. Therefore, higher relative wages and lower relative prices of agricultural inputs encourage the use of labor-saving, input-intensive technologies. To manage the margins of production choices, we introduce a nested choice structure based on a generalized Fréchet distribution. This parsimonious formulation allows for a different elasticity for each margin of adjustment, i.e. across crops and across technologies within crops, and a field-level productivity shifter for every pair of crop and technology. The two elasticities govern the curvature of the PPF in every field along the dimension of crops and technologies, and the productivity shifters tightly map to field-level measures of agricultural productivity from FAO-GAEZ data. Country-level supply of crops is the endogenous outcome of the aggregation of field-level productions within a country.

We take the model to data in two steps. First, we estimate country-level parameters related to trade and production following standard practices in the literature using data from 2015. Second, we construct PPFs at the field level. We calibrate heterogeneous field-level productivity shifters for every crop-technology pair based on the FAO-GAEZ data, and estimate the elasticities that govern the curvature of PPFs using method of moments. We construct our moments based on field-level variations in yields and cross-country variations in input use.[1] The field-level variation in actual yields identifies the elasticity that governs choices of crops, and cross-country variation in the cost share of agricultural inputs identifies the elasticity that governs choices of technology.

Equipped with our quantified model, we turn to answering our questions. To this end, we simulate a counterfactual in which we bring trade costs in agricultural outputs and inputs back to their estimated level in 1980, and compare the resulting equilibrium with that

[1] To construct our moments, we also use field-level data on actual yields from the United States, as we explain later.

in the baseline of 2015. Our results highlight that the effects of globalization in agricultural inputs operate through a distinct channel compared to that in the output side. While reductions in trade costs on the output side bring efficiency gains that stem from comparative advantage forces, reductions in the trade costs of agricultural inputs induce a notable increase in the share of land allocated to modern technologies.

We find that due to general equilibrium effects and differences in country characteristics, the resulting welfare gains were largely heterogeneous across countries. For instance, countries that fell behind in the process of globalization, that is, countries with a low reduction in trade costs relative to others, ended up facing relatively higher prices of agricultural inputs. Consequently, in these countries incentives for adopting modern technologies were lower, creating a barrier for their economic development. Yet, at the global level, due to the overall lower trade costs of agricultural outputs and inputs, yields rose by 2.4% to 9.0% across crops, real consumption of agriculture increased by 2.83%, and overall welfare rose by 1.63%.

To shed light on the effects of input-side mechanisms on agricultural productivity, we consider counterfactuals in which trade costs fall only in agricultural outputs or only in agricultural inputs. We find that globalization in agricultural inputs account for around half of global welfare gains from globalization in both outputs and inputs of agriculture. Not accounting for input-side mechanisms leads to a large underestimation of the gains from globalization in agriculture.

This paper relates to different strands of literature. First, our theoretical framework builds on two papers that model crop choices using tools from Eaton and Kortum (2002): (i) Sotelo (2020) who uses regional-level data from Peru to study the effects of domestic trade costs on agricultural productivity, and (ii) Costinot, Donaldson, and Smith (2016) who combine country and field level data to evaluate the impact of climate change on agricultural productivity.[2] Both of these papers allow for a single technology choice for each crop and estimate their frameworks using FAO-GAEZ data. We extend their framework by introducing choices of crop and technology using a nested choice structure that allows for different elasticities governing each type of choice.[3] In addition to bringing a new mechanism driving

[2]Gouel and Laborde (2018) extend Costinot, Donaldson, and Smith (2016) to a wider set of agricultural activities — more crops as well as livestock— incorporating a flexible demand structure with elasticities estimated in the literature on agricultural economics.

[3]Our nested choice structure relates to a several studies using generalized extreme value distributions of productivities in trade models (Arkolakis, Ramondo, Rodríguez-Clare, and Yeaple, 2018; Lind and Ramondo, 2018; Lashkaripour and Lugovskyy, 2018).

changes in agricultural productivity related to technology choice, our modeling approach can incorporate additional layers of field-level data on agricultural productivity related by technology type in a theoretically consistent manner.

More broadly, this paper provides different contributions to research on agricultural trade and economic development (Tombe, 2015; Gafaro and Pellegrina, 2019; Porteous, 2016; Costinot and Donaldson, 2014; Fajgelbaum and Redding, 2019; Allen and Atkin, 2016; Donaldson, 2018; Pellegrina, 2019; Baldos, Hertel, and Moore, 2019; Bergquist, Faber, Fally, Hoelzlein, Miguel, and Rodriguez-Clare, 2019; Porteous, 2020; McArthur and McCord, 2017). First, we quantify the effects of globalization on agricultural productivity across countries, highlighting the critical role of access to foreign intermediates for the adoption of modern agricultural technologies. Second, we introduce the role of technology choice into general equilibrium model.[4] By examining the role of trade in agricultural inputs, we complement papers that incorporate input-output structures into quantitative trade and economic geography models, such as Caliendo and Parro (2015).[5] In contrast to this literature, which generally assumes exogenous cost shares of inputs based on input-output tables, we allow differences in these cost shares to reflect endogenous technology choices.

This paper relates to a large and varied research in agricultural economics that examines farmers' response to government policies in their choices of land use or crop supply, e.g. see Lee and Helmberger (1985) for a pioneer study, Hertel (2002) and Hertel (2013) for a review of recent literature on the application of computation general equilibrium models in agriculture, and Anderson (2010, 2016) for a discussion of the causes of globalization in agricultural trade. We provide two contributions to this literature. First, we quantify the effects of globalization on agricultural productivity using a framework that allows for farmers' crop and technology responses at fine levels of geographic disaggregation. Second, we show that international trade in agricultural inputs has been a key component of the effects of globalization, which reinforces findings in recent papers studying the role of commodity trade as crucial inputs to downstream sectors, including Farrokhi (2020) and Fally and Sayre (2018).

Lastly, our paper relates to a rich literature in macroeconomics examining the role of

[4] In a recent paper, Bergquist, Faber, Fally, Hoelzlein, Miguel, and Rodriguez-Clare (2019) analyze the general equilibrium effects of scaling up policy interventions using farm-level data from Uganda. In their model, they also allow farmers to choose between modern and traditional technologies. There are two differences between their approach and ours. First, using the method of exact hat algebra they do not require data on agricultural productivity, but their method requires detailed farm-level data which is not available at a global scale. Second, they do not allow for different elasticities of substitution between crops and technologies.

[5] See Costinot and Rodríguez-Clare (2014) for a survey on tools and applications in this literature.

agriculture productivity in the process of economic development, e.g. Caselli (2005). Among recent studies, Lagakos and Waugh (2013) and Gollin, Lagakos, and Waugh (2014) explore sources of cross-country labor productivity differences in agriculture, and Donovan (2017) studies the role of insurance markets in agriculture input use across countries. Closer to our paper is Gollin, Parente, and Rogerson (2007), who emphasize labor-saving, input-intensive technologies as a key mechanism driving the gains in agriculture productivity and structural change. Our contribution to these studies is two-fold. First, we put the analysis of agricultural productivity into a global perspective. Second, we connect the macro-level analysis to micro-level heterogeneity intrinsic in conditions of land and climate across the world geography.

The rest of this paper is organized as follows. In Section 2, we present the data and empirical patterns that motivate our model, which we develop in Section 3. We estimate our model in Section 4, and run quantitative exercises in Section 5.

2 Data & Empirical Patterns

This section describes the data used in our analysis and presents empirical patterns about agricultural production and trade that inform the formulation of our model.

2.1 Data

We describe our data in two parts: at the aggregate level of countries, and at the disaggregated level of fields. Here, we highlight main features of our data, and leave a thorough description to the appendix.

2.1.1 Country-level data

We have collected information from several sources to construct a panel of country-level data on gross output, bilateral trade, and expenditure of crops, agricultural inputs, and nonagricultural goods (see Table A.1). As for agricultural inputs, we focus on fertilizers, pesticides, and agricultural machinery, which are exclusively used as inputs into agriculture production.[6] We complement this panel data with information on employment, cropland, share of

[6]Data on these three input categories are available with sufficient coverage and quality. Furthermore, they are frequently cited as the most crucial inputs. For example, FAO-STAT provides data only for these three agricultural inputs.

expenditure on agriculture, value added in agriculture and non-agriculture sectors as well as standard macroeconomic indicators such as GDP and population.

Our final dataset contains the 65 countries with the largest agricultural value added in the world (excluding countries with missing data) plus one region that aggregates remaining countries which we refer to as the rest of the world (ROW). We include 10 major crops in our analysis: banana, cotton, corn, palm oil, potato, rice, soybean, sugarcane, tomato, and wheat.

Table 1 reports summary of statistics for our country-level data. For each variable, the table reports aggregate values in year 2015 for eight regions that cover the world geography and the growth at the global level between 1980 and 2015. We normalize GDP per capita and agricultural value-added per worker such that the GDP per capita in North America is set at unity. As it has been documented in Gollin, Lagakos, and Waugh (2014), in our data the valued added per worker in agriculture is typically lower than its economy-wide counterpart, and the gap between the two decreases with countries' income per capita.

Table 1: Summary Statistics

	Values for 2015								Global
	North America	East Asia & Pacific	East Europe	Latin America	MENA	South Asia	SSA	West Europe	Growth Rate 1980-2015
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GDP per capita	1.00	0.21	0.16	0.17	0.10	0.04	0.03	0.63	281.3
VA per worker in ag	0.53	0.14	0.06	0.06	0.04	0.01	0.01	0.28	314.4
Import share of inputs	0.33	0.10	0.46	0.56	0.29	0.32	0.62	0.65	89.2
- Machinery	0.30	0.06	0.65	0.27	0.35	0.17	0.75	0.62	49.7
- Fertilizer	0.33	0.12	0.20	0.64	0.20	0.39	0.51	0.53	120.2
- Pesticide	0.42	0.15	0.64	0.68	0.51	0.30	0.75	0.82	57.5
Number of countries	2	5	6	12	6	9	12	13	

Notes: Value added data are normalized such that economy-wide value added per worker in North America is set at unity. SSA stands for Sub-Saharan Africa. MENA stands for Middle East and North Africa region. For every region the reported number is the aggregate value for all countries in that region. Import share of inputs is total imports divided by total expenditure.

An inspection of the import share of agricultural inputs in Table 1 reveals large variations across countries, and across input categories within a country. Most countries depend on international trade to procure at least one of fertilizers, pesticides, and farm machinery, which reflects the high geographic concentration in the production of agricultural inputs. The production of fertilizers is concentrated in a few countries that have the required natural resources, and the production of pesticides and farm machinery requires chemical- and

Another important input category is seeds, but we do not have production data to include them in our empirical analysis.

machinery-related technologies that might be unavailable to low-income countries. For instance, countries in the Middle East and North Africa (MENA) and in the East Europe have large endowments of raw fertilizers, and, therefore, present a small import share of fertilizers, but imports in these countries account for a large share of their expenditure on farm machinery and pesticides. Import shares of all the input categories are typically the largest among Sub-Saharan African countries and the lowest in North America and East Asia & Pacific. For most European and Latin American countries imports account for about a half of their expenditure on agricultural inputs. In summary, these figures suggest that international trade plays a crucial role in the use of agricultural inputs.

2.1.2 Field-level data

The field-level data is given at the level of agro-ecological zones (AEZs). An AEZ is a 5 minute by 5 minute latitude/longitude grid, which encompasses an area of approximately 10 km by 10 km. We integrate two field-level datasets, one on potential yield of crops that represent the agricultural suitability of crops by technology type, the other on actual yields based on national censuses for a subset of countries.

The field-level data on potential yields of crops is taken from Global Agro-Ecological Zones (GAEZ) project, organized by the Food and Agriculture Organization (FAO) and the International Institute for Applied Systems Analysis. The "potential yield" of a crop in a field is defined as the maximum attainable yield of that crop (measured in tons per hectare) if the entire field were allocated to that crop. Potential yields are estimated based on agronomic models that use information on local conditions of land such as soil types, elevation, and land gradient and of climate such as rainfall, humidity, and temperature. These measures capture local agro-ecological characteristics of a field, but not local market conditions related to crop and input prices. We use data on potential yields for low and high input technologies. The low-input technology corresponds to a traditional farming activity where production is labor intensive and there is minimum to no use of agricultural inputs. The high-input technology corresponds to a modern system where production is input-intensive. Hereafter, we thus call low and high input technologies, respectively, "traditional" and "modern".[7] In addition, we define "modern potential yield premium" as the

[7] According to the definition from FAO-GAEZ, the low-input technology represents a production regime with "no application of nutrients, no use of chemicals for pest and disease control" and the high-input technology a production that is "fully mechanized with low labor intensity and uses optimum applications of nutrients and chemical pest, disease and weed control."

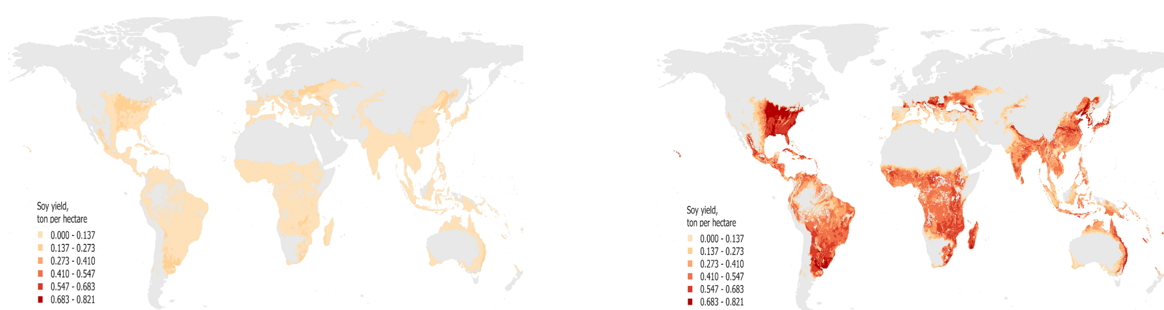
ratio of potential yield of modern to that of traditional technology. Our data on potential yields contain nonzero records of traditional and modern potential yields of at least one crop for 1,300,427 fields. To illustrate the data, Figure 1 plots potential yields of soybean based on traditional (low-input) and modern (high-input) technologies across the world geography.

We also collect field-level data from EarthStat.[8] These data contain field-level information on actual yields of crops circa year 2000 and are constructed based on national, state, and county level census statistics. A key limitation here is that the data does not always capture the field-level heterogeneity that is required for our analysis. In regions where data from agricultural censuses are not available at fine levels of disaggregation, aggregate data from national level are downscaled to the AEZ level. In contrast, for countries where detailed census information are available, the mapping is between disaggregated units, such as counties in the USA. For these reasons, we take the EarthStat field-level data only for the United States, which is constructed based on disaggregated information at the county level with a sufficiently large coverage of crops.[9]

Figure 1: Potential Yield of Soybean: Traditional (low-input) vs Modern (high-input)

a. Traditional

b. Modern



Notes: FAO-GAEZ potential yields are the maximum attainable yields of soybean in every field if the entire area of the field is allocated to soybean, using traditional (low-input) technology, and modern (high-input) technology.

[8]EarthStat is a collaboration between the Global Landscape Initiative at The University of Minnesota's Institute on the Environment and the Land Use and Global Environment lab at the University of British Columbia to construct field-level dataset on agriculture at the global level.

[9]There is a growing body of research in agriculture and climate sciences using EarthStat data at a global scale. For examples, see Deryng, Conway, Ramankutty, Price, and Warren (2014), Niedertscheider, Kastner, Fetzel, Haberl, Kroisleitner, Plutzar, and Erb (2016), Foley, Ramankutty, Brauman, Cassidy, Gerber, Johnston, Mueller, OConnell, Ray, West, et al. (2011) and Mueller, Gerber, Johnston, Ray, Ramankutty, and Foley (2012). For a detailed description of the construction of the dataset, see Monfreda, Ramankutty, and Foley (2008).

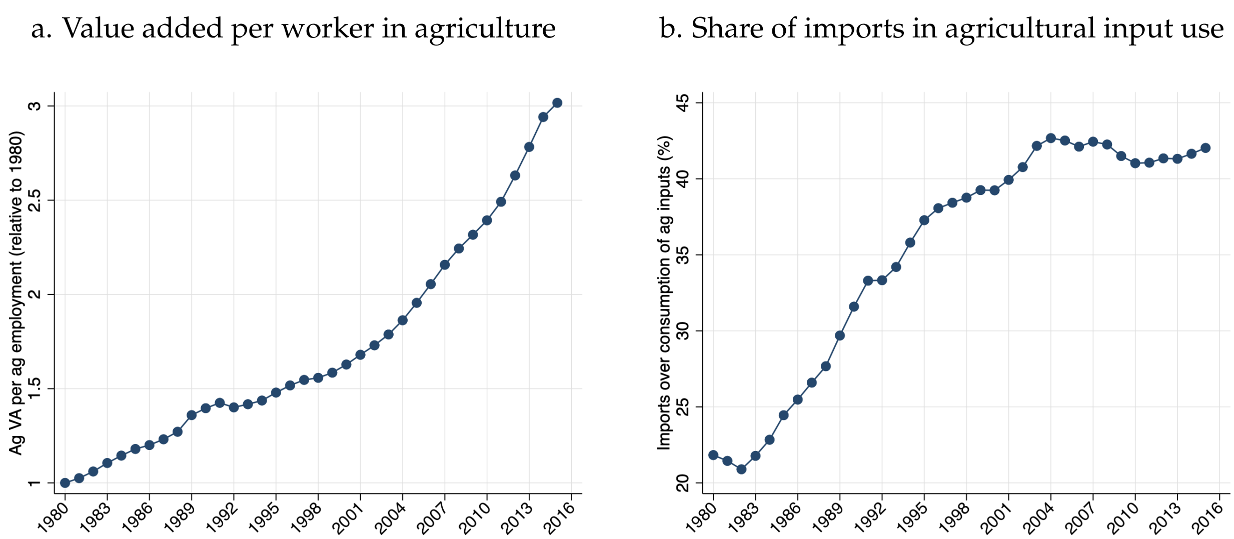
2.2 Empirical Patterns

We present three empirical patterns that motivate the particular structure that we impose in our model. The first two empirical patterns highlight the importance of agricultural input use and their imports over time and across countries. The last empirical pattern presents variations in the field level data which we incorporate into our model.

Empirical Pattern 1. At the global level, agricultural value added per worker and the share of imports in agricultural input use have increased substantially between 1980 and 2015.

Agricultural value added per worker has approximately tripled at the global level by between 1980 and 2015 (Figure 2). This substantial growth in agricultural productivity was accompanied by a striking growth of the share of imports in agricultural input use. Figure 2 panel (b) illustrates the evolution of import share in the use of agricultural inputs at the global scale. This share grew from 20% in 1980 to 40% in 2015. Moreover, the extent to which imports account for agricultural input use varies largely across countries and across the three categories of inputs. We illustrate these cross-country and cross-input variations in the appendix. This heterogeneity illustrates that many countries largely depend on imports to procure all or at least one of the agricultural input categories.

Figure 2: Value Added per Worker & Share of Imports in Agricultural Input Use at the Global Level (1980-2015).



Notes: Panel (a) presents global value added per worker in agriculture using value added data at constant prices. Panel (b) shows global imports of agricultural inputs across countries in value divided by total expenditure on agricultural inputs.

Empirical Pattern 2. Across countries, measures of agricultural input intensity are strongly correlated with GDP per capita.

In the previous empirical pattern, we highlighted the increasing role of trade in the use of agricultural inputs. We now show that agricultural input intensity across countries is strongly associated with the level of development.

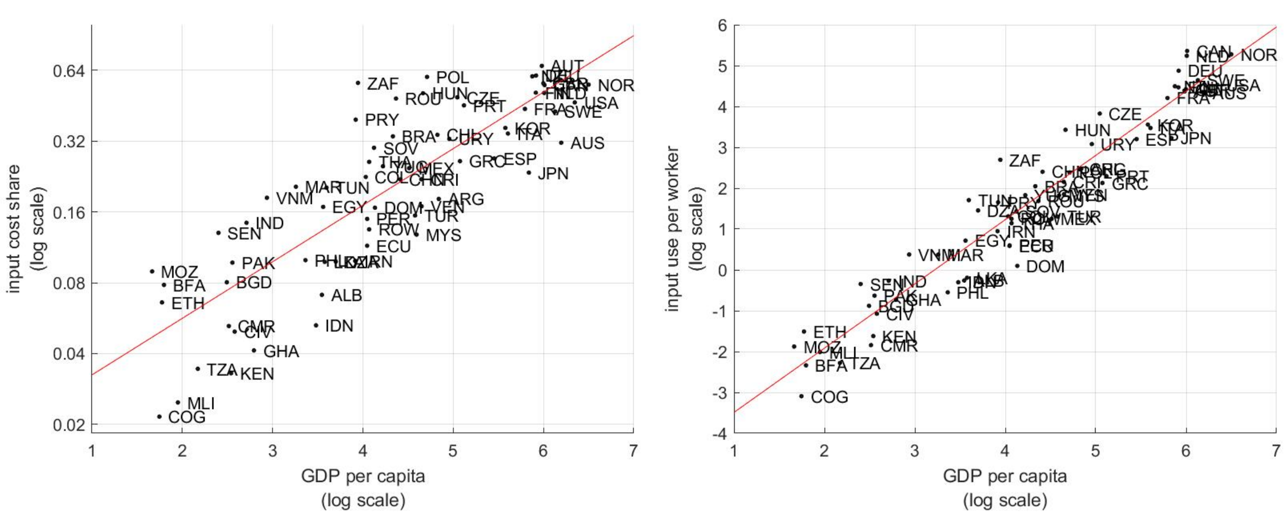
Figure 3-a shows a strong correlation between the cost share of inputs in agriculture against GDP per capita across countries. Figure 3-b shows that agriculture input use per worker is also strongly correlated with GDP per capita. The cross-country variations in measures of development and input intensity in agriculture are enormous. For example, GDP per capita in the United States is 35 times larger than that in Ghana, with the cost share of inputs being around 46% and 4% in the respective countries.

The combination of Empirical Patterns 1 and 2 suggests that the increasing access to foreign agricultural inputs is an important channel through which globalization affects agricultural productivity. With this takeaway in mind, we turn to an analysis of our field-level data.

Figure 3: Agricultural Input Intensity versus GDP per capita.

a. Cost share of inputs

b. Input expenditure per worker



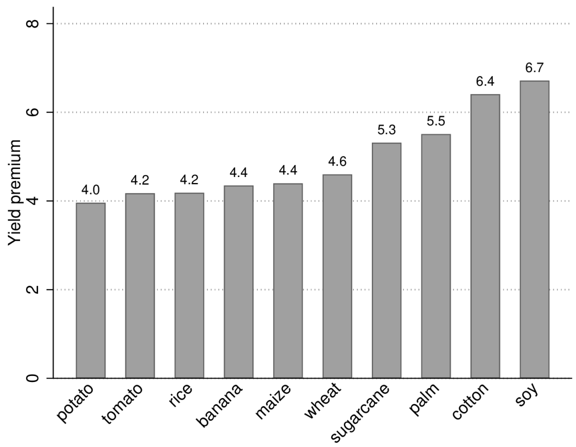
Notes: This figure plots for countries in year 2015 their agricultural cost share of inputs (panel a) and agricultural input expenditure per worker (panel b) against GDP per capita.

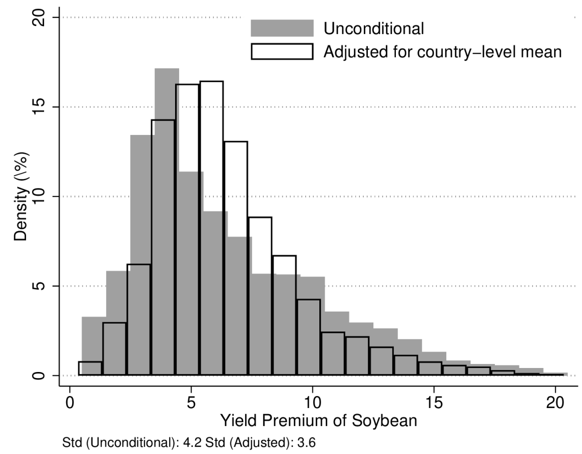
Empirical Pattern 3. The potential yield of the high-input (modern) technology over the low-input

Figure 4: Potential yield premium

a. Average across crops

b. Distribution across fields for soybean





Notes: Panel (a) shows the average premium of the modern technology across fields in the world. Panel (b) shows the distribution of the premium in the case of soybeans. Adjusted for country mean is computed as the premium at the field level plus the global average premium minus the the country-level average premium.

(traditional) technology is typically large and varies substantially across fields and crops.

Using field-level data on potential yields, we document that substantial productivity gains can be made by shifting the production towards input-intensive technologies. Figure 4 panel (a) shows the average yield premium of modern (high-input) technology over traditional (low-input) technology across fields in the world. Modern potential yield premia are, on average, in the range of four to seven across crops. For example, an average field around the world would yield 4.6 times more wheat using modern than traditional technology if the entire field was allocated to wheat. This average global premium is at minimum 4 in the case of potato, and at maximum 6.7 for soybean.

Focusing on the case of soybeans, Figure 4 panel (b) shows that modern potential yield premia vary substantially across the world geography. The global average premium for soybean hides the vast heterogeneity across fields and the relatively fat right tail of the distribution. Specifically, the 5th, 50th, and 95th percentile are 1.9, 5.5, and 14.9. Figure 4 panel (b) also shows that, even if we adjust the premium by the average in every country to control for between-country variations,[10] remarkable heterogeneity remains in the premia across fields.

[10]Our adjustment is given by: (field-level - country-level mean) + global mean.

From the empirical patterns to the theory

Our empirical patterns show that the global growth in agricultural productivity in the past few decades has coincided with a remarkable globalization in agricultural inputs, and that there are potentially large yield premia for the use of high-input agricultural technologies across fields. These observations motivate the formulation of a model in which adoption of input-intensive technologies can increase agricultural productivities and access to internationally supplied inputs encourage the adoption of these input-intensive technologies. In addition to the facts presented here, our modeling choices are motivated by two additional empirical patterns that are well-documented in the literature: the non-homotheticty in food consumption and the importance of land heterogeneity to study crop specialization.[11]

3 Model

3.1 Environment

The global economy consists of multiple countries, indexed by i or $n \in \mathcal{N}$. Each country n is endowed by a given supply of labor N_n, land L_n, and raw fertilizer V_n. Consumption combines sector-level bundles of nonagriculture and agriculture. The nonagriculture bundle consists of an outside good defined by a singleton $\mathcal{O} \equiv \{0\}$. The agriculture bundle comprises multiple crops, indexed by $k \in \mathcal{K}$. Every crop can be produced using a technique characterized by input and factor intensities. Specifically, technique is either traditional that uses only land and labor, or modern that uses labor, land, and multiple agricultural inputs indexed by $j \in \mathcal{J}$. We denote by $\mathcal{G}$ the set of all goods in the economy consisting of nonagriculture good, agricultural inputs, and crops,

$$\mathcal{G} \equiv \mathcal{O} \cup \mathcal{J} \cup \mathcal{K} = \Big\{ \underbrace{0}_{\text{nonagriculture}}, \underbrace{1, ..., J}_{\text{agricultural inputs } j \in \mathcal{J}}, \underbrace{J+1, ..., J+K}_{\text{crops } k \in \mathcal{K}} \Big\}$$

A set $\mathcal{F}_n$ of fields f, each with area L_n^f, characterizes the total land in country n, where $\sum_{f \in \mathcal{F}_n} L_n^f = L_n$. Our setup allows for differences in agroclimatic conditions at the level of fields, meaning that yields associated with producing a crop-technique pair (k, τ) are heterogeneous across fields $f \in \mathcal{F}_n$. Labor is homogeneous and freely mobile across productive

[11] See Comin, Lashkari, and Mestieri (2015) for an analysis of non-homothetic demands and Sotelo (2020) for evidence on the relationship between agricultural suitability and yields.

activities within countries. Endowments of fertilizers are used as an input in the production of processed fertilizers. All goods $g \in \mathcal{G}$ are tradeable, and markets are perfectly competitive.

3.2 Consumption and Trade

Every good $g \in \mathcal{G}$ is differentiated by the origin of production. We denote by $C_{ni,g}$ the consumption of good g in country n originated from country i, and by $C_{n,g}$ the aggregate consumption of good g as a CES combination of varieties across origin countries,

$$C_{n,g} = \left[\sum_{i \in \mathcal{N}} (b_{ni,g})^{1/\sigma_g} \left(C_{ni,g} \right)^{(\sigma_g - 1)/\sigma_g} \right]^{\sigma_g/(\sigma_g - 1)} \tag{1}$$

Here, $b_{ni,g}$ is a demand shifter, and $\sigma_g > 0$ is the elasticity of substitution of good g across countries (e.g. US corn vs Mexican corn). Sector-level bundles of consumption, C_n^s, with $s = 0$ for nonagriculture and $s = 1$ for agriculture, are

$$C_n^s = \begin{cases} C_{n,0} & \text{if } s = 0 \\ \left[\sum_{k \in \mathcal{K}} (b_{n,k})^{1/\kappa} (C_{n,k})^{(\kappa-1)/\kappa} \right]^{\kappa/(\kappa-1)} & \text{if } s = 1 \end{cases} \tag{2}$$

Here, $C_{n,0}$ and $C_{n,k}$ are given by equation (1), $b_{n,k}$ is a demand shifter, and $\kappa > 0$ is the elasticity of substitution across crops (e.g. corn vs wheat). The representative consumer in country n receives utility from the aggregate consumption, C_n, defined implicitly by the following non-homothetic CES representation,[12]

$$\sum_{s \in \{0,1\}} \left(b_n^s \right)^{\frac{1}{\eta}} \left(C_n \right)^{\frac{\varepsilon^s - \eta}{\eta}} \left(C_n^s \right)^{\frac{\eta - 1}{\eta}} = 1, \tag{3}$$

where b_n^s is a demand shifter for sector $s \in \{0, 1\}$; $\eta > 0$ is the elasticity of substitution across the consumption of nonagriculture C_n^0 and agriculture C_n^1 given by equation (2). $\varepsilon^s > 0$ is the elasticity of income with respect to sector s. If $\eta < 1$, agriculture and nonagriculture are complements; otherwise, they are substitutes. Sector s is a luxury if $\varepsilon^s > 1$, and a necessity if $\varepsilon^s < 1$. When $\varepsilon^s = 1$ for all s, the system collapses to CES preferences.

[12] This system of preferences has several appealing features, discussed in details in Comin, Lashkari, and Mestieri (2015).

International trade in every good $g \in \mathcal{G}$ is subject to iceberg trade costs. To deliver one unit of g from origin i to destination n, $d_{ni,g} \geq 1$ units must be shipped under triangle inequality. Price of g originated from i and destined at n is $p_{ni,g} = p_{i,g} d_{ni,g}$, where $p_{i,g}$ denotes the producer price.

3.3 Production

Every field $f \in \mathcal{F}_i$ consists of a continuum of plots $\omega \in f$. The agriculture production in field f involves allocating crops $k \in \mathcal{K}$, using techniques $\tau \in \mathcal{T}$, to plots $\omega \in f$. The production technology is given to a representative agricultural producer by

$$Q^f_{i,k\tau}(\omega) = \bar{q}_{k\tau} \left(z^f_{i,k\tau}(\omega) L^f_{i,k\tau}(\omega) \right)^{\gamma^L_{k\tau}} \left(N^f_{i,k\tau}(\omega) \right)^{\gamma^N_{k\tau}} \left(M^f_{i,k\tau}(\omega) \right)^{\gamma^M_{k\tau}} \tag{4}$$

Here, $\bar{q}_{k\tau}$ is a constant scalar,[13] $z^f_{i,k\tau}(\omega)$ is the land productivity of plot ω for producing crop k using technique τ, and $L^f_{i,k\tau}(\omega)$, $N^f_{i,k\tau}(\omega)$, and $M^f_{i,k\tau}(\omega)$ are respectively the use of land, labor, and material inputs. In addition, setting up every plot ω for agricultural use requires a fixed cost $z^f_{i,0}(\omega)$ paid in units of nonagriculture good. $\gamma^N_{k\tau} \in [0,1]$, $\gamma^M_{k\tau} \in [0,1]$, and $\gamma^L_{k\tau} = 1 - \gamma^N_{k\tau} - \gamma^M_{k\tau} \in [0,1]$ are, respectively, intensity parameters of labor, inputs, and land in production of crop k using technique τ. These intensity parameters characterize techniques which are either traditional $\tau = 0$ or modern $\tau = 1$.[14] The traditional technique, characterized by $\gamma^M_{k0} = 0$, is intensive in the use of labor and does not employ material inputs (i.e., $M^f_{i,k0}(\omega) = 0$). The modern technique, characterized by $\gamma^M_{k1} > 0$, employs material inputs. The aggregate input use $M^f_{i,k\tau}(\omega)$ is a Cobb-Douglas combination of agricultural inputs,

$$M^f_{i,k\tau}(\omega) = \prod_{j \in \mathcal{J}} \left(M^{j,f}_{i,k\tau}(\omega) \right)^{\lambda^j_k} \tag{5}$$

where $M^{j,f}_{i,k\tau}(\omega)$ is the use of input j with $\lambda^j_k \in [0,1]$ as the share parameter, and $\sum_{j \in \mathcal{J}} \lambda^j_k = 1$.

We now derive the rental price of every plot of land ω, which we denote by $r^f_{i,k\tau}(\omega)$. Let the producer price of good g in origin i be $p_{i,g}$, the consumer price index of g in destination i be $P_{i,g}$, and wage in country i be w_i. The price index of the bundle of agricultural inputs

[13] $\bar{q}_{k\tau} \equiv (\gamma^L_{k\tau})^{-\gamma^L_{k\tau}} (\gamma^N_{k\tau})^{-\gamma^N_{k\tau}} (\gamma^M_{k\tau})^{-\gamma^M_{k\tau}}$

[14] Our modeling allows for any arbitrary number of techniques. The choice of two is made only due to data availability.

in destination i is given by $m_{i,k} = \prod_{j \in \mathcal{J}} (P_{i,j})^{\lambda_k^j}$. By cost minimization, the unit cost of crop k using technique τ, $c_{i,k\tau}^f(\omega)$, equals

$$c_{i,k\tau}^f(\omega) = \Big(\frac{r_{i,k\tau}^f(\omega)}{z_{i,k\tau}^f(\omega)}\Big)^{\gamma_{k\tau}^L} \Big(w_i\Big)^{\gamma_{k\tau}^N} \Big(m_{i,k}\Big)^{\gamma_{k\tau}^M}$$

Since markets are perfectly competitive, net profits in every plot are pushed down to zero. Combining profit maximization and zero profit condition requires $c_{i,k\tau}^f(\omega) = p_{i,k}$. This delivers the gross rental price of land in plot ω, $r_{i,k\tau}^f(\omega)$, if assigned to crop-technique (k, τ),

$$r_{i,k\tau}^f(\omega) = z_{i,k\tau}^f(\omega) h_{i,k\tau} \tag{6}$$

$$\text{where} \quad h_{i,k\tau} = p_{i,k} \underbrace{\Big(\frac{w_i}{p_{i,k}}\Big)^{-\gamma_{k\tau}^N/\gamma_{k\tau}^L} \Big(\frac{m_{i,k}}{p_{i,k}}\Big)^{-\gamma_{k\tau}^M/\gamma_{k\tau}^L}}_{\widetilde{h}_{i,k\tau}} \tag{7}$$

Returns to crop-technique (k, τ) depends on land productivity $z_{i,k\tau}^f(\omega)$, and a component which we call $h_{i,k\tau}$ that summarizes the effect from market prices. The price-inclusive component, $h_{i,k\tau}$, rises in the output price $p_{i,k}$, and falls in the effective relative input price $\widetilde{h}_{i,k\tau}$. The latter term depends on wages and prices of material inputs relative to price of output, $w_i / p_{i,k}$ and $m_{i,k} / p_{i,k}$, with the extent of the relationship governed by intensities of labor and input use relative to land.

Fixed costs are investments in units of nonagriculture bundle, with price index P_i^0. We denote the net rental price of land in ω by $n_{i,k\tau}^f(\omega)$

$$n_{i,k\tau}^f(\omega) = z_{i,k\tau}^f(\omega) h_{i,k\tau} - z_{i,0}^f(\omega) P_i^0 \tag{8}$$

The optimal allocation in every plot $\omega \in f$ maximizes returns to plot ω by selecting among crop-technique pairs (k, τ), that is the one with the highest rents or by leaving the plot idle if no crop-technique pair delivers positive net rents,

$$\max \Big\{ z_{i,k\tau}^f(\omega) h_{i,k\tau} \text{ for all } (k, \tau),\ z_{i,0}^f(\omega) P_i^0 \Big\}$$

The vector of investment requirement and land productivities, $\mathbf{z}_i^f(\omega) \equiv [z_{i,k\tau}^f(\omega)$ for all $(k, \tau) \in$

$\mathcal{K} \times T$, $z^f_{i,0}(\omega)]$ is randomly drawn across plots $\omega \in f$ from a nested Fréchet distribution,

$$\Pr(\mathbf{z}^f_i(\omega) \leq \mathbf{z}^f_i) = \exp\left\{ -\bar{\phi}\left[\left(\Gamma_0(z^f_{i,0})\right)^{-\theta_1} + \sum_{k\in\mathcal{K}} \left(\Gamma_k(\mathbf{z}^f_{i,k})\right)^{-\theta_1}\right]\right\}$$
$$\text{where} \quad \Gamma_0(z^f_{i,0}) = \left(\frac{z^f_{i,0}}{a^f_{i,0}}\right), \quad \Gamma_k(\mathbf{z}^f_{i,k}) = \left[\sum_{\tau\in\mathcal{T}} \left(\frac{z^f_{i,k\tau}}{a^f_{i,k\tau}}\right)^{-\theta_2}\right]^{-\frac{1}{\theta_2}} \text{ for all } k \in \mathcal{K}$$

Here, $\bar{\phi} \equiv \left[\Gamma(1 - 1/\theta_1)\right]^{-\theta_1}$ is a normalization to ensure that $\mathbb{E}[z^f_{i,0}(\omega)] = a^f_{i,0}$, and $\mathbb{E}[z^f_{i,k\tau}(\omega)] = a_{i,k\tau}$. Our formulation generalizes a standard Fréchet distribution as the one in Eaton and Kortum (2002) by relaxing the assumption that productivity draws across alternatives are independent. We achieve this extension by building on tools from the literature on discrete choice based on generalized extreme value distributions, as studied in detail in McFadden (1981). We present a detailed derivation in the appendix, and explain the intuition below.

This generalized Fréchet distribution allows productivity draws to be correlated in a structured way. In the upper nest, θ_1 controls the dispersion of land productivity draws across crops. The higher θ_1, the less heterogeneous the land productivity draws across crops within a field. Consequently, producers will be more responsive in substituting across crops when relative returns to crops change. In the lower nest, θ_2 controls the dispersion of productivity draws across techniques within every crop. The larger θ_2 relative to θ_1 is, the larger the correlation between draws are across techniques within a crop. Given a choice of crop, at a higher θ_2 producers are more responsive in adopting a technology when returns to that technology rise. All together, θ_1 and θ_2 govern the pattern of specialization respectively regarding which crop to grow and with which technique to grow it.

When $\theta_2 > \theta_1 \geq 1$, then productivity draws between corn-traditional and corn-modern are more similar compared to draws between corn and wheat. Setting $\theta_1 = \theta_2$ brings the model back to a one-nest Fréchet distribution where the correlation between draws across techniques within a crop is zero. Then, for example, draws between corn-modern and corn-traditional are equally dissimilar to draws between corn-modern and wheat-traditional. In the other special case, where $\theta_2 \to \infty$, there will be perfect correlation between draws across technologies within a crop. As a result, every crop will be produced using only one technology. [15]

[15] In agriculture-related studies, this one-nest version has been used for Roy-type models in labor markets (Lagakos and Waugh, 2013), and in land allocation problems (Costinot, Donaldson, and Smith, 2016; Sotelo, 2020). In the trade literature, recent applications that allow for correlations include Lashkaripour and Lugovskyy (2018) and Lind and Ramondo (2018).

Lastly, we specify the production technology of non-crop goods consisting of nonagriculture and agricultural inputs. Among agricultural inputs, production of processed fertilizer, denoted by $v \in \mathcal{J}$, uses the domestic endowments of raw fertilizers, V_i. The production of other non-crop goods (nonagriculture and non-fertilizer inputs) uses labor. Specifically,

$$Q_{i,g} = \begin{cases} A_{i,v} V_i, & \text{fertilizer, } g = v \\ A_{i,g} N_{i,g}, & \text{nonagriculture \& other agricultural inputs, } g \in \mathcal{O} \cup \mathcal{J}, g \neq v \end{cases} \tag{9}$$

where production features constant returns to scale, and $A_{i,g}$ is a vector of productivity shifters.

3.4 Equilibrium

3.4.1 Prices and Expenditures

Let E_n be total expenditure in country n. Price indexes of consumption aggregates $C_{n,g}$, C_n^s, and C_n for all $g \in \mathcal{G} \equiv \mathcal{O} \cup \mathcal{J} \cup \mathcal{K}$, $s = \{0, 1\}$, $n \in \mathcal{N}$ are given by

$$P_{n,g} = \Big[\sum_{i \in \mathcal{N}} b_{ni,g} (p_{i,g} d_{ni,g})^{1-\sigma_g} \Big]^{\frac{1}{1-\sigma_g}} \tag{10}$$

$$P_n^s = \begin{cases} P_{n,0}, & \text{if } s = 0 \\ \Big[\sum_{k \in \mathcal{K}} b_{n,k} (P_{n,k})^{1-\kappa} \Big]^{\frac{1}{1-\kappa}}, & \text{if } s = 1 \end{cases} \tag{11}$$

$$P_n = \Big[\sum_{s \in \{0,1\}} b_n^s (E_n / P_n)^{\varepsilon^s - 1} (P_n^s)^{1-\eta} \Big]^{\frac{1}{1-\eta}} \tag{12}$$

We denote by $\beta_{ni,g}$ the share of expenditure by country n on good $g \in \mathcal{G}$ originated from i, by $\beta_{n,k}$ the share of expenditure by country n on crop $k \in \mathcal{K}$ relative to aggregate agriculture expenditure there, and by β_n^s the share of expenditure by country n on sector-level bundles

We complement these studies by illustrating how to apply the tools to Roy-type or land-use problems, and in addition we derive several new expressions related to productivity distributions conditional on selection.

of nonagriculture and agriculture,

$$\beta_{ni,g} = \frac{b_{ni,g}(p_{i,g}d_{ni,g})^{1-\sigma_g}}{(P_{n,g})^{1-\sigma_g}} \tag{13}$$

$$\beta_{n,k} = \frac{b_{n,k}(P_{n,k})^{1-\kappa}}{(P_n^1)^{1-\kappa}} \tag{14}$$

$$\beta_n^s = \frac{b_n^s(E_n/P_n)^{\varepsilon^s-1}(P_n^s)^{1-\eta}}{(P_n)^{1-\eta}} \tag{15}$$

The price effects operate through substitutions in the upper tier between nonagriculture and agriculture through $(P_n^s/P_n)^{1-\eta}$, in the middle tier between crops within agriculture through $(P_{n,k}/P_n^1)^{1-\kappa}$, and in the lower tier between varieties of different origin countries within a crop through $(p_{ni,k}/P_{n,k})^{1-\sigma_k}$.

The income effect operates through $(E_n/P_n)^{\varepsilon^s-1}$ in the upper tier with respect to nonagriculture ($s = 0$) and agriculture ($s = 1$) bundles of consumption. Equation (15) shows that expenditure shares β_n^0 and β_n^1 depend on real total expenditure, E_n/P_n. In the empirically relevant case, where $\varepsilon^0 > \varepsilon^1$, a rise in E_n/P_n increases the share of expenditures on nonagriculture.

Our measure of welfare is utility C_n received by total consumption as implicitly defined by equation (3). Our derivation ensures that welfare is given by $C_n = E_n/P_n$. The overall price index, P_n, is itself a function E_n/P_n. Therefore, equation (12) implicitly defines the price index, P_n, to be solved at any level of income and sector-level prices. The pair of equations (12) and (15) characterize the non-homotheticity in demand, i.e. how the price index and expenditure shares change by income.

3.4.2 Agricultural Output and Land Allocation at the Field Level

For every field f, we denote the fraction of land allocated to crop-technique (k, τ) by $\pi_{ni,k}^f$. Further, let $\alpha_{i,k}^f$ be the fraction of land allocated to crop k, and $\alpha_{i,k\tau}^f$ be the fraction of land within crop k allocated to technique τ. The land shares are given by

$$\pi_{i,k\tau}^f = \alpha_{i,k}^f \times \alpha_{i,k\tau}^f \tag{16}$$

where

$$\alpha^f_{i,k\tau} = \frac{\left(a^f_{i,k\tau} h_{i,k\tau}\right)^{\theta_2}}{(H^f_{i,k})^{\theta_2}} \tag{17}$$

$$\alpha^f_{i,k} = \frac{(H^f_{i,k})^{\theta_1}}{(a^f_{i,0} P^0_i)^{\theta_1} + \sum_{k \in \mathcal{K}} (H^f_{i,k})^{\theta_1}} \tag{18}$$

Here, $h_{i,k\tau}$ is the price-inclusive component of returns to crop-technique pair (k, τ), given by equation (7), and aggregate returns to crop k, $H^f_{i,k}$, equals

$$H^f_{i,k} = \left[\sum_{\tau \in \mathcal{T}} \left(a^f_{i,k\tau} h_{i,k\tau}\right)^{\theta_2}\right]^{\frac{1}{\theta_2}} \tag{19}$$

Equations (16)–(19) connect the dispersion parameters of the Fréchet distribution to elasticities of land use. Specifically, θ_2 appears as the elasticity of substitution across techniques within a crop choice, and θ_1 as the elasticity of substitution in land use across crops (and no agriculture). The opportunity cost of agriculture production, $a^f_{i,0} P^0_i$, pins down the share of cropland. Within the cropland, land share of crop k increases in its average returns $H^f_{i,k}$, with the extent of the relationship governed by θ_1. Within the land allocated to crop k, the land share of technique τ rises in average returns to technique τ, $a^f_{i,k\tau} h_{i,k\tau}$, with the extent of the relationship disciplined by θ_2.

Let $\Omega^f_{i,k\tau}$ be the set of plots ω in field f to which crop-technique (k, τ) is optimally allocated. Conditional on optimal selections, the average productivity of crop-technique (k, τ) in field f equals

$$\mathbb{E}[z^f_{i,k\tau}(\omega) \mid \omega \in \Omega^f_{i,k\tau}] = a^f_{i,k\tau} (\alpha^f_{i,k})^{-\frac{1}{\theta_1}} (\alpha^f_{i,k\tau})^{-\frac{1}{\theta_2}} \tag{20}$$

The conditional mean productivity of crop-technique (k, τ), given by equation (20), is greater than the unconditional mean productivity, $\mathbb{E}[z^f_{i,k\tau}(\omega)] = a^f_{i,k\tau}$, due to the selection of a crop-technique pair (k, τ) if its productivity draw is sufficiently large. A higher share of land is allocated to crop k if returns to crop k, $H^f_{i,k}$, rise relative to those of other crops and opportunity cost of agriculture. This margin of adjustment, governed by θ_1, determines the pattern of specialization across crops.

In addition, conditional on selecting crop k, higher returns to technique τ, $a^f_{i,k\tau} h_{i,k\tau}$, increase the share of land for which technique τ is adopted among the competing techniques to produce crop k. Using equation (7) and equation 17, the relative share of modern to traditional technique, conditional on producing crop k, satisfies

$$\left(\frac{\alpha^f_{i,k1}}{\alpha^f_{i,k0}}\right) = \left[\underbrace{\frac{a^f_{i,k1}}{a^f_{i,k0}}}_{\text{relative avg productivity}} \underbrace{\frac{(w_i/p_{i,k})^{-\gamma^N_{k1}/\gamma^L_{k1}}(m_{i,k}/p_{i,k})^{-\gamma^M_{k1}/\gamma^L_{k1}}}{(w_i/p_{i,k})^{-\gamma^N_{k0}/\gamma^L_{k0}}(m_{i,k}/p_{i,k})^{-\gamma^M_{k0}/\gamma^L_{k0}}}}_{\text{relative input price}} \right]^{\theta_2}.$$

The term in the brackets equals the unconditional expected return of modern technique relative to traditional, which rises when productivity of modern technique rises on average relative to traditional, and when relative prices of inputs fall. The extent to which these changes imply a change in relative land share of modern technique is governed by θ_2. The larger θ_2 is, the greater the extent of adopting modern technique in response to changes in relative productivities and input prices.

Discrete choices of crop-technique pairs for every plot ω implies that

$$Q^f_{i,k\tau}(\omega) = \begin{cases} (\gamma^L_{k\tau})^{-1}\widetilde{h}_{i,k\tau} z^f_{i,k\tau}(\omega), & \omega \in \Omega^f_{i,k\tau} \\ 0, & \omega \notin \Omega_{i,k\tau} \end{cases} \tag{21}$$

The optimal allocation requires each plot $\omega \in f$ either not to be used for agriculture or fully used for the production of one crop using one technique.[16] At the field level, aggregate output of crop k using technique τ in field f within country i, $Q^f_{i,k\tau}$, equals land use, $\pi^f_{i,k\tau} L^f_i$, times average production per plot, $\mathbb{E}[Q^f_{i,k\tau}(\omega) \mid \omega \in \Omega^f_{i,k\tau}]$. Using equations (16), (20), (21),

$$\begin{aligned} Q^f_{i,k\tau} &= \pi^f_{i,k\tau} L^f_i \times \mathbb{E}\left[Q^f_{i,k\tau}(\omega) \mid \omega \in \Omega^f_{i,k\tau}\right] \\ &= L^f_i (\gamma^L_{k\tau})^{-1} \widetilde{h}_{i,k\tau} a^f_{i,k\tau} (\alpha^f_{i,k})^{\frac{\theta_1 - 1}{\theta_1}} (\alpha^f_{i,k\tau})^{\frac{\theta_2 - 1}{\theta_2}} \end{aligned} \tag{22}$$

3.4.3 Agricultural Output at the Country Level

Aggregate output of crop k in country i is then the sum across techniques and fields there,

$$Q_{i,k} = \sum_{f \in \mathcal{F}_i} \sum_{\tau \in \mathcal{T}} Q^f_{i,k\tau}. \tag{23}$$

[16] In addition, returns to land in plot ω, $h_{i,k\tau} z^f_{i,k\tau}(\omega)$, are only a fraction $\gamma^L_{k\tau}$ of output, with the other fraction paid to labor and material inputs.

Aggregate quantity of nonagriculture good required for fixed costs of setting up plots is denoted by S_i and equals

$$S_i = \sum_{f\in\mathcal{F}_i} L_i^f a_{i,0}^f \Big[1 - \Big(1 - \sum_{k\in\mathcal{K}} \alpha_{i,k}^f\Big)^{(\theta_1-1)/\theta_1}\Big]. \tag{24}$$

3.4.4 Market Clearing and General Equilibrium

Labor market clearing in every country $i \in \mathcal{N}$ requires labor supply N_i to equal labor demand from production of nonagriculture, non-fertilizer agricultural inputs, and crops,

$$w_i N_i = \sum_{g\in\mathcal{O}\cup\mathcal{J}, g\neq v} p_{i,g} Q_{i,g} + \sum_{k\in\mathcal{K}} \sum_{f\in\mathcal{F}_i} \sum_{\tau\in\mathcal{T}} \gamma_{k\tau}^N p_{i,k} Q_{i,k\tau}^f \tag{25}$$

Goods market clearing for nonagriculture, agricultural inputs $j \in \mathcal{J}$ (including fertilizers), and crops $k \in \mathcal{K}$ require supply at the origin country to equal world demand,

$$p_{i,0} Q_{i,0} = \sum_{n\in\mathcal{N}} \beta_{ni,0} \beta_n^0 E_n + P_i^0 S_i \tag{26}$$

$$p_{i,j} Q_{i,j} = \sum_{f\in\mathcal{F}_i} \sum_{n\in\mathcal{N}} \sum_{k\in\mathcal{K}} \beta_{ni,j} \lambda_k^j \gamma_{k1}^M p_{n,k} Q_{n,k1}^f \tag{27}$$

$$p_{i,k} Q_{i,k} = \sum_{n\in\mathcal{N}} \beta_{ni,k} \beta_{n,k} \beta_n^1 E_n \tag{28}$$

Finally, by national accounting of flows total expenditure in country i, E_i, equals the sum of factor rewards and trade deficits,

$$E_i = \sum_{k\in\mathcal{K}} \sum_{f\in\mathcal{F}_i} \sum_{\tau\in\mathcal{T}} (\gamma_{k\tau}^N + \gamma_{k\tau}^L) p_{i,k} Q_{i,k\tau}^f - P_i^0 S_i + \sum_{g\in\mathcal{O}\cup\mathcal{J}} p_{i,g} Q_{i,g} + D_i \tag{29}$$

The first term net of the second term in the RHS equals payments to labor and land in agriculture. The third term is payments to labor in nonagriculture and agricultural inputs as well as revenues from fertilizer sales, and the fourth term is trade deficits. Equations 25-29 guarantee that trade deficits sum up to zero, $\sum_{i\in\mathcal{N}} D_i = 0$, and land market clearing condition holds.

We close the layout of our model by defining the global economy and general equilibrium.

Definition. For all countries $n, i \in \mathcal{N}$, fields $f \in \mathcal{F}_n$, goods $g \in \mathcal{G}$ consisting of nonagriculture, agricultural inputs $j \in \mathcal{J}$, and crops $k \in \mathcal{K}$, sectors $s \in \{0,1\}$, and techniques $\tau \in \mathcal{T}$, a **global economy** is characterized by

- Endowments $\mathcal{E} \equiv \{L_n^f, N_n, V_n, D_n\}$;
- Production elasticity parameters $\Theta_{\mathcal{A}} \equiv \{\theta_1, \theta_2\}$;
- Consumption elasticity parameters $\Theta_{\mathcal{B}} \equiv \{\varepsilon^0, \varepsilon^1, \eta, \kappa, \sigma_g\}$;
- Production shifters $\mathcal{A} = \{\gamma_{k\tau}^L, \gamma_{k\tau}^M, \gamma_{k\tau}^N, \lambda_k^j, a_{n,0}^f, a_{n,k\tau}^f, A_{n,g}\}$;
- Consumption shifters $\mathcal{B} = \{b_n^s, b_{n,k}, b_{ni,g}, d_{ni,g}\}$

Definition. Given a global economy characterized by $\{\mathcal{E}, \mathcal{A}, \mathcal{B}, \Theta_{\mathcal{A}}, \Theta_{\mathcal{B}}\}$, a **general equilibrium** consists of prices $\{p_{n,g}\}$ for all $n \in \mathcal{N}$, $g \in \mathcal{G}$ such that equations 7–29 hold.

3.5 The Production Possibility Frontier in each Field

In this section, we show that the allocation of land at the field level predicted by our model is equivalent to an alternative problem in which landowners in each field produce efficiency units of land subject to a production possibility frontier. For a given field f in country i, the aggregate problem of a landowner is given by:

$$\max_{\{\widetilde{L}_{i,k\tau}^f\}_{k,\tau}\ ,\ \{\widetilde{L}_{i,k}^f\}_k} \quad \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}} h_{i,k\tau} \widetilde{L}_{i,k\tau}^f$$

$$\text{subject to} \quad \left[\sum_{\tau \in \mathcal{T}} (\widetilde{L}_{i,k\tau}^f / a_{i,k\tau}^f)^{\frac{\theta_2}{\theta_2 - 1}} \right]^{\frac{\theta_2 - 1}{\theta_2}} \leq \widetilde{L}_{i,k}^f \tag{30}$$

$$\left[\sum_{k \in \mathcal{K}} (\widetilde{L}_{i,k}^f)^{\frac{\theta_1}{\theta_1 - 1}} \right]^{\frac{\theta_1 - 1}{\theta_1}} \leq L_i^f \tag{31}$$

Here, $\widetilde{L}_{i,k\tau}^f$ and $\widetilde{L}_{i,k}^f$ are efficiency units of land at the level of crop-technique $k\tau$ and crop k.[17] The landowner maximizes the sum of returns across uses of land given price-inclusive term $h_{i,k\tau}$ described by equation (7), technology coefficients $a_{i,k\tau}^f$, and endowment L_i^f.[18]

[17] Efficiency units $\widetilde{L}_{i,k\tau}^f$ immediately deliver production quantities $Q_{i,k\tau}^f$ according to: $Q_{i,k\tau}^f = (1/\gamma_{k\tau}^L)\widetilde{h}_{i,k\tau}\widetilde{L}_{i,k\tau}^f$, where as defined by equation (7), $\widetilde{h}_{i,k\tau} = (w_i/p_{i,k})^{-\gamma_{k\tau}^N/\gamma_{k\tau}^L}(m_{i,k}/p_{i,k})^{-\gamma_{k\tau}^M/\gamma_{k\tau}^L}$ is the effective relative input price.

[18] For the sake of exposition, here we have set the value of the outside option at zero.

We illustrate this problem with diagrams for two crops, which we call rice and wheat. To save on notation, we drop country and field indicators. The production possibility frontiers are represented in two tiers. The lower tier reflects substitution possibilities across techniques within a crop, and the upper tier disciplines substitution possibilities between crops. Figure 5 illustrates the frontier along the dimension of technology within a crop (Panel a), and between crops (Panel b).

In Panel (a), we show for every crop k the optimal choices of output in units of land efficiency that are produced using traditional ($\tau = 0$) and modern ($\tau = 1$) techniques. The maximum that could be achieved if all resources for the production of crop k was allocated to technology τ is given by $a_{k\tau}\widetilde{L}_k$. This maximum value depends on technology coefficients $a_{k\tau}$ as well as aggregate efficiency units allocated to crop k, $\widetilde{L}_k$, that is a choice variable in the upper tier. The slope of the frontier curve at point $(\widetilde{L}_{k0}, \widetilde{L}_{k1})$ is proportional to $(\widetilde{L}_{k0}/\widetilde{L}_{k1})^{1/(\theta_2-1)}$, governed by $\theta_2 \in (1, \infty)$. The smaller θ_2, the greater the curvature, the less elastic choices of technology for a given change in market conditions.[19] The slope of the iso-value line in turn equals h_{k0}/h_{k1}, which incorporates the effects from relative wages and input prices adjusted by relative labor and input intensities.

In Panel (b), we show the upper tier of production choices that represents the substitution possibilities between rice and wheat. The slope of the frontier at point $(\widetilde{L}_{rice}, \widetilde{L}_{wheat})$ equals $(\widetilde{L}_{rice}/\widetilde{L}_{wheat})^{1/(\theta_1-1)}$, that is governed by $\theta_1 \in (1, \infty)$. A smaller θ_1 means a greater curvature, hence a lower sensitivity in substitution across crops if relative prices change.[20] In addition, the slope of the iso-value line is given by $(-H_{rice}/H_{wheat})$. Reproducing H_k from equation (19), it is a generalized mean of $a_{k\tau}h_{k\tau}$ across technologies within every crop, $H_k = \left[\sum_\tau (a_{k\tau}h_{k\tau})^{\theta_2}\right]^{\frac{1}{\theta_2}}$. Therefore, crop-level returns that are taken into account in the upper tier depend on optimal decisions made in the lower tier. Moreover, the maximum efficiency units of land that can be allocated to crop k equals total area of land. This maximum value is not greater than total land area because the selection margin raises average land productivity only if a fraction of land, not the entire area of it, is allocated to a crop.

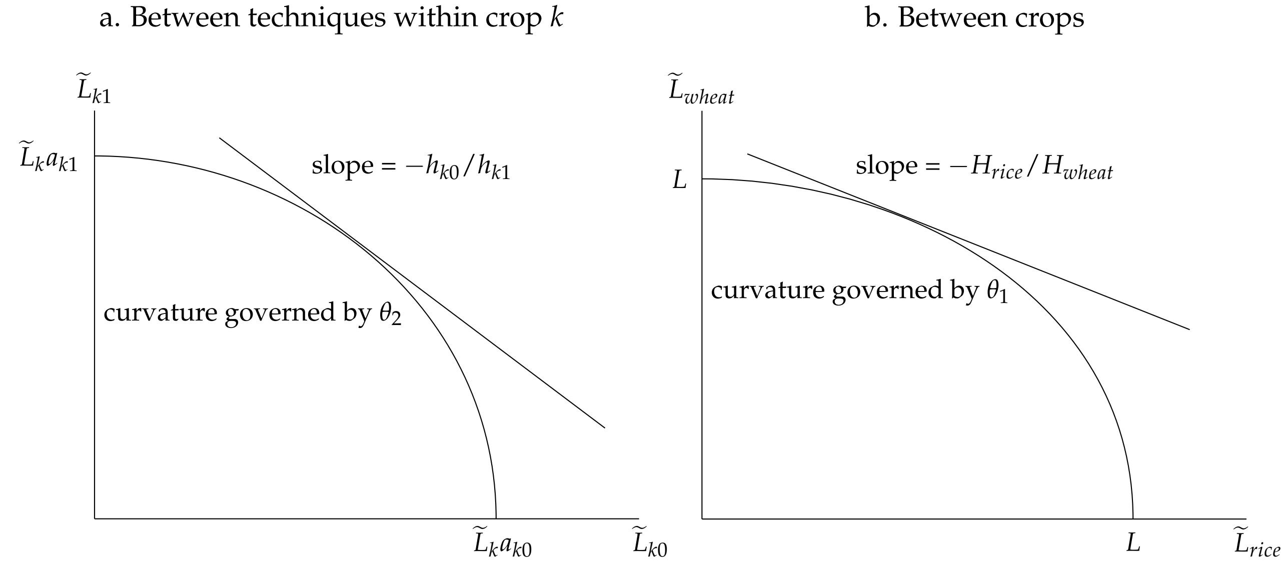
Lastly, it is remarkable that shadow prices of this aggregate problem replicate land rents predicted by our micro-founded model. Specifically, we derive in the appendix that the Lagrange multiplier associated with the slack constraints (30) and (31) are respectively given

[19] In one extreme where $\theta_2 \to \infty$, the frontier is a straight line, and the problem has a corner solution reflecting that choices of technology are extremely sensitive to relative prices. In the other extreme where $\theta_2 \to 1$, the frontier collapses to a right angle, and the optimal choice becomes insensitive to prices.

[20] Similarly, if $\theta_1 \to \infty$, the producer problem has a corner solution, and if $\theta_1 \to 1$, the optimal choice of $(\widetilde{L}_{rice}, \widetilde{L}_{wheat})$ becomes insensitive to price changes.

by H_k and $[\sum_k H_k^{\theta_1}]^{1/\theta_1}$. That is, the shadow price of the land allocated to crop k is H_k, which is the average returns to land used for production of crop k, and the shadow price of the entire cropland is given by $[\sum_k H_k^{\theta_1}]^{1/\theta_1}$, which is precisely the average rents of cropland.[21]

Figure 5: Production Possibility Frontier



Notes: Panel (a) shows the lower-tier production possibility frontier within crop k between two technologies, 1 as modern and 0 as traditional. Panel (b) shows the upper-tier production possibility frontier between two crops. $\{\widetilde{L}_{k\tau}, \widetilde{L}_k\}$ are in units of land efficiency. In Panel (a) the slope of the curve is proportional to $-(\widetilde{L}_{k0}/\widetilde{L}_{k1})^{1/(\theta_2-1)}$, and the maximum quantity of $\widetilde{L}_{k\tau}$ is $a_{k\tau}\widetilde{L}_k$ where $\widetilde{L}_k$ is the choice variable in the upper tier. In Panel (b), the slope of of the curve equals $-(\widetilde{L}_{rice}/\widetilde{L}_{wheat})^{1/(\theta_1-1)}$, and $H_k = \left[\sum_\tau (a_{k\tau}h_{k\tau})^{\theta_2}\right]^{\frac{1}{\theta_2}}$ for $k \in \{rice, wheat\}$. The maximum quantity of $\widetilde{L}_k$ is L.

4 Taking the Model to Data

In this section, we take our model to data. We first quantify country-level parameters following standard practices in the literature. We then estimate parameters that control the shape of field-level production possibility frontiers using FAO-GAEZ data on potential yields of crop-technique pairs for more than 1.3 million fields across countries.

[21] For more details and full derivations for this aggregate problem, see Appendix (C.5).

4.1 Country-level Production and Trade Parameters

At the country-level, we calibrate demand elasticities (κ, and σ_g for $g \in \mathcal{G}$), demand shifters ($b_{n,k}$ for $k \in \mathcal{K}$, and $b_{ni,g} d_{ni,g}^{1-\sigma_g}$ for $g \in \mathcal{G}$), productivities in non-agricultural and agricultural inputs ($A_{i,g}$ for $g \in \{0, 1, ..., J\}$) and upper-tier demand elasticities and shifters (η, ε^s, b_n^s for $s = \{0, 1\}$).

Demand Elasticities and Shifters for Crops (κ, σ_k, $b_{ni,k} d_{ni,k}^{1-\sigma_k}$). Following Costinot, Donaldson, and Smith (2016), we estimate the elasticity of substitution between crops (σ_k) from the following gravity-type equation,

$$\log\left(\frac{X_{ni,k}}{X_{n,k}}\right) = \delta_{n,k} + (1 - \sigma_k) \log p_{i,k} + \epsilon_{ni,k}, \tag{32}$$

where $\delta_{n,k} \equiv -\log[\sum_i b_{ni,k}(p_{i,k} d_{ni,k})^{1-\sigma_k}]$, $\epsilon_{ni,k} = \log b_{ni,k} d_{ni,k}^{1-\sigma_k}$, $X_{ni,k}$ is the purchases of n from country i of crop k, and $X_{n,k}$ is total purchases of country n of crop k. Without loss of generality we set $\sum_{i=1}^{N} \epsilon_{ni,k} = 0$. We recover $b_{ni,k} d_{ni,k}^{1-\sigma_k}$ from the residuals.[22] As in Costinot, Donaldson, and Smith (2016), we impose the same elasticity of substitution between crops ($\sigma_k = \sigma$). Due to potential correlations between demand shocks $\epsilon_{ni,k}$ and prices $\log p_{n,k}$, we instrument $\log p_{i,k}$ with the average agricultural suitability of the exporting country, using FAO-GAEZ data. With our estimate of σ_k, we recover demand shifters of crops from the residuals of equation (32) and construct the country-level price index of every crop $P_{n,k}$. With this price index in hand, we take logs of equation (14) and estimate the elasticity of substitution between crops, κ, based on

$$\log\left(\frac{X_{n,k}}{X_n^1}\right) = \delta_n + (1 - \kappa) \log P_{n,k} + \epsilon_{n,k} \tag{33}$$

where $\epsilon_{n,k} = \log b_{n,k}$, $\sum_{k \in \mathcal{K}} \epsilon_{n,k} = 0$, and X_n^1 is aggregate purchases of all crops. To address the endogeneity of $\log P_{n,k}$, we instrument the price index using the average agricultural suitability of each country.

Demand Elasticities and Shifters for Non-agriculture and Agricultural Inputs (σ_g, $b_{ni,g} d_{ni,g}^{1-\sigma_g}$, $g = 0, 1, ..., J$). We set $\sigma_g = 4$ for non-agriculture and the three categories of agricultural inputs

[22] We do not separately identify trade costs from preference shifters. All we require to compute our general equilibrium is the combined bilateral frictions deterring trade between countries. However, for the purpose of our counterfactual exercise in Section 5, we will also estimate trade costs.

based on the literature, and estimate

$$\log\left(\frac{X_{ni,g}}{X_{n,g}}\right) - (1-\sigma_g)\log w_i = \delta_{n,g} + \delta_{i,g} + \epsilon_{ni,g}, \tag{34}$$

where $\delta_{i,g}$ and $\delta_{n,g}$ are origin and destination fixed effects, and demand shifters are given by $b_{ni,g}d_{ni,g}^{1-\sigma_g} = \exp(\epsilon_{ni,g})$.

Productivities in Non-agriculture and in Agricultural Inputs ($A_{i,g}$). We recover productivity parameters $A_{n,g}$ from the origin fixed effect in equation (34).

Upper-tier Demand Parameters (η, ε^0, ε^1, b_n^0, b_n^1). We set income elasticities at $\varepsilon^0 = 1.5$ and $\varepsilon^1 = 0.5$, and the substitution elasticity between agriculture and nonagriculture at $\eta = 0.5$ in line with estimates from Comin, Lashkari, and Mestieri (2015). These parameters imply that agriculture is a necessity good whereas nonagriculture is a luxury, and that agriculture and nonagriculture are complements in consumption. Within our estimation, as we explain below, we recover upper tier of demand shifters (b_n^0 and b_n^1) consistent with model-implied price indices of nonagriculture P_n^0 and agriculture (P_n^1).

4.2 Field-level Agriculture Production Technology

We now quantify parameters of agriculture production technology at the field level. We calibrate factor and input shares ($\gamma_{k\tau}^L, \gamma_{k\tau}^N, \gamma_{k\tau}^M$ and λ_k^j), then estimate the production elasticity of substitution between crops (θ_1) and between technologies within a crop (θ_2). Within this estimation we calibrate field-level land productivity shifters for crop-technology pairs ($a_{i,k\tau}^f$), and the investment parameter ($a_{n,0}^f$). We first explain our calibration, then our estimation.

4.2.1 Calibration

Factor and Input Shares ($\gamma_{k\tau}^L, \gamma_{k\tau}^N, \gamma_{k\tau}^M$ and λ_k^j). We set the factor shares in the agricultural production function using a few data sources together with equations implied by our model. Since the traditional technique uses no material input, we set $\gamma_{k0}^M = 0$. Due to data limitations, we make the assumption that factor shares of land, labor, and material inputs for a technology, $\gamma_{k\tau}^L, \gamma_{k\tau}^N, \gamma_{k\tau}^M$ are common across crops. In addition, with an approximation, our model implies that

$$(\gamma_0^L/\gamma_1^L) = \frac{(\bar{\gamma}_j^M \widetilde{\alpha}_j) - (\bar{\gamma}_i^M \widetilde{\alpha}_i)}{\bar{\gamma}_j^M - \bar{\gamma}_i^M}$$

where i and j refer to any two countries, $\widetilde{\alpha}_i = (1 - \bar{\alpha}_{i,1})/\bar{\alpha}_{i,1}$ is relative land share of traditional to modern technology at the aggregate in country i, and $\bar{\gamma}_i^M$ is the aggregate cost share of agricultural inputs in country i. Using data for these aggregate variables for the United States and Brazil, we calibrate the factor share of land in traditional relative to modern technology, γ_0^L/γ_1^L. In addition to $\bar{\gamma}_{USA}^M$, we collect data from the USDA Commodity Costs and Returns for aggregate cost shares of land $\bar{\gamma}_{USA}^L$ and of labor $\bar{\gamma}_{USA}^L$, which we use to recover all factor and inputs shares. We also obtain the cost share of fertilizers, pesticides, and farm machinery across crops from USDA Commodity Costs and Returns. Table (A.2) reports our calibrated factor and input shares, $(\gamma_\tau^L, \gamma_\tau^N, \gamma_\tau^M)_\tau$ and the median of λ_k^j across crops for the three categories j of agricultural inputs.

Productivity Parameters at the Field Level ($\alpha_{i,k}^f$ and $a_{i,0}^f$). Given demand elasticities and shifters $(\Theta_{\mathcal{B}}, \mathcal{B})$, intensity parameters $(\gamma_{k\tau}^L, \gamma_{k\tau}^M, \gamma_{k\tau}^N, \lambda_k^j)$, and production elasticities (θ_1, θ_2), we calibrate $(a_{i,k\tau}^f, a_{i,0}^f)$ such that all equilibrium relationships hold, and the model exactly matches calibration targets which we describe below. This calibration depends on the choice of (θ_1, θ_2). As such, our calibration of $(a_{i,k\tau}^f, a_{i,0}^f)$ will be nested within our estimation of (θ_1, θ_2).[23]

We first explain how the field-level data on agricultural productivity maps into our model. Define $y_{i,k\tau}^{f,FAO}$ as the potential yield in FAO-GAEZ data. By construction, $y_{i,k\tau}^{f,FAO}$ is the yield of crop k using technique τ if the land of the entire field f were allocated to crop-technique (k, τ). According to our model, this is the quantity we would obtain by setting $\alpha_{i,k}^f = \alpha_{i,k\tau}^f = 1$ in equation (22) divided by the entire field area L_i^f, which equals $(\gamma_{k\tau}^L)^{-1}\widetilde{h}_{ik\tau}a_{i,k\tau}^f$. Remember that here $\widetilde{h}_{ik\tau}$ captures the role of input and output prices in shaping land productivity, which in the FAO-GAEZ data are *not* meant to reflect country-level market conditions. The FAO-GAEZ data can be interpreted in our model as the potential yield of a field if there were no geographic trade costs and therefore no spatial differences in output and input prices. Given these remarks, we assume the following structure for $y_{i,k\tau}^{f,FAO}$:

$$y_{i,k\tau}^{f,FAO} = (\gamma_{k\tau}^L)^{-1}\widetilde{h}_{k\tau}^{FAO}a_{i,k\tau}^f$$

where $\widetilde{h}_{k\tau}^{FAO}$ is consistent with some vector of global prices and $a_{i,k\tau}^f$ is the productivity shifter that we seek to recover to construct our PPFs. Isolating $a_{i,k\tau}^f$ in the equation above, we write:

$$a_{i,k\tau}^f = \delta_{k\tau} \times y_{i,k\tau}^{f,FAO}$$

[23] This procedure is computationally intensive, and results in this section are currently based on five percent sample of fields randomly drawn for each country. We will update our results for the full sample in the next version of our paper.

where $\delta_{k\tau} \equiv \gamma^L_{k\tau}/\widetilde{h}^{FAO}_{k\tau}$ is an unobserved scale parameter. We calibrate $\delta_{k\tau}$ according to aggregate data on agricultural production to recover $a^f_{i,k\tau}$. Noticed that, by doing so, the ratio of productivity shifters in our model will reflect directly the ratios of productivities in the FAO-GAEZ data, as $a^f_{i,k\tau}/a^{f'}_{i,k\tau} = y^{f,FAO}_{i,k\tau}/y^{f',FAO}_{i,k\tau}$ for any field f and f'. We consider two calibration targets to pin down the $2 \times K$ unobserved scale parameters $\delta_{k\tau}$: the supply quantity of every crop k in the USA, and the aggregate share of agricultural land in the USA allocated to modern technology for every crop k.

In addition to the agricultural productivity parameters, we also calibrate investment intensity parameter ($a^f_{n,0}$) by matching field-level data on the share of cropland which we obtain from the EarthStat data.

In what follows, we express the calibration problem as $\mathbf{c}(\Theta_{\mathcal{A}}) = 0$, reflecting that the calibration of production technology parameters depends on $\Theta_{\mathcal{A}}$.

4.2.2 Estimation

We estimate production elasticities $\Theta_{\mathcal{A}} = (\theta_1, \theta_2)$ using generalized method of moments conditional on the calibration problem given by $\mathbf{c}(\Theta_{\mathcal{A}}) = 0$. In the calibration problem, we solve, at a given value of $\Theta_{\mathcal{A}}$, the general equilibrium and calibrate parameters of production technology as explained in the previous section. We continue by describing the moments that we construct for our estimation.

First, we construct a moment condition based on model predictions for yields and data on actual yields. Let $\hat{y}^f_{i,k}$ be the actual yield of crop k in field f in country i. Our specification allows for measurement error in the actual yield data:

$$y^f_{i,k} = \hat{y}^f_{i,k}\epsilon^f_{i,k},$$

where $y^f_{i,k}$ is the true value of the actual yield, and $\epsilon^f_{i,k}$ is an error term. One of our identification assumptions requires that log of the error term is orthogonal to predicted log of land share, $\mathbb{E}[\ln \epsilon^f_{i,k} \times \ln \alpha^f_{i,k}] = 0$. We construct this moment using data on actual yields across fields within the United States, where we have more reliable data. The intuition behind this orthogonality condition is that, controlling for land productivity shifters $(a^f_{i,k\tau})$, land shares are sufficient statistics for yields. Our identification assumption requires that deviations of predicted yields from data on actual yields are not explained by land shares. In other words, all information that is required for the model to predict yields is already summarized in land shares. Since we use field-level data on actual yields from the United States where modern technology is by far the dominant choice, the informative variations are expected to be for choices across crops. As such, this moment condition is tightly related to the identification of θ_1. We define the resulting moment as m_1.

We construct two additional moment conditions based on cost shares of agricultural inputs across countries. As we presented in the empirical pattern 2, countries with a higher GDP per capita have higher cost shares of agricultural inputs (Fig. 3-a). Note that, if we assumed a Cobb-Douglas technology with the same factor shares across countries, that would force all countries in the model to have the same cost share of inputs, sharply contradicting this empirical pattern. Our second set of moments is designed to capture this cross-country heterogeneity in the cost share of inputs. In our model, the elasticity of substitution between technologies (θ_2) controls how responsive agricultural producers are to relative wages and input prices. Higher wages and lower input prices induce agricultural producers to shift toward modern technologies, which raises the share of input costs in agriculture. If θ_2 is higher, the model generates larger differences in aggregate costs shares of agricultural inputs across low-income and high-income countries. Alternatively, if θ_2 approaches 1, then the cost share of inputs across countries will be negligible. Given this intuition, we use the mean cost share of agricultural inputs in the two upper deciles and the two lower decides of income per capita across countries as targeted moments. We call these moments m_2.

Stacking the moment conditions, we have $g(\Theta_{\mathcal{A}}) = [m_1(\Theta_{\mathcal{A}}), m_2(\Theta_{\mathcal{A}})] - [m_1^{data}, m_2^{data}]$. Our estimation procedure is then based on the following moment condition

$$\mathbb{E}(g(\Theta_{\mathcal{A}})) = 0$$

We seek values of $\widehat{\Theta}_{\mathcal{A}} = (\widehat{\theta}_1, \widehat{\theta}_2)$ that achieves

$$\begin{aligned} &\widehat{\Theta}_{\mathcal{A}} = \arg\min_{\Theta_{\mathcal{A}}} \; g(\Theta_{\mathcal{A}}) g(\Theta_{\mathcal{A}})' \\ &\text{subject to} \quad \mathbf{c}(\Theta_{\mathcal{A}}) = 0. \end{aligned}$$

4.3 Estimation Results

Table 2 presents results from the estimation of the model. On the demand side, we have estimated the elasticity of substitution for crops across supplying countries, σ_k, at 5.1; and the elasticity of substitution across crops at 3.1. We have taken the rest of demand elasticities

Table 2: Estimation of the Key Elasticities of the Model

Description	Parameter	Method	Value
A. Demand Side			
- Elast of subst across crops	κ	IV	3.1
- Elast of subst across origins within a crop	σ_k	IV	5.1
B. Supply Side			
- Elast of subst across crops	θ_1	GMM	2.05
- Elast of subst across technologies within a crop	θ_2	GMM	4.38

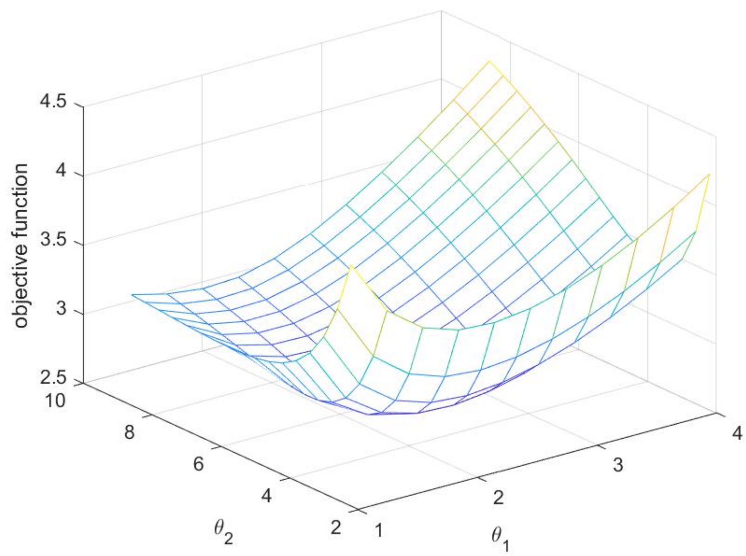
Notes: This table shows estimation results for the demand side and the supply side. See Section 4.1 and 4.2 for the details on our estimation procedure.

from the literature, as explained in Section 4.1.

On the supply side, our estimation implies a production elasticity of substitution across crops, θ_1, equal to 2.05, and a production elasticity of substitution across technologies within a crop, θ_2, equal to 4.38. Our estimate of θ_1 is in the range suggested by the literature. Using variations in crop outputs across countries, Costinot, Donaldson, and Smith (2016) estimate this elasticity at 2.6. Using variations in land shares and prices across regions within Peru, Sotelo (2020) estimate this elasticity at 1.6. Using farm-level data from Uganda, Bergquist, Faber, Fally, Hoelzlein, Miguel, and Rodriguez-Clare (2019) estimate a range of elasticities between 1.8 and 2.9. To the best of our knowledge, we are the first to estimate θ_2, so we do not have a benchmark value for comparison. Our estimates imply that agricultural producers are more responsive in substituting between technologies within a choice of crops, than substituting between crops.

To illustrate the identification of θ_1 and θ_2, we conduct a grid search exercise. Specifically, we show in Figure 6 the surface of our GMM objective function evaluated at a wide range of (θ_1, θ_2). The figure demonstrates a concave function with our estimates as its *global* minimum.

Figure 6: Identification of Productivity Dispersion Parameters (θ_1 and θ_2).



Notes: This figure shows the objective function of the GMM procedure used to estimate θ_1 and θ_2, where θ_1 captures the dispersion of productivities between crops and θ_2 captures the dispersion of productivities between technologies within a crop. The values at which the objective function is minimized are $\theta_1 = 2.05$ and $\theta_2 = 4.38$.

4.4 Goodness of Fit

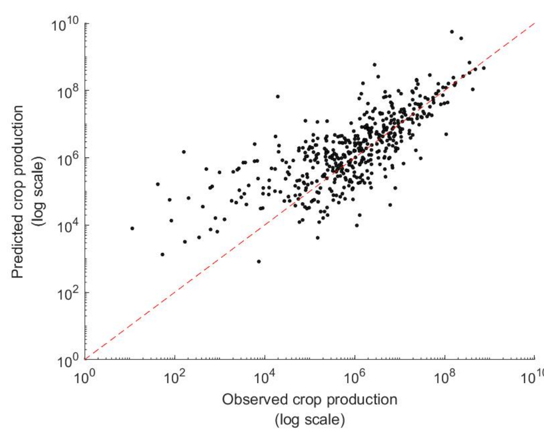
Different from quantitative frameworks that perfectly fits production data by treating productivities as residuals, we have the opportunity to construct PPFs from field-level on potential yields without forcing a perfect fit of the model with data. In other words, we let our model fail to fit the baseline data on agricultural production. Before using our model for counterfactual analyses, we therefore check whether the model has a good fit with endogenous variables that will be central in our counterfactual analysis, such as production quantities, employment and trade in agriculture. In doing so, we also discuss model's predictions about technology use.

Figure 7 compares the log quantity of crop outputs at the level of countries predicted by the model versus data. In our calibration, we calibrated land productivities ($a^f_{i,k1}$ and $a^f_{i,k0}$) using potential yields from FAO-GAEZ, where we used data on production of crops only from the United States. Therefore, the main factors driving the predictions in Figure 7 are our calibrated land productivities as well our estimated elasticities. The model predictions line up reasonably well with the data.

Figure 8(a) shows a scatter plot of the aggregate share of land allocated to modern technology across countries against their GDP per capita. The model generates substantial variation in the share of land allocated to modern technologies. Low-income countries such as Ethiopia, Ghana, and Burkina Faso have shares of their land in modern technologies that are very low, which is consistent with the extremely low use of inputs observed in the data. The share of land in modern technology is typically in the range of 40% to 90% in middle-income countries such as Brazil, Thailand, and Turkey, and around 90% to 100% in high-income countries such as USA, Japan, and West and North European countries. Our calibration requires the US share to be at 95% in line with US agriculture data on the use of agricultural inputs, but the remaining points on the scatter plot are predictions of the model. There are two key mechanisms driving these cross-country differences in technology choices. First, since traditional technologies are labor-intensive, low-income countries tend to use traditional technologies due to lower wages. Second, due to import barriers, low-income countries tend to face higher input prices, discouraging the use of modern technologies.

Figure 8(b) shows the share of labor in agriculture predicted by the model against data. Note that agricultural employment is entirely an out-of-sample variable since we did not target it in our calibration or estimation. The variation and the slope that is generated by the model is reasonably close to the data. The model under-predicts the agricultural employment share compared to the data.

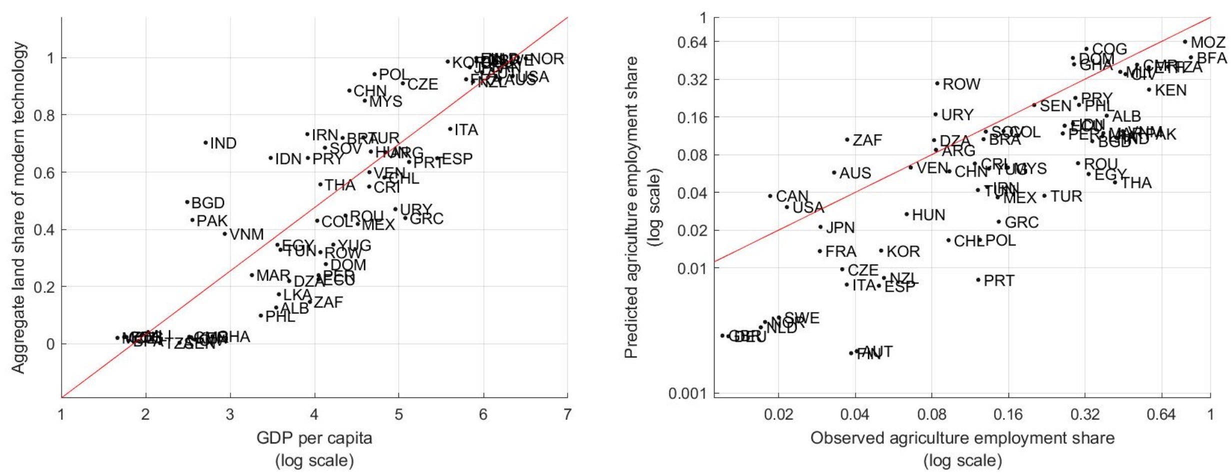
Figure 7: Fit of the Model with respect to Crop Output



Notes: This figure shows the model fit of crop outputs at the level of countries. The red line shows the 45 degree line. Both variables are represented in logs. Each data point represents a country-crop pair.

Figure 8: Model Predictions for Technology and Employment Share

a. Share of Land Allocated to Modern Technology b. Share of Agricultural Employment



Notes: This figure shows the predictions of the model with respect to share of land allocated to modern technologies and share of employment in agriculture.

5 Impact of Globalization on Agricultural Productivity

Equipped with our estimated model that allows for adjustments in crop specialization and technology choice, we now examine the quantitative contribution of different mechanisms through which globalization shaped agricultural productivity. In particular, we ask what would be the effects of bringing trade costs between countries to their 1980 levels on agricultural productivity, food consumption, and welfare around the world. We separately examine the reductions in trade costs in agricultural output and in agricultural inputs. We present our results as counterfactuals relative to baseline. Before presenting our simulations, we first describe our procedure to estimate changes to trade costs between 1980 and 2015.

5.1 Estimating Changes to Trade Costs between 1980 and 2015

To estimate changes to trade costs, we follow a common approach in the trade literature (see Head and Mayer (2014)). We assume that the trade cost component of the demand shifters ($d_{ni,g}$) is symmetric and normalize the trade cost of a country with itself to one, $d_{ii,g} = 1$. We estimate these trade costs for agricultural outputs by combining data on crops, and for agricultural inputs by combining data on fertilizers, pesticides, and farm machinery. Putting these assumption together, our model gives

$$\log\left(\frac{X_{in,g}}{X_{ii,g}} \times \frac{X_{in,g}}{X_{nn,g}}\right) = \underbrace{2(1-\sigma_g)\log(d_{ni,g})}_{\delta_{in,g}} + \epsilon_{in,g}$$

where $\epsilon_{in,g} = \log(b_{ni,g}b_{in,g}/b_{ii,g}b_{ni,g})$. We estimate $\delta_{in,g}$ as fixed effects using bilateral expenditure and trade data, and then recover $d_{ni,g}$, for agricultural outputs and inputs in 1980 and 2015.

Figure A.2 shows percentage changes to trade costs between 1980 and 2015 for agricultural outputs and inputs, aggregated by regions. We find an average reduction of trade costs (weighted by trade flows of 2015) of 21% for outputs and by 35% for inputs. The fall in trade costs of agricultural inputs is typically larger than that of agricultural output. For both cases changes to trade costs are substantially heterogeneous across countries. For example on the input side, while the average trade costs in the sub-Saharan Africa fell by 5%, in Latin America they fell by 15% and in East Europe by 45%.

5.2 Counterfactual Simulations of the Impact of Globalization

We simulate a counterfactual in which trade costs of agricultural outputs and inputs are set at their 1980 levels, and all other remaining parameters are kept unchanged. Specifically, let $\Delta_{ni,g}$ be the percentage change in trade cost $d_{ni,g}$ from 2015 to 1980. We compute counterfactual demand shifters as $b_{ni,g}(\Delta_{ni,g}d_{ni,g})^{(1-\sigma_g)}$ which we use to simulate a counterfactual equilibrium of our model.

Table A.3 shows the general equilibrium effects from these changes on yields, and Table A.4 presents the resulting effects on selected variables including food consumption and welfare. Yields fall by $2.4 - 9.0\%$ across crops at the global level. The loss in yields is associated with a fall in the imported share of expenditure on crops and on agricultural inputs, which falls respectively by 16.08% and 52.26%, at the global scale. In addition, the global share of land allocated to modern, input intensive technologies decreases by 7.49%. As a result of these changes, the consumer price index of agriculture rises by a global average of 8.19%, contributing to 2.83% reduction in food consumption and 1.63% reduction in welfare at the global level.

The welfare effects from moving to the counterfactual with trade costs of 1980 are largely heterogeneous across countries. Countries that lose the most include some of those in East Europe such as Hungary and Poland, in Latin America such as Colombia and Mexico, in Asia such as Bangladesh and Malaysia. In contrast, a number of countries in Sub-Saharan Africa, Middle East, and Latin America appear to be the ones that lose the least or even gain in terms of food consumption and welfare.

To study the channels at work, we report results for additional counterfactual simulations where we change trade costs only in agricultural outputs and only in agricultural inputs. Table 3 shows a summary of results these counterfactuals at the global level. First, we find that changes in trade costs in agricultural output and in agricultural input have comparable effects on global welfare (-0.87% and -0.75% respectively). Therefore, not accounting for imports of agricultural inputs and their effect on technology choice would largely underestimate productivity gains from globalization in agriculture. The channels through which these two types of reductions in trade cost operate are completely different. Table 3 indicates that globalization in agricultural inputs has a minor contribution to changes in international crop specialization, as reflected by the changes in the import share of agricultural outputs. Moreover, the table also shows that globalization in agricultural output has a minor effect on input intensification, as captured by the changes in the share of land allocated to modern agricultural technologies.

Table 3: Welfare Effects of the Absence of Globalization at the Global Scale

	Counterfactual effects from changes in trade costs of		
	Both ag outputs and inputs	only ag outputs	only ag inputs
Import share on ag output	−16.1%	−15.4%	−1.3%
Share of land in modern tech	−7.5%	0.0%	−7.5%
Welfare	−1.63%	−0.87%	−0.75%

Notes: This table shows the percentage change to imported share of expenditure on agricultural outputs (crops), share of land allocated to modern technology, and welfare at the global scale, from the baseline of 2015 to the counterfactual with trade costs of 1980 for both agricultural outputs and inputs, only agricultural outputs, and only agricultural inputs.

We continue to explain these distinct channels of effects that operate through the output side (comparative advantage) and input side (technology) in more details.

Globalization Only in Agricultural Outputs

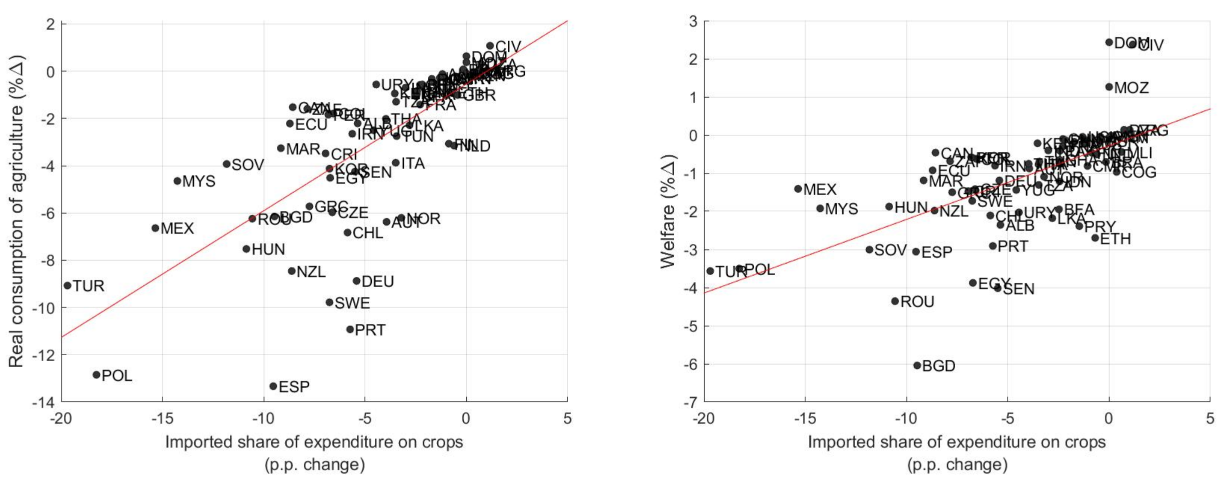
As it has been discussed in the trade literature, the effects of trade openness in agricultural output on food consumption and welfare operate through comparative advantage forces and their implications for international crop specialization. To isolate the role of this mechanism, we run a counterfactual in which we only bring trade costs in agricultural output to their levels in 1980, while leaving trade costs in agricultural inputs at their levels in 2015.

For a wide class of trade models, including the Armington structure that we adopt here, changes to the share of expenditures on domestic goods (or equivalently, on imports) is a sufficient statistic for changes to welfare in response to a trade shock (see Arkolakis, Costinot, and Rodriguez-Clare (2012)). Figure 9 shows that percentage changes in food consumption and welfare are indeed strongly associated with percentage changes to the imported share of expenditure on crops. In addition, these figures show large differences in effects across countries. The magnitudes of welfare loss from the move to the counterfactual with trade costs of 1980 can be as large as 13% for Poland in terms of food consumption, and 6% for Bangladesh in terms of welfare. At the global level, there would be 1.13% reduction in real agricultural consumption and 0.87% reduction in welfare.

Figure 9: Welfare effects of globalization in agricultural outputs against imported share of expenditures on crops

a. Food Consumption

b. Welfare



Notes: This figure shows percentage changes to food consumption and welfare across countries from the baseline of 2015 to the counterfactual in which trade costs of agricultural outputs are set to their 1980 levels.

Globalization Only in Agricultural Inputs

We now explore the efficiency gains generated by globalization in agricultural inputs. In doing so, we show its distinct way of generating efficiency gains compared to the output side mechanisms. In particular, gains from globalization in agricultural inputs are associated with a higher adoption of modern, input-intensive technologies in agricultural production. To spell out this mechanism, we run a counterfactual in which we bring trade costs in only agricultural inputs to their levels in 1980, while leaving trade costs in agricultural output at their levels in 2015. In a sequence of figures, we report the chain of effects from agricultural input prices to agricultural productivity, and from there to consumption and welfare.

First, as a result of the overall increase in trade costs of agricultural inputs, the price index of agricultural inputs at the location of use would rise for many countries. A higher price of inputs makes agricultural producers allocate a larger share of land to traditional technologies. Figure (10) shows that the larger the increase in the price of agricultural inputs, the larger the extent to which land is allocated to traditional technologies in the counterfactual relative to the baseline. Due to general equilibrium effects, the agricultural input price index does not necessarily rise for all countries. To understand this result, consider the case of Malaysia and Iran. Between 1980 and 2015, due to the global fall of trade costs in agricultural inputs, the global demand rose for agricultural inputs, increasing the price of these inputs at the *location of production*. In this period, trade costs in inputs reduced on average by around 38% for Malaysia, but only by an average of 7% for Iran. Putting these together, the price of agricultural inputs in the counterfactual compared to the baseline is on average 38% *higher* for Malaysia, but on average 9% *lower* for Iran. Consequently, the aggregate share of land allocated to modern technology is 16.5% smaller for Malaysia, and 4.7 percentage point larger for Iran. A takeaway from this example is that countries that fell behind in the process of globalization ended up facing higher prices of agricultural inputs, creating a barrier for their shift toward modern, input-intensive technologies.

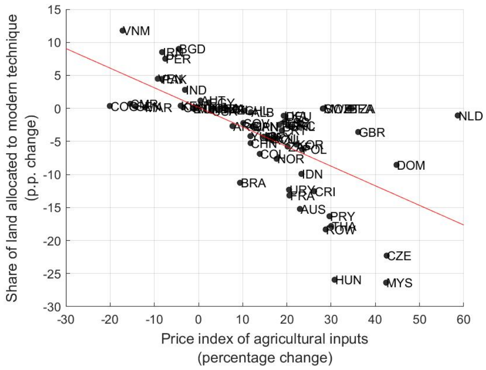
We now turn our attention to examining how shifting from traditional to modern technologies can bring about higher land productivities. Figure (10) shows, for selected crops, the predicted change to yields against change to the share of land allocated to modern technologies. By and large across crops and countries, land productivities (yields) are smaller in the counterfactual with trade costs of 1980 compared to the baseline. Across countries within a crop, yields are systematically lower, the smaller the land share of modern technology in the counterfactual outcome compared to the baseline. As reported in Table (A.3), at the global scale the counterfactual yields are 3 to 8 percent lower across crops, indicating notable global gains in agricultural land productivity as a result of globalization in agricultural inputs.

In addition, Figure (11) shows counterfactual changes for measures of input intensity in agricultural production. The smaller the land share of modern technology in the counterfactual compared to the baseline, the smaller the cost share of inputs and the smaller agricultural output per worker. Hence, globalization in agricultural inputs contributes to agricultural input-intensity through the choice of modern technologies.

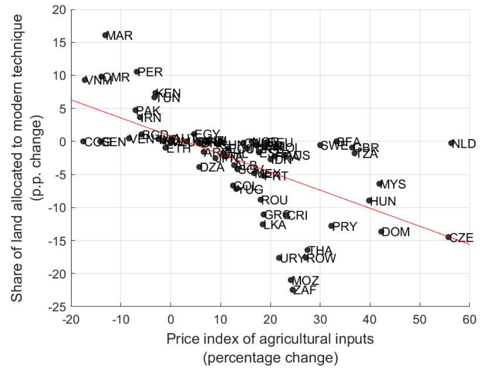
Lastly, we report the subsequent changes to agricultural prices, consumption, and welfare. Figure (11) presents the scatter plot of percentage changes to real agricultural consumption (Panel c) and welfare (Panel d). Comparing the counterfactual with trade costs of 1980 to the baseline, real consumption of agriculture tends to be systematically smaller across countries, the lower the land share of modern technology. Due to lower agricultural productivities in the counterfactual outcome, consumers in most countries face higher price index of agricultural good. As a result of this price effect, consumers would consume less food. However, since real income tends to be lower in the counterfactual outcome, the income effect requires a higher demand for food through non-homothetic preferences. Overall, these interactions would result in 1.75% reduction in the global food consumption and 0.75% reduction in the global welfare.

Figure 10: Effects of Globalization of Agricultural Inputs on Selected Crops

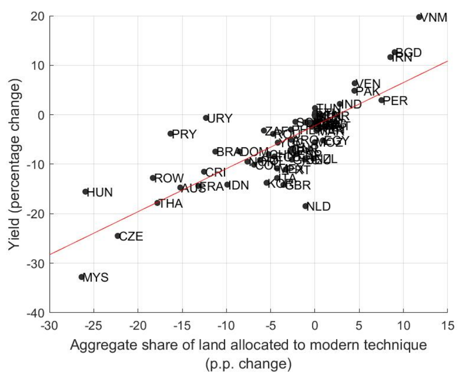
a. Corn



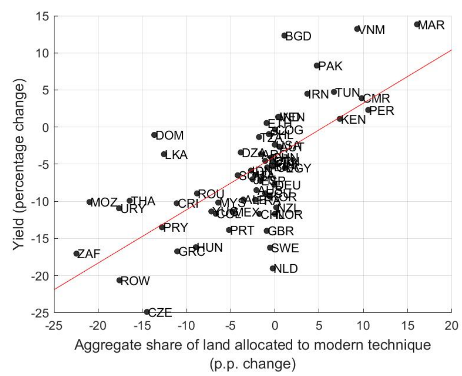
b. Wheat



c. Corn

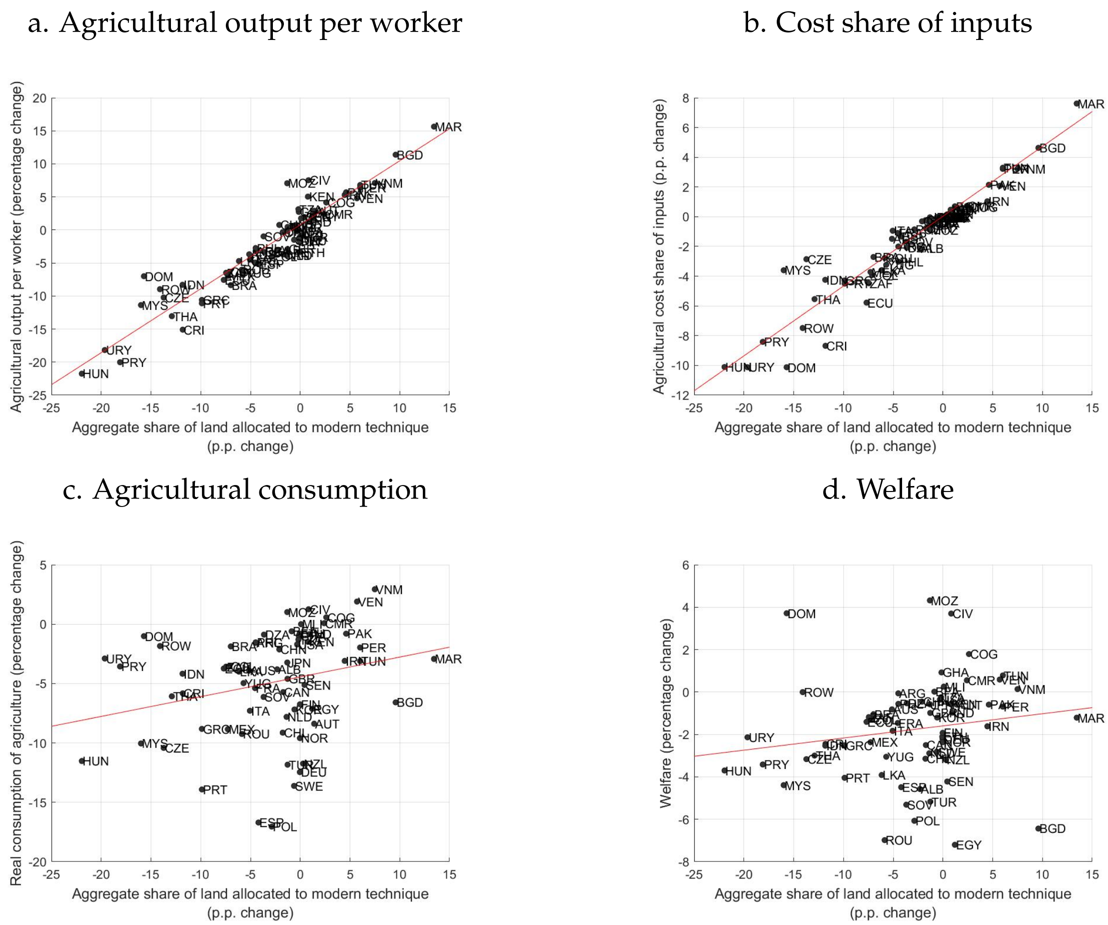


d. Wheat



Notes: This figure shows how changes in key variables for selected crops if there had been no reductions in trade cost in agricultural inputs between 1980 and 2015.

Figure 11: Land share allocated to modern technique vs Selected measures of productivity and welfare



Notes: This figure shows the relationship between changes in agricultural consumption and the share of land allocated to modern technology, as well as the relationship between changes in welfare and the share of land allocated to modern technology in the absence of the reductions in trade cost in agricultural inputs between 1980 and 2015.

6 Conclusion

In this paper, we evaluated the impact of globalization on agricultural productivity across the world geography. We developed a quantifiable, multi-country general equilibrium model of trade with an extremely rich spatial structure that incorporates choices of crops and technologies in agricultural production. We connected our model to data on agricultural productivity for crops and technologies at the level of fields covering the surface of the earth. We found large productivity gains at the global scale, with notable distributional gains across countries. We found that in addition to well-studied welfare effects from the output side of agriculture, the effects from the fall of trade costs in agricultural inputs on agricultural productivity have been remarkable.

Previous research has either evaluated the effects of international trade on agricultural productivity without considering the role of input use and technology choice, or the effects of input use and technology choice on agricultural productivity without considering the role of international trade. In this paper we showed that, because agricultural inputs are largely procured in international markets, international trade has a crucial role in shaping agricultural productivity through facilitating the adoption of modern agricultural technologies.

Finally, new and richer high-resolution datasets are becoming available at the intersections of natural and social sciences. We take a step forward in incorporating such data into a theoretical framework that can be used for a wide range of applications. Integrating these types pf micro-level data into economic models appears as a promising direction for future research, particularly with applications to resources and environment.

References

ALLEN, T., AND D. ATKIN (2016): "Volatility and the Gains from Trade," *Working Paper*.

ANDERSON, K. (2010): "Globalization's effects on world agricultural trade, 1960–2050," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554), 3007–3021.

——— (2016): *Agricultural trade, policy reforms, and global food security*. Springer.

ARKOLAKIS, C., A. COSTINOT, AND A. RODRIGUEZ-CLARE (2012): "New trade models, same old gains," *American Economic Review*, 102(1), 94–130.

ARKOLAKIS, C., N. RAMONDO, A. RODRÍGUEZ-CLARE, AND S. YEAPLE (2018): "Innovation and production in the global economy," *American Economic Review*, 108(8), 2128–73.

BALDOS, U. L., T. W. HERTEL, AND F. C. MOORE (2019): "Understanding the spatial distribution of welfare impacts of global warming on agriculture and its drivers," *American Journal of Agricultural Economics*, 101(5), 1455–1472.

BERGQUIST, L. F., B. FABER, T. FALLY, M. HOELZLEIN, E. MIGUEL, AND A. RODRIGUEZ-CLARE (2019): "Scaling Agricultural Policy Interventions: Theory and Evidence from Uganda," *Working Paper*.

CALIENDO, L., AND F. PARRO (2015): "Estimates of the trade and welfare effects of NAFTA," *Review of Economic Studies*, 82(1), 1–44.

CASELLI, F. (2005): "Accounting for cross-country income differences," *Handbook of economic growth*, 1, 679–741.

COMIN, D. A., D. LASHKARI, AND M. MESTIERI (2015): "Structural change with long-run income and price effects," Discussion paper, National Bureau of Economic Research.

COSTINOT, A., AND D. DONALDSON (2014): "How Large Are the Gains from Economic Integration: Theory and Evidence from U.S. Agriculture, 1880-1997," *Working paper*.

COSTINOT, A., D. DONALDSON, AND C. SMITH (2016): "Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world," *Journal of Political Economy*, 124(1), 205–248.

COSTINOT, A., AND A. RODRÍGUEZ-CLARE (2014): "Trade theory with numbers: Quantifying the consequences of globalization," in *Handbook of international economics*, vol. 4, pp. 197–261. Elsevier.

DERYNG, D., D. CONWAY, N. RAMANKUTTY, J. PRICE, AND R. WARREN (2014): "Global crop yield response to extreme heat stress under multiple climate change futures," *Environmental Research Letters*, 9(3), 034011.

DONALDSON, D. (2018): "Railroads of the Raj: Estimating the impact of transportation infrastructure," *American Economic Review*, 108(4-5), 899–934.

DONOVAN, K. (2017): "Agricultural risk, intermediate inputs, and cross-country productivity differences," *Working Paper*.

EATON, J., AND S. KORTUM (2002): "Technology, geography, and trade," *Econometrica*, 70(5), 1741–1779.

FAJGELBAUM, P., AND S. J. REDDING (2019): "External integration, structural transformation and economic development: Evidence from argentina 1870-1914," Discussion paper, Working Paper.

FALLY, T., AND J. SAYRE (2018): "Commodity Trade Matters," *Working Paper*.

FARROKHI, F. (2020): "Global sourcing in oil markets," *Journal of International Economics*, p. 103323.

FEENSTRA, R. C., R. E. LIPSEY, H. DENG, A. C. MA, AND H. MO (2005): "World trade flows: 1962-2000," Discussion paper, National Bureau of Economic Research.

FOLEY, J. A., N. RAMANKUTTY, K. A. BRAUMAN, E. S. CASSIDY, J. S. GERBER, M. JOHNSTON, N. D. MUELLER, C. OCONNELL, D. K. RAY, P. C. WEST, ET AL. (2011): "Solutions for a cultivated planet," *Nature*, 478(7369), 337.

GAFARO, M., AND H. PELLEGRINA (2019): "Trade, Farmers' Heterogeneity and Agricultural Productivity: Evidence from Colombia," *Working Paper*.

GOLLIN, D., D. LAGAKOS, AND M. E. WAUGH (2014): "The Agricultural Productivity Gap," *The Quarterly Journal of Economics*, 129(2), 939–993.

GOLLIN, D., S. L. PARENTE, AND R. ROGERSON (2007): "The food problem and the evolution of international income levels," *Journal of Monetary Economics*, 54(4), 1230–1255.

GOUEL, C., AND D. LABORDE (2018): "The crucial role of international trade in adaptation to climate change," Discussion paper, National Bureau of Economic Research.

HEAD, K., AND T. MAYER (2014): "Gravity equations: Workhorse, toolkit, and cookbook," in *Handbook of international economics*, vol. 4, pp. 131–195. Elsevier.

HERTEL, T. (2013): "Global applied general equilibrium analysis using the global trade analysis project framework," in *Handbook of Computable General Equilibrium Modeling*, vol. 1, pp. 815–876. Elsevier.

HERTEL, T. W. (2002): "Applied general equilibrium analysis of agricultural and resource policies," *Handbook of agricultural economics*, 2, 1373–1419.

LAGAKOS, D., AND M. E. WAUGH (2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *The American Economic Review*, 103(2), 948–980.

LASHKARIPOUR, A., AND V. LUGOVSKYY (2018): "Scale economies and the structure of trade and industrial policy," *Working Paper*.

LEE, D. R., AND P. G. HELMBERGER (1985): "Estimating supply response in the presence of farm programs," *American Journal of Agricultural Economics*, 67(2), 193–203.

LIND, N., AND N. RAMONDO (2018): "Trade with correlation," Discussion paper, National Bureau of Economic Research.

MCARTHUR, J. W., AND G. C. MCCORD (2017): "Fertilizing growth: Agricultural inputs and their effects in economic development," *Journal of development economics*, 127, 133–152.

MCFADDEN, D. (1981): "Econometric models of probabilistic choice," *Structural analysis of discrete data with econometric applications*, 198272.

MONFREDA, C., N. RAMANKUTTY, AND J. A. FOLEY (2008): "Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000," *Global biogeochemical cycles*, 22(1).

MUELLER, N. D., J. S. GERBER, M. JOHNSTON, D. K. RAY, N. RAMANKUTTY, AND J. A. FOLEY (2012): "Closing yield gaps through nutrient and water management," *Nature*, 490(7419), 254.

NIEDERTSCHEIDER, M., T. KASTNER, T. FETZEL, H. HABERL, C. KROISLEITNER, C. PLUTZAR, AND K.-H. ERB (2016): "Mapping and analysing cropland use intensity from a NPP perspective," *Environmental Research Letters*, 11(1), 014008.

PELLEGRINA, H. S. (2019): "Trade and the Spatial Organization of Agriculture: Evidence from Brazil ," *Working Paper*.

PORTEOUS, O. (2016): "High trade costs and their consequences: An estimated dynamic model of African agricultural storage and trade," *American Economic Journal: Applied Economics*.

——— (2020): "Trade and agricultural technology adoption: Evidence from Africa," *Journal of Development Economics*, p. 102440.

SOTELO, S. (2020): "Domestic Trade Frictions and Agriculture," *Forthcoming in Journal of Political Economy*.

TOMBE, T. (2015): "The Missing Food Problem: Trade, Agriculture, and International Productivity Differences," *American Economic Journal: Macroeconomics*, 7(3), 226–258.

Appendix A Additional Figures and Tables

Table A.1: List of main variables and their data sources

Variable	Source
Country-level Data	
Employment in agriculture and nonagriculture	ILO
Value added in agriculture and nonagriculture	UN National Accounts
Agricultural land	FAO
Trade and production of crops	FAO
Producer price of crops	FAO
Trade and production of fertilizer, pesticide, farm machinery	UNIDO-IDSB, FAO, BACI
Trade and production of nonagriculture	BACI
Expenditure share on agriculture	World Bank-Global Consumption Database, WIOD
Gravity variables	CEPII
Population and GDP	ILO, Penn World Tables
Field-level Data	
Crop potential yield of low-input and high-input technologies	FAO-GAEZ
Crop actual yield in the USA	EarthStat
Crop land share worldwide	EarthStat
Total land area	FAO, EarthStat

Notes. ILO: International Labor Organization, FAO: Food and Agriculture Organization, UNIDO: United Nations Industrial Development Organization, IDSB: Industrial Demand-Supply Balance Database at 4-digit ISIC, BACI: World trade database developed by the CEPII based on UN Comtrade, WIOD: World Input-Output Database, GAEZ: Global Agro-ecological zone

Table A.2: Factor Shares in Agricultural Production Technology

	land (N)	labor (L)	material (M)
Traditional	0.51	0.49	0.00
Modern	0.34	0.18	0.48
	fertilizer in M	pesticide in M	machinery in M
Median across crops	0.22	0.25	0.53

Table A.3: Percentage change to crops yields across countries

	banana	cotton	corn	palm oil	potato	rice	soybean	sugarcane	tomato	wheat
Albania		-1.9%	-1.6%		-11.3%		-1.0%		-0.2%	-9.8%
Argentina	-4.2%	-6.9%	-5.1%		-6.4%	-6.3%	-5.7%	-3.9%	-4.9%	-3.6%
Australia	-12.1%	-8.9%	-14.7%	-14.3%	-9.9%	-15.0%	-15.7%	-10.9%	-15.5%	-8.5%
Austria			-2.2%		0.0%		4.2%		-4.1%	-2.6%
Burkina Faso		0.4%	-0.1%			-0.7%	-3.9%	-1.4%	13.1%	
Bangladesh	14.6%	20.0%	12.6%	8.6%	14.8%	17.3%	18.6%	7.6%	15.8%	12.3%
Brazil	-8.1%	-6.9%	-7.4%	-8.3%	-4.9%	-8.3%	-3.1%	-1.5%	-6.9%	-5.3%
Canada			-7.1%		-2.0%		-14.1%		-1.7%	-4.6%
Chile			-2.4%		-10.9%		6.5%		-9.4%	-11.7%
China	-5.1%	-4.7%	-8.0%	-6.3%	-4.1%	-6.4%	-3.6%	-4.2%	-6.5%	-4.1%
Cote dIvoire	-3.8%	-2.3%	-1.9%	-6.7%		-1.5%	-6.5%	-2.0%	-6.4%	
Cameroon	-3.3%	-3.6%	-0.2%	0.0%	10.6%	-0.4%	6.3%	7.7%	-1.5%	3.9%
Congo	-2.2%	2.7%	-1.7%	-5.6%	1.0%	-1.6%	-2.1%	-1.7%	-1.7%	-0.3%
Colombia	-9.1%	-10.1%	-10.1%	-8.6%	-8.7%	-10.2%	-9.2%	-4.2%	-12.3%	-11.7%
Costa Rica	-14.9%	-7.9%	-11.5%	-11.1%	-11.4%	-16.4%	-15.0%	-5.6%	-14.1%	-10.3%
Czech Republic			-24.5%		-22.7%		-29.9%		-26.1%	-24.9%
Germany			-9.1%		-8.7%		1.7%		-5.1%	-7.7%
Dominican Republic	-14.7%	-3.7%	-7.4%	-16.5%	-19.2%	-24.2%	-14.5%	-15.1%	-14.2%	-1.0%
Algeria		-0.2%	-1.6%		-3.0%		-1.3%	-5.8%	-0.5%	-3.4%
Ecuador	-7.4%	-9.6%	-8.4%	-9.1%	-12.2%	-10.7%	-11.4%	-6.7%	-7.9%	-8.9%
Egypt		-4.1%	-5.2%		-3.2%		-5.5%		-5.3%	-5.5%
Spain		-4.5%	-7.9%		-9.9%		-4.7%		-3.2%	-7.2%
Ethiopia	1.0%	2.0%	0.4%		2.1%	0.3%	6.8%	0.8%	2.5%	0.6%
Finland					-6.4%					-5.0%
France		-8.0%	-14.4%		-9.4%	-14.3%	-15.9%		-11.4%	-9.8%
United Kingdom	-2.7%	-14.7%	-14.2%	-5.0%	-14.3%		-16.8%	-20.2%	-18.5%	-14.0%
Ghana	-2.2%	1.0%	-2.0%	-1.8%		-1.7%	1.0%	-3.5%	-2.0%	
Greece		-14.7%	-9.0%		-14.3%		-11.9%		-11.8%	-16.7%
Hungary			-15.5%		-21.8%		-18.4%		-15.9%	-16.2%
Indonesia	-14.5%	-6.3%	-14.1%	-13.1%	-8.9%	-13.1%	-10.0%	-5.1%	-10.8%	-5.8%
India	2.1%	0.0%	2.2%	3.4%	1.0%	1.1%	0.1%	0.8%	1.0%	1.4%
Iran		5.4%	11.6%		4.3%	16.8%	6.8%		4.5%	4.5%
Italy	-9.9%	-12.8%	-12.8%		-9.3%	-7.8%	-11.7%	-2.7%	-7.5%	-6.4%
Japan	-5.3%	-8.5%	-7.3%		-7.1%	-6.5%	-7.1%	-6.2%	-5.9%	-7.0%
Kenya	-1.7%	-0.2%	-2.9%		-0.4%	-1.8%	6.7%	-3.2%	-1.6%	1.1%
South Korea		-4.1%	-13.7%		-8.8%	-9.6%	-7.0%		-8.6%	-9.3%
Sri Lanka	-4.7%	3.5%	-1.5%	-1.9%	-7.8%	-3.2%	-5.9%	-1.2%	-7.0%	-3.6%
Morocco	15.7%	-3.7%	-3.2%	-3.3%	8.4%	5.9%	-3.4%	0.2%	-1.4%	13.8%
Mexico	-11.6%	-14.6%	-10.8%	-14.0%	-10.7%	-11.1%	-15.5%	-9.0%	-10.1%	-11.4%
Mali		-1.0%	-0.4%			-0.4%	-6.8%	0.2%	0.4%	
Mozambique	-2.2%	-1.0%	-5.6%	0.6%	-9.7%	-5.3%	3.0%	-6.9%	-5.5%	-10.1%
Malaysia	-25.0%	-20.1%	-32.8%	-15.3%	-11.4%	-21.2%	-3.3%	-28.5%	-13.2%	-10.2%
Netherlands	-17.1%	-21.4%	-18.4%	-18.4%	-18.8%		-24.5%	-25.1%	-18.9%	-19.0%
Norway	-11.2%	-7.7%	-9.5%	-2.0%	-11.6%		-6.9%	-6.0%	-10.5%	-11.7%
New Zealand	-4.9%	-1.3%	-8.9%		-9.2%		-5.8%	-2.6%	-9.9%	-10.8%
Pakistan	1.9%	3.5%	4.9%		11.0%	11.2%	-3.2%	7.8%	-0.2%	8.3%
Peru	1.8%	2.1%	2.9%	-0.2%	4.8%	2.2%	0.7%	-1.8%	2.9%	2.3%
Philippines	0.0%	-1.3%	-3.0%	-0.8%	-11.0%	-6.6%	-3.8%	-0.8%	-1.6%	-0.9%
Poland			-9.1%		-12.9%				-16.2%	-5.4%
Portugal		-4.0%	-11.1%		-13.7%		-8.8%		-13.7%	-13.9%
Paraguay	-11.5%	-6.4%	-3.8%	-9.9%	-14.7%	-7.9%	-10.0%	-2.5%	-10.1%	-13.5%
Romania		2.4%	-3.9%		-10.1%		-2.9%		-4.8%	-8.9%
RoW	-14.2%	-8.1%	-12.8%	-11.8%	-18.4%	-15.7%	-16.5%	-7.7%	-18.9%	-20.6%
Senegal		-3.2%	-0.5%			-4.7%	-5.1%		-0.6%	
fmr USSR	-1.9%	-8.3%	-1.4%	-4.0%	-8.0%	-7.5%	-6.0%	4.4%	-7.8%	-6.5%
Sweden					-16.3%					-16.2%
Thailand	-15.3%	-9.2%	-17.8%	-13.5%	-7.4%	-15.6%	-7.4%	1.1%	-2.9%	-9.9%
Tunisia		-0.9%	1.4%		4.6%		0.7%		-1.8%	4.7%
Turkey		-4.8%	-6.8%		-4.1%	-7.0%	-0.6%		-4.9%	-4.5%
Tanzania	-2.2%	-3.2%	-2.0%	-1.5%	1.1%		-2.3%	5.9%	-2.1%	-1.3%
Uruguay	13.3%	8.5%	-0.6%		-14.0%	-2.4%	-11.5%	3.1%	-5.1%	-10.9%
United States	-2.9%	-2.6%	-1.9%		-2.8%	-2.0%	-2.6%	-1.8%	-2.0%	-2.4%
Venezuela	5.1%	3.9%	6.4%	6.2%	4.2%	6.0%	-0.8%	5.4%	4.6%	1.4%
Vietnam	13.3%	16.2%	19.7%	15.1%	5.4%	13.8%	7.8%	3.4%	11.9%	13.2%
fmr Yugoslavia		4.9%	-5.6%		-11.0%	-3.6%	-8.3%		-5.6%	-11.4%
South Africa	-3.0%	2.7%	-3.2%		-13.9%	-2.5%	-3.8%	-2.4%	-1.6%	-17.0%
World	-6.1%	-4.2%	-6.5%	-6.8%	-9.0%	-3.7%	-2.4%	-3.1%	-7.3%	-8.1%

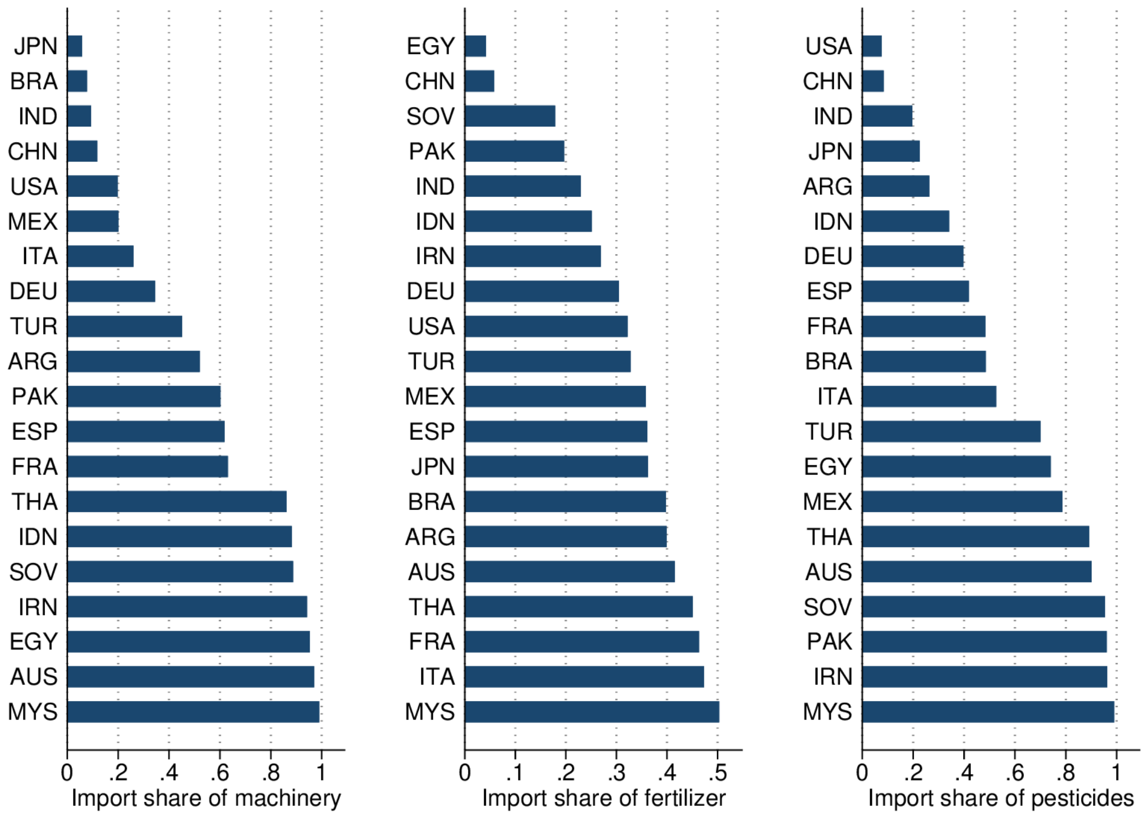
Notes: This table shows the percentage change to yields (production quantity per unit of land) across crops and countries from the baseline of 2015 to the counterfactual with trade costs of 1980 for agricultural outputs and inputs.

Table A.4: Percentage changes to agriculture-related variables and welfare

countries	ISE agr output	ISE agr input	share of land in modern tech	agriculture price	agriculture consumption	welfare
Albania	-17.80%	-51.38%	-18.30%	10.38%	-3.80%	-4.57%
Argentina	32.53%	40.91%	-6.70%	5.31%	-1.57%	-0.06%
Australia	-14.14%	29.64%	-5.53%	9.24%	-3.92%	-0.82%
Austria	-5.38%	-37.88%	1.47%	20.17%	-8.38%	-0.60%
Burkina Faso	-20.87%	41.47%	-85.13%	2.46%	-0.60%	0.02%
Bangladesh	-30.49%	18.35%	19.36%	19.48%	-6.58%	-6.44%
Brazil	-2.29%	40.72%	-9.70%	1.73%	-1.87%	-1.04%
Canada	-20.93%	21.75%	-1.77%	8.31%	-5.72%	-2.50%
Chile	-7.90%	-17.64%	-3.03%	21.57%	-9.13%	-3.14%
China	-52.14%	-10.97%	-2.37%	7.19%	-2.13%	-0.45%
Cote dIvoire	10.14%	93.79%	58.27%	6.61%	1.25%	3.70%
Cameroon	-36.35%	103.28%	99.73%	1.57%	0.08%	0.55%
Congo	8.56%	104.29%	112.95%	2.49%	0.57%	1.79%
Colombia	-22.39%	-8.07%	-16.54%	9.62%	-3.53%	-1.30%
Costa Rica	-9.25%	3.92%	-21.49%	12.59%	-5.83%	-2.45%
Czech Republic	-3.01%	-49.26%	-15.07%	22.37%	-10.41%	-3.16%
Germany	-5.61%	-71.28%	-0.04%	28.22%	-12.45%	-2.22%
Dominican Republic		80.21%	-56.41%	12.09%	-1.01%	3.71%
Algeria	-2.30%	22.17%	-16.63%	1.39%	-0.88%	-0.53%
Ecuador	-50.82%	62.83%	-34.15%	10.94%	-3.73%	-1.41%
Egypt	-12.75%	-108.18%	3.45%	16.47%	-7.13%	-7.20%
Spain	-10.42%	-131.38%	-6.47%	41.49%	-16.70%	-4.49%
Ethiopia	-46.43%	95.05%	-3.03%	-2.48%	-0.79%	-2.11%
Finland	-1.32%	-162.27%	-0.03%	12.10%	-6.75%	-1.92%
France	-2.07%	-31.63%	-4.92%	10.26%	-5.39%	-1.46%
United Kingdom	1.21%	-60.49%	-1.29%	8.56%	-4.60%	-0.99%
Ghana	-12.21%	48.21%	-5.57%	6.18%	-0.99%	0.92%
Greece	-10.04%	-103.05%	-22.50%	19.77%	-8.82%	-2.51%
Hungary	-11.03%	-92.38%	-32.67%	24.19%	-11.52%	-3.70%
Indonesia	-28.43%	2.25%	-18.21%	11.60%	-4.17%	-2.53%
India	-63.40%	-33.25%	1.31%	1.11%	-0.84%	-0.97%
Iran	-28.19%	-417.35%	6.09%	8.03%	-3.09%	-1.62%
Italy	-4.29%	-79.35%	-6.73%	14.68%	-7.29%	-1.83%
Japan	1.68%	38.36%	-1.35%	6.89%	-3.23%	-0.58%
Kenya	-28.33%	65.24%	35.04%	8.47%	-1.49%	-0.57%
South Korea	-10.23%	-123.62%	-0.59%	17.06%	-7.18%	-1.20%
Sri Lanka	-5.69%	-50.26%	-35.63%	10.00%	-4.00%	-3.91%
Morocco	-40.83%	64.68%	56.08%	7.38%	-2.93%	-1.21%
Mexico	-26.66%	-64.26%	-17.38%	21.68%	-8.84%	-2.37%
Mali	0.31%	72.04%	2.82%	1.64%	0.00%	0.25%
Mozambique		98.78%	-64.30%	6.37%	1.02%	4.32%
Malaysia	-20.09%	-32.78%	-18.83%	20.45%	-10.04%	-4.39%
Netherlands	-0.11%	-41.32%	-1.38%	13.66%	-7.81%	-2.89%
Norway	-3.44%	-60.32%	-0.04%	20.28%	-9.59%	-2.37%
New Zealand	-12.98%	-161.23%	0.31%	27.47%	-11.74%	-3.19%
Pakistan	-29.29%	39.99%	10.66%	2.16%	-0.80%	-0.59%
Peru	-35.87%	58.65%	25.11%	5.28%	-1.97%	-0.69%
Philippines	-21.38%	35.63%	-45.58%	5.52%	-1.53%	-0.56%
Poland	-23.72%	-103.56%	-3.05%	43.26%	-17.04%	-6.07%
Portugal	-5.78%	-93.84%	-15.55%	33.19%	-13.92%	-4.04%
Paraguay	-24.24%	84.22%	-27.87%	1.56%	-3.56%	-3.41%
Romania	-16.30%	-83.89%	-13.07%	22.03%	-9.24%	-6.98%
RoW	-17.04%	64.15%	-44.07%	7.19%	-1.84%	0.00%
Senegal	-9.65%	18.81%	131.46%	14.15%	-5.12%	-4.22%
fmr USSR	-39.69%	-57.02%	-5.38%	15.97%	-6.13%	-5.31%
Sweden	-6.22%	-26.03%	-0.62%	32.53%	-13.62%	-2.81%
Thailand	-4.68%	-35.73%	-23.16%	13.92%	-6.07%	-3.00%
Tunisia	-9.44%	-5.63%	18.34%	10.87%	-3.11%	0.78%
Turkey	-38.95%	-148.90%	-1.76%	33.65%	-11.83%	-5.17%
Tanzania	-22.98%	51.58%	-83.63%	7.23%	-1.31%	-0.25%
Uruguay	-51.34%	62.47%	-41.75%	2.05%	-2.89%	-2.13%
United States	-10.72%	-15.40%	-0.34%	3.60%	-1.71%	-0.36%
Venezuela	-9.22%	129.68%	9.51%	-4.12%	1.90%	0.59%
Vietnam		117.13%	19.50%	-8.26%	2.94%	0.14%
fmr Yugoslavia	-8.61%	-23.16%	-16.44%	10.72%	-4.96%	-3.05%
South Africa	-26.76%	-4.60%	-51.01%	6.52%	-3.58%	-1.18%
World	-16.08%	-52.26%	-7.49%	8.19%	-2.83%	-1.63%

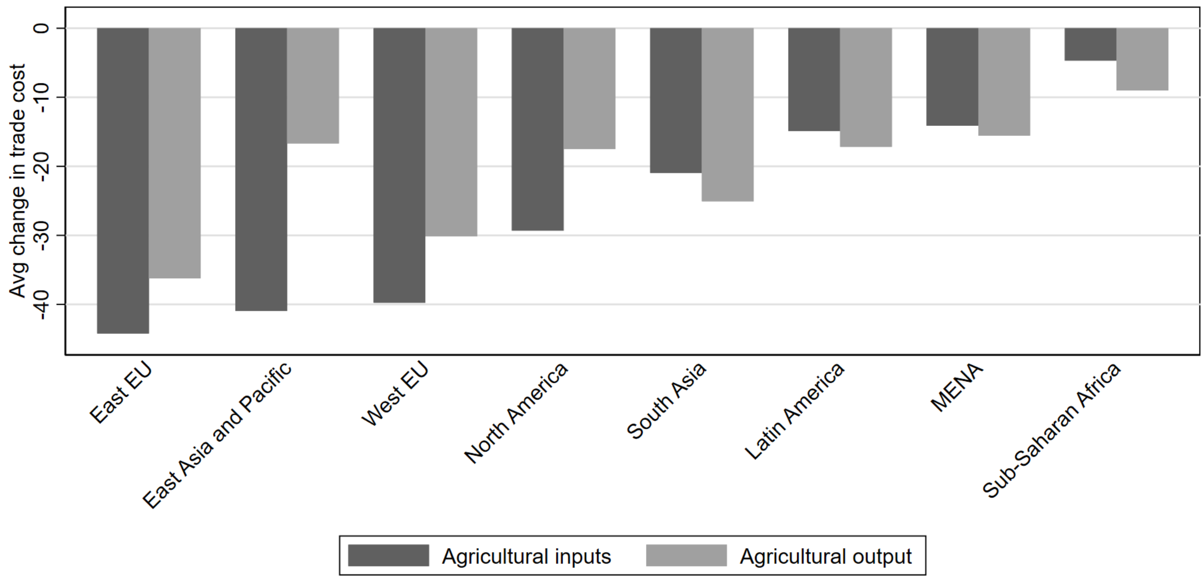
Notes: This table shows the percentage change to agriculture-related variables across countries from the baseline of 2015 to the counterfactual with trade costs of 1980 for agricultural outputs and inputs. ISE stands for imported share of expenditures.

Figure A.1: Import content of agricultural input use



Notes: This figure shows the share of imports in the use of fertilizers, pesticides, and agricultural machinery in 20 largest countries in terms of agricultural output for year 2015

Figure A.2: Trade costs changes between 1980 and 2015, for agricultural outputs and inputs, aggregated by region



Notes: This figure shows the percentage change to trade costs of agricultural outputs (as the aggregate of crops) and agricultural inputs (as the aggregate of fertilizers, pesticides, and farm machinery) between 1980 and 2015.

Appendix B Data

Potential Yields. The data on potential yields (also called "maximum attainable yields") comes from Global Agro-Ecological Zones project, which is produced by the International Institute for Applied System Analysis (IIASA) and the Food and Organization of the United Nations (FAO). The goal of the project is to generate global datasets about agriculture at the disaggregated level of fields to promote studies of the conditions affecting agricultural development and food security. The first version of the dataset was published in 2000.

Among the different datasets produced by FAO-GAEZ, we use for our analysis data on agro-climatically attainable biomass and yield for specific land utilization types (LUTs) by crop. The different types of land utilization corresponds to what we denote by different technologies in our model. The estimation of the maximum attainable yield is based on a function that maps rich climate data into maximum attainable yields. The parameters of this function depend on each LUT and crop. Importantly, local socio-economic conditions do not enter as an input in the estimation of maximum attainable yields. As such, variations in maximum attainable yields across fields should only capture differences in climatic conditions rather than levels of development of each country. Indeed, we find little to no systematic variation between maximum attainable yields and gdp per capita in our data once we control for a parsimonious set of geographic characteristics of a field.

The land utilization types in the data are divided into three groups. First, there is a low level of input use type, which corresponds to a farming system that is largely subsistence based. This dataset represents the maximum attainable yield if farmers use traditional cultivars and, importantly, no application of nutrients, no use of chemicals and minimum conservation measures. Therefore, we denote this technology as traditional in our analysis. Second, there is an intermediate level of input use type, which corresponds to a faming system that is partly market oriented. We do not directly use this type of technology because we do not have enough data to identify an additional set of parameters for factor-intensity in our model. Third, the high level of input use type, which corresponds to a farming system that is mainly market oriented. In this case, production is fully mechanized and uses optimum applications of nutrients and chemical pest, disease and weed control.

Actual Yields and Agricultural Land. The data on actual yields come from Earthstat. Earthstat also provides geographic datasets at the disaggregated level of fields providing information about agriculture. The project is a collaboration between the Global Landscapes Initiative at the University of Minnesota's Institute on the Environment and the Land Use and Global Environment Lab at the University of British Columbia.

Among the several datasets organized by Earthstat, we use information on actual yields by crop and on the share of land in agriculture. The data on actual yields use data from agricultural census and survey information on the areas from the smallest political units from each country. The level of disaggregation of the source data, however, varies substantially across countries. For example, in some countries, the data on actual yields is provided in terms of smaller areas that correspond to counties in the US, whereas in other countries data is provided at the state and province levels. Therefore, in our estimation we restricted our sample to the data constructed only from sources with sufficiently disaggregated information.

In addition to the yield level data, we also bring data from Earthstat on the share of a field dedicated to the production of crops. This dataset, different from the yields one, is constructed using a combination of agricultural datasets and satellite imagery. In our calibration, we use the entire sample of cropland shares coming from Earthstat.

Trade and Gross Output for Agriculture and Non-Agricultural Goods. For bilateral trade flows in non-agriculture goods, agricultural machinery, pesticides, fertilizers and agriculture as a whole, we use data from BACI for the years after 1995 and from Feenstra, Lipsey, Deng, Ma, and Mo (2005) before then. Both of these datasets are constructed based on Comtrade. To select the industry codes associated with agricultural inputs, we follow closely the guidelines specified in FAO-STAT. We adjust aggregate values for each good (i.e., non-agriculture goods, agricultural machinery, pesticides, fertilizers and agriculture as a whole) in Feenstra, Lipsey, Deng, Ma, and Mo (2005) to ensure consistency with aggregate values in BACI.

For the non-agriculture sector, we construct data on gross output using data the 2 digit level data from UNIDO. In a few cases, we complement our data with information from other sources as follows. First, we construct domestic trade shares of country i in non-agriculture ($\pi_{i,nag}$) using $\pi_{i,nag} = (Y_{i,nag} - X_{i,nag})/(Y_{i,nag} - X_{i,nag} + M_{i,nag})$, where $Y_{i,nag}$ is the gross output, $X_{i,nag}$ is non-agriculture export, and $M_{i,nag}$ is non-agriculture imports. If domestic trade shares are not within 0 and 1, we then bring data on domestic trade shares from the World Input-Output Database (WIOD) or from the trade and production dataset from CEPII, depending on availability, which are constructed to satisfy these bounds. Finally, using our data on imports and exports, we then infer the gross output that is consistent with these domestic trade shares using $Y_{i,nag} = X_{i,nag} + M_{i,nag}\pi_{i,nag}/(1 - \pi_{i,nag})$.

For the agricultural sector as a whole (i.e., not disaggregated by crop), we construct gross output based mostly on data from FAO-STAT. When data from FAO-STAT is not available, we bring information on gross output in agriculture from the STAN dataset and, when the dataset from STAN is not available, we construct gross output combining value added from United Nations with data on total expenditure in agricultural inputs. For a few cases, our domestic trade shares are not within 0 and 1 and we then apply the same procedure that we adopted for the non-agriculture sector to construct gross output.

Our data on gross output and trade flows by crop comes from FAO-STAT. We have to make minimal adjustments to crop names and codes to ensure consistency between the values coming from these two datasets. For each crop, we pick the codes associated with trade in less processed goods. For example, for palm production we do not include data on bilateral trade flows in palm oil. Our final data on trade flows and gross output crop includes only data after 1991.

To construct our data on agricultural input sales and expenditures, we collected information on value added per sector from United Nations, on apparent consumption by industry from UNIDO and on exports and imports from Comtrade and BACI. We focus on three agricultural inputs: pesticide, machinery and fertilizer. The construction of our data for expenditure on fertilizers follows a slightly different approach since richer data on quantities is available from FAO. We next explain our procedure to construct the data for machinery, which is the same that we use to the data on pesticide.

To construct our data on gross-output in agricultural machinery, we combine the 4 digit level data from UNIDO with our data on trade flows. Our procedure is as follows. We first compute the ratio of exports in agricultural machinery relative to total exports in non-agriculture goods using our trade flow data and the ratio of gross output in agricultural machinery relative to gross output in non-agriculture goods. We then run a regression of the relative gross output agains the relative exports with country-fixed effects, which gives us a R^2 of 0.82, and use the gross output predicted from this regression. For countries without data available in UNIDO, we run a regression without country-fixed effects, which gives an adjusted R^2 of 0.51 and use the gross output predicted from this regression. This procedure provides an extremely good within sample fit: a regression of the log of the gross output in agricultural machinery in UNIDO against the log of our constructed gross output yields a R^2 of 0.92.

To construct our data on gross-output for fertilizers, we combine data on exports, production, consumption and imports in fertilizer quantity from FAO-STAT with our data on trade flows in values. The data on fertilizers from FAO-STAT comes disaggregated by nutrients, i.e., nitrogren N, phosphate P and potassium K. To simplify our analysis we summed the weight of the total amount of nutrients coming from each type. This summation tends to give the same sample proportion of each ingredient per country, which is often referred to as all-purpose fertilizer. Using the data from FAO-STAT, we construct the domestic share of consumption by diving imports in quantity by total consumption in quantity. Using this domestic share of consumption, which we call $\pi_{i,F}$, we construct gross in values output using $Y_{i,F} = X_{i,F} + M_{i,F}\pi_{i,F}/(1-\pi_{i,F})$. Noticed here that we rely on the assumption that domestic shares of consumption in quantity are equivalent to domestic share of consumption in values. This is the case when the price of imported fertilizers are in average the same as the price of fertilizers consumed from domestic source. This assumption is consistent with the Eaton and Kortum (2002) framework, where the average price of goods coming from any source given destination is the same.

Consumption Share and Labor Employment in Agriculture. To construct our data on consumption share in agricultural goods, we collect data from different sources. For developing countries, we use data from the Global Consumption database organized by the World Bank to construct the consumption shares in agricultural goods. For the United States, we use data from the consumer expenditure survey. For Canada, we use data from household surveys available from Queen's University of Canada. For European countries, we bring data from Eurostat. To construct labor employment, we use data from UN-ILO. When data from UN-ILO was not available, we infer the share of workers in agriculture using data on the share of workers in rural areas from the World Bank.

Appendix C Model

C.1 Costs and Output

Unit cost. Focusing on production in a plot given a choice of agriculture activity, we drop country-field-crop-technique indicators, and write down the cost minimization problem:

$$\min_{L\geq 0, N\geq 0, M\geq 0} \; rL + wN + mM \quad s.t. \quad \bar{q}\Big(zL\Big)^{\gamma^L}\Big(N\Big)^{\gamma^N}\Big(M\Big)^{\gamma^M} = 1$$

where

$$\bar{q} \equiv (\gamma^L)^{-\gamma^L}(\gamma^N)^{-\gamma^N}(\gamma^M)^{-\gamma^M}$$

The Lagrangian function is:

$$\mathcal{L} = rL + wN + mM - \mu\Big[\bar{q}\Big(zL\Big)^{\gamma^L}\Big(N\Big)^{\gamma^N}\Big(M\Big)^{\gamma^M} - 1\Big]$$

First order conditions are:

$$\begin{aligned} r &= \mu\bar{q}\gamma^L z^{\gamma^L} L^{\gamma^L - 1} N^{\gamma^N} I^{\gamma^M} \\ w &= \mu\bar{q}\gamma^N z^{\gamma^L} L^{\gamma^L} N^{\gamma^N - 1} I^{\gamma^M} \\ m &= \mu\bar{q}\gamma^M z^{\gamma^L} L^{\gamma^L} N^{\gamma^N} I^{\gamma^M - 1} \end{aligned}$$

The employment of labor and land relative to inputs are then given by:

$$L = \frac{\gamma^L}{\gamma^M}\frac{mM}{r}, \quad N = \frac{\gamma^N}{\gamma^M}\frac{mM}{w}$$

Replace L and N into the production equation, $\bar{q}\Big(z\frac{\gamma^L}{\gamma^M}\frac{mM}{r}\Big)^{\gamma^L}\Big(\frac{\gamma^N}{\gamma^M}\frac{mM}{w}\Big)^{\gamma^N}\Big(M\Big)^{\gamma^M} = 1$, delivers:

$$M = (\bar{q})^{-1} z^{-\gamma^L}(\gamma^L)^{-\gamma^L}(\gamma^N)^{-\gamma^N}(\gamma^M)^{1-\gamma^M} r^{\gamma^L} w^{\gamma^N} m^{\gamma^M - 1}$$

which then results:

$$\begin{aligned} M &= (r/z)^{\gamma^L} w^{\gamma^N} m^{\gamma^M} \times \frac{\gamma^M}{m} \\ L &= (r/z)^{\gamma^L} w^{\gamma^N} m^{\gamma^M} \times \frac{\gamma^L}{r} \\ N &= (r/z)^{\gamma^L} w^{\gamma^N} m^{\gamma^M} \times \frac{\gamma^N}{w} \end{aligned}$$

Using these optimal choices of inputs, the unit cost of production equals

$$c = rL + wN + mM = (r/z)^{\gamma^L} w^{\gamma^N} m^{\gamma^M}$$

Rent. Combining zero profit condition and returns to land,

$$c = p \Rightarrow (r/z)^{\gamma^L} w^{\gamma^N} m^{\gamma^M} = p$$

which results:

$$r = zp^{\frac{1}{\gamma^L}} w^{-\frac{\gamma^N}{\gamma^L}} m^{-\frac{\gamma^M}{\gamma^L}}$$

Output. The size of each plot of land is w.l.o.g. normalized to one, and it is optimal to use the entire plot as long as profits are non-negative. Therefore, land use L equals one. It follows that:

$$\begin{aligned} N &= \frac{rL}{w}\frac{\gamma^N}{\gamma^L} = zp^{\frac{1}{\gamma^L}} w^{-\frac{\gamma^N}{\gamma^L}} m^{-\frac{\gamma^M}{\gamma^L}} \frac{\gamma^N}{w\gamma^L} \\ M &= \frac{rL}{m}\frac{\gamma^M}{\gamma^L} = zp^{\frac{1}{\gamma^L}} w^{-\frac{\gamma^N}{\gamma^L}} m^{-\frac{\gamma^M}{\gamma^L}} \frac{\gamma^M}{m\gamma^L} \end{aligned}$$

Replace N, M, and $L = 1$ into the production equation gives output at the plot level:

$$Q = \bar{q}\Big(zL\Big)^{\gamma^L}\Big(N\Big)^{\gamma^N}\Big(M\Big)^{\gamma^M} = \bar{q}\Big(z\Big)^{\gamma^L}\Big(zp^{\frac{1}{\gamma^L}} w^{-\frac{\gamma^N}{\gamma^L}} m^{-\frac{\gamma^M}{\gamma^L}}\Big)^{\gamma^N+\gamma^M}\Big(\frac{\gamma^N}{w\gamma^L}\Big)^{\gamma^N}\Big(\frac{\gamma^M}{m\gamma^L}\Big)^{\gamma^M}$$

Since $\bar{q} \equiv (\gamma^L)^{-\gamma^L}(\gamma^N)^{-\gamma^N}(\gamma^M)^{-\gamma^M}$, and $\gamma^L + \gamma^N + \gamma^M = 1$,

$$Q = z\Big(\gamma^L\Big)^{-1}\Big(\frac{w}{p}\Big)^{-\gamma^N/\gamma^L}\Big(\frac{m}{p}\Big)^{-\gamma^M/\gamma^L}$$

C.2 Quantity of fixed costs

The unconditional mean of investment intensity draw, $s_i^f(\omega)$, is given by

$$\mathbb{E}\Big[a_{i,0}^f(\omega)\Big] = a_{i,0}^f$$

Let Ω_i^f be the set of plots within field f which are selected for agriculture use. The share of land allocated to all agricultural uses is denoted by α_i^f,

$$\alpha_i^f \equiv \Pr(\omega \in \Omega_i^f) = \sum_{k\in\mathcal{K}} \alpha_{i,k}^f$$

The mean of $a_{i,0}^f(\omega)$ conditional on plot ω not being selected for agriculture is

$$\mathbb{E}\Big[a_{i,0}^f(\omega) \mid \omega \notin \Omega_i^f\Big] = a_{i,0}^f(1-\alpha_i^f)^{-1/\theta_1}$$

The conditional mean is greater than the unconditional mean because when the investment intensity of a plot is too large, it will be less likely to select that plot for agriculture. By relating conditional and unconditional means and rearranging the resulting terms,

$$\mathbb{E}\left[a_{i,0}^{f}(\omega)\right] = \mathbb{E}\left[a_{i,0}^{f}(\omega) \mid \omega \in \Omega_i^f\right] \Pr(\omega \in \Omega_i^f) + \mathbb{E}\left[a_{i,0}^{f}(\omega) \mid \omega \notin \Omega_i^f\right] \Pr(\omega \notin \Omega_i^f)$$

$$\mathbb{E}\left[a_{i,0}^{f}(\omega) \mid \omega \in \Omega_i^f\right] = \frac{1}{\Pr(\omega \in \Omega_i^f)}\left[\mathbb{E}\left[a_{i,0}^{f}(\omega)\right] - \mathbb{E}\left[a_{i,0}^{f}(\omega) \mid \omega \notin \Omega_i^f\right] \Pr(\omega \notin \Omega_i^f)\right]$$

$$\mathbb{E}\left[a_{i,0}^{f}(\omega) \mid \omega \in \Omega_i^f\right] = \frac{1}{\alpha_i^f}\left[a_{i,0}^f - a_{i,0}^f(1-\alpha_i^f)^{-1/\theta_1}(1-\alpha_i^f)\right]$$

$$\mathbb{E}\left[a_{i,0}^{f}(\omega) \mid \omega \in \Omega_i^f\right] = \frac{a_{i,0}^f}{\alpha_i^f}\left[1-(1-\alpha_i^f)^{(\theta_1-1)/\theta_1}\right]$$

The field-level quantity required for fixed investments in agriculture, S_i^f, equals the average fixed cost requirement conditional on plots being used for agriculture times the number of plots used for agriculture, $S_i^f = \mathbb{E}\left[a_{i,0}^{f}(\omega) \mid \omega \in \Omega_i^f\right]\alpha_i^f L_i^f$. Replacing in this equation the above one reproduces equation (24),

$$S_i^f = a_{i,0}^f L_i^f \left[1-(1-\alpha_i^f)^{(\theta_1-1)/\theta_1}\right]$$

C.3 Discrete Choice, Generalized Extreme Value, and Choice Probabilities

C.3.1 McFadden's Theorem

We reformulate Theorem 5.2 in "Econometric Models of Probabilistic Choice" by McFadden (1981). Consider the following discrete choice problem

$$\max_{i\in\Omega} \; -q_i + u_i$$

where Ω is the set of alternatives, q_i is the non-stochastic component of the objective function, and u_i is the stochastic term. For example, if $q_i = -\mathbf{b}'\mathbf{z}_i$, and u_i is a random variable drawn independently from type I extreme value distribution, $F(u) = \exp(-e^{-u})$, then the choice probabilities are given by

$$\pi_i = \frac{e^{-q_i}}{\sum_{j\in\Omega} e^{-q_j}} = \frac{e^{\mathbf{b}'\mathbf{z}_i}}{\sum_{j\in\Omega} e^{\mathbf{b}'\mathbf{z}_j}}$$

Theorem. Given $\Omega = \{1, ..., m\}$, *consider* $H(\mathbf{y})$ *with* $\mathbf{y} = (y_1, ..., y_m)$ such that

1. $H(\mathbf{y})$ is nonnegative, and it is homogeneous of degree one.

2. $H(\mathbf{y}) \to \infty$ as $y_i \to \infty$ for all $i \in \Omega$.

3. The mixed partial derivatives of H exist and are continuous, with non-positive even and non-negative odd mixed partial derivatives.

Then,

(I) The following function

$$F(\boldsymbol{u}) = \exp\left[- H\Big(e^{-u_1}, ..., e^{-u_m}\Big)\right]$$

is a multivariate extreme value distribution.

(II) Choice probabilities satisfy

$$\pi_i(\mathbf{q}) = -\frac{\partial}{\partial q_i} \ln H\Big(e^{-q_1}, ..., e^{-q_m}\Big)$$

We will use this theorem in our derivations below. For illustrative purposes, we first begin with applying the theorem to a choice structure with one nest. Then, we focus on a two-nest structure, that is the one in our framework.

C.3.2 Discrete choices with one nest

Suppose H is given by

$$H(\mathbf{y}) = \Big[\sum_{i \in \Omega} y_i^{\rho}\Big]^{1/\rho}$$

With $\rho = \frac{1}{1-\sigma}$, as long as $0 \leq \sigma < 1$, the conditions in Mc Fadden's theorem are satisfied. Then, according to result (I) of the theorem, the following is is a multivariate EV distribution:

$$F(\mathbf{u}) = \exp\left[- \Big(e^{-\rho u_1} + ... + e^{-\rho u_K}\Big)^{1/\rho}\right] \tag{A.1}$$

where σ is the correlation parameter between $(u_j, u_{j'})$. According to result (II) of the theorem, choice probabilities are:

$$\pi_i = -\frac{\partial}{\partial q_i} \ln \Big(e^{-\rho q_1} + ... + e^{-\rho q_K}\Big)^{1/\rho} = \frac{e^{-\rho q_i}}{e^{-\rho q_1} + ... + e^{-\rho q_K}} \tag{A.2}$$

By a change of variables, we can specify draws based on Type II EV (Fréchet) rather than Type I EV. Recall that the discrete choice problem as originally formulated in McFadden's theorem was:

$\max_{i\in\Omega} -q_i + u_i$. This problem is equivalent to

$$\max_{i\in\Omega} h_i z_i,$$

where $q_i = -\theta \ln h_i a_i$, and $u_i = \theta \ln(z_i/a_i)$. Here, h_i is the non-stochastic component and z_i is a realization of draw from a distribution. Replacing z_i for u_i in (A.1), the probability distribution of $\mathbf{z}(\omega) = (z_1(\omega), ..., z_K(\omega))$ is:

$$\Pr(z_1(\omega) \leq z_1, ..., z_K(\omega) \leq z_K) \equiv F(z_1, ..., z_K) = \exp\left[-\Big(\sum_{k=1}^{K} (z_k/a_k)^{-\theta\rho}\Big)^{\frac{1}{\rho}}\right], \tag{A.3}$$

which is a Fréchet (Type II EV) distribution. Replacing for $q_i = -\theta \ln h_i a_i$ in (A.2), choice probabilities are:

$$\pi_i = \frac{(h_i a_i)^{\theta\rho}}{\sum_{k=1}^{K}(h_k a_k)^{\theta\rho}} \tag{A.4}$$

The case of Eaton and Kortum with independent draws is a special case in which $\rho = 1$ (or equivalently, $\sigma = 0$), and so, $z_1(\omega)$, ..., $z_K(\omega)$ are independent. The probability distribution simplifies to

$$F(z_1, ..., z_K) = \exp\left[-\Big(\sum_{k=1}^{K} (z_k/a_k)^{-\theta}\Big)\right].$$

Thanks to independence, distribution of $z_k(\omega)$ equals

$$\Pr(z_k(\omega) \leq z_k) \equiv F_k(z_k) = F(\infty, ...\infty, z_K, \infty, ...\infty) = \exp\left[- (z_k/a_k)^{-\theta}\right],$$

which is the distribution used in EK. In addition, setting $\rho = 1$ implies choice probabilities: $\pi_i = \frac{(h_i a_i)^{\theta}}{\sum_{k=1}^{K}(h_i a_i)^{\theta}}$.

C.3.3 Discrete choices with two nests

The following function H satisfies the conditions in McFadden's theorem,

$$H(\mathbf{y}) = \sum_{k\in K}\Big[\sum_{i\in\Omega_k} y_{ik}^{\rho}\Big]^{1/\rho}$$

Using result (I) of the theorem, the following is a multivariate EV distribution

$$F(\mathbf{u}) = \exp\left[-\sum_{k\in K}\Big[\sum_{i\in\Omega_k} e^{-\rho u_{ik}}\Big]^{1/\rho}\right] \tag{A.5}$$

and, choice probabilities are as follows, based on result (II) of the theorem,

$$\pi_{ik} = -\frac{\partial}{\partial q_{ik}} \ln \left[\sum_{k \in K} \left[\sum_{i \in \Omega_k} e^{-\rho q_{ik}} \right]^{1/\rho} \right] = \frac{e^{-\rho q_{ik}}}{\sum_{i \in \Omega_k} e^{-\rho q_{ik}}} \times \frac{\left[\sum_{i \in \Omega_k} e^{-\rho q_{ik}} \right]^{1/\rho}}{\sum_{k \in K} \left[\sum_{i \in \Omega_k} e^{-\rho q_{ik}} \right]^{1/\rho}} \tag{A.6}$$

These changes of variables convert the formulation from EV type I to EV type II distribution: $q_{ik} = -\theta \ln(a_{ik} h_{ik})$ and $u_{i,k} = \theta \ln(z_{i,k}/a_{i,k})$. Replacing these in (A.5) and (A.6) delivers the distribution function of $\mathbf{z} = \{z_{i,k}\}_{i,k}$ and choice probabilities:

$$F(\mathbf{z}) = \exp\left[-\sum_{k \in K} \left[\sum_{i \in \Omega_k} (z_{i,k}/a_{i,k})^{-\theta\rho} \right]^{1/\rho} \right] \tag{A.7}$$

$$\pi_{ik} = \frac{h_{ik}^{\theta\rho}}{\sum_{i \in \Omega_k} h_{ik}^{\theta\rho}} \times \frac{a_k \left[\sum_{i \in \Omega_k} h_{ik}^{\theta\rho} \right]^{1/\rho}}{\sum_{k \in K} \left[\sum_{i \in \Omega_k} h_{ik}^{\theta\rho} \right]^{1/\rho}} \tag{A.8}$$

C.4 Expected Value conditional on selection

C.4.1 One Nest

Reproducing equations (A.3) and (A.4),

$$\Pr(z_1(\omega) \leq z_1, ..., z_K(\omega) \leq z_K) \equiv F(z_1, ..., z_K) = \exp\left[-\left(\sum_{k=1}^{K} (z_k/a_k)^{-\theta\rho} \right)^{\frac{1}{\rho}} \right]$$

$$\pi_i = \frac{(h_i a_i)^{\theta\rho}}{\sum_{k=1}^{K} (h_k a_k)^{\theta\rho}}$$

For notaional simplicity, and w.l.o.g. we focus on the choice of 1st alternative. Let $\Omega_j = \{\omega : h_j z_j = \max_i h_i z_i\}$. Define

$$F^1(z_1, ..., z_K) \equiv \frac{\partial}{\partial z_1} F(z_1, ..., z_K)$$

which equals

$$F^1(z_1, ..., z_K) = \theta a_1^{\theta\rho} z_1^{-\theta\rho-1} \left(\sum_{k=1}^{K} (z_k/a_k)^{-\theta\rho} \right)^{\frac{1}{\rho}-1} \exp\left[-\left(\sum_{k=1}^{K} (z_k/a_k)^{-\theta\rho} \right)^{\frac{1}{\rho}} \right]$$

The probability distribution of $z_1(\omega)$ conditional on selecting the 1st alternative, $\omega \in \Omega_1$,

$$\begin{aligned}
\widetilde{F}_1(z) &\equiv \Pr\Big(z_1(\omega) \le z \mid \omega \in \Omega_1\Big) \\
&= \frac{1}{\Pr(\omega \in \Omega_1)} \Pr\Big(z_1(\omega) \le z,\ h_1 z_1(\omega) \ge h_j z_j(\omega)\Big) \\
&= \frac{1}{\pi_1} \Pr\Big(z_1(\omega) \le z,\ z_j(\omega) \le \frac{h_1}{h_j} z_1(\omega)\Big) \\
&= \frac{1}{\pi_1} \int_{z_1=0}^{z} \int_{z_2=0}^{\frac{h_1}{h_2}z} \int_{z_K=0}^{\frac{h_1}{h_K}z} f(z_1, z_2, ..., z_K) dz_K ... dz_2 dz_1 \\
&= \frac{1}{\pi_1} \int_{z_1=0}^{z} F^1(z, \frac{h_1}{h_2}z, ..., \frac{h_1}{h_K}z) dz_1 \\
&= \frac{1}{\pi_1} \int_{z_1=0}^{z} \theta a_1^{\theta\rho} z^{-\theta\rho-1} \Big((\frac{z}{a_1})^{-\theta\rho} + \sum_{k=2}^{K} (\frac{h_1 z}{h_k a_k})^{-\theta\rho} \Big)^{\frac{1}{\rho}-1} \exp\left[-\Big((\frac{z}{a_1})^{-\theta\rho} + \sum_{k=2}^{K} (\frac{h_1 z}{h_k a_k})^{-\theta\rho} \Big)^{\frac{1}{\rho}} \right] dz_1 \\
&= \frac{1}{\pi_1} \int_{z_1=0}^{z} \theta a_1^{\theta} z^{-\theta-1} \Big(1 + \sum_{k=2}^{K} (\frac{h_1 a_1}{h_k a_k})^{-\theta\rho} \Big)^{\frac{1}{\rho}-1} \exp\left[- z^{-\theta} a_1^{\theta} \Big(1 + \sum_{k=2}^{K} (\frac{h_1 a_1}{h_k a_k})^{-\theta\rho} \Big)^{\frac{1}{\rho}} \right] dz_1 \\
&= \frac{1}{\pi_1} \int_{z_1=0}^{z} \theta a_1^{\theta} z^{-\theta-1} \Big((h_1 a_1)^{-\theta\rho} \sum_{k=1}^{K} (h_k a_k)^{\theta\rho} \Big)^{\frac{1}{\rho}-1} \exp\left[- z^{-\theta} a_1^{\theta} \Big((h_1 a_1)^{-\theta\rho} \sum_{k=1}^{K} (h_k a_k)^{\theta\rho} \Big)^{\frac{1}{\rho}} \right] dz_1 \\
&= \int_{z_1=0}^{z} \theta a_1^{\theta} z^{-\theta-1} \Big(\frac{1}{\pi_1} \Big)^{\frac{1}{\rho}} \exp\left[- z^{-\theta} a_1^{\theta} \Big(\frac{1}{\pi_1} \Big)^{\frac{1}{\rho}} \right] dz_1
\end{aligned}$$

which is a Fréchet distribution with location parameter $a_1^{\theta}\pi_1^{-1/\rho}$ and dispersion parameter θ. It is straightforward to show that the expected value of a Fréchet distribution with location parameter T and dispersion parameter θ equals $\Gamma(1-1/\theta)T^{1/\theta}$. Putting together, the expected value of $z_1(\omega)$ conditional on $\omega \in \Omega_1$ equals

$$\mathbb{E}\Big(z_1(\omega) |\ \omega \in \Omega_1\Big) = \Gamma(1-1/\theta) a_1 \pi_1^{-1/\theta\rho}$$

To make a closer connection to the notation we adopted in the main text, let $\theta_2 \equiv \theta\rho$, and $\theta_1 \equiv \theta$. Then,

$$\Pr(z_1(\omega) \le z_1, ..., z_K(\omega) \le z_K) = \exp\left[-\Big(\sum_{k=1}^{K} (z_k/a_k)^{-\theta_2} \Big)^{\frac{\theta_1}{\theta_2}} \right]$$

And, the conditional expected value is given by

$$\mathbb{E}\Big(z_1(\omega) |\ \omega \in \Omega_1\Big) = \Gamma(1-1/\theta_1) a_1 \pi_1^{-1/\theta_2}$$

C.4.2 Two Nests

We first reproduce equations (A.7) and (A.6),

$$F(z_{11},...,z_{1K},z_{21},...,z_{2K}) = \exp\left[-\left\{\Big(\sum_{k=1}^{K}(z_{1k}/a_{1k})^{-\theta\rho}\Big)^{\frac{1}{\rho}} + \Big(\sum_{k=1}^{K}(z_{2k}/a_{2k})^{-\theta\rho}\Big)^{\frac{1}{\rho}}\right\}\right]$$

$$\pi_{sk} = \frac{(h_{sk}a_{sk})^{\theta\rho}}{H_s^{\theta\rho}}\frac{H_s^{\theta}}{H_1^{\theta}+...+H_S^{\theta}}, \quad \text{where} \quad H_s = \left[(h_{s1}a_{s1})^{\theta\rho}+...+(h_{sK}a_{sK})^{\theta\rho}\right]^{\frac{1}{\theta\rho}}$$

Here, for the sake of illustration, we are considering a choice structure with $S=2$ alternatives in in the upper nest and K sub-trees in the lower nests. For notational simplicity and w.o.l.g, we focus on the choice of $(s,k)=(1,1)$. Let $F^{11}(z_{11},...,z_{1K},z_{21},...,z_{2K}) \equiv \frac{\partial}{\partial z_{11}}F(z_{11},...,z_{1K},z_{21},...,z_{2K})$. Then,

$$\begin{aligned} F^{11} &= \theta a_{11}^{\theta\rho}z_{11}^{-\theta\rho-1}\Big(\sum_{k=1}^{K}(z_{1k}/a_{1k})^{-\theta\rho}\Big)^{\frac{1}{\rho}-1} \\ &\quad \times \exp\left[-\left\{\Big(\sum_{k=1}^{K}(z_{1k}/a_{1k})^{-\theta\rho}\Big)^{\frac{1}{\rho}} + \Big(\sum_{k=1}^{K}(z_{2k}/a_{2k})^{-\theta\rho}\Big)^{\frac{1}{\rho}}\right\}\right] \end{aligned}$$

The probability distribution of $z_{11}(\omega)$ conditional on $\omega \in \Omega_{11}$,

$$\begin{aligned} \widetilde{F}_{11}(z) &\equiv \Pr\Big(z_{11}(\omega) \le z \mid \omega \in \Omega_{11}\Big) \\ &= \frac{1}{\Pr(\omega \in \Omega_{11})}\Pr\Big(z_{11}(\omega) \le z,\ h_{sk}z_{sk}(\omega) \le h_{11}z_{11}(\omega)\Big) \\ &= \frac{1}{\pi_{11}}\Pr\Big(z_{11}(\omega) \le z,\ z_{sk}(\omega) \le \frac{h_{11}}{h_{sk}}z_{11}(\omega)\Big) \\ &= \frac{1}{\pi_{11}}\int_{z_{11}=0}^{z} F^{11}(z, \frac{h_{11}}{h_{12}}z, ..., \frac{h_{11}}{h_{1K}}z, \frac{h_{11}}{h_{21}}z, \frac{h_{11}}{h_{22}}z, ..., \frac{h_{11}}{h_{2K}}z)dz_{11} \\ &= \frac{1}{\pi_{11}}\int_{z_{11}=0}^{z} \theta a_{11}^{\theta\rho}z^{-\theta\rho-1}\Big((z/a_{11})^{-\theta\rho} + \sum_{k=2}^{K}(\frac{h_{11}z}{h_{1k}a_{1k}})^{-\theta\rho}\Big)^{\frac{1}{\rho}-1} \\ &\quad \times \exp\left[-\left\{\Big((z/a_{11})^{-\theta\rho} + \sum_{k=2}^{K}(\frac{h_{11}z}{h_{1k}a_{1k}})^{-\theta\rho}\Big)^{\frac{1}{\rho}} + \Big(\sum_{k=1}^{K}(\frac{h_{11}z}{h_{2k}a_{2k}})^{-\theta\rho}\Big)^{\frac{1}{\rho}}\right\}\right]dz_{11} \\ &= \frac{1}{\pi_{11}}\int_{z_{11}=0}^{z} \theta a_{11}^{\theta}z^{-\theta-1}\Big(1 + \sum_{k=2}^{K}(\frac{h_{11}a_{11}}{h_{1k}a_{1k}})^{-\theta\rho}\Big)^{\frac{1}{\rho}-1} \\ &\quad \times \exp\left[-\left\{z^{-\theta}a_{11}^{\theta}\Big(1 + \sum_{k=2}^{K}(\frac{h_{11}a_{11}}{h_{1k}a_{1k}})^{-\theta\rho}\Big)^{\frac{1}{\rho}} + z^{-\theta}a_{11}^{\theta}\Big(\sum_{k=1}^{K}(\frac{h_{11}a_{11}}{h_{2k}a_{2k}})^{-\theta\rho}\Big)^{\frac{1}{\rho}}\right\}\right]dz_{11}, \end{aligned}$$

which can be further simplified to:

$$
\begin{aligned}
\widetilde{F}_{11}(z) &\equiv \Pr\Big(z_{11}(\omega) \leq z \mid \omega \in \Omega_{11}\Big) \\
&= \frac{1}{\pi_{11}} \int_{z_{11}=0}^{z} \theta a_{11}^{\theta} z^{-\theta-1} \Big((h_{11}a_{11})^{-\theta\rho} \Big[\sum_{k=1}^{K} (h_{1k}a_{1k})^{\theta\rho}\Big]\Big)^{\frac{1}{\rho}-1} \\
&\qquad \times \exp\Bigg[- z^{-\theta} a_{11}^{\theta} \Bigg\{ \Big((h_{11}a_{11})^{-\theta\rho} \sum_{k=1}^{K} (h_{1k}a_{1k})^{\theta\rho}\Big)^{\frac{1}{\rho}} + \Big((h_{11}a_{11})^{-\theta\rho} \sum_{k=1}^{K} (h_{2k}a_{2k})^{\theta\rho}\Big)^{\frac{1}{\rho}} \Bigg\}\Bigg] dz_{11} \\
&= \int_{z_{11}=0}^{z} \theta a_{11}^{\theta} z^{-\theta-1} \Big((h_{11}a_{11})^{-\theta}(H_1^{\theta} + H_2^{\theta})\Big) \\
&\qquad \times \exp\Bigg[- z^{-\theta} a_{11}^{\theta} \Big((h_{11}a_{11})^{-\theta}(H_1^{\theta} + H_2^{\theta})\Big)\Bigg] dz_{11}
\end{aligned}
$$

This is a Fréchet distribution with location parameter $a_{11}^{\theta}\Big((h_{11}a_{11})^{-\theta}(H_1^{\theta} + H_2^{\theta})\Big)$ and dispersion parameter θ. Note that probability of choosing (s,k) equals probability of choosing k conditional on s times the probability of choosing s, $\pi_{sk} = \alpha_{k|s}\alpha_s$. The "inverse of location parameter" is

$$
a_{11}^{-\theta} \frac{(h_{11}a_{11})^{\theta}}{H_1^{\theta} + H_2^{\theta}} = a_{11}^{-\theta} \Big(\frac{(h_{11}a_{11})^{\theta\rho}}{H_1^{\theta\rho}}\Big)^{1/\rho} \frac{H_1^{\theta}}{H_1^{\theta} + H_2^{\theta}} = a_{11}^{-\theta} \alpha_{1|1}^{1/\rho} \alpha_1
$$

Similarly, the distribution of $\Pr\Big(z_{11}(\omega) \leq z \mid \omega \in \Omega_{11}\Big)$ is Frećhet, and its inverse of location parameter is $(a_{sk}^{-\theta_1} \alpha_{k|s}^{\theta_1/\theta_2} \alpha_s)$ for $\theta_2 = \theta\rho$ and $\theta_1 = \theta$. Expected value of a Fréchet distributed random variable with location parameter T and dispersion parameter θ equals $\bar{\gamma} T^{1/\theta}$ with $\bar{\gamma} \equiv \Gamma(1 - 1/\theta)$. Thus, here the expected value conditional on selection equals

$$
\mathbb{E}(z_{sk}(\omega) | \omega \in \Omega_{sk}) = \bar{\gamma} \Big[a_{sk}^{-\theta_1} \alpha_{k|s}^{\theta_1/\theta_2} \alpha_s\Big]^{-1/\theta_1} = \bar{\gamma} \Big(a_{sk}\Big) \Big(\alpha_{k|s}\Big)^{-1/\theta_2} \Big(\alpha_s\Big)^{-1/\theta_1}
$$

C.5 Derivations for recasting the micro to macro problem

In this section, we recast the land use problem onto crop supply. We show (i) that the following problem reproduces equation (22), and (ii) the Lagrange multipliers reproduce returns to land. Using

$Q^f_{i,k\tau} = (1/\gamma^L_{k\tau})\widetilde{h}_{i,k\tau}\widetilde{L}^f_{i,k\tau}$, the problem of the landowner in Section 3.5 can be written as:

$$\begin{aligned}
&\max_{\{Q^f_{i,k\tau}\}_{k,\tau}\ ,\ \{\widetilde{Q}^f_{i,k}\}_k} && \sum_{\tau\in\mathcal{T}}\sum_{k\in\mathcal{K}}\gamma^L_{k\tau}p_{i,k}Q^f_{i,k\tau} \\
&\text{subject to} && \left[\sum_{\tau\in\mathcal{T}}\left(\frac{Q^f_{i,k\tau}}{v^f_{i,k\tau}}\right)^{\frac{\theta_2}{\theta_2-1}}\right]^{\frac{\theta_2-1}{\theta_2}} \leq \widetilde{Q}^f_{i,k} \\
& && \left[\sum_{k\in\mathcal{K}}(\widetilde{Q}^f_{i,k})^{\frac{\theta_1}{\theta_1-1}}\right]^{\frac{\theta_1-1}{\theta_1}} \leq L^f_i
\end{aligned}$$

where

$$v^f_{i,k\tau} = \widetilde{h}_{i,k\tau}a^f_{i,k\tau}(\gamma^L_{k\tau})^{-1}.$$

The Lagrangian function is:

$$\mathcal{L} = \sum_\tau\sum_k\gamma^L_{k\tau}p_{i,k}Q^f_{i,k\tau} - \lambda^f_{i,k}\left\{\left[\sum_\tau\left(\frac{Q^f_{i,k\tau}}{v^f_{i,k\tau}}\right)^{\frac{\theta_2}{\theta_2-1}}\right]^{\frac{\theta_2-1}{\theta_2}} - \widetilde{Q}^f_{i,k}\right\} - \mu^f_i\left\{\left[\sum_k(\widetilde{Q}^f_{i,k})^{\frac{\theta_1}{\theta_1-1}}\right]^{\frac{\theta_1-1}{\theta_1}} - L^f_i\right\}$$

Provided that the solution is interior, and quantities are all positive, the first order conditions require that:

$$\gamma^L_{k\tau}p_{i,k} = \lambda^f_{i,k}(v^f_{i,k\tau})^{-\frac{\theta_2}{\theta_2-1}}(Q^f_{i,k\tau})^{\frac{1}{\theta_2-1}}(\widetilde{Q}^f_{i,k})^{-\frac{1}{\theta_2-1}} \tag{A.9}$$

$$\lambda^f_{i,k} = \mu^f_i(\widetilde{Q}^f_{i,k})^{\frac{1}{\theta_1-1}}(L^f_i)^{-\frac{1}{\theta_1-1}} \tag{A.10}$$

Using equation (A.9), and $v^f_{i,k\tau} = \widetilde{h}_{i,k\tau}a^f_{i,k\tau}(\gamma^L_{k\tau})^{-1}$,

$$Q^f_{i,k\tau} = (\lambda^f_{i,k})^{-(\theta_2-1)}(\gamma^L_{k\tau})^{-1}(p_{i,k})^{\theta_2-1}(a^f_{i,k\tau}\widetilde{h}_{i,k\tau})^{\theta_2}\widetilde{Q}^f_{i,k}$$

or, equivalently,

$$\frac{Q^f_{i,k\tau}}{v^f_{i,k\tau}} = (\lambda^f_{i,k})^{-(\theta_2-1)}(p_{i,k})^{\theta_2-1}(a^f_{i,k\tau}\widetilde{h}_{i,k\tau})^{\theta_2-1}\widetilde{Q}^f_{i,k} \tag{A.11}$$

Recall the definition of $H^f_{i,k}$ from equation (19),

$$H^f_{i,k} = \left[\sum_\tau(a^f_{i,k\tau}p_{i,k}\widetilde{h}_{i,k\tau})^{\theta_2}\right]^{\frac{1}{\theta_2}}$$

Using equation (A.11),

$$\underbrace{\Big[\sum_{\tau}\Big(\frac{Q^f_{i,k\tau}}{v^f_{i,k\tau}}\Big)^{\frac{\theta_2}{\theta_2-1}}\Big]^{\frac{\theta_2-1}{\theta_2}}}_{\widetilde{Q}^f_{i,k}} \quad = \quad (\lambda^f_{i,k})^{-(\theta_2-1)}\widetilde{Q}^f_{i,k}(H^f_{i,k})^{\theta_2-1}$$

which delivers the shadow price of crop k $\lambda^f_{i,k}$ precisely equal to $H^f_{i,k}$,

$$\lambda^f_{i,k} = H^f_{i,k} \tag{A.12}$$

Using equation (A.10),

$$\widetilde{Q}^f_{i,k} = (\lambda^f_{i,k})^{\theta_1-1}(\mu^f_i)^{-(\theta_1-1)}L^f_i \tag{A.13}$$

which we use to derive the following relationship:

$$\underbrace{\Big[\sum_{k}(\widetilde{Q}^f_{i,k})^{\frac{\theta_1}{\theta_1-1}}\Big]^{\frac{\theta_1-1}{\theta_1}}}_{L^f_i} = (\mu^f_i)^{-(\theta_1-1)}L^f_i\Big[\sum_k(\lambda^f_{i,k})^{\theta_1}\Big]^{\frac{\theta_1-1}{\theta_1}}$$

Replacing $\lambda^f_{i,k} = H^f_{i,k}$ we find the shadow price of cropland, μ^f_i,

$$\mu^f_i = \Big[\sum_k(H^f_{i,k})^{\theta_1}\Big]^{\frac{1}{\theta_1}} \tag{A.14}$$

Plug μ^f_i from (A.14) into (A.13),

$$\widetilde{Q}^f_{i,k} = (\lambda^f_{i,k})^{\theta_1-1}\Big[\sum_k(\lambda^f_{i,k})^{\theta_1}\Big]^{-\frac{\theta_1-1}{\theta_1}}L^f_i = \left[\frac{(H^f_{i,k})^{\theta_1}}{\sum_k(H^f_{i,k})^{\theta_1}}\right]^{\frac{\theta_1-1}{\theta_1}}L^f_i$$

Putting things together,

$$Q^f_{i,k\tau} = (\gamma^L_{k\tau})^{-1}a^f_{i,k\tau}\widetilde{h}_{i,k\tau}\left[\frac{(a^f_{i,k\tau}\widetilde{h}_{i,k\tau})^{\theta_2}}{(H^f_{i,k})^{\theta_2}}\right]^{\frac{\theta_2-1}{\theta_2}}\left[\frac{(H^f_{i,k})^{\theta_1}}{\sum_k(H^f_{i,k})^{\theta_1}}\right]^{\frac{\theta_1-1}{\theta_1}}L^f_i$$

which is the same as equation (22).