

WP/18/145

IMF Working Paper

The Globalization of Farmland: Theory and Empirical Evidence

by Rabah Arezki, Christian Bogmans, and Harris Selod

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

I N T E R N A T I O N A L M O N E T A R Y F U N D

Research Department

The Globalization of Farmland: Theory and Empirical Evidence¹

Prepared by Rabah Arezki, Christian Bogmans, and Harris Selod

Authorized for distribution by Gian Maria Milesi-Ferretti

June 2018

Abstract

This paper is the first to provide both theoretical and empirical evidence of farmland globalization whereby international investors directly acquire large tracts of agricultural land in other countries. A theoretical framework explains the geography of farmland acquisitions as a function of cross-country differences in technology, endowments, trade costs, and land governance. An empirical test of the model using global data on transnational deals shows that international farmland investments are on the aggregate likely motivated by re-exports to investor countries rather than to world markets. This contrasts with traditional foreign direct investment patterns where horizontal as opposed to vertical FDI dominates.

JEL classification: F1 Trade; F2 International Factor Movements and International Business; Q1. Agriculture; Q17 Agriculture in International Trade.

Keywords: Large-Scale Land Acquisitions, Food Independence, Platform FDI, Land Governance.

Authors' E-Mail Addresses: rarezki@worldbank.org, cbogmans@imf.org, hselod@worldbank.org.

¹ We are grateful to Vanessa Diaz Montelongo, Ziqi Li and Amjad Khan for excellent research assistance, to Brian Blankespoor for guidance on the construction and use of geographic data, and to Jeffrey Frankel, Davide Malacrino, Akito Matsumoto, François de Soyres, and the participants of the EAERE 2017 Conference in Athens, the 2018 Land Governance and Poverty Conference held at the World Bank, and the IMF Research Department seminar for very useful comments.

1 Introduction

Transnational acquisitions of land raise important questions regarding food security and economic development. Looking back in history, the surge in land acquisitions over the past decade brings back memories from the colonial era. In the 15th century, Portugal and Spain, the then most advanced maritime powers, conquered a large number of countries around the world and subsequently secured control over these countries' land resources and their subjugated populations. Subsequent colonial empires included the French, British, Dutch and Japanese, who also imposed their rules on foreign territories and controlled vast swathes of land. These invasions led to a dramatic expansion of global trade in natural resources and agricultural products, as well as in slaves and indentured servants to work on plantations.

Most recently, following the 2007-2008 spike in food prices, there has been a booming interest in the direct acquisition of farmland in developing and emerging markets, often involving multinational companies and foreign governments. The recent expansion in the globalization of farmland has led to a polarized debate between those welcoming foreign investments, hoping they will help raise agricultural land yields and alleviate poverty, and those who see the phenomenon as a "land grab" (Financial Times, 2016 and Bloomberg, 2017). The present paper is the first to provide both theoretical and empirical evidence of this new wave of investments, which marks a new trend towards the globalization of farmland.

The increased interest of international investors to acquire farmland is part of a broader set of developments that are changing the nature of commercial agriculture at the global scale. These include the increased importance of multinational companies and foreign direct investment in promoting sectoral growth, and a more prominent role for global value chains in expanding food supply (Maertens and Swinnen, 2015). In this ongoing process of agricultural globalization, the volume of international trade in agricultural commodities increased almost five-fold in a period of three decades, from about \$200 billion in 1980 to almost \$1.1 trillion in 2010, the largest growth recorded by any sector in that period.

Interestingly, the acquisition of large tracts of land by private investors and sovereign wealth funds coincides with the rising demand for food associated with rising incomes and growing populations. This rising interest in farmland suggests that globalization has entered a new phase, one that is defined by the integration of pristine land in developing countries into the world economy. According to Collier and Venables (2012), this transition has parallels with China's process of economic development. Like the increase in foreign investment into China driven by its abundance of labor, the sudden shift in the foreign acquisition of land in Africa and elsewhere may have been triggered by large spatial differences in factor productivity. But there are also significant differences between the globalization of China's cheap labor and the globalization of farmland. Based on an early empirical account of the phenomenon, Arezki et al. (2015) provide evidence that investors tend to favor countries where local populations have

insecure property rights, echoing the concerns of many global observers. Surprisingly, the logic underpinning these investments thus seems to run counter not only to the standard findings in the foreign direct investment literature, but also to the parallel story of China opening up to the rest of the world following institutional improvement (Song et al., 2011 and di Giovanni et al., 2014).

In this paper, we present a multi-country model of international trade with two sectors of production (agriculture and manufacturing) and two factors of production (labor and land). We assume that investors can use a home-country specific technology to engage in multinational production abroad. Agricultural varieties are both host-country *and* investor-country specific *à la* Armington (1969). The rationale behind this double Armington assumption is that whereas differences in agro-ecological conditions leads to differentiated agricultural products around the globe, agricultural products also differ given the diversity of agricultural technologies used by different investors or because of their different branding. In each country, there is an exogenous and fixed land price, which influences the quantity of land supplied to the commercial agricultural sector.

Using this framework, we derive a bilateral gravity-type specification of demand for land. We first do this in a model with investor differentiation that has both horizontal and vertical FDI motives, and then for an alternative model without investor differentiation that only has the vertical motive. Under both motives, bilateral investment in land is shown to depend on endowments and technology, as well as access to world markets (or its inverse, “remoteness”). In the first model, we find that host-country remoteness negatively affects investment as investors prefer to produce close to export markets. In the second model, however, due to the absence of investor differentiation, host country-remoteness implies less investor competition and therefore encourages investment.

To test the predictions of the model, we combine data from the Land Matrix, an online database of large-scale land acquisitions, with country-level data on land endowments, land governance and agricultural productivity per worker. We proxy for remoteness, that is, for the inverse of market potential, in three different ways. Furthermore, following Silva and Tenreyro (2006), Fally (2015), Anderson and Yotov (2016) and others, we estimate our gravity-type specification using Poisson Pseudo Maximum Likelihood (PPML), to deal with heteroskedasticity and a large proportion of zero observations.

Our novel result is that host-country remoteness—a lack of host country access to agricultural consumer markets—has been a positive and sizeable determinant of large-scale bilateral land acquisitions. Whereas an increase in the host country’s suitable land endowment by one standard deviation from the mean is found to increase the number of bilateral deals by 86.1%, an increase in the remoteness of the host-country by one standard deviation from the mean is expected to increase the numbers of deals per country pair by as much as 115.8%, depending on the remoteness proxy that is considered. In sum, host-country remoteness appears to be a relevant determinant

of investment, and its importance is comparable to or even exceeds that of agricultural land endowments in explaining large-scale land acquisitions. This finding is fully consistent with the food independence motive underpinning such investments.

Our finding that host-country remoteness plays a positive role in driving bilateral investment may reflect a relative scarcity of land at the global scale. Indeed, in our model, under low supply but high demand for land, investors are competing with one another and look to invest in remote countries to shield themselves from competition. To a certain extent, our finding contradicts the intuition formulated by Collier and Venables (2012) in the case of Africa that land markets are “land-abundant, investor-scarce”, with demand very small relative to potential supply. The two observations can be reconciled by noticing that although land in general may be abundant, what matters is the amount of available and suitable land, which is available in smaller quantities.

The remainder of this paper is as follows. Section 2 presents a multi-country model of trade and land acquisitions. Section 3 uses the model to derive a gravity style equation for domestic and transnational demand for land. Section 4 presents the econometric specifications for our cross-country analysis. Section 5 describes the data used. Section 6 estimates the determinants of transnational land deals using bilateral regressions, and focusses on the role of remoteness. Section 7 concludes.

2 Modeling international trade and land acquisitions

2.1 The general setting

In this section we present a multi-country model that accounts for international investments in land and for trade of agricultural products. The world consists of N countries. Each country produces an agricultural good (A) and a manufactured good (M) using labor (L) and land (T). In each country $i = 1 \dots N$, factors of production are available in quantities L_i and T_i . Although labor can be used in both the manufacturing and agricultural sector, land is used for the production of agricultural goods only. Labor is domestically mobile (between the agricultural and the manufacturing sector within each country) and capital is internationally mobile.

In-country and multinational production The manufacturing and the agricultural sectors produce goods that are differentiated by the country of origin. For the agricultural sector, this is consistent with trade being motivated by the consumption of different varieties of agricultural products (see Costinot et al., 2016). Because agricultural potential varies across countries, producers grow different crop varieties of the same crop or different crops across countries. Our model, however, offers an extra layer of differentiation for agricultural goods as investors from a country i can also engage in production of an agricultural variety in another country l , something we refer to as *multinational production*. Differentiation according to both the country of

production and the origin of the investor is a realistic feature of the model as consumers may for instance differentiate between coffee sold by a national producer of a Latin American country (i.e., a “fair trade” producer) and coffee produced in the same country by a multinational firm. We will detail in the next subsection how these two types of differentiations affect consumer utilities. For the time being, simply note that on the production side, the agricultural technologies Z_{li} used by investor country i differ for each host country l and are given by a vector of Total Factor Productivity (TFP) terms, $Z_i \equiv \{Z_{li}, l = 1 \dots N\}$. With these notations, Z_{li} is the technology for own-country production of the agricultural variety by country l . For simplicity, we abstract from multiple crop choices by assuming that an investor country produces only one type of differentiated crop in a given country.¹ Also note that, as standard in the literature, trade is subject to iceberg costs: in order to sell one unit of good to consumers in importing country n , a firm producing in country l must ship $t_{nl} \geq 1$ units of its product.²

Platform and food-independence motivated investments in land Our representation of multinational production in the agricultural sector revolves around the transfer of technology from an investor country to a host country (see Ramondo and Rodríguez-Clare, 2013). As in the Ramondo and Rodríguez-Clare framework, investors can engage in multinational production and use a host specific variety of their home-country technology to produce in close proximity to other markets (thereby saving on trade costs t_{nl} as in Brainard (1997), Markusen and Venables (1998), Helpman et al. (2004) and others) or in order to benefit from lower production costs in the host country. In the main version of our model, both features of the the Ramondo and Rodríguez-Clare framework are present and producers invest abroad with the purpose of re-exporting to world markets, a pattern of investment to which we, following Ekholm et al. (2003), refer as “*platform*” FDI. Hence, in the platform FDI version of our model, investors are competing with other investors for (i) access to resources (i.e., labor and land) in the host countries, as well as (ii) for access to agricultural consumer markets in importing countries. In addition, we also consider an alternative model in which country i invests in country l in order to benefit from lower local production costs while “reimporting” the agricultural good to its own domestic market only (see section (3.2)). As we will show in more detail in section (3.2), this pattern is obtained by shutting down the host-country market access channel. We will interpret investment that follows this *vertical* FDI pattern as *food-independence* driven. In the rest of this section, we present the general structure of our framework, which is common to both the platform FDI and the food independence models.³

¹We also abstract from considering multinational production in the manufacturing sector, an issue which is outside the scope of this paper.

²For simplicity and without loss of generality, we do not differentiate the trade costs of the manufacturing and the agricultural goods. Introducing sector-specific trade costs would only make notations burdensome without changing the intuition of the model.

³While FDI typically involves the investment of physical capital, here investors bring “knowledge capital” , i.e., agricultural technology, and hire local workers to work the land they rent.

2.2 Preferences and consumption

Each country n has a representative agent who consumes all varieties of the agricultural and manufacturing goods. We assume that the upper-tier utility function is Cobb-Douglas:

$$u_n = (Q_n^A)^\gamma (Q_n^M)^{1-\gamma} \quad (1)$$

where Q_n^A and Q_n^M are the quantities of the composite agricultural good and composite manufacturing good consumed in country n (see equations (2) and (3) below for the exact formulas), and $\gamma \in]0, 1[$ and $1 - \gamma$ represent the expenditure shares on agricultural and manufactured goods respectively.

To account for the idea that consumers differentiate between goods based on the country of origin, we specify agricultural consumption and manufacturing consumption in country n as CES aggregates over the N number of discrete varieties that are differentiated by country of production, with

$$Q_n^A = \left(\sum_{l=1}^N (q_{nl}^A)^{\frac{\varepsilon_A - 1}{\varepsilon_A}} \right)^{\frac{\varepsilon_A}{\varepsilon_A - 1}} \quad (2)$$

and

$$Q_n^M = \left(\sum_{l=1}^N (q_{nl}^M)^{\frac{\varepsilon_M - 1}{\varepsilon_M}} \right)^{\frac{\varepsilon_M}{\varepsilon_M - 1}} \quad (3)$$

where q_{nl}^A (respectively q_{nl}^M) is the consumption of agricultural goods (respectively manufactured goods) produced in country l and consumed by the representative consumer in country n , and $\varepsilon_A > 1$ (respectively $\varepsilon_M > 1$) is the elasticity of substitution between agricultural varieties (respectively manufacturing varieties) produced in different countries.

Because our model allows for differentiated multinational production of the agricultural good, we further specify the consumption in country n of the agricultural goods produced in country l as a CES aggregate over the N number of discrete varieties from different investor countries, with

$$q_{nl}^A = \left(\sum_{i=1}^N (q_{nli}^A)^{\frac{\sigma - 1}{\sigma}} \right)^{\frac{\sigma}{\sigma - 1}} \quad (4)$$

where q_{nli}^A represent country n 's consumption of the agricultural variety that is produced in country l by investors from country i , and $\sigma > 1$ is the elasticity of substitution between agricultural varieties produced by different investors. Observe that if $\sigma \rightarrow \infty$, then the agricultural varieties are not differentiated by investor country.

In this “double Armington” setting for the consumption of agricultural goods, observe that (2) accounts for product differentiation by country of origin (*where* the good is produced), whereas (4) accounts for product differentiation by investor country (*by whom* the good is produced).

2.3 Production

Firms in the manufacturing sector produce a manufacturing good under constant returns to scale using labor L_i^M only:

$$y_i^M = B_l L_i^M \quad (5)$$

where we have assumed that in order to produce one unit of the manufactured good a producer in country l requires $1/B_l$ units of labor. Production of agricultural commodities by an investor from country i in host country l requires the input of land T_{li}^A and labor L_{li}^A , which are acquired locally in host country l , under a constant-return to scale Cobb Douglas production technology:

$$y_{li}^A = Z_{li} \left(\frac{L_{li}^A}{\alpha} \right)^\alpha \left(\frac{T_{li}^A}{1-\alpha} \right)^{1-\alpha} \quad (6)$$

where α and $1 - \alpha$ are the output elasticities with respect to labor and land. To account for adjustment costs in the export of agricultural technology or adaptation to foreign contexts, we assume that agricultural TFP may decay with distance between the investor and the host country in a multiplicative way.⁴ We thus have

$$Z_{li} = Z_i F(t_{il}) \quad (7)$$

where Z_i is the intrinsic investor productivity and F is a decreasing function of transport costs t_{il} .

2.4 Equilibrium

Assuming perfect competition for each variety, the producer price of each manufacturing variety equals its unit cost. Based on the manufacturing production function (5), this implies that the price for the manufactured good in country l is:

$$p_i^M = \frac{1}{B_l} w_l \quad (8)$$

⁴Our assumption is based on the ideas of Diamond (1999), who hypothesized that the diffusion of agricultural knowledge and technology is inhibited by geography.

where w_l is the return to labor in country l . Similarly, under perfect competition, the unit cost function associated with the agricultural production function (6) of investor i in country l is:

$$c_{li}^A = \frac{(w_l)^\alpha (f_l)^{1-\alpha}}{Z_{li}} \quad (9)$$

where f_l is the price of agricultural land in country l .

Consumers maximize utility by choosing how much to consume of each manufacturing variety and each agricultural variety. We can solve the utility maximization problem of consumers in three stages. First, at the top-tier, consumers maximize (1) by choosing agricultural and manufacturing consumption subject to the budget constraint, $I_n = P_n^A Q_n^A + P_n^M Q_n^M$, where I_n , P_n^A and P_n^M are defined as national income and the respective agricultural and the manufacturing price indices in country n . We obtain expenditures D_n^A and D_n^M on agricultural and manufactured goods consumed in country n respectively:

$$D_n^A = \gamma I_n \quad (10)$$

and

$$D_n^M = (1 - \gamma) I_n \quad (11)$$

Second, consumers maximize agricultural utility (2) and manufacturing utility (3) by allocating total expenditures across the agricultural and manufacturing varieties from different countries. Let p_{nl}^A be the consumer price in country n of the agricultural varieties produced in country l . Then X_{nl}^A and X_{nl}^M , the respective consumptions by country n of the agricultural and manufacturing varieties produced in country l , are given by the following Marshallian demand functions:

$$X_{nl}^A = \left(\frac{p_{nl}^A}{P_n^A} \right)^{1-\varepsilon_A} \gamma I_n \quad (12)$$

$$X_{nl}^M = \left(\frac{t_{nl} p_{nl}^M}{P_n^M} \right)^{1-\varepsilon_M} (1 - \gamma) I_n \quad (13)$$

where

$$P_n^A \equiv \left(\sum_{h=1}^N (p_{nh}^A)^{1-\varepsilon_A} \right)^{\frac{1}{1-\varepsilon_A}} \quad (14)$$

and

$$P_n^M \equiv \left(\sum_{h=1}^N (t_{nh} p_{nh}^M)^{1-\varepsilon_M} \right)^{\frac{1}{1-\varepsilon_M}} \quad (15)$$

are the respective consumer price indexes of the composite agricultural good and the composite manufacturing good consumed in country n , also commonly referred to as the inward multilateral resistance indices (see Anderson and van Wincoop, 2003).

Third, consumers in country n maximize agricultural subutility (4) by choosing the level of consumption of the various investor-country differentiated varieties produced in country l , taking total expenditures on varieties from country l as given. The resulting Marshallian demand function, X_{nli}^A , that is, country n consumption of the variety produced in country l by investor country i , is:

$$X_{nli}^A = \left(\frac{t_{nl} c_{li}^A}{p_{nl}^A} \right)^{1-\sigma} \left(\frac{p_{nl}^A}{P_n^A} \right)^{1-\varepsilon_A} \gamma I_n \quad (16)$$

where

$$p_{nl}^A = \left(\sum_{j=1}^N (t_{nl} c_{lj}^A)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \quad (17)$$

is the CES price index of the various agricultural varieties produced by different investors in country l and consumed in country n . Observe that the producer (“factory-gate”) price index for the agricultural sector in country l is

$$p_l^A \equiv \left(\sum_{j=1}^N (c_{lj}^A)^{1-\sigma} \right)^{1/(1-\sigma)} = \frac{p_{nl}^A}{t_{nl}} \quad (18)$$

As can be seen from equations (16) and (13), imports decrease with distance from exporters and with producer price.

Next, let us define total revenues of sector $s \in \{A, M\}$ in producer country l as the sum of imports from all countries $n \in [1, \dots, N]$ of goods produced in l (including “own imports”), so that $V_l^A \equiv \sum_{n=1}^N X_{nl}^A$ and $V_l^M \equiv \sum_{n=1}^N X_{nl}^M$. With these notations, we also define global nominal outputs as $V^A \equiv \sum_{l=1}^N V_l^A$ and $V^M \equiv \sum_{l=1}^N V_l^M$ for the agricultural and manufacturing sectors respectively. Using (12) and (13), substituting for X_{nl}^A and X_{nl}^M in the formulas for V_l^A and V_l^M gives:

$$V_l^A = (p_l^A \Omega_l^A)^{1-\varepsilon_A} V^A \quad (19)$$

$$V_l^M = (p_l^M \Omega_l^M)^{1-\varepsilon_M} V^M \quad (20)$$

where

$$\Omega_l^A \equiv \left[\sum_{n=1}^{n=N} \left(\frac{t_{nl}}{P_n^A} \right)^{1-\varepsilon_A} \frac{\gamma I_n}{V^A} \right]^{1/(1-\varepsilon_A)} \quad (21)$$

and

$$\Omega_l^M \equiv \left[\sum_{n=1}^{n=N} \left(\frac{t_{nl}}{P_n^M} \right)^{1-\varepsilon_M} \frac{(1-\gamma)I_n}{V^M} \right]^{1/(1-\varepsilon_M)} \quad (22)$$

are outward multilateral resistance indices in the agricultural and manufacturing sectors respectively. They are also the inverse of the agricultural and manufacturing market potentials of country l , that is, the weighted sum of the market sizes of all trade partners of country l .⁵ Note here that revenues (or nominal outputs) by sector decrease with “factory gate” prices and increase with market potential. Rearranging (19) to substitute for the power transform of factory-gate prices, $(p_l^A)^{1-\varepsilon_A}$, into equation (14), we can write the inward multilateral resistance terms P_n^A as:

$$P_n^A = \left(\sum_{h=1}^{h=N} \left(\frac{t_{nh}^A}{\Omega_h^A} \right)^{1-\varepsilon_A} \frac{V_h^A}{V^A} \right)^{1/(1-\varepsilon_A)} \quad (23)$$

To close the model, we must now specify how the amount of agricultural land $T_l^A \equiv \sum_{i=1}^{i=N} T_{li}$ is determined. For simplicity, we assume that the government leases land to agricultural investors at a fixed price, $f_l = \bar{f}_l$, that is determined outside of our model, and supplies a total quantity of land that meets investors’ demand at this price. Hence, in our framework f_l is exogenous and T_l^A is endogenous, and some land, in the amount of $T_l - T_l^A$, is not used for commercial agriculture.⁶

Next, we note that national income in each country is given by the sum of labor earnings and government revenue from supplying land to the agricultural sector:⁷

$$I_l = w_l L_l + f_l T_l^A \quad (24)$$

Before defining the equilibrium, let us now define aggregate demands for labour and land in each country. From equations (16) and (13), we obtain the agricultural revenues of investors from country i in country l : $V_{li}^A \equiv \sum_{n=1}^N X_{nli}^A = \left(\frac{c_{li}^A}{p_l^A} \right)^{1-\sigma} (p_l^A \Omega_l^A)^{1-\varepsilon_A} V^A$. Next, recognizing that by virtue of the Cobb-Douglas production technology in (17) a fraction $1-\alpha$ of agricultural revenues are paid out to agricultural land, we obtain a formula for land demanded by agricultural investor i in country l ,

$$T_{li}^A = \frac{(1-\alpha)V_{li}^A}{f_l} = \frac{(1-\alpha) \left(\frac{c_{li}^A}{p_l^A} \right)^{1-\sigma} (p_l^A \Omega_l^A)^{1-\varepsilon_A} V^A}{f_l} \quad (25)$$

⁵Note that these measures of market potential attach more weight to nearby markets t (smaller t_{nl}) and markets with higher prices (larger P_n^s).

⁶Note that it is straightforward to consider an alternative setting in which the government sets the supply of land to be allocated to agricultural use, $T_l^A = \bar{T}_l^A$, and lets the price f_l adjust. Alternatively, one could consider a setting in which both f_l and T_l^A are endogenous, resulting from an upward-sloping land supply curve, but this would make the model much more complicated.

⁷Our setting with perfect competition in all markets implies zero profits.

From equation (25) it then follows that the aggregate demand for land in host country l (the sum over all investors' demand for land in country l , $T_l^A = \sum_{i=1}^N T_{li}^A$) can be written as $T_l^A = \frac{1-\alpha}{f_l} (p_l^A \Omega_l^A)^{1-\varepsilon_A} V^A \leq T_l$.

Recognizing that a fraction α of agricultural revenues are paid out to agricultural labor and that all manufacturing revenues are paid out to manufacturing labor, the aggregate demand for labor (the sum over all agricultural investors' demand for labor in country l , $L_l^A = \sum_{i=1}^N L_{li}^A$, plus the demand for labor by manufacturing producers L_l^M) can be written as $L_l = \frac{\alpha}{w_l} (p_l^A \Omega_l^A)^{1-\varepsilon_A} V^A + \frac{1}{w_l} (p_l^M \Omega_l^M)^{1-\varepsilon_M} V^M$.

Note that all terms on the right-hand side of the aggregate labor demand and the aggregate land demand equations can be written as functions of the vectors $\mathbf{w} = w_1, \dots, w_N$ and $\mathbf{T}^A = T_1^A, \dots, T_N^A$, such that we can think of these equations as a system of $2N$ equations in \mathbf{w} and \mathbf{T}^A .⁸ Solving this system of equations in the unknowns \mathbf{w} and \mathbf{T}^A can also be interpreted as finding the zeros of an excess demand system $G(\mathbf{w}, \mathbf{T}^A) = \langle G^L(\mathbf{w}, \mathbf{T}^A), G^T(\mathbf{w}, \mathbf{T}^A) \rangle$, where $\mathbf{G}^L = (G_1^L, G_2^L, \dots, G_N^L)$ and $\mathbf{G}^T = (G_1^T, G_2^T, \dots, G_N^T)$ are respectively the N -dimensional excess labor demand and excess land demand vectors:

$$G_l^L(\mathbf{w}, \mathbf{T}^A) = \frac{\alpha}{w_l} (p_l^A \Omega_l^A)^{1-\varepsilon_A} V^A + \frac{1}{w_l} (p_l^M \Omega_l^M)^{1-\varepsilon_M} V^M - L_l \quad (26)$$

$$G_l^T(\mathbf{w}, \mathbf{T}^A) = \frac{1-\alpha}{T_l^A} (p_l^A \Omega_l^A)^{1-\varepsilon_A} V^A - f_l \quad (27)$$

Finally, note that we derive the country's balanced trade (BT) condition by adding up the excess demand for labor and land conditions and rearranging terms:

$$I_l = (p_l^A \Omega_l^A)^{1-\varepsilon_A} V^A + (p_l^M \Omega_l^M)^{1-\varepsilon_M} V^M \quad (28)$$

We are now ready to define the model's equilibrium.

Definition 1. An equilibrium is a vector $\langle \mathbf{w}, \mathbf{T}^A \rangle \in \mathbb{R}_{++}^{2n}$ such that $G_l^L = 0$ and $G_l^T = 0$ for $l = 1, \dots, N$, with c_{li}^A , p_l^A , P_n^A and Ω_l^A satisfying equations (9), (18), (14) and (21) respectively, and p_l^M , P_n^M and Ω_l^M satisfying equations (8), (15) and (22) respectively.

To prove that an equilibrium as in Definition 1 exists, we extend Theorem 2 in Alvarez and Lucas (2007)—which is an existence theorem for an exchange economy—to our Armington economy with

⁸Our reasoning is as follows. First, note that unit costs, prices and national incomes can be written as functions of wages and agricultural land supplies. Unit costs c_{li}^A depend on w_l , so that prices p_l^A and p_{nl}^A depend on w_l , implying that P_n^A depends on \mathbf{w} . Similarly, p_l^M depends on w_l and P_n^M depends on \mathbf{w} . National income I_l depends on w_l and T_l^A . Second, it follows that $V_l^A, V^A, V_l^M, V^M, \Omega_l^A$ and Ω_l^M are functions of the vectors \mathbf{w} and \mathbf{T}^A .

two sectors and two factors of production. We assume that all countries own a sufficient quantity of land so that an interior equilibrium always exists, i.e., $T_l^A < T_l$ for all $l = 1, \dots, N$.

Proposition 1. *There is at least one vector of wages and quantities of agricultural land, $\langle \mathbf{w}, \mathbf{T}^A \rangle \in \mathbb{R}_{++}^{2n}$, such that the markets for agricultural land and labour clear in every country: $G = 0$.*

Proof. See Appendix C. □

Now that we have established the foundations of our framework, we next derive testable hypotheses that can be used in conjunction with data on flows of cross-border land acquisitions, land availability and land governance, among other determinants.

3 The determinants of large-scale land acquisitions: Platform FDI or food independence?

In this section, we use our theoretical framework to characterize patterns of land investments and derive gravity-style equations for transnational investment in land between any given investor-host country pair. We consider two alternative models rooted in the same general framework; in the first one (presented in Section 3.1), investors produce abroad with the purpose of exporting to other countries (as in platform FDI), whereas in the second one (presented in Section 3.2), land acquisitions are driven by a food independence motive and the resulting pattern of investment resembles vertical FDI, i.e., investors from a given country will produce abroad and sell their entire production to their own domestic market. As we will show in subsection (3.2), the second version of the model is obtained making different assumptions which result in shutting down the horizontal, market-seeking FDI mechanism, leaving only the resource-seeking FDI mechanism.

By comparing the gravity equations from the two different models, we are able to formulate a number of hypotheses that allow us to look for evidence of food-independence and platform-driven FDI. We present the bilateral investment equations associated with reexport (platform) or food-independence patterns in the two subsections below. As we will show, host-country remoteness plays a different role in the two models as a determinant of FDI patterns: whereas host-country remoteness deters bilateral investment under the platform FDI model, it attracts bilateral investment in the food independence model.

3.1 Land acquisitions as platform-driven investments

We describe here the additional assumption made to obtain the platform FDI model. Following Ramondo and Rodríguez-Clare (2013), we assume $\varepsilon_A = \sigma$ and finite.⁹ We can thus rewrite

⁹This simplifying assumption equates the elasticity of substitution between agricultural varieties produced in different countries and the elasticity of substitution between agricultural varieties produced by different investors. We discuss the implication of this assumption for the empirical analysis in section (4.1).

equation (25), the amount of land in country l acquired by producers from country i , as:

$$T_{li}^A = \frac{(1 - \alpha) (c_{li}^A)^{1-\sigma}}{f_l (p_l^A)^{1-\sigma}} (p_l^A \Omega_l^A)^{1-\varepsilon_A} V^A = (1 - \alpha) \frac{(c_{li}^A \Omega_l^A)^{1-\sigma}}{f_l} V^A \quad (29)$$

Equation (29) predicts the quantity of land acquired in a bilateral relationship between any given host country l and investor country i under the platform FDI motive.

Proposition 2. *Under the platform FDI motive, the quantity of land leased by investor country i in host country l is decreasing in host-country remoteness from agricultural consumer markets Ω_l^A , decreasing in the return to labor w_l and land price f_l in host-country l (via c_{li}^A), and increasing in the agricultural productivity of the investor-country, Z_{li} (via c_{li}^A).*

Proof. The results follow directly from inspection of equations (9) and (29). □

The most interesting result summarized in Proposition 2 is the relationship between bilateral investment and host-country remoteness (or its inverse, market potential). Under the platform FDI motive, investors tend to invest less in countries that are remote from large agricultural consumer markets, as it implies a lower aggregate demand for their product. In the next subsection, we turn to another version of our model in which the relationship between investment and remoteness is reversed.

3.2 Land acquisitions for food independence

In the model presented in the previous subsection, land investment is driven by a platform FDI motive. To see this, we observe from equation (29) that land investment is increasing in market potential, reminiscent of horizontal (export market-seeking) FDI, and decreasing in the local cost of production, which is characteristic of vertical (import resource-seeking) FDI. However, by changing our assumptions, we can shut down the horizontal export market-seeking mechanism and restrict the pattern of land investment to reflect a vertical FDI motive only. Intuitively, vertical FDI occurs when each investor country has an absolute advantage in producing abroad and serving its home market over all the other investor countries. Mathematically, this can be warranted under two conditions. The first condition is that varieties are no longer differentiated by the country of the investor (i.e., when $\sigma \rightarrow \infty$), which ensures that in every host country there is only one investor country that produces (i.e., the most efficient one). Without imposing a further condition, however, the pattern of agricultural production would be indeterminate. The second condition we impose is that each investor country holds an absolute (agricultural TFP) advantage to produce abroad and reimport to its own country over all other investors.¹⁰

¹⁰Denoting Z_{nli} and t_{nli} the agricultural TFP and bilateral trade cost of investor i in host country l exporting to consumer country n , the condition requires $t_{nln}/Z_{nln} < t_{nli}/Z_{nli}$, for all $n, l, i \neq n$.

This can be the case, for instance, when the production or supply “technology” is broadly considered to incorporate marketing and logistics costs, which will typically be lower to serve the home market.¹¹ Taken together, these two assumptions *de facto* shut down competition among investors and ensure that each consumer market is served by an investor from the same country. In the Appendix we show that this set of assumptions results in the following specification for bilateral investment:

$$T_{li}^A = \left(\frac{t_{il}}{\widetilde{Z}_{li}} \right)^{1-\varepsilon_A} \frac{\gamma I_i T_l^A}{\widetilde{\Omega}_l^A \widetilde{P}_i^A} \quad (30)$$

where the following terms are redefined, with $\widetilde{\Omega}_l^A \equiv \left(\sum_{i=1}^N \left(\frac{t_{il}}{\widetilde{Z}_{li}} \right)^{1-\varepsilon_A} \frac{\gamma I_i}{V^A} \right)$, $\widetilde{P}_i^A \equiv \left(\sum_{l=1}^N \left(\frac{t_{il}}{\widetilde{Z}_{li}} \right)^{1-\varepsilon_A} \frac{\widetilde{V}_l^A}{V^A} \right)$ and $\widetilde{V}_l^A \equiv \left(w_l^\alpha f_l^{1-\alpha} \widetilde{\Omega}_l^A \right)^{1-\varepsilon_A} V^A$.

Proposition 3. *Under the food independence motive, the quantity of land leased by investor country i in host country l is increasing in host-country remoteness from agricultural consumers Ω_l^A , investor-country remoteness from agricultural producers P_i^A , agricultural productivity of the investor-country Z_{li} , expenditures on agricultural goods γI_i , agricultural land T_l^A , and decreasing in bilateral trade costs t_{il} .*

Proof. The results follow directly from inspection of equations (30). \square

Equation (30) reflects that, conditional on the amount of agricultural land that is made available to investors in equilibrium, country pairs with remote hosts and remote investors feature more deals. This is similar to the role of importer-country and exporter-country remoteness in the determination of bilateral flows in trade models, as in Anderson and van Wincoop (2003). To see this, consider the role of host-country remoteness for instance. All things equal, if the host country’s access to consumer markets in the rest of the world is weaker by virtue of the higher prices of its agricultural product in those markets, then the investor country will be able to acquire more land in the host country.

Hence, under the food independence motive, we see that the market potential of the destination country is now a *negative* determinant of bilateral land acquisitions: in equilibrium, investor countries that are far away from agricultural producers (P_i^A high) invest more in nearby countries (t_{il} low) that are far away from agricultural consumers (Ω_l^A small). In contrast, under the platform FDI configuration, investor-country remoteness from agricultural producers had no impact on land acquisitions and host-country remoteness played a negative role.

¹¹See Matsuyama (2007), who analyzes the general equilibrium consequences of making technologies of supplying goods dependent on whether the destination is home or abroad.

3.3 Differentiating between food-independence and platform-driven investments in land

In the previous subsections we used the two versions of our model to derive two different gravity-style equations that predict bilateral patterns of land investment. These equations are presented in Table 1.

Table 1: Gravity equations for land-acquisition	
Bilateral	
<i>Platform</i>	$T_{li}^A = (1 - \alpha) \frac{(c_{li}^A \Omega_i^A)^{1-\sigma}}{f_i} V^A$
<i>Food independence</i>	$T_{li}^A = \left(\frac{\frac{t_{li}^A}{Z_{li}}}{\widetilde{\Omega}_i^A P_i^A} \right)^{1-\varepsilon_A} \frac{\gamma I_i T_i^A}{V^A}$

From the overview of equations in Table (1), we observe that there are both commonalities as well as differences between the gravity equation under the platform FDI and food independence models. For example, from Table 1, we learn that if bilateral investment in land is decreasing in host-country remoteness and if bilateral investment does not depend on investor-country GDP nor on investor-country remoteness from agricultural producers, then these findings would be suggestive of land deals as a type of platform FDI.¹² If bilateral investment, however, is increasing in investor-country GDP and investor-country remoteness from agricultural producers, and if host-country remoteness from agricultural consumers stimulates investments, then this would be suggestive of food-independence-motivated FDI.

It should be noted that the remoteness terms are specified in slightly different ways in each model as can be seen by comparing the t_{nl} and $\frac{t_{li}^A}{Z_{li}}$ terms in equations (21) and (30) respectively. The reason for this is that under platform FDI, investors from different countries do not compete directly, as their home-country specific technologies are used to produce an investor-specific and country-specific variety, while in the food independence version of our model they do compete directly as investors all produce the same country-specific variety, and as such their agricultural TFP Z_{li} has a direct impact on their competitiveness in acquiring land relative to other investors. Hence the occurrence of the term $\frac{t_{li}^A}{Z_{li}}$ in the host-country remoteness term of the food independence model instead of t_{nl} in the host-country remoteness term of the platform FDI model (30). However, we show in Appendix C that because Z_{li} can be written as a multiplicative function of investor-country agricultural productivity and a bilateral distance term, then the remoteness terms $\widetilde{\Omega}_i^A$ and \widetilde{P}_i^A in the food independence model will in general be “magnified” versions of the simpler remoteness terms Ω_i^A and P_i^A in the platform-FDI model. This is because distance

¹²Note that the absence of a role for investor-country GDP and for investor-country remoteness is consistent with absence of imports.

enters twice in the remoteness formulas in the food independence model as opposed to only once in the remoteness formulas in the platform-FDI model. In economic terms, this is because in the food independence model, in addition to discouraging agricultural trade, remoteness also deters investment by inhibiting the transfer of technology. Observe, moreover, that if we further assume $F(t_{il}) = 1/\phi t_{il}$ in the distance decay function (7), the empirical proxy that we will use to measure remoteness will be exactly the same for the two models (see section 4.2 below and the Appendix). For these reasons, we will only consider empirical proxies of the remoteness term Ω_i^A as defined in the platform-FDI model. In the next section, we will use Table (1) to generate an empirical specification and hypotheses that are directly linked to our model. By testing these hypotheses, we are able to seek empirical support for the two different motives potentially driving large-scale land acquisitions.

4 Econometric Approach

4.1 Specification

In view of our theoretical framework, we are interested in two sets of hypotheses that explain transnational investments. First, we want to test the extent to which endowment of land, population size and agricultural productivity in host and investor countries affect bilateral land deals. Second, we want to test the role of remoteness in stimulating or discouraging land deals, which will also shed light on the motives for those investments (whether platform of food-independence driven). We present our theory-inspired proxies for the host-country and investor-country remoteness terms in subsection (5.1). To ensure our empirical results are robust, we rely on different approaches from the trade literature to proxy for remoteness. Based on the bilateral land investment equations presented in Table (1), we propose the following empirical specification for bilateral investment in land:

$$T_{li}^A = \gamma_0 + \sum_{m=1}^{m=4} \gamma_{1,m} \ln(DIST_{li,m}) + \gamma_2 COL_{ij} + \gamma_3 \ln(T_l) + \gamma_4 \ln(w_l) + \gamma_5 \ln(q_l) \quad (31)$$

$$+ \gamma_6 \ln(Z_i) + \gamma_7 \ln(L_i) + \gamma_8 \ln(\Omega_l^A) + \gamma_9 \ln(P_i^A) + \varepsilon_{li}$$

where T_{li}^A represents investments of country i in country l (i.e., the number of deals or total land covered under those deals) over the time period considered (for details of the relevant time periods considered in the empirical analysis, see section (5.1)).

In this bilateral gravity equation, $DIST_{li,m}$ is the physical distance between investor country i and host country l if it falls in interval m , and zero otherwise. Following, Eaton and Kortum (2002) and Anderson et al. (2015), this construction allows to account for non-linearity by decomposing the effect of physical distance into four intervals $m = 1, 2, 3$ and 4, which correspond

to distances below 3,000 kilometers, between 3,000 km (included) and 7,000 km, between 7,000 km (included) and 10,000 km, and equal to or greater than 10,000 km. COL_{ij} is a dummy variable for historical colonial ties between investor country and host country. Observe that both physical distance and colonial ties are proxies for trade costs between country i and country l .

T_l is a measure of host-country land that is both suitable and “available” (see section 5.1 for details regarding definition and construction), w_l is host-country GDP per capita (proxying for labor costs), q_l is the same host-country index that measures the recognition of preexisting land rights and associated level of tenure security in Arezki et al. (2015) (see section 5.1 of the present paper), Z_i is investor-country agriculture value-added per worker (proxying for agricultural TFP), L_i is investor-country population size (proxying for expenditures on agricultural consumption), Ω_l^A and P_i^A are measures of host-country remoteness from agricultural markets and investor-country remoteness from agricultural producers respectively (see section (4.2)), and ε_i is the error term.

Observe that we propose a single specification for equation (31) but that the expected sign of some of the regression coefficients will be model specific. Also observe that the host-country agricultural productivity is absent from the bilateral regression specification. From an estimation perspective, omitting host-country agricultural productivity is justified because it is likely correlated with other host-country regressors such as GDP per capita (which we include in our specification). From a theory perspective, it is also warranted. Indeed, in the food independence version of the model, it is the investor who brings its productivity to the host-country, and in the platform-FDI model, the irrelevance of host-country agricultural productivity is a direct consequence of our assumption that $\varepsilon_A = \sigma$ following Ramondo and Rodríguez-Clare (2013).¹³

In light of Propositions 2 and 3, our expectations are as follows: populous countries acquire more land ($\gamma_7 > 0$), all the more if they are agriculturally productive ($\gamma_6 > 0$). Investors tend to target countries with abundant suitable agricultural land ($\gamma_3 > 0$) as in Arezki et al. (2015) and Lay and Nolte (2018). They are also more interested to acquire land in countries with cheaper labor costs ($\gamma_4 < 0$). In line with the finding of Arezki et al. (2015), we expect investments to be made more difficult when preexisting land rights are protected ($\gamma_5 < 0$) and when distance between the investor and the host country is large ($\gamma_{1,m} < 0$ for all m). Finally, observe that host-country remoteness from agricultural markets can play in different ways, depending on whether the motive for investment is food independence ($\gamma_8 > 0$) or platform FDI ($\gamma_8 < 0$). If the motive is food independence, we also expect more investments by investor countries that are remote from food producing countries, i.e., that have what Anderson and van Wincoop (2003) refer to as high outward multilateral resistance ($\gamma_9 > 0$).

¹³To see the latter point, note that host-country productivity, producer price, and unit cost can be written as $\widetilde{Z}_l \equiv \left(\sum_{j=1}^{j=N} \left(\frac{1}{Z_{lj}} \right)^{1-\sigma} \right)^{-\frac{1}{1-\sigma}}$, $p_l = \frac{1}{\widetilde{Z}_l} w_l^\alpha f_l^{1-\alpha}$ and $c_{li}^A = \frac{1}{\widetilde{Z}_{li}} w_l^\alpha f_l^{1-\alpha}$ respectively, so that the bilateral investment equation can be written as $T_{li}^A = \frac{(1-\alpha) \left(\frac{1}{\widetilde{Z}_{li}} \right)^{1-\sigma} \left(w_l^\alpha f_l^{1-\alpha} \right)^{1-\varepsilon_A} \left(\frac{1}{\widetilde{Z}_l} \right)^{\varepsilon_A - \sigma} \left(\Omega_l^A \right)^{1-\varepsilon_A}}{f_l} V^A$. Assuming $\varepsilon_A = \sigma$ and given that $Z_{li} = Z_i F(t_{il})$, it is easy to see T_{li}^A is a function of Z_i but not \widetilde{Z}_l .

4.2 Estimation

We estimate equation (31) using Poisson Pseudo-Maximum Likelihood (PPML), which provides a consistent estimator for our bilateral gravity equation that is robust to measurement error, heteroskedasticity and the inclusion of zeros (see Silva and Tenreyro, 2006). An important challenge in estimating our gravity-style equations is how to properly construct a measure of remoteness. Anderson and van Wincoop (2003) have shown that if gravity equations do not account for these terms, the other coefficients of interest (such as e.g. trade costs) may become biased (Baldwin and Taglioni, 2006). To address this issue, we alternatively consider four different approaches to measure or account for remoteness. We present them sequentially below.

First, we use linear approximations to the remoteness terms *a la* Baier and Bergstrand (2009).¹⁴ Namely, we account for host-country l 's remoteness with the GDP-weighted average distance of that country to all other countries, which accounts for agricultural export potential. We measure investor country i 's remoteness as the agricultural GDP-weighted average distance to all host countries, which accounts of agricultural import potential. Hence, we use:

$$\Omega_l^A = \prod_{n=1, n \neq l}^N t_{nl}^{\theta_n} \quad (32)$$

$$P_i^A = \prod_{l=1, l \neq i}^N t_{il}^{\xi_l} \quad (33)$$

where t_{nl} (t_{il}) is the physical distance between consumer country n (investor country i) and host country l , θ_n is the importer's share of global GDP, and ξ_l is the exporter's share of global agricultural production.

Note that both remoteness measures eliminate the first-order endogeneity between GDP and bilateral investment by removing the internal distance component $t_{li}^{\theta_l}$. The linear approximations suggested by Baier and Bergstrand (2009) are a particularly attractive solution to the puzzle of measuring host-country and investor-country remoteness, as it allows us to compute the remoteness measures directly using observable information on distance, GDP and agricultural GDP. A disadvantage of this approach, however, is that one would preferably use a direct measures of trade cost t_{nl} instead of physical distance to calculate these remoteness terms.

Second, another approach therefore improves upon the proxy suggested by Baier and Bergstrand (2009) by using estimates of bilateral agricultural trade costs instead of distance. To come up with trade-cost estimates, we use a two-stage structural gravity approach recently proposed by Anderson and Yotov (2016) and Anderson et al. (2015). In the first step of the two-stage

¹⁴Based on Monte Carlo analysis, Baier and Bergstrand (2009) show that their reduced form approach to remoteness (which we use in our paper) results in regression coefficients that are virtually identical to those estimated in the structural approach of Anderson and van Wincoop (2003).

procedure we estimate a dynamic panel version of the bilateral agricultural trade equation (12) in multiplicative form using PPML. This estimation allows us to obtain estimates of the bilateral fixed effects $\hat{\mu}_{nl}$ for the country pairs for which agricultural imports are observed:

$$X_{nl,t}^A = \exp[\pi_{l,t} + \chi_{n,t} + \mu_{nl} + \beta_1 RTA_{nl,t}] \times \varepsilon_{nl,t} \quad (34)$$

where $\pi_{l,t}$, $\chi_{n,t}$ and μ_{nl} respectively represent the producer/exporter time-varying fixed effect, the consumer/importer time-varying fixed effect, and the pair fixed effect, and $RTA_{nl,t}$ is a dummy variable that indicates the presence of a regional trade agreement between countries l and n at time t —the inclusion of which is necessary to obtain estimates of the bilateral fixed costs—and $\varepsilon_{nl,t}$ is a remainder error term. Since we do not observe agricultural trade flows for all possible country-pair combinations, we add an additional step to our procedure. In that second stage, we use the estimates of the pair fixed effects $\hat{\mu}_{nl}$ as a dependent variable and regress them on a set of standard determinants of trade costs and importer and exporter fixed effects. This can be written:

$$\exp[\hat{\mu}_{nl}] = \exp \left[\pi_l + \chi_n + \sum_{m=1}^{m=4} \eta_{1,m} \ln(DIST_{ln,m}) + \eta_2 CONTIG_{nl} + \eta_3 COMLANG_{nl} + \eta_4 COL_{nl} \right] \times \varepsilon_{nl} \quad (35)$$

where $CONTIG_{nl}$ and $COMLANG_{nl}$ are dummy variables for respectively a common border and common language between countries n and l . The predicted values from this second stage regression, that is,

$$\hat{t}_{nl}^{1-\varepsilon_A} = \exp[\hat{\pi}_l + \hat{\chi}_n + \hat{\eta}_1 \ln(DIST_{nl}) + \hat{\eta}_2 CONTIG_{nl} + \hat{\eta}_3 COMLANG_{nl} + \hat{\eta}_4 COL_{nl}],$$

are then used to complete the full bilateral trade cost matrix by substituting for the predicted values for the country-pairs that are missing.¹⁵

To recover the trade cost estimate \hat{t}_{nl} from the exponentiated term $\hat{t}_{nl}^{1-\varepsilon_A}$, we need to pick a value for the agricultural trade elasticity of substitution ε_A . We follow Tombe (2015) who finds an agricultural trade elasticity of substitution of 4.06. This value sits almost exactly in between the value of 5.4 and 2.82 that Costinot et al. (2016) obtain for respectively the trade elasticity of substitution at the crop level and the elasticity of substitution across different crops, and thus appears to be a reasonable pick. Armed with a bilateral trade cost matrix, we then calculate the agricultural multilateral resistance indices, Ω_l^A and \widetilde{P}_i^A . As before, we rely on the approximations of Baier and Bergstrand (2009), eqs. (32)-(33), but use our trade cost estimates instead of physical distance to substitute for bilateral trade cost terms.

¹⁵Note that to estimate the two-stage regression model expressed by eqs. (34)-(35), we rely on the Stata command “ppml-panel.sg”, developed by Thomas Zylkin, which is the only PPML command in Stata that allows for fast estimation of panel specifications with a large number of fixed effects (see Zylkin et al. (2017) for an application and a technical companion to this application).

Third, an alternative approach to come up with a proxy for remoteness builds on Frankel and Romer (1999). It consists in running a PPML regression of the bilateral agricultural trade equation (12) using the exogenous regressors $DIST_{nl}$, $CONTIG_{nl}$, $COMLANG_{nl}$ and COL_{nl} . It is then possible to use the predicted values of this regression, that is, agricultural trade as predicted by the beforementioned geographical determinants, as a proxy for the *inverse* of remoteness.

Fourth, another approach suggested by Redding and Venables (2004) and Feenstra (2004) does not build a proxy for remoteness but consists in running as a benchmark PPML regressions with investor-country and host-country fixed effects that fully absorb the remoteness terms. However, accounting for remoteness in this way precludes the researcher from obtaining coefficient estimates of the effects of remoteness.

5 Data

5.1 Sources

To measure bilateral investments in land, we use the Land Matrix (as of June 2016), which is an online database that contains extensive information on land deals that have undergone ground verification by NGOs affiliated with the International Land Coalition (Anseeuw et al., 2012) and provides the most complete dataset available.¹⁶ The 2016 version that we use is much improved compared to the 2011 version previously used in Arezki et al. (2015). The data includes information on 2,152 transnational deals negotiated between 2000 and 2016, including the origin country of the investor, and the total area covered by the investment. For the few land deals that involve multiple locations or multiple investor countries, we split them into several subprojects, which leaves us with a sample of 2,601 bilateral land deals. Removing the projects outside agriculture or biofuels, such as those associated with tourism, industry and other renewable energy, we are left with a total of 2,122 bilateral deals. With this information we calculate the cumulative number of projects (and cumulative quantity of land associated with these deals) by investor-country / host-country pair. We sum all deals by country pair over an unrestricted period (2000-2016) and a restricted period (2006-2013). Our main empirical analysis focusses on the restricted period. The reason for this is twofold. First, monitoring of large-scale land deals did not consistently take place before the onset of the global economic crisis of 2008. As retroactive information on deals in early years may be hard to come by, it is therefore possible

¹⁶In this dataset, a deal is defined as an intended, concluded, or failed attempt to acquire land through purchase, lease, or concession that meets the following criteria: It (1) entails a transfer or rights to use, control, or ownership of land through sale, lease, or concession; (2) occurred after the year 2000; (3) covers an area of 200 hectares or more; and (4) implies the potential conversion of land from smallholder production, local community use, or important ecosystem service provision to commercial use. The vast majority of the deals in the database (80%) are listed based on two or more independent sources, e.g., research papers, policy reports and governments sources, and only 6% of the deals are based on media reports only (Nolte et al. 2016).

that the selective inclusion of early years increases the likelihood of mismeasurement. Second, there is reason to believe that the intense scrutiny of these deals by civil society was most intense in the aftermath of the 2008 global crisis so that the inclusion of later years also increases the possibility of mismeasurement. As we discuss in section (6.2), results are nevertheless similar across the restricted and unrestricted periods.

To measure the determinants of land acquisitions at the national level, we compile data from a variety of sources, including: the physical distance between countries and a dummy variable for a former colonial relationship from the GeoDist database (Mayer and Zignago, 2011); a land governance index at the national level which measures the level of recognition and associated tenure security of preexisting rights in rural areas (see Arezki et al., 2015);¹⁷ a dummy variable for regional trade agreements for the period 1995-2004 that is borrowed from Mario Larch’s Regional Trade Agreements Database from Egger and Larch (2008) and which includes all 468 multilateral and bilateral trade agreements as notified to the World Trade Organization for the last 66 years from 1950 to 2015; and a set of variables from the 2009 World Development Indicators database (World Bank) that includes agricultural value added per worker (in constant 2010 US \$), GDP per capita (in current US \$) and population size.

Furthermore, for gross food production (current US \$) and bilateral food trade (current US \$), we use the CEPII TradeProd database (a trade and production dataset originally constructed by de Sousa et al. (2012) that covers 26 sectors using the International Standard Industrial Classification (ISIC) revision 2, and 151 importing and exporting countries). For our purpose of estimating trade costs as explained in the previous section, we sum the trade flow data of the categories food (ISIC 311) and beverages (ISIC 313) for each year in the period 1995-2004.¹⁸ Because proper estimation of (34)-(35) requires the availability of *intranational* trade flow data, we use the CEPII TradeProd dataset which already contains these flows for many countries and impute missing values by using the Anderson and Yotov (2010) method.¹⁹

Finally, we construct measures of available suitable land, using the Global Agro-Ecological Zones (GAEZ) data jointly developed by the Food and Agriculture Organization (FAO) and the International Institute for Applied System Analysis (IIASA) (<http://gaez.fao.org/>). For each of five crops (wheat, maize, oil palm, sugarcane, soybean) we calculate the maximum of the suitability indexes (which has been rescaled and is comprised between 0 and 100) under rainfed and baseline climate (1961-1990) conditions in each 5-arc minute (~10km by 10km) grid cell. We then calculate the total area of all the cells for which this maximum for at least one crop is greater than 70 (which corresponds to high or very high suitability under rainfed conditions),

¹⁷The index is the first component of a principal component analysis on land governance variables contained in the 2009 Institutional Profiles Database (de Crombrughe et al. 2009).

¹⁸We also experimented with using agricultural trade flow data from the UN Comtrade database and found comparable estimates of trade costs.

¹⁹To do this, we first determine for each country the ratio of aggregate intranational trade over total goods expenditures for all sectors for which the data is available. Second, we multiply total expenditures on food and beverages with this ratio and take this as our estimate of intranational trade for those countries for which the data is missing.

after having excluded cultivated land, forests, and protected land. The variable is expressed in 1,000 hectares.²⁰

5.2 Descriptive statistics

Table (4) provides an overview of land that is available and suitable for agricultural production for different regions in the world. It reveals that little land remains in the Middle East and North Africa and South-East Asia, while Sub-Saharan Africa and Latin America still have significant endowments of unexploited land. Forests constitute almost 63 percent of all available, non-protected land. Note that in our regressions, we use a definition of suitable available land that excludes forests and protected areas.

Table (5) presents basic characteristics of deals using data from the Land Matrix. All projects in the Land Matrix database were initiated between 2000 and 2016, although almost 80% of the deals were negotiated between 2006 and 2013. As of June 2016, the Land Matrix has information on 2,152 (trans)national deals, with a cumulative size of 58.4 million hectares, affecting 88 host countries worldwide. This roughly corresponds to an area the size of France or Ukraine, or 13 times the size of the Netherlands.

As one can tell from Table (5), Sub-Saharan Africa (884 deals) and East-Asia (611 deals) have been the most important target regions for investment, followed by Latin-America (368 deals). Few deals have been recorded outside of these regions. In addition, the size distribution of these land deals is very uneven, with only 52 deals covering a staggering 40.9 percent of all land traded. Approximately 66 percent of all deals are smaller than 10,000 hectares (or 100 km²). Deal implementation has been slow, both at the extensive and intensive margin: to date, 51.9 percent of the projects have been implemented, but only 37.6 percent of the 43.54 million hectares of land associated with these implemented deals has been confirmed under production. We note that the fraction of deals not implemented is significantly smaller in Sub-Saharan Africa (37 percent).

According to our calculations, 76.5 percent of all recorded deals have been linked to agricultural and biofuel related projects, see Table (5). This fact fits the narrative that large scale land acquisitions have been partly motivated by private sector expectations of higher food and (bio)fuel prices and government concerns over food independence.

²⁰Using this approach, we estimate that the total quantity of suitable available land (defined here as uncultivated, non-forest, non-protected land suitable for agriculture outside urbanized areas) is about 403 million hectares for the whole world.

Table 2: Descriptive Statistics - Large-Scale Land Acquisitions Dataset

Variables	(1) N	(2) mean	(3) sd	(4) min	(5) max
<i>project_b</i>	79,135	0.0208	0.504	0	54
<i>dist_b.log</i>	50,176	8.809	0.838	-0.00487	9.901
<i>colony_b</i>	50,176	0.00961	0.0975	0	1
<i>suitable_nf_np.log</i>	29,573	5.786	2.463	1.085	10.94
<i>suitability_index1</i>	63,063	3,748	2,577	0	9,194
<i>land_governance</i>	34,098	-0.0196	2.279	-5.048	4.194
<i>agri_value_added_o.log</i>	37,645	8.412	1.556	5.283	11.53
<i>population_o.log</i>	45,264	15.29	2.333	9.191	21.01
<i>gdp_cap.log</i>	42,354	8.457	1.524	5.249	11.53
$\log(P_i^A)$ (BB)	42,348	8.754	0.392	6.735	9.519
$\log(\Omega_i^A)$ (BB)	42,341	8.755	0.392	6.735	9.519
$\log\left(\sum_{l=1, l \neq i}^{l=N} \hat{X}_{il}^A\right)$	63,066	23.59	0.430	22.57	24.44
$\log\left(\sum_{l=1, l \neq n}^{l=N} \hat{X}_{nl}^A\right)$	63,067	23.59	0.430	22.57	24.44
$\log(P_i^A)$ (AY)	42,348	1.985	0.456	0.835	3.098
$\log(\Omega_i^A)$ (AY)	42,341	1.283	0.245	0.595	2.156

Notes: *project_b* measures the number of deals concluded per investor-host country pair between 2006-2013, *dist_b.log* measures the log of physical distance (in km) between investor and host country, *colony_b* represents a dummy variable that equals 1 if investor and host country have been in a colonial relationship and equals 0 when they have not, *suitable_nf_np.log* represents the log of the quantity of non-forest, non-protected land (in ha) that is suitable for agriculture, *land_governance* is an index of tenure security enjoyed by existing land users, *agri_value_added_o.log* is the log of value added per worker in the agricultural sector of the investor country (in US dollars), *population_o.log* is the log of population size, *gdp_cap.log* is the log of GDP per capita (in current US dollar), $\log(P_i^A)$ (BB) is the level of investor-country remoteness a la Baier and Bergstrand (2009), $\log(\Omega_i^A)$ (BB) is the log of the level of host-country remoteness a la Baier and Bergstrand (2009), $\log\left(\sum_{l=1, l \neq i}^{l=N} \hat{X}_{il}^A\right)$ is the log of predicted level of investor-country agricultural exports a la Frankel and Romer (1999), $\log\left(\sum_{n=1, n \neq l}^{n=N} \hat{X}_{nl}^A\right)$ is the log of the predicted level of host-country agricultural exports a la Frankel and Romer (1999), $\log(P_i^A)$ (AY) is the log of the level of investor-country remoteness a la Anderson and Yotov (2016), $\log(\Omega_i^A)$ (AY) is the log of the level of host-country remoteness a la Anderson and Yotov (2016).

Table 3: Descriptive Statistics - Large Scale-Land Acquisitions Dataset

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
project_b	79,135	0.0208	0.504	0	54
dist_b_log	50,176	8.809	0.838	-0.00487	9.901
colony_b	50,176	0.00961	0.0975	0	1
suitable_nf_np_log	29,573	5.786	2.463	1.085	10.94
land_governance	27,333	2.680	1.066	1	4
agri_va_o_log	37,645	8.412	1.556	5.283	11.53
population_o_log	45,264	15.29	2.333	9.191	21.01
gdp_cap_log	42,354	8.457	1.524	5.249	11.53
$\log(P_i^A)$ (BB)	42,348	8.754	0.392	6.735	9.519
$\log(\Omega_i^A)$ (BB)	42,341	8.755	0.392	6.735	9.519
$\log\left(\sum_{l=1, l \neq i}^{l=N} \hat{X}_{il}^A\right)$	63,066	23.59	0.430	22.57	24.44
$\log\left(\sum_{n=1, n \neq l}^{n=N} \hat{X}_{nl}^A\right)$	63,067	23.59	0.430	22.57	24.44
$\log(P_i^A)$ (AY)	42,348	1.983	0.456	0.835	3.098
$\log(\Omega_i^A)$ (AY)	42,341	1.282	0.245	0.594	2.154

Table 4: Availability of Suitable Land for Agriculture (all in Mn ha), by Region.

	All	MENA	SSA	LAC	ECA	EAP	SA	NAM	WNE
Available uncultivated, non-forest, non-protected land	403.2	1.3	208.2	130.9	24.8	22.1	0.4	7.2	8.1
Available uncultivated, forest, non-protected land	669.8	0.1	152.4	242.2	113.2	49.5	7.6	103.2	1.4
Total available uncultivated, non-protected land	1073	1.4	360.6	373.1	138	71.6	8.0	110.4	9.5

Notes: Region abbreviations refer to Middle East and North Africa (MENA), Sub-Saharan Africa (SSA), Latin America and the Caribbean (LAC), East Asia and the Pacific (EAP), Europe and Central Asia (ECA), South Asia (SA) and Western Europe (WNE).

Table 5: Key Characteristics of Land Projects by Region, 2000 - 2016, by Region.

	All	MENA	SSA	LAC	ECA	EAP	SA	NAM
Size of projects								
Host countries (#)	88	6	38	18	9	13	4	N/A
Total area (mn ha)	58.9	0.9	23.7	10.2	10.6	13.4	0.1	N/A
Projects ≥ 1 mn ha (#)	3	0	2	0	0	1	0	N/A
Projects ≥ 250 k ha and < 1 mn ha (#)	49	1	18	9	15	6	0	N/A
Projects ≥ 10 k ha and < 250 k ha (#)	675	4	284	144	54	188	1	N/A
Projects < 10 k ha (#)	1,425	16	580	215	97	416	101	N/A
Total # projects	2,152	21	884	368	166	611	102	N/A
Intended use (percent)								
Agriculture	62.3	96.2	59.8	55.8	89.9	65.9	13.9	N/A
Biofuels	14.2	0.0	19.3	12.3	2.1	13.0	7.0	N/A
Forestry, Industry and Other	14.3	3.8	8.5	17.1	6.3	16.7	73.0	N/A
“Green”	7.3	0.0	10.8	9.6	1.7	3.1	6.1	N/A
Unknown	1.9	0.0	1.6	5.2	0.0	1.3	0.0	N/A

Notes: Region abbreviations refer to Middle East and North Africa (MENA), Sub-Saharan Africa (SSA), Latin America and the Caribbean (LAC), East Asia and the Pacific (EAP), Europe and Central Asia (ECA) and South Asia (SA).

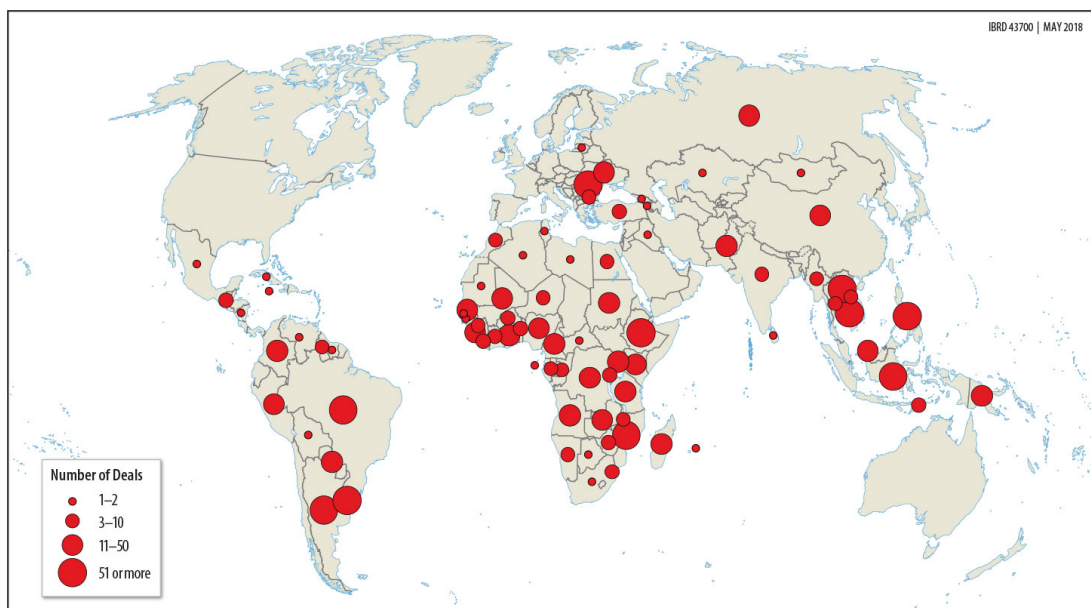


Figure 1: Number of deals by target country

6 Results

6.1 Main results

Table (6) presents the results from our bilateral regressions. With PPML, the coefficient estimates of regressors that are expressed in logs and levels can be interpreted as elasticities and semi-elasticities respectively. In column (1) we present the results of our baseline regression that includes all the regressors as in eq. (31), except for the remoteness terms.

In line with our hypotheses, we find that countries with more available suitable land are a more frequent target in land deals, while the role of land governance is not significant. Let us consider the role of land endowments first. The coefficient on the variable of interest suggest that all things else equal the number of bilateral deals increases by 3.5 percent for a 10 percent increase in the quantity of available land (non-cultivated, non-forested and non-protected) in the host country. This elasticity is relatively stable and varies between 2.2-4.7 percent once we control in various ways for multilateral resistance in columns (2)-(7). As for the role of land governance, while the coefficient enters with a negative sign in most columns as in Arezki et al. (2015), it is not found to be statistically significant in any of the regressions, which contrasts with the previous findings in Arezki et al. (2015) in which Land Matrix data from 2011 was used. In addition to our baseline empirical specification being different to Arezki et al. (2015), the non-significance of the land governance variable could indicate that “land grabs” have become more difficult in the recent period, possibly due to more intense civil society monitoring and increased transparency

from initiatives such as the Land Matrix.²¹

Our results also confirm Arezki et al. (2015)'s finding that investor countries with larger populations acquire more land. An increase in investor-country population size by 10 percent raises the expected number of deals roughly by 7.5 percent. The new finding here is that value added per worker in the agricultural sector of the investor country, a proxy for investor agricultural productivity, is positively associated with the number of land deals as hypothesized in our theoretical framework: If agricultural value added per worker in the origin country increases by 10 percent, the expected number of bilateral deals increases by 4.1-5.2 percent.

Next, we turn to an examination of the role of trade costs and remoteness. We observe that the impact of physical distance—our main proxy for trade costs—is constant across the various distance intervals, with the distance coefficients on the four intervals not being statistically different from one another. This contrasts with the impact of distance on bilateral trade flows, see e.g., Eaton and Kortum (2002) and Anderson and Yotov (2016), who find that long distances impede trade more strongly than short distances. Next, we find that the inclusion of country fixed effects (FE) seems to magnify the role of trade costs in driving land deals; in column 2 (host-country and investor-country FE), column 3 (investor-country FE) and column 4 (host-country FE), we see that the coefficients on distance and the former colony-colonizer dummy variable increase in magnitude. More specifically, we find that in the FE regressions, the absolute magnitude of each of the four distance coefficients is substantially larger than 1, which means that (i) distance is a stronger deterrent of cross-border land acquisitions than it is of trade in goods because, for trade, scholars have often reported absolute coefficient values close to 1, and (ii) that distance between investor country and host country has a much stronger negative effect on the number of land deals than was initially found by Arezki et al. (2015).

So what explains the increased importance of distance in the fixed effects regressions? Our theory of land acquisitions and previous results from the trade gravity literature (see Head and Mayer, 2015) suggest that this is first and foremost because of remoteness (or market potential). Based on the bilateral specifications from Table (1), we hypothesized that host-country remoteness from agricultural consumers, Ω_i^A , and investor-country remoteness from agricultural producers, P_i^A , played a key role in driving land deals. Under the food independence motive, both multilateral resistance terms were expected to increase land deals ($\gamma_8 > 0$, $\gamma_9 > 0$; see eq. (31)), whereas under the platform FDI motive, host-country remoteness from agricultural consumers was hypothesized to play a negative role ($\gamma_8 < 0$) while investor-country remoteness from agricultural producers was hypothesized not to play any role ($\gamma_9 = 0$).

In columns 5, 6 and 7 we incorporate the remoteness terms into our baseline regression using each one of three proxy approaches that we presented in section (4.2). As theory suggests, the first two

²¹Furthermore, we note that our land governance variable is correlated substantially with host-country GDP per capita, so that the inclusion of host-country GDP per capita may mask the effect of land governance on bilateral land investment. As such, the insignificant coefficient on the land governance variable should not be interpreted as conclusive evidence that land governance is irrelevant to investment in the agricultural sector.

Table 6: Bilateral regressions of the number of deals. The role of distance and market potential.

VARIABLES	(1) project_b	(2) project_b	(3) project_b	(4) project_b	(5) project_b	(6) project_b	(7) project_b
dist_1.log	-0.983*** (0.137)	-2.011*** (0.164)	-1.429*** (0.124)	-1.361*** (0.144)	-0.938*** (0.134)	-1.244*** (0.144)	
dist_2.log	-1.092*** (0.115)	-1.995*** (0.136)	-1.499*** (0.103)	-1.396*** (0.116)	-1.081*** (0.112)	-1.303*** (0.118)	
dist_3.log	-1.035*** (0.107)	-1.939*** (0.128)	-1.484*** (0.101)	-1.364*** (0.111)	-1.065*** (0.104)	-1.260*** (0.111)	
dist_4.log	-0.999*** (0.103)	-1.894*** (0.125)	-1.431*** (0.101)	-1.335*** (0.108)	-1.042*** (0.1000)	-1.241*** (0.107)	
colony_b	0.686** (0.268)	0.845*** (0.256)	0.910*** (0.348)	1.063*** (0.278)	0.939*** (0.234)	1.039*** (0.320)	
agri.va.o.log	0.411*** (0.0542)			0.521*** (0.0512)	0.468*** (0.0783)	0.485*** (0.0540)	0.114** (0.0540)
population.o.log	0.748*** (0.0646)			0.740*** (0.0485)	0.751*** (0.0694)	0.674*** (0.0568)	0.403*** (0.0459)
gdp.cap.log	-0.559*** (0.106)		-0.628*** (0.0897)		-0.357*** (0.0899)	-0.668*** (0.0833)	-0.467*** (0.0832)
suitable.nf.np.log	0.346*** (0.0656)		0.472*** (0.0535)		0.221*** (0.0617)	0.270*** (0.0451)	0.287*** (0.0537)
land.governance	0.00483 (0.0734)		-0.0381 (0.0632)		1.77e-05 (0.0642)	-0.0333 (0.0735)	0.0500 (0.0710)
$\log(P_i^A)$ (BB)					-0.158 (0.165)		
$\log(\Omega_i^A)$ (BB)					3.332*** (0.465)		
$\log\left(\sum_{l=1, l \neq i}^{l=N} \hat{X}_{il}^A\right)$						-0.505*** (0.182)	
$\log\left(\sum_{n=1, n \neq l}^{n=N} \hat{X}_{nl}^A\right)$						-2.455*** (0.305)	
agritradecost.log							-2.995*** (0.175)
$\log(P_i^A)$ (AY)							-0.0664 (0.155)
$\log(\Omega_i^A)$ (AY)							3.035*** (0.295)
N	15960	15960	15960	15960	15960	15960	15960
r2_p	0.404	0.745	0.578	0.628	0.458	0.481	0.437
Investor FE		YES	YES				
Host FE		YES		YES			

Notes: the dependent variable project_b measures the number of deals concluded per investor-host country pair between 2006-2013. Estimates are obtained with the PPML estimator. Constant and fixed effects are included but not reported. Robust standard errors, clustered by country-pair, are reported in parentheses. *** p<0.01; ** p<0.05; * p<0.1

approaches in column 5 and 6 bring the coefficients on distance and former colony-colonizer link somewhat closer to those obtained under PPML with fixed effects, thereby providing support to our approach. Furthermore, more support is provided by the fact that the coefficient on our empirical estimate of bilateral agricultural trade costs (=3.0) in column 7 is much larger in absolute magnitude than any of the coefficients on distance (which are between 1.0-1.1) in column 1.

First, in column 5, we present the results of our PPML regression that incorporates our remoteness proxies for remoteness *a la* Baier and Bergstrand (2009). We find that if the host country is more remote from agricultural consumers in the rest of the world—implying a weaker agricultural market potential—the expected number of bilateral deals increases. In quantitative terms the coefficient on host-country remoteness in column 5 implies that if the income-weighted average physical distance to agricultural markets increases by 10 percent, then the expected number of bilateral deals increase by as much as 33 percent. This finding is consistent with the food independence motive for land acquisitions. In contrast, the coefficient on investor-country remoteness (from agricultural producers) is not statistically significant.

Second, in column 6, we report the baseline regression from column 1 to which we added proxies for agricultural trade *a la* Frankel and Romer (1999). Interestingly, we find that *predicted* agricultural exports of both the investor country and the host country are highly significant and enter the regression with a negative sign. Recall that these *predicted* agricultural exports can be interpreted as *inverse* measures of remoteness (see Anderson et al., 2015). Hence, land deals appear to be more frequent for the country pairs for which both the investor country and the host country are predicted to have a weak export potential (or strong “natural” import requirements). In other words, the negative coefficients imply that bilateral investment is positively associated with both investor-country and host-country remoteness. This finding is consistent with the food independence motive for land investments (see Proposition 3). So the results in column 5 and 6 are consistent as far as host-country remoteness is concerned.

In quantitative terms, we notice that the coefficient on predicted agricultural exports for the host country is larger in magnitude than that for the investor country. For every 10 percent decrease in predicted exports (or every 10 percent increase of predicted imports) of the host country, the expected number of bilateral deals increases by 24.6 percent. For the investor country this number is significantly smaller; for every 10 percent increase of predicted imports the expected number of bilateral deals increases by 5.1 percent. The notion that remote investor countries, i.e., as measured by a high level of predicted agricultural imports, have a higher tendency to invest in land fits the food independence motive, and the narrative that a number of wealthy Gulf states, who are dependent on imports for 80 to 90 percent of their food consumption, have invested large sums of money over the last decade to buy cheap farmland in Tanzania, Ethiopia and other African nations in order to feed their domestic populations.

Third, in column 7, we extend our baseline regression from column 1 by adopting the structural

gravity approach pioneered by Anderson and Yotov (2016), for which the remoteness terms are computed using data on GDP and agricultural GDP and our trade-flow inferred estimates of bilateral trade costs (see section 4.2). Remember that under the third approach remoteness is not merely a concept based on physical distance, but it is inferred from the observed global pattern of agricultural trade flows. According to this inferred remoteness measure, countries that do not participate much in global food trade are more remote from other countries in the world.²²

Like the first and second approach to proxy for remoteness, the results of our third approach also point towards an important role for remoteness in driving large-scale land acquisitions. Even after controlling for land governance and agricultural productivity in the host country and agricultural productivity and population size in the investor country, we find that bilateral investments increase by 30.4 percent in response to a 10 percent increase in host-country remoteness. The coefficient on investor-country remoteness is significantly smaller and not statistically significant.

The regression results in columns (5), (6) and (7) all point towards a positive and statistically significant role for host-country remoteness in attracting large-scale land acquisitions. This robust result is consistent with the food independence motive for land acquisitions, with investor countries acquiring more land in countries that did not before participate extensively in global agricultural trade. However, for investor-country remoteness, the results across columns (5), (6) and (7), i.e., the coefficients on P_i^A (the Baier and Bergstrand measure of remoteness), $\log\left(\sum_{l=1, l \neq i}^{l=N} \hat{X}_{il}^A\right)$ (the Frankel and Romer inverse measure of remoteness) and P_i^A (the Anderson and Yotov measure of remoteness), suggest there is either no role or a negative role for investor-country remoteness. Consequently, we cannot unambiguously point towards either the food-independence or the platform-FDI motive as a driver of investment.

The insignificant role of investor-country remoteness in column 6 could stem from negative effects that work against the positive investor-country remoteness effect that is present in the food independence version of our model. For instance, it is possible that a large land investment requires the investor to tap into a variety of scarce production factors (e.g., human capital and financial capital), which are less abundant in remote investor countries. The overall role of investor-country remoteness would then be ambiguous. All in all, the results of Table (6) confirm the positive role of host-country remoteness in attracting large-scale land investments.

To put these results in perspective, we can look in more detail at the role of land endowments and host-country remoteness in explaining land acquisitions in Angola, a country that has recorded an above average number of 19 deals between 2000 and 2016. According to our data, Angola offers 236 percent more available suitable agricultural land (not cultivated, non-protected, and non-forested) than the sample average, that is, 6.71 million hectares versus 2.00 million hectares. Based on the coefficient on available suitable land in our baseline regression of Table (6), we

²²In Appendix (7), we report the estimation results of regression (35) and check that the bilateral trade costs estimates from our first-stage fixed effects regression are partially explained by a standard set of geographical variables that includes physical distance and a dummy variable for common language.

expect that an investor-host country pair (featuring Angola as the host) would have, by virtue of this difference alone, $2.36 \times 0.34 \times 100 = 80$ percent more land acquisitions than an investor-host country pair with the average country in our sample as the host.

Next, consider the impact of remoteness. Because the average country in our dataset and Angola are located at respective GDP-weighted average distance of approximately 6,340 km and 8,554 km from importing countries, Angola is 35 percent more remote than the average country. Applying the coefficient on the Baier and Bergstrand (2009) remoteness proxy from column (5) of Table (6), this would translate into $0.35 \times 3.32 \times 100 = 116$ percent more bilateral deals for any given investor-host country pair in which Angola features as the host country than one in which the average country plays this role.

To shed more light on the relative importance of cross-country variation in land endowments and remoteness in explaining bilateral land deals, we consider a more general thought experiment. For example, while an increase in the host country’s suitable land endowment by one standard deviation from the mean will increase the number of bilateral deals by $0.346 \times (7580/3047) \times 100 = 86.1$ percent, an increase in Baier and Bergstrand’s distance based remoteness of the host-country of one standard deviation from the mean will increase the number of bilateral deals by

$3.33 \times (2361/6790) \times 100 = 115.8$ percent. Using remoteness based on agricultural trade costs, the latter number changes to $3.04 \times (1.04/2.90) \times 100 = 109.0$ percent. In sum, we find that host-country remoteness is a relevant determinant of bilateral investment, and its importance is comparable to or even exceeds that of agricultural land endowments in explaining large scale land acquisitions.

Interestingly, our results in column (5), (6) and (7) of Table (6) are more consistent with the food independence motive for land acquisitions than the platform FDI motive. All in all, the evidence presented in Table (6) sketches a clear picture of (i) large-scale land acquisitions frequently targeting countries that have abundant quantities of available suitable land and that are not yet well integrated into the global food trading system, as observed by their significant degree of remoteness from existing food importers, and of (ii) productive, populous investor countries being the driver behind the rush for land. The positive relationship between agricultural productivity and population size in investor countries on their investment activity suggests that productive investors from existing centers of agricultural production in Western-Europe, North-America and South-East Asia are looking to accommodate the growth in demand for food by investing away from these centers of demand in potentially, new emerging centers of food production.

6.2 Robustness Checks

To verify the robustness of our results, we ran several alternative regressions. First, one may object that the relationship between land acquisitions and target-country remoteness is purely driven by an “Africa effect” as more than 41 percent of all land deals target countries in Sub-Saharan Africa (see Table 5 and Figure (1)). To control for this, we extended the regressions in

Table (6) with an African continent dummy variable. The dummy variable is highly significant and leaves all other results qualitatively unaffected. The inclusion of the dummy variable does reduce the magnitude of the remoteness regression coefficients in some cases, but this is to be expected: the majority of African countries is substantially above-average in terms of remoteness, so among other factors the Africa dummy variable is expected to pick up this effect.²³

Second, we re-estimated the bilateral regressions of Table (6) by constraining the dependent variable, that is, the number of land deals by country pair, to strictly positive values only. Because the majority of the observations in our data consists of zero's, we are left with a total of 409 bilateral flows²⁴

Running our regressions on this restricted sample, Table (A.1) shows that host-country remoteness, as measured by our Baier and Bergstrand proxy (column 5) and Frankel and Romer's predicted food exports (column 6), remains statistically significant at the 1% level, even though some other typical gravity regressors are not significant. While the elasticity on trade-cost-induced remoteness (column 7) still enters the regression with a positive sign, indicating that host-country remoteness is associated with more not less deals, it is no longer statistically significant. We also note that the elasticities on host-country remoteness in columns (5), (6) and (7) are smaller in magnitude than in the unconstrained regression. This makes sense as one could interpret the unconstrained elasticity as the combination of an extensive margin effect (i.e., whether to acquire land) and intensive margin effect (i.e., how many deals). As our results are by and large qualitatively unaffected and because gravity type regressions are typically not run on so few observations, we view the results of this robustness exercise as reaffirming the paper's main message.

Third, we run the same set of bilateral regression with the quantity of land under each contract as dependent variable. Recall that this variable measures the cumulative contract size associated with all deals per investor-country host-country pair. In Table (A.2), we observe that our remoteness proxies in column (5) and (7) but not in column (6) remain statistically significant. The magnitudes of the coefficients on bilateral distance, as well as host-country remoteness, are smaller than in the baseline case where the number of deals per country-pair serves as the dependent variables. This is intuitive: distance and remoteness should have larger effects on the decision whether to acquire land in another country in the first place (extensive margin) than on the decision about the quantity of land to acquire in each deal once investments have started (intensive margin).

Fourth and finally, expanding our horizon to all land deals between 2000 and 2016 instead of 2006 and 2013 leaves our results unaffected too. In Table (A.3) we see that again all three proxies for host-country remoteness are positive and statistically significant, as in Table (5), while the

²³For space considerations we do not include the regressions with the African dummy variable. The full set of results is available upon request from the authors.

²⁴We remind the reader that the PPML estimator is generally well behaved even if the proportion of zeros in the dataset is very large (Silva and Tenreyro, 2011).

magnitude of the coefficients little changes. These results confirm that, even across a larger time span starting well before the 2008 crisis, the observed pattern of large-scale land acquisitions is consistent with the notion that investors aimed to secure land with the desire to reexport the produce to their home countries.

7 Conclusions

There has been an increased interest in the direct acquisition of farmland in low income and middle income countries by international investors in the last decade or so. In this paper we have shed light on this issue by presenting a novel model of international trade and large-scale land acquisitions. Our framework explains how cross-country differences in technology, endowments and land governance drive these acquisitions. Using this framework, we derived a bilateral gravity-type specification of demand for land. In the first model, which features both horizontal and vertical FDI motives, we found that host-country remoteness negatively affects investment as investors prefer to produce close to export markets. In the second model, however, where investor differentiation is absent, and which features only a vertical FDI motive, we found that host country-remoteness implies less investor competition and therefore in fact encourages investment.

Using global data on large-scale land acquisitions, we find that more populous and technologically advanced investors all things equal preferred to target land-abundant countries that until recently had been characterized by a limited participation in the global food trading system. This finding that host-country remoteness plays a positive role in bilateral investment, supported by all of our three distinct approaches to measuring remoteness, is consistent, on the aggregate, with the food independence motive to large-scale land acquisitions, but not with the platform FDI motive. Hence, our evidence suggests that land investment is more appropriately characterized as food-independence driven vertical FDI than profit-seeking platform FDI. At the same time, we also note that the high importance of host-country remoteness as a driver of bilateral investment suggests that demand for land may have been relatively large compared to the supply of land (“investor-abundant, land-scarce”). The reason for this is that all things equal investors prefer to invest in remote countries in order to shield themselves from competing with other foreign investors. This stands in stark contrast with the intuition formulated by Collier and Venables (2012), who argued that host-countries found themselves in a situation with abundant land but investor scarcity. We argued that while land in general may be abundant, what matters is the amount of suitable and available land, which is available in smaller quantities.

While agricultural prices have declined since in recent years, they remain elevated compared to pre-crisis levels, with the future of food prices being dependent on, among other factors, the relative pace of population growth versus technological progress (Baldos and Herel, 2016). Expanding populations and a growing middle class could continue to fuel global interest in large scale land acquisitions in the coming decades. In this context, the debate on the benefits, costs

and risks of large-scale land acquisitions in developing countries in a context where farmland has become truly globalized seems far from settled.

Acknowledgments

We are grateful to Vanessa Diaz Montelongo, Ziqi Li and Amjad Khan for excellent research assistance, to Brian Blankespoor for guidance on the construction and use of geographic data, and to Jeffrey Frankel, Davide Malacrino, Akito Matsumoto, François de Soyres, and the participants of the EAERE 2017 Conference in Athens, the 2018 Land Governance and Poverty Conference held at the World Bank, and the IMF Research Department seminar for very useful comments.

References

- [1] Anderson, J.E., and Yoto V. Yotov (2010), “Specialization: Pro- and Anti-globalizing, 1990-2002,” NBER Working Papers 16301, National Bureau of Economic Research, Inc.
- [2] Anderson, J.E., van Wincoop, E. (2003), “Gravity with Gravitas: a solution to the border puzzle”, *The American Economic Review*, vol. 69, pp.106-116.
- [3] Anderson, J.E., Larch, M. and Yoto V. Yotov (2015), “Growth and Trade with Frictions: A Structural Estimation Framework”, NBER Working Papers 21377, National Bureau of Economic Research, Inc.
- [4] Anderson, J.E., Yoto V. Yotov (2016), “Terms of trade and global efficiency effects of free trade agreements, 1990–2002,” *Journal of International Economics*, Volume 99, pp. 279-298, ISSN 0022-1996, <https://doi.org/10.1016/j.jinteco.2015.10.006>. (<http://www.sciencedirect.com/science/article/pii/S0022199615001531>)
- [5] Antras, P., Staiger, R. (2012), “Offshoring and the Role of Trade Agreements”, *American Economic Review*, vol. 102(7), pp. 3140-83.
- [6] Anseeuw, Ward, Liz Alden Wily, Lorenzo Cotula, and Michael Taylor (2012). Land Rights and the Rush for Land: Findings of the Global Commercial Pressures on Land Research Project. Rome: International Land Coalition.
- [7] Arezki, Rabah, Deininger, Klaus and Harris Selod (2015), “What Drives the Global ”Land Rush”?”, *World Bank Economic Review*, vol. 29 (2): 207-233 first published online October 21, 2013 doi:10.1093/wber/lht034
- [8] Baier, Scott L. & Bergstrand, Jeffrey H., (2009), “Bonus vetus OLS: A simple method for approximating international trade-cost effects using the gravity equation,” *Journal of International Economics*, Elsevier, vol. 77(1), pp.77-85, February.

- [9] Baldos, U. L. C., Hertel, T. W. (2016), “Debunking the ‘new normal’: Why world food prices are expected to resume their long run downward trend”, *Global Food Security*, vol. 8, pp.27-38.
- [10] Bloomberg (2017), “Farming the World: China’s Epic Race to Avoid a Food Crisis”, available at: <https://www.bloomberg.com/graphics/2017-feeding-china/>
- [11] Brainard, S. Lael (1997), ”An empirical assessment of the proximity-concentration trade-off between multinational sales and trade”, *The American Economic Review*, vol. 87 (4), pp. 520-544.
- [12] Collier, Paul and Venables, Anthony J. (2012), “Land Deals in Africa: Pioneers and Speculators”, *Journal of Globalization and Development*, vol. 3(1), article 3, DOI: 10.1515/1948-1837.1228
- [13] Costinot, A., Donaldson, D., & Smith, C. (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*, vol. 124(1), 205-248.
- [14] Cotula, L. (2011) Land deals in Africa: What is in the contracts?, IIED, London, 49 pages.
- [15] de Crombrughe, D, K. Farla, N. Meisel, C. de Neubourg, J. Ould Aoudia, and A. Szirmai. 2009. ”Institutional Profiles Database III. Presentation of the Institutional Profiles Database 2009 (IPD 2009).” Documents de Travail de la DGTPE, No.2009/14. Treasury Directorate General of the French Ministry of the Economy, Industry and Employment, Paris.
- [16] Deininger, K., D. Byerlee, J. Lindsay, A. Norton, H. Selod, and M. Stickler (2011) Rising Global Interest in Farmland, Washington, D.C.: The World Bank.
- [17] Deininger, K., H. Selod, and A. Burns (2012) The Land Governance Assessment Framework, Washington, D.C.: The World Bank.
- [18] Diamond, J. (1997). Guns, germs, and steel: the fates of human societies. NY: WW Norton & Company, 14.
- [19] Dixit, Avinash K., Stiglitz, Joseph E., (1977). ’Monopolistic Competition and Optimum Product Diversity’, *American Economic Review*, vol. 67(3), pp 297-308.
- [20] di Giovanni, J., Levchenko, A., Zhang, J. (2014), “The Global Welfare Impact of China: Trade Integration and Technological Change”, *American Economic Journal: Macroeconomics*, vol. 6:3, pp.153-183.
- [21] Eaton, Jonathan, and Samuel Kortum (2002), “Technology, Geography and Trade”, *Econometrica*, Vol.70(5).

- [22] Economist Intelligence Unit (2016). Global Food Security Index. Accessed date, May 7, 2016. <http://foodsecurityindex.eiu.com/Downloads>
- [23] Egger, Peter H., Larch, Mario (2008), "Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis", *Journal of International Economics*, vol. 76(2), pp. 384-399.
- [24] Ekholm, K., Forslid, R. and J. R. Markusen (2007), "Export-platform foreign direct investment", *Journal of the European Economic Association*, vol. 5 (4), pp. 776-795.
- [25] Feenstra, R.C., (2004), 'Advanced International Trade: Theory and Evidence'. Princeton University Press, Princeton, New Jersey.
- [26] Financial Times (2016), The Great Land Rush, <http://www.ft.com/great-land-rush> , accessed on June 30, 2016.
- [27] Frankel, Jeffrey A., and David Romer (1999), "Does trade cause growth?", *American Economic Review*, pp. 379-399.
- [28] Head, K., Mayer, T. (2015), 'Gravity Equations: Workhorse, Toolkit, and Cookbook'. In: Elhanan Helpman, Kenneth Rogoff and Gita Gopinath, editor, *Handbook of International Economics*, Vol 4, Oxford: BV, pp. 131-195.
- [29] Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple (2004), "Export versus FDI with heterogeneous firms", *American economic review*, vol 94 (1), pp. 300-316.
- [30] Land Matrix Global Observatory. Accessed Date, May 7, 2016. <http://landmatrix.org/en/get-the-detail/>
- [31] Larch, Mario, Wanner, Joschka, Yotov, Yoto V., Zylkin Thomas (2017), "The Currency Union Effect: A PPML Re-assessment with High-dimensional Fixed Effects", Drexel University School of Economics Working Paper 2017-07.
- [32] Markusen, James R., and Anthony J. Venables (1998), "Multinational firms and the new trade theory", *Journal of international economics*, vol. 46 (2), pp.183-203.
- [33] Jann Lay, Kerstin Nolte (2017), "Determinants of foreign land acquisitions in low- and middle-income countries", *Journal of Economic Geography*, lbx011. doi: 10.1093/jeg/lbx011
- [34] Levchenko, Andrei A. (2007), 'Institutional Quality and International Trade', *Review of Economic Studies*, vol. 74 (3), pp. 791-819. MacKenzie, I., Ohndorf, M. (2013), 'Restricted Coasean Bargaining', *Journal of Public Economics*, vol. 97, pp. 296-307.
- [35] Mayer, T., Zignago, S. (2011), "Notes on CEPII's distances measures: the GeoDist Database", CEPII Working Paper 2011-25.

- [36] Matsuyama, K. (2007), “Beyond icebergs: towards a theory of biased globalization”, *The Review of Economic Studies*, vol. 74(1), pp.237-253.
- [37] Nolte, Kerstin; Chamberlain, Wytske; Giger, Markus (2016). International Land Deals for Agriculture. Fresh insights from the Land Matrix: Analytical Report II. Bern, Montpellier, Hamburg, Pretoria: Centre for Development and Environment, University of Bern; Centre de coopération internationale en recherche agronomique pour le développement; German Institute of Global and Area Studies; University of Pretoria; Bern Open Publishing.
- [38] Oxfam (2012) ‘Our Land, Our Lives’ Time out on the Global Land Rush, Oxfam Briefing Note, October, 26 pages. https://www.oxfam.org/sites/www.oxfam.org/files/bn-land-lives-freeze-041012-en_1.pdf
- [39] Ramondo, N., Rodríguez-Clare, A. (2013), “Trade, Multinational Production, and the Gains from Openness”, *Journal of Political Economy*, vol. 121 (2), pp. 273 - 322.
- [40] Silva, J. M.C.S., and S. Tenreyro. (2006), “The Log of Gravity”, *Review of Economics and Statistics*, vol. 88 (4), pp. 641–58.
- [41] Song, Zheng, Storesletten, K., Zilibotti, F. (2011), “Growing Like China”, *American Economic Review*, vol. 101(1), pp. 196–233.
- [42] de Sousa, José, Mayer, Thierry and Zignago, Soledad (2012), “Market access in global and regional trade”, *Regional Science and Urban Economics*, vol. 42 (6), pp.1037-1052.
- [43] Tombe, T. (2015), “The missing food problem: Trade, agriculture, and international productivity differences”, *American Economic Journal: Macroeconomics*, vol. 7(3), pp.226-258.

Appendix A. Trade cost regression

Agricultural bilateral trade costs are obtained by estimating equation (34) using data on agricultural imports between 1995-2004. For those country-pairs for which data is missing, we use the estimates of the pair fixed effects $\hat{\mu}_{nl}$ as a dependent variable and regress them on a set of standard trade cost related covariates and importer and exporter fixed effects. The results from this estimation are reported below in the estimated equation:

$$\exp[\hat{\mu}_{nl}] = \exp\left[-\underset{(0.050)}{0.764} \ln DIST_{nl,1} - \underset{(0.041)}{0.781} \ln DIST_{nl,2} - \underset{(0.040)}{0.784} \ln DIST_{nl,3} - \underset{(0.038)}{0.755} \ln DIST_{nl,4}\right] \quad (\text{A.1})$$

$$- \underset{(0.116)}{0.024} CONTIG_{nl} + \underset{(0.068)}{0.670} COMLANG_{nl} + \underset{(0.133)}{1.134} COL_{nl}] \quad (\text{A.2})$$

The predicted values that we obtain from this regression are then used to complete the full bilateral trade cost matrix. With this trade cost matrix we are then able to compute the multilateral resistance terms. We note that all the regression estimates have the expected sign, reasonable magnitude and are statistically significant; distance impedes trade with elasticities that sit around 0.77, while common language and colonial ties between importer and exporter promote trade with elasticities of respectively 0.67 and 1.13. The effect of a contiguous border between importer and exporter is found to be slightly negative, but the result is statistically insignificant.

Appendix B. Additional regression tables

Table A.1: Bilateral regressions of the number of deals, 2006-2013. Only includes country-pairs with a positive number of deals.

VARIABLES	(1) project_b	(2) project_b	(3) project_b	(4) project_b	(5) project_b	(6) project_b	(7) project_b
dist_1_log	-0.293** (0.122)	-0.411** (0.178)	-0.172 (0.152)	-0.438*** (0.123)	-0.225* (0.119)	-0.333*** (0.118)	
dist_2_log	-0.363*** (0.100)	-0.497*** (0.146)	-0.266** (0.128)	-0.500*** (0.0950)	-0.330*** (0.0969)	-0.398*** (0.0972)	
dist_3_log	-0.329*** (0.0961)	-0.485*** (0.143)	-0.273** (0.123)	-0.454*** (0.0939)	-0.316*** (0.0911)	-0.384*** (0.0935)	
dist_4_log	-0.326*** (0.0921)	-0.491*** (0.137)	-0.278** (0.118)	-0.468*** (0.0904)	-0.325*** (0.0870)	-0.390*** (0.0894)	
colony_b	-0.188 (0.239)	0.159 (0.212)	0.00307 (0.243)	0.279 (0.215)	-0.0387 (0.220)	0.0572 (0.212)	
agri_va_o_log	-0.00766 (0.0474)			0.0172 (0.0425)	0.0318 (0.0601)	0.0232 (0.0471)	-0.198*** (0.0708)
population_o_log	0.189*** (0.0463)			0.198*** (0.0344)	0.224*** (0.0524)	0.188*** (0.0413)	-0.0309 (0.0634)
gdp_cap_log	0.0462 (0.0776)		0.0275 (0.0699)		0.154** (0.0765)	-0.0531 (0.0652)	0.0580 (0.0756)
suitable_nf_np_log	0.0499 (0.0358)		0.155*** (0.0493)		0.0297 (0.0315)	0.0442 (0.0287)	0.0494 (0.0350)
land_governance	-0.0207 (0.0536)		0.0143 (0.0445)		-0.0478 (0.0493)	-0.0299 (0.0510)	0.0121 (0.0537)
$\log(P_i^A)$ (BB)					-0.0121 (0.130)		
$\log(\Omega_i^A)$ (BB)					1.766*** (0.348)		
$\log\left(\sum_{l=1, l \neq i}^{l=N} \hat{X}_{il}^A\right)$						-0.0200 (0.138)	
$\log\left(\sum_{n=1, n \neq i}^{n=N} \hat{X}_{nl}^A\right)$						-1.513*** (0.233)	
agritradecost_log							-0.967*** (0.148)
$\log(P_i^A)$ (AY)							-0.225 (0.226)
$\log(\Omega_i^A)$ (AY)							0.593 (0.442)
N	409	409	409	409	409	409	409
r2_p	0.162	0.483	0.365	0.404	0.226	0.256	0.158
Investor FE		YES	YES				
Host FE		YES		YES			

Notes: the dependent variable project_b measures the number of deals concluded per investor-host country pair between 2006-2013. Estimates are obtained with the PPML estimator. Constant and fixed effects are included but not reported. Robust standard errors, clustered by country-pair, are reported in parentheses. *** p<0.01; ** p<0.05; * p<0.1

Table A.2: Bilateral regressions of cumulative deal size, 2000-2016.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	contract_int_b	contract_int_b	contract_int_b	contract_int_b	contract_int_b	contract_int_b	contract_int_b
dist_1_log	-0.785*** (0.164)	-1.909*** (0.195)	-1.124*** (0.197)	-0.844*** (0.236)	-0.769*** (0.163)	-0.845*** (0.185)	
dist_2_log	-0.966*** (0.136)	-1.919*** (0.170)	-1.272*** (0.165)	-0.984*** (0.193)	-0.952*** (0.136)	-1.007*** (0.148)	
dist_3_log	-0.881*** (0.130)	-1.815*** (0.164)	-1.222*** (0.167)	-0.925*** (0.206)	-0.879*** (0.133)	-0.937*** (0.144)	
dist_4_log	-0.935*** (0.120)	-1.872*** (0.157)	-1.256*** (0.148)	-0.993*** (0.176)	-0.944*** (0.120)	-1.007*** (0.135)	
colony_b	0.541 (0.385)	-0.247 (0.672)	0.290 (0.659)	0.447 (0.420)	0.670* (0.403)	0.853* (0.446)	
agri_va_o_log	0.454*** (0.0594)			0.531*** (0.0608)	0.546*** (0.103)	0.496*** (0.0637)	0.200*** (0.0648)
population_o_log	0.811*** (0.110)			0.791*** (0.0974)	0.854*** (0.0993)	0.723*** (0.118)	0.447*** (0.0790)
gdp_cap_log	-0.226* (0.128)		-0.227 (0.169)		-0.114 (0.121)	-0.248** (0.117)	-0.193* (0.116)
suitable_nf_np_log	0.521*** (0.106)		0.586*** (0.107)		0.461*** (0.107)	0.461*** (0.0989)	0.438*** (0.0744)
land_governance	-0.0320 (0.0898)		-0.0794 (0.0949)		-0.0165 (0.0806)	-0.0286 (0.0887)	0.0574 (0.0929)
$\log(P_i^A)$ (BB)					0.256 (0.291)		
$\log(\Omega_i^A)$ (BB)					0.969* (0.533)		
$\log\left(\sum_{i=1, i \neq l}^N \hat{X}_{il}^A\right)$						-0.744* (0.400)	
$\log\left(\sum_{n=1, n \neq l}^N \hat{X}_{nl}^A\right)$						-0.423 (0.549)	
agritradecost_log							-3.022*** (0.283)
$\log(P_i^A)$							0.124 (0.229)
$\log(\Omega_i^A)$ (AY)							3.753*** (0.543)
N	15960	15960	15960	15960	15960	15960	15960
r2_p	0.400	0.764	0.583	0.602	0.411	0.416	0.445
Investor FE		YES	YES				
Host FE		YES		YES			

Notes: the dependent variable contract_int_b measures the cumulative quantity of land (in 1000 ha) associated with all deals concluded per investor-host country pair between 2000-2016. Estimates are obtained with the PPML estimator. Constant and fixed effects are included but not reported. Robust standard errors, clustered by country-pair, are reported in parentheses. *** p<0.01; ** p<0.05; * p<0.1

Table A.3: Bilateral regressions of the number of deals, 2000-2016.

VARIABLES	(1) project_int_b	(2) project_int_b	(3) project_int_b	(4) project_int_b	(5) project_int_b	(6) project_int_b
dist_1_log	-1.032*** (0.137)	-1.523*** (0.117)	-1.490*** (0.140)	-0.981*** (0.136)	-1.306*** (0.139)	
dist_2_log	-1.137*** (0.116)	-1.576*** (0.0961)	-1.510*** (0.114)	-1.121*** (0.114)	-1.357*** (0.116)	
dist_3_log	-1.079*** (0.107)	-1.555*** (0.0938)	-1.467*** (0.108)	-1.103*** (0.106)	-1.312*** (0.109)	
dist_4_log	-1.041*** (0.103)	-1.497*** (0.0934)	-1.431*** (0.104)	-1.077*** (0.102)	-1.289*** (0.104)	
colony_b	0.702*** (0.263)	0.968*** (0.336)	1.073*** (0.276)	0.950*** (0.230)	1.086*** (0.303)	
agri_va_o_log	0.402*** (0.0548)		0.513*** (0.0484)	0.452*** (0.0789)	0.470*** (0.0536)	0.136** (0.0551)
population_o_log	0.747*** (0.0606)		0.755*** (0.0458)	0.748*** (0.0652)	0.678*** (0.0546)	0.432*** (0.0486)
gdp_cap_log	-0.540*** (0.104)	-0.599*** (0.0942)		-0.347*** (0.0897)	-0.634*** (0.0819)	-0.437*** (0.0750)
suitable_nf_np_log	0.366*** (0.0661)	0.478*** (0.0515)		0.247*** (0.0630)	0.293*** (0.0476)	0.295*** (0.0522)
land_governance	-0.0242 (0.0690)	-0.0727 (0.0630)		-0.0249 (0.0616)	-0.0649 (0.0708)	0.0190 (0.0662)
$\log(P_i^A)$ (BB)				-0.186 (0.164)		
$\log(\Omega_i^A)$ (BB)				3.208*** (0.432)		
$\log\left(\sum_{l=1, l \neq i}^{l=N} \hat{X}_{il}^A\right)$					-0.524*** (0.175)	
$\log\left(\sum_{l=1, l \neq n}^{l=N} \hat{X}_{nl}^A\right)$					-2.284*** (0.288)	
agritradecost_log						-3.119*** (0.172)
$\log(P_i^A)$						0.125 (0.169)
$\log(\Omega_i^A)$ (AY)						2.864*** (0.313)
N	15960	15960	15960	15960	15960	15960
r2_p	0.432	0.600	0.643	0.482	0.503	0.463
Investor FE		YES				
Host FE			YES			

Notes: the dependent variable project_int_b measures the number of deals concluded per investor-host country pair between 2000-2016. Estimates are obtained with the PPML estimator. Constant and fixed effects are included but not reported. Robust standard errors, clustered by country-pair, are reported in parentheses. *** p<0.01; ** p<0.05; * p<0.1

Appendix C. Technical Appendix.

Proposition 1.

Proof. We assume f_l is fixed in each country and the quantity of land T_l^A adjusts to clear the market for land. To prove existence of at least one vector $\langle \mathbf{w}, \mathbf{T}^A \rangle \in \mathbb{R}_{++}^{2n}$, we verify that $G(\mathbf{w}, \mathbf{T}^A)$ has the following five properties:

- (i) $G(\mathbf{w}, \mathbf{T}^A)$ is continuous,
- (ii) $G(\mathbf{w}, \mathbf{T}^A)$ is homogenous of degree zero,
- (iii) $\langle \mathbf{w}, \mathbf{T}^A \rangle \cdot G(\mathbf{w}, \mathbf{T}^A) = 0$ for all $\langle \mathbf{w}, \mathbf{T}^A \rangle \in \mathbb{R}_{++}^{2n}$ (Walras' law),
- (iv) for $k = \max(\max_h L_h, \max_h f_h)$, $G_l^L(\mathbf{w}, \mathbf{T}^A) > -k$ and $G_l^T(\mathbf{w}, \mathbf{T}^A) > -k$ for all $l = 1, \dots, N$ and $\langle \mathbf{w}, \mathbf{T}^A \rangle \in \mathbb{R}_{++}^{2n}$, and
- (v) if $\langle \mathbf{w}, \mathbf{T}^A \rangle \rightarrow \langle \mathbf{w}^0, \mathbf{T}^{A,0} \rangle$, where $\langle \mathbf{w}^0, \mathbf{T}^{A,0} \rangle \neq 0$ and $w_l^0 = 0$ and/or $T_l^{A,0} = 0$ for some l , then $\max_h G_h(\mathbf{w}, \mathbf{T}^A) \rightarrow \infty$.

The existence result then follows from Proposition 17.C.1. of Mas-Collel et al. (1995, p. 585). The latter is an application of Kakutani's Fixed-Point Theorem.

(i) Since all the prices $c_{ii}^A, p_i^A, P_n^A, p_i^M$ and P_n^M are continuous, Ω_i^A and Ω_i^M are continuous too, and so is \mathbf{G} .

(ii) The unit variety prices c_{ii}^A and p_i^A are homogenous of degree α in $\langle \mathbf{w}, \mathbf{T}^A \rangle$, while p_i^M is homogenous of degree one in $\langle \mathbf{w}, \mathbf{T}^A \rangle$. It is then immediate from (21), (22), (14) and (15) that $\frac{1}{T_i^A} (p_i^A \Omega_i^A)^{1-\varepsilon_A} V^A$, $\frac{1}{w_i} (p_i^A \Omega_i^A)^{1-\varepsilon_A} V^A$ and $\frac{1}{w_i} (p_i^M \Omega_i^M)^{1-\varepsilon_M} V^M$ are homogenous of degree zero, and so \mathbf{G} has the same property.

(iii) We note that the requirement of $\langle \mathbf{w}, \mathbf{T}^A \rangle \cdot G(\mathbf{w}, \mathbf{T}^A) =$

$\sum_{l=1}^{l=N} w_l G_l^L(\mathbf{w}, \mathbf{T}^A) + \sum_{l=1}^{l=N} T_l^A G_l^T(\mathbf{w}, \mathbf{T}^A) = 0$ corresponds to balanced trade in each country:

$$\begin{aligned} \sum_{l=1}^{l=N} w_l G_l^L(\mathbf{w}, \mathbf{T}^A) + \sum_{l=1}^{l=N} T_l^A G_l^T(\mathbf{w}, \mathbf{T}^A) &= \sum_{l=1}^{l=N} \left(\sum_{n=1}^{n=N} \left(\frac{p_l^A t_{nl}}{P_n^A} \right)^{1-\varepsilon_A} \gamma I_n \right) \\ &+ \sum_{l=1}^{l=N} \left(\sum_{n=1}^{n=N} \left(\frac{p_l^M t_{nl}}{P_n^M} \right)^{1-\varepsilon_M} (1-\gamma) I_n \right) - \sum_{l=1}^{l=N} I_l \end{aligned}$$

where we substituted for Ω_i^A and Ω_i^M into the excess demand functions. We observe that $\langle \mathbf{w}, \mathbf{T}^A \rangle \cdot G(\mathbf{w}, \mathbf{T}^A) = 0$ follows from the fact that in each country consumers spend all their in-

come on the consumption of agricultural and manufacturing varieties, that is, $\left(\sum_{l=1}^{l=N} \left(\frac{p_l^A t_{nl}}{P_n^A}\right) \gamma I_n\right) + \left(\sum_{l=1}^{l=N} \left(\frac{p_l^M t_{nl}}{P_n^M}\right)^{1-\varepsilon_M} (1-\gamma) I_n\right) - I_n = 0$.

(iv) Inspection of equations (26) and (27) shows that each excess demand function has a lower bound that is equal to the inelastically supplied amount of labour or the exogenous price of land, that is, $G_l^L(\mathbf{w}, \mathbf{T}^A) > -L_l$ and $G_l^T(\mathbf{w}, \mathbf{T}^A) > -f_l$ for all $\langle \mathbf{w}, \mathbf{T}^A \rangle \in \mathbb{R}_{++}^{2n}$. Hence, it holds that $G_l^L(\mathbf{w}, \mathbf{T}^A) > -\max(\max_h L_h, \max_h f_h)$ and $G_l^T(\mathbf{w}, \mathbf{T}^A) > -\max(\max_h L_h, \max_h f_h)$ for all $l = 1, \dots, N$.

(v) Suppose $\langle \mathbf{w}^m, \mathbf{T}^{A,m} \rangle \rightarrow \langle \mathbf{w}^0, \mathbf{T}^{A,0} \rangle$, where $\langle \mathbf{w}^0, \mathbf{T}^{A,0} \rangle \neq 0$ and $w_l^0 = 0$ or $T_l^{A,0} = 0$ for some l . For any $\langle \mathbf{w}, \mathbf{T}^A \rangle \in \mathbb{R}_{++}^{2n}$ we have

$$\begin{aligned} \max_k G_k^L &\geq \max_k \left(\sum_{n=1}^{n=N} \left(\frac{p_k^A t_{nk}}{P_n^A} \right)^{1-\varepsilon_A} \frac{\gamma I_n}{w_k} + \sum_{n=1}^{n=N} \left(\frac{p_k^M t_{nk}}{P_n^M} \right)^{1-\varepsilon_M} \frac{(1-\gamma) I_n}{w_k} \right) - \max_k L_k \\ &\geq \max_{k,n} \left(\left(\frac{p_k^A t_{nk}}{P_n^A} \right)^{1-\varepsilon_A} \frac{\gamma I_n}{w_k} + \left(\frac{p_k^M t_{nk}}{P_n^M} \right)^{1-\varepsilon_M} \frac{(1-\gamma) I_n}{w_k} \right) - \max_k L_k \end{aligned}$$

and

$$\max_k G_k^T \geq \max_k \sum_{n=1}^{n=N} \left(\frac{p_k^A t_{nk}}{P_n^A} \right)^{1-\varepsilon_A} \frac{\gamma I_n}{T_k^A} - \max_k f_k \geq \max_{k,n} \left(\frac{p_k^A t_{nk}}{P_n^A} \right)^{1-\varepsilon_A} \frac{\gamma I_n}{T_k^A} - \max_k f_k$$

Then $\max_h G_h(\mathbf{w}, \mathbf{T}^A) \rightarrow \infty$ will be proved if we can show that either (i) $\max_{k,n} \left(\frac{p_k^M t_{nk}}{P_n^M} \right)^{1-\varepsilon_M} \frac{(1-\gamma) I_n}{w_k} \rightarrow \infty$, (ii) $\max_{k,n} \left(\frac{p_k^A t_{nk}}{P_n^A} \right)^{1-\varepsilon_A} \frac{\gamma I_n}{w_k} \rightarrow \infty$ or (iii) $\max_{k,n} \left(\frac{p_k^A t_{nk}}{P_n^A} \right)^{1-\varepsilon_A} \frac{\gamma I_n}{T_k^A} \rightarrow \infty$ for the factor price sequence $\langle \mathbf{w}^m, \mathbf{T}^{A,m} \rangle$. This is straightforward since (i) and (ii) are verified by (a) $\max_n w_n^m \rightarrow \max_n w_n^0 > 0$ and $\min_k w_k^m \rightarrow \min_k w_k^0 = 0$, while (iii) is verified by (b) $\max_n T_n^{A,m} \rightarrow \max_n T_n^{A,0} > 0$ and $\min_k T_k^{A,m} \rightarrow \min_k T_k^{A,0} = 0$. Since we had assumed that either (a) or (b) holds, this verifies (v). Finally, it is straightforward to show that all five properties hold for the case where countries sell a fixed quantity of land to agricultural producers, $T_l^A = \bar{T}_l^A$, with the solution to the excess demand system given by a vector $\langle \mathbf{w}, \mathbf{f} \rangle \in \mathbb{R}_{++}^{2n}$. This completes the proof. \square

Land Acquisitions under Food Independence Motive

With $\sigma \rightarrow \infty$ and assuming multinational production is the cheapest option for each consumer country to obtain every producer's country unique agricultural product, equation (17) becomes

$p_{il}^A = t_{il}c_{li}$, where we now use subscript i instead of n to refer to the consumer country, since under the food independence motive the investor and consumer country are the same. Using the same steps as before, we can then write $\widetilde{V}_l^A \equiv \sum_{i=1}^N X_{il}^A$ as

$$\widetilde{V}_l^A \equiv \left(w_l^\alpha f_l^{1-\alpha} \widetilde{\Omega}_l^A \right)^{1-\varepsilon_A} V^A$$

with

$$\widetilde{\Omega}_l^A \equiv \left(\sum_{i=1}^N \left(\frac{t_{il}}{\widetilde{Z}_{li}} \right)^{1-\varepsilon_A} \frac{\gamma I_i}{V^A} \right)$$

and where $\widetilde{P}_i^A \equiv \left(\sum_{i=1}^N (t_{il}c_{li}^A)^{1-\varepsilon_A} \right)^{\frac{1}{1-\varepsilon_A}}$. Next, by substituting for the power transform $(w_l^\alpha f_l^{1-\alpha})^{1-\varepsilon_A} = \frac{V^A}{(\widetilde{\Omega}_l^A)^{1-\varepsilon_A} \widetilde{V}_l^A}$ into \widetilde{P}_i^A and into equation (12), and noting that $T_{il}^A = (1-\alpha) \frac{1}{f_l} V_{il}^A$, we can write

$$T_{li}^A = \left(\frac{t_{il}}{\widetilde{Z}_{li} \widetilde{\Omega}_l^A \widetilde{P}_i^A} \right)^{1-\varepsilon_A} \frac{\gamma I_i T_l^A}{V^A} \quad (\text{B.1})$$

and

$$\widetilde{P}_i^A \equiv \left(\sum_{l=1}^N \left(\frac{t_{il}}{\widetilde{Z}_{li}} \right)^{1-\varepsilon_A} \frac{\widetilde{V}_l^A}{V^A} \right)$$

where we made use of the fact that $\widetilde{V}_l^A = \frac{f_l T_l^A}{1-\alpha}$.

The Equivalence of Remoteness under Platform FDI and Food Independence FDI

Next, we show that the remoteness terms associated with the two different versions of our model (Platform FDI and food-independence FDI) can be expressed in terms of the same observables. To this end, since we assume that Z_{ln} can be written as a multiplicative function of investor country agricultural productivity and a bilateral distance term, i.e., $Z_{li} = Z_i F(t_{il})$. We then note that by substituting for $Z_{li} = Z_i F(t_{il})$ into the various terms of eq. (B.1), we can simplify all $\frac{t_{il}}{Z_{li}}$ -terms to $\frac{t_{il}}{F(t_{il})}$, as we can eliminate the Z_i term. Since $F'(t_{il}) < 0$ it follows that the Baier and Bergstrand approximation to the Food Independence host-country remoteness term, $\prod_{i=1, i \neq l}^N \left(\frac{t_{il}}{F(t_{il})} \right)^{\theta_i}$, is simply a more magnified version of the remoteness term under the Platform FDI model (to the extent that distance enters twice in the formula) and that thus relies on the same observables (i.e., distance and population or agricultural production shares).

Furthermore, now suppose $F(t_{il}) = 1/\phi t_{il}$ with $\phi > 0$, then it follows that the host-country remoteness term in the Food Independence version of our model can be written as:

$$\widetilde{\Omega}_l^A \equiv \left(\sum_{i=1}^N \left(\frac{\phi t_{il}^2}{\widetilde{P}_i^A} \right)^{1-\varepsilon_A} \frac{\gamma I_i}{V^A} \right)$$

Using the logic of Baier and Bergstrand's linear approximation method, this leads to the following proxy: $\widetilde{\Omega}_l^A = \prod_{i=1, i \neq l}^N (\phi t_{il}^2)^{\theta_i} = \phi \left(\prod_{i=1, i \neq l}^N (t_{il})^{\theta_i} \right)^2 = \phi (\Omega_l^A)^2$. Taking the natural log we can write this approximation as $\ln \phi + 2 \left(\sum_{i=1, i \neq l}^N \theta_i \ln(t_{il}) \right)$; in the regressions the $\ln \phi$ term will be absorbed by the regression constant and the factor 2 is an irrelevant scaling factor. Hence, under this reasonable functional form the two host-country remoteness terms from the two models are indistinguishable from one another. The same argument holds for the investor-country remoteness terms.