APPENDIX A:

The Role of Geography in Determining the Global Land Use Impacts of Biofuels

Nelson Villoria

nvillori@purdue.edu

Thomas Hertel

hertel@purdue.edu

Department of Agricultural Economics

Purdue University

This version: September 24, 2009

Abstract

Indirect land use changes (iLUC) triggered by the expansion of crop-based biofuels has became a critical issue in the debate over biofuels policy. In their Renewable Fuel Standard Programs both US-EPA and the State of California have required measurement of iLUC. Yet the estimates are highly contentious. One critical point of discussion is where the iLUC will take place and this in turn depends on the assumed relationships between international trade and global output distribution. The geography of iLUC matters due to differences in yields, which determine how much land conversion is necessary, and land cover types, which determine the greenhouse-gas emissions of the converted land. This paper tests which of the two trade assumptions used in the leading economic models employed to study iLUC has better support from the historical data on trade and harvested areas of coarse grains.

1 Introduction

Crop-based biofuels have the potential to sequester carbon in the land employed in their production and thus counteract the greenhouse gases (GHG) they release when burned (McCarl and Schneider, 2000). However, an influential article by Searchinger et al. (2008) which combines results from a physical model of land emissions with the results of a large-scale economic model (FAPRI), reports that when the indirect land use changes (iLUC) triggered by the expansion of biofuels crops are taken into account, current-technology crop-based biofuels are likely to actually increase GHG emissions in the medium as well as in the long run. The idea brought to light by Searchinger et al. (2008) is simple. Following the price incentives of U.S. biofuels policies, the output of currently productive areas in the U.S. is diverted to biofuels, thus leaving a gap in the supply of food. Hence, to fill in this gap, crop production must increase either in the U.S. or elsewhere. To the extent that new lands are brought into production, their clearing will release GHG according to the existing land cover and the area of land converted.

Although the idea is simple, the mechanisms for implementing it in a modeling framework are not. The fact that GHG estimates depend on the existing land cover makes it necessary to project where production will occur, how much land will be needed, and what type of land will be brought into production. In principle, the increases in production can come from a combination of increasing yields, a change in the use of current croplands, or the incorporation of new areas. Keeney and Hertel (2009) use a general equilibrium model with explicit subnational markets for heterogeneous lands to examine the relative importance and uncertainty surrounding each of these processes. As discussed below, their conclusions differ in important ways from those of Searchinger et al. (2008). As noted by Keeney and Hertel (2009), much of this discrepancy is due to different assumptions about how international trade shapes the geographic distribution of production. While Searchinger et al. (2008) assume the existence of an integrated world market (IWM), Keeney and Hertel (2009) use an Armington formulation which assumes that products are differentiated by national origin.

Starting with the recent decision of the California Air Resources Board to factor indirect land use into the state's renewable fuels standard, the conflicting results of different assumptions are no longer restricted to academic debate. More generally, the U.S. Energy Independence Act of

2007 (EISA) has instructed governmental bodies to develop analytical capabilities for assessing the GHG emissions of biofuels, including those stemming from indirect land use effects. In the case of corn-based ethanol, the exercises carried out by the U.S. Energy Protection Agency (EPA) indicate that, although ethanol is less GHG intensive than fossil oil, corn-based ethanol falls short of the necessary reductions in GHG to be classified as a renewable fuel. Based on the uncertainty surrounding iLUC estimates, the recommendations from EPA stemming from its findings have been controversial among ethanol interests (ACE, 2009). More generally, due to the uncertainty in the assumptions of the leading economic models used in the iLUC predictions, some scientists remain skeptical about their accuracy (Wang and Haq, 2008).

This paper seeks to further inform the current debate by examining which of the two competing views of international trade in homogeneous products has better support in the data of cross-country land allocations for the production of coarse grains. The focus on coarse grains is justified by the fact that the products in this category (corn, oats, barley, sorghum) substitute closely for the corn destined for ethanol in production and consumption both in the U.S. and overseas.

An overview of the main players and dynamics in the market for coarse grains is discussed in the next section. The rest of the paper combines analytical and simulation elements to show how the different trade assumptions mentioned above lead to important differences in the conclusions drawn about the location of land conversion, and by extension, about the size of associated GHG emissions. To test which assumption has better support from the historical data, a standard production function is used to derive a land equation that connects domestic returns to land to a unique world price, and thus, is consistent with the idea of an integrated-world market. Then, using an export demand equation as in Armington (1969) an alternative pricing mechanism that takes into account competition between countries in international markets is derived. Our findings on global land use changes under the two different pricing mechanisms highlight the importance of the role of bilateral trade flows in shaping global cropping patterns.

2 Background on the Global Coarse Grains Sector

Corn, sorghum, barley, oats, and rye comprise the trade category known as coarse grains. While maize (corn) is by far the largest traded product in this category (Allen and Lutman, 2009), the composition of the sector varies widely across countries. This can be seen in Figure 7, which shows the relative contribution of each individual crop to the total national area under coarse grains for a sample of countries (the same employed in the econometric exercise below). This area constitutes approximately 75% (average 1975-2002) of the global area harvesting coarse grains. These countries are sorted by their average (1992-2002) hectarage (indicated within parentheses next to each country's name) which reveal that the U.S. has the largest area, followed by India (36.42 million ha), China (27.2), Brazil (12.34), Mexico (9.05), and Canada (7.21). Note that for the U.S. and China, the coarse grains are mostly maize, while in India, sorghum and millet dominate. Combining its dominance in area and higher yields, the U.S. is the largest producer of coarse grains in the world.

During the period 1975-2002, an average of 49% of global exports originated in the U.S. (GTAP database documented in Gehlhar, 2005). Allen and Lutman (2009) indicated that more recently (2003-2008), the U.S. share of world exports averaged 60%. According to these authors:

> While the United States dominates world corn trade, exports account for only a relatively small portion of demand for U.S. corn — about 15 percent. This means that corn prices are largely determined by supply-and-demand relationships in the U.S. market, and the rest of the world must adjust to prevailing U.S. prices. This makes world corn trade and prices dependent on weather in the U.S. Corn Belt. However, Argentina, the second-largest corn exporter in most years, is in the Southern Hemisphere. Farmers there plant their corn after the size of the U.S. crop is known, providing a quick, market-oriented supply response to short U.S. crops.

As discussed in the next section, the "large" country nature of the U.S. is the basic argument justifying the identification strategy of the links between price events in the U.S. and land use decisions elsewhere.

Figure 7 compares trends in area harvested and production in the U.S. and the rest of the world during the period 1975-2002. The figure broadly reveals similar trends where areas have a slow declining trend while production is increasing steadily. This is the result of yield improvements during the last three decades. Figure 7 plots the log of total U.S. corn exports (in bushels) against time. As can be seen, while the volatility of production translates into a high export volatility, the average level of exports seems to have been relatively constant over the last three decades. Indeed, regressing the total U.S. corn exports on a trend variable (period 1975-2007) gives a statistically insignificant trend coefficient of 0.002 (the straight line in the plot).

Figure 7 shows the time distribution (1975-2002) of imports and exports of total coarse grains in the countries in the sample (excluding the U.S.). The dots represent mean flow values, while the segments extend a 95% confidence interval. Imports are plotted to the left of zero and exports to the right. The countries are sorted by their net trade position (mean exports - mean imports). Several features of the international market for coarse grains are apparent. First, with the exception of China (the second largest country in terms of coarse grains area), for the countries that trade, specialization is the norm. According to the figure, France, Argentina, Australia, Canada, China, and to a lesser extent, Hungary, South Africa, and Thailand, are the most important exporters. Japan is by far the largest world importer, followed by Mexico, Germany, Spain and other countries in western Europe. Second, most countries in the sample do not participate significantly in international trade, regardless of their size (see for example, India). Likewise, countries with large croplands such as Mexico, Brazil and China have sizable imports. Overall, the 38 countries in the sample are responsible for 41% of global exports and 64% of global imports (average 1975-2002). When the U.S. is added, almost all the global exports are covered. The U.S. is a marginal importer with average annual imports of coarse grains of less than 2%.

3 Trade Assumptions and Global Land Use Predictions

Under the assumption of an integrated world market, when the relative prices charged by different suppliers change, countries easily adjust their trading patterns by shifting to the lowest cost provider. This arbitrage equalizes global prices. Due to differences in border frictions, area, and yield response, the existence of one global price does not translate into a uniform supply response across countries. However, the assumption of a uniform price signal causes countries with large areas to tend to have the greatest changes in land use.

The IWM assumption is used in the FAPRI model employed by Searchinger et al. (2008), arguably the most influential study of global iLUC associated with biofuels expansion. These authors find that out of the new 10,817 thousand hectares (26,729 '000 acres) needed to accommodate a 55.92 billion liter (14.77 billion gallons) increase in U.S. corn ethanol, 21% are located in the U.S., 26% in Brazil, 10% in China, and 11% in India. Although the FAPRI model does take into account cross-country differences in price transmission, supply elasticity, and yields (Fabiosa et al., 2008), a glance at the results above reveals that the greatest changes are predicted to happen in the largest producing countries of the world,and that the pre-shock trading behavior is largely irrelevant. (While the U.S. and Brazil are active traders, China and India have historically been relatively closed to trade.)

While the IWM assumption seems sensible for homogeneous products such as coarse grains, observed trade patterns tend to be highly persistent over time and space. This persistence of trade patterns is associated with the difficulties of changing supply sources and is behind the idea that products are differentiated by place of origin, as proposed by Armington (1969). Because of this differentiation, changes in quantities (consumed, produced, and traded) of each product following a change in relative prices will depend on the ease with which consumers can substitute among the different origins of coarse grains, as captured by a constant elasticity of substitution (CES).

Econometric estimates from Hertel et al. (2007) based on the cross-country variability of bilateral trade costs give a CES value for coarse grains of 2.6 (std. error 1.1), suggesting that coarse grains behave indeed as differentiated products. This estimate is much lower than those obtained by the same authors for other commodities, such as rice, wheat, and oilseeds. Drawing on this elasticity, a recent analysis of the importance of the different assumptions employed in assessing the iLUC of biofuels by Keeney and Hertel (2009) reveals that out of the land needed to satisfy an increase of 1 billion gallons of ethanol in the U.S., 56% is located in the U.S., 4% in Brazil, 6% in China, and a negligible amount in India.

Due to differences in modeling strategies, the results of Keeney and Hertel (2009) are not directly comparable to those of Searchinger et al. (2008); however, the results contrast enough to hint that ignoring spatial elements of competition leads to very different conclusions about global iLUC changes. To better understand the consequences of using one or the other assumption, consider the following *compensated* CES demand facing country i in each market j (including i itself)[1]:

$$X_{ij} = B_{ij}^{\sigma_o} X_j \left(\frac{P_{ij}}{P_j} \right)^{-\sigma_o}$$

where σ_o is the elasticity of substitution that, depending on the specification, could be different for substitution between domestic products and imports (σ_D) and among import sources (σ_M). X_{ij} are exports from country i to market j, B_{ij} is a constant preference weight, X_j is total demand in the market j, P_{ij} is the price charged by country i's suppliers in market j, and P_j is the average price level in market j.

The analysis of changes in X_{ij} is facilitated by linearizing in percentage changes the CES demand equation above[2]. In the rest of this paper, percentage changes are denoted using the variable names in lower cases (i.e., $x_{ij} = \frac{dX_{ij}}{X_i}$):

$$x_{ij} = x_j - \sigma_o \left(p_{ij} - p_j \right) \tag{1}$$

The first term represents the growth of the overall import market for X_{ij}. The second represents the percentage change in X_{ij}'s share of the market, or the substitution effect. Note that the larger the elasticity of substitution σ_M, the smaller the difference in price changes needed to cause a large swing in x_{ij}. In the limit, when a product is completely homogeneous such that $\sigma_M \to \infty$, the slightest deviation of the changes in country i's prices above the deviations in the changes of the price index of country j, eliminates exporter i from j's market.

With a finite σ_o, there is some stickiness in the trade relationship as the substitution effect depends on the market shares. Appendix A shows that $p_j = \sum_i^n \delta_{ij} p_{ij}$ where $\delta_{ij} = \frac{P_{ij} X_{ij}}{P_j X_j}$ is the market share of X_{ik} in value terms. Thus the effects on X_{ij} in prices of products competing in

[1] The derivation of the CES demand in the Armington context and other properties of the Armington model are in Appendix A.

[2] Ibidem 1

the j^{th} market depend not only on the substitution (and income and price) elasticities of demand (Armington, 1969, p.174-175), but also on market shares. Substituting the change in the price index p_j in (1):

$$x_{ij} = x_j - \sigma_o \left(p_{ij} - \sum_i^n \delta_{ij} p_{ij} \right) \tag{2}$$

shows that the larger the share of X_{ij}, δ_{ij}, in market j, the smaller the percentage change from the substitution associated with a given change in its price and the larger the percentage change in demand for all the other competitors in the market. This is the main idea underlying the econometric exercise below.

In order to further illustrate the differences between assumptions in a comparable framework, a numerical simulation of the effects of a sudden increase of 15% in the industrial domestic demand for coarse grains within the U.S. is carried out using a partial equilibrium closure in the standard GTAP model.[3] The increased demand for coarse grains in the U.S. drives domestic prices up, discouraging exports. The lower presence of the U.S. in international markets offers an opportunity for other countries to increase their market shares. One of the main markets of the U.S. is Japan, which absorbs around 26% of U.S. coarse grains exports. With very low domestic production in Japan, the U.S. supplies around 74% of Japan's market.

Figure 7 plots the percentage changes of exports to Japan (on the vertical axis) from the U.S., Argentina, Brazil, China, and India under different elasticities of substitution (on the horizontal axis). When the standard GTAP parameter of 2.6 is employed, U.S. exports decline by 0.61%. China, Argentina, and Brazil partially fill the gap left by the U.S. through increasing their exports around 1.4%. As the elasticity of substitution increases, the decline in U.S. exports becomes larger. With a very large elasticity of substitution ($\sigma_M = 100$), the price increase causes a reduction of U.S. exports to Japan of around 14% and dramatically increases the exports of China (by 38.45%), Argentina (21.62%), and Brazil (30.77%).

[3]The standard GTAP model (Hertel, 1999) and the GTAP data V.6 (Dimaranan, 2006) are used. To isolate the substitution effects between domestic and imported sources, regional incomes, output in the non-coarse grain sectors out of the U.S., and endowment prices are kept constant. This partial equilibrium closure has the effect of eliminating income effects, decreasing the responsiveness of the coarse-grains sector by keeping resources locked in the other sectors, and discouraging producers from substituting among production factors — and between these factors and intermediate inputs— responding only to the expansionary effect of the increase in demand.

The vertical axis of the lower panel of Figure 7 shows the percentage changes in output associated with the increase in U.S. industrial demand for coarse grains. Using the standard GTAP elasticity ($\sigma_M = 2.6$), the largest percentage change (+11%) is in the U.S., the source of the shock. As the elasticity of substitution increases, the response of the U.S. is mitigated by the fact that it is displaced with increasing strength by its competitors in international markets; in other words, a higher substitutability implies that the U.S. is more sensitive to its relative price disadvantage. Notice that with the lowest elasticity of 2.6, Argentina has the largest output response (+0.32%) after the U.S., a result consistent with the fact that Argentina is an important exporter of coarse grains and competes with the U.S. in many of the same foreign markets. With large elasticities of substitution, the proportional supply response across countries becomes more even. This reasoning explains the pattern of results obtained by Searchinger et al. (2008) who show that the U.S. has only a moderately larger supply response than Brazil, China, or India.

To illustrate what happens to countries with low shares, consider the case of Argentina, which has only 3.8% of the Japanese market share. Figure 7 compares the declines in Argentinean and U.S. exports to Japan under comparable domestic price increases (percentage changes shown in the upper horizontal axis)[4] across several values of the elasticity of substitution (in the lower horizontal axis). Starting from the lowest value of σ_M (2.6), the declines in the exports to Japan are larger for the Argentina (-1.98%) than for the U.S. (-0.61%). As discussed above, this is due to the low market share of Argentina in the Japanese market, which is associated with the low participation of Argentina in the Japanese CES price index. As the goods become more homogeneous (σ_M=100), the same price increase implies a much larger loss of market share for Argentina (-46.75%) than for the U.S. (-14.49%).

The interaction between elasticities and market shares leads to the idea of a geographic pattern of trade in which countries trade more intensively with some partners than with others. Once the shares are established, and depending on the substitutability of a given product, changes in prices are not enough to shift suppliers. As hinted above, in the iLUC-biofuels literature, the difference in trade assumptions leads to two different sets of results that can have important implications

[4]For comparison purposes, the Argentinean market price of coarse grains is shocked by the same percentage changes in the U.S. price obtained under the demand shocks discussed above.

for the life-cycle analysis of biofuels. First, to the extent that the U.S. is the main supplier of the U.S. market, the Armington model tends to predict an iLUC effect in the U.S. larger than what an IWM model would predict. Second, in the IWM framework, prices are equalized so that, even after correcting by differences in price transmission and supply elasticities, the relatively uniform price signal received by different countries causes a relatively uniform proportional change in production. In absolute terms the largest change in land use is assigned to the largest countries, irrespective of their exposure to the source of the shock in international markets. To see which of these two assumptions has better support in the historical data, the next section proposes a conceptual framework to use in statistically investigating the cross country relationship between areas harvested for coarse grains and developments in the U.S. supply sector.

4 Conceptual Framework

This section uses the perspective of a profit maximizing farmer to link land decisions to the price of grain. Next, the Armington assumption is used to develop an alternative price transmission mechanism driven by the intensity of competition in international markets. The section closes by proposing a regression model based on a derived demand for land that nests the two alternative price mechanisms.

4.1 The Derived Demand for Land

Let c index the crop choices available to a farmer endowed with a fixed amount of land $\bar{A}$, facing output prices P_c, crop-specific input-bundle prices W_c, and a technology $f_c(V_c, A_c : \mathbf{Z})$ that combines inputs and land to produce crop output X_c subject to a set of exogenous shifters $\mathbf{Z}$ (e.g., weather). In order to maximize profits, she allocates A_c hectares of her land to each crop c and chooses the crop-specific input-bundles V_c which solve the following problem:

$$\max_{V_c, A_c} \left(\sum_c P_c f_c(V_c, A_c : \mathbf{Z}) - \sum_c V_c W_c : \sum_c A_c = \bar{A} \right) \tag{3}$$

The solution of (3) gives the optimal demands for land A_c^* as a function of the $(1 \times c)$ vectors of output prices $\mathbf{P}_c$, input bundle prices $\mathbf{W}_c$, the total endowment of land, and the vector of exogenous shifters (Shumway, Pope, and Nash, 1984):

$$A_c^* = l(\mathbf{P}_c, \mathbf{W}_c, \mathbf{Z}, \bar{A}) \tag{4}$$

Although this derivation is strictly valid for individual farmers, this study follows the tradition of most empirical work by assuming that the optimal land demands also apply at an aggregated level (e.g., Abler and Pick, 1993; Russo, Green, and Howitt, 2008). The interest is in the cross-country land decisions for a single crop composite of coarse grains production, so below, i will be used to index producing countries and the subscript c will be dropped. As is also traditional in this literature, a lagged area term A_{it-1} is added to the right hand side of the land equation to allow for a lagged adjustment in reaching the optimal allocation A_{it} at each period t. This lagged adjustment can be associated with the costs of crop rotation, clearing new land, and other year-to-year adjustment costs. Although theory is mute about functional form, the relationship between land allocations and its determinants is assumed to be log-linear and can be formalized by the following behavioral equation:

$$\ln(A_{it}) = \ln(P_{it})\beta + \mathbf{Z_{it}}\mathbf{\Gamma} \tag{5}$$

where P_{it} is a measure of the relative return to land capturing the price of coarse grains relative to the prices of competing crops, β is the elasticity of area to changes in relative return to lands, and $\mathbf{\Gamma}$ is the vector of parameters relating the rest of the exogenous variables contained in matrix $\mathbf{Z}$ which is now augmented with the exogenous variables lagged area $\ln(A_{it-1})$ and total area $\ln(\bar{A}_i)$.

National prices P_{it} are not observable[5]. The strategy then is to proxy domestic returns to land in each country i with a world reference price denoted by P_{kt}. The relationship between domestic

[5]The FAO has some data on domestic producer prices; however, the series are available only for limited periods of time.

prices and the world price can be formalized using:

$$P_{it} = RER_{ikt} P_{kt} \tau_{it} \tag{6}$$

where RER_{ikt} is the real exchange rate in local currency units (LCU/$) in country i, and τ_{it} is a term (in power-of-the-tax form) that collects the myriad of transaction costs that foreign products face when entering i's markets — these costs will be discussed in more detail below. Expression (6) is a classical price transmission equation of the type used to study the Law of One Price and has been widely tested in agricultural and commodity markets (e.g. Mundlak and Larson, 1992).

An additional assumption is that farmers have naive expectations such that they expect prices in the current season to be equal to the reference price of the previous year. With this in mind, substituting (6) into (5) yields the land response equation in terms of the one year lagged reference price:

$$\ln(A_{it}) = \ln(P_{k,t-1})\beta + [\mathbf{Z_{it}}, \mathbf{Z_{i,t-1}}]\mathbf{\Gamma} \tag{7}$$

where the matrix of exogenous covariates $\mathbf{Z}$, and its corresponding parameter vector $\mathbf{\Gamma}$, now include the one year lag of exchange rates and trade costs in (6).

Equation (7) is consistent with integrated world markets in which farmers uniformly respond, other things held constant, to a unique world price. This equation is the basis of the econometric estimation below. To accomplish the goal of contrasting the IWM and the Armington view, an alternative theory of international price formation based on product differentiation is next developed.

4.2 International Trade and Price Transmission

At any point in time t, the disposition of *domestically produced* output X_i is either destined for the domestic market or for exports; from now on, these markets are indexed by j. Thus, X_i can be decomposed into $X_{i,j=i}$, domestic consumption, and X_{ij}, exports originating in country i destined to country j. There are n countries with potentially $m = n - 1$ foreign partners. Therefore:

$$X_i = \sum_{j}^{n} X_{ij}, \quad i \in j \tag{8}$$

To guide the estimation below, some structure on how the X_{ij}s behave is needed. We use the Armington assumption. We also impose equal elasticities of substitution between domestic and imported and among goods from different countries. Therefore, in what follows σ_D=σ_M=σ and (1) is rewritten as:

$$x_{ij} = x_j - \sigma(p_{ij} - \sum_{k}^{n} \delta_{kj} p_{kj}) \tag{9}$$

Taking the domestic price i out of the summation yields:

$$x_{ij} = x_j - \sigma \left[p_{ij}(1 - \delta_{ij}) - \sum_{k \neq i}^{m} \delta_{kj} p_{kj} \right] \tag{10}$$

using the fact that $(1 - \delta_{ij}) = \sum_{k \neq i}^{m} \delta_{kj}$, (10) is further simplified as:

$$x_{ij} = x_j - \sigma \left[\sum_{k \neq i}^{m} \delta_{kj} (p_{ij} - p_{kj}) \right]. \tag{11}$$

In other words, the percentage change in i's sales to market j due to the substitution effects is a function of the weighted sum of the changes in the prices charged by i relative to those charged by its competitors (indexed by k), using as weights the market share of producer k in market j. For a given change in relative prices that favors producer k, the larger the participation of k in market j, the more pronounced the decline of X_{ij} — a fact that reinforces the discussion of the previous section.

Summing (11) over all the markets j yields:

$$x_i = \sum_{j}^{n} \theta_{ij} x_j - \sigma \sum_{j}^{n} \theta_{ij} \left[\sum_{k \neq i}^{m} \delta_{kj} (p_{ij} - p_{kj}) \right], \tag{12}$$

where $\theta_{ij} = \frac{P_{ij} X_{ij}}{P_i X_i}$ is the revenue share that producer i obtains from its sales to market j. Expanding terms and rearranging them, Expression (12) can be rewritten as:

$$x_i = \sum_{j}^{n} \theta_{ij} x_j - \sigma \sum_{k}^{m} \sum_{j}^{n} \theta_{ij} \delta_{kj} (p_{ij} - p_{kj}). \tag{13}$$

The first term in the right hand side (RHS) is the expansion effect: a revenue-share weighted sum of the growth in each of the individual markets j catered by producers from country i. The first summation in the second term of the RHS is over the m producers that compete with producer i. The second summation is over destination markets, including producer i's market. The main message from (13) is that the percentage change in country i's total output will depend on the market share of its competitors in the destination market (δ_{kj}), and on the importance that the destination market has for i's total production of coarse grains (θ_{ij}). To focus on the effects of competition between the U.S. and producer i we can rewrite (13) in the following way:

$$x_i = \sum_{j}^{n} \theta_{ij} x_j - \sigma \sum_{j}^{n} \theta_{ij} \delta_{us,j} (p_{ij} - p_{us,j}) - \sigma \sum_{k \neq US}^{m} \sum_{j}^{n} \theta_{ij} \delta_{kj} (p_{ij} - p_{kj}). \tag{14}$$

The first term in the RHS of 14 can be further decomposed into country i's own-expansion as well as expansion in the markets that it sells to. In the second term, the summation over n reflects that both i and the U.S. compete in each other's market. This allows one to define weights that allow one to detect competition in producer i's own market, in the US market, and in other markets indexed by j. The third term captures the effects of competition with other suppliers. This entails rewriting (14) as:

$$\begin{aligned} x_i = {} & \theta_{ii} x_i + \sum_{j \neq i}^{m} \theta_{ij} x_j \\ & - \sigma \left[\sum_{j \neq (i,US)}^{n-2} \theta_{ij} \delta_{us,j} (p_{ij} - p_{us,j}) + \theta_{i,us} \delta_{us,us} (p_{i,us} - p_{us,us}) + \theta_{ii} \delta_{us,i} (p_{ii} - p_{us,i}) \right] \\ & + \epsilon_i \end{aligned} \tag{15}$$

where the first term in the RHS is the own-expansion effect, the second term is the expansion effect from trading partners, the terms within the brackets respectively capture competition between the

U.S. and country i in country i's own market, in the US market, and in third markets, and the last term captures the effects on output associated with competition with other suppliers.

To summarize, the Armington assumption structured in a CES utility framework allows expressing changes in total output as a weighted sum of changes in each of the markets in which a given producer sells. To the extent that, in the short run, yields of coarse grains can be treated as constant, changes in total output map one-to-one with changes in area. This means that area decisions are a function of market growth (the expansion effect) and differences in relative prices (the substitution effect), and as such, global land use decisions are a function of the relative importance of foreign markets *vis à vis* other competitors.

4.3 Regression Model

To test whether the elasticity of harvested area to changes in the reference price varies with the degree of competition in international markets, it is useful to note that both (7) and (15) propose that changes in area harvested are a function of relative returns to land. The important difference is that while (7) postulates a unique channel of transmission between the reference price and the area decisions, (15) proposes that the price transmission will depend on competition intensity, and also on whether competition takes place in the domestic or in third markets.

To facilitate notation, the competition indexes in (15) are denoted by:

$$\begin{aligned} \omega_i &= \theta_{ii}\delta_{us,i}, \\ \omega_{us} &= \theta_{i,us}\delta_{us,us}, \quad and \\ \omega_j &= \sum_{j\neq(i,US)}^{n-2} \theta_{ij}\delta_{us,j}, \end{aligned} \tag{16}$$

which capture competition between country i and the U.S. in country i's own market (ω_i), in the U.S. (ω_{us}), and in third markets (ω_j). To make the integrated world market and Armington equations (7)-(15) compatible, we focus only on changes in the U.S. price, thus imposing the restriction $p_{ij} = 0$.

The empirical model is given by combining the ω indexes with the US price:

$$\begin{aligned} \ln(A_{it}) =& \beta \ln(P_{us,t-1}) \\ &+ \alpha_1[\omega_{i,t} \ln(P_{us,t-1})] + \alpha_2[\omega_{us,t} \ln(P_{us,t-1})] + \alpha_3[\omega_{j,t} \ln(P_{us,t-1})] \\ &+ [\mathbf{Z_{it}}, \mathbf{Z_{i.t-1}}]\mathbf{\Gamma} + \eta_{it}, \end{aligned} \tag{17}$$

where β, α_1, α_2, α_3, and $\mathbf{\Gamma}$ are parameters to be estimated. The error term η_{it} is a disturbance that is further decomposed in a country-specific fixed effect (μ_i) and a remainder disturbance ε_{it} assumed to be identically and independently distributed, centered around zero, uncorrelated over time, and with constant variances. The matrices $\mathbf{Z}_{it}, \mathbf{Z}_{i,t-1}$ contain all the covariates from the structural equations behind the empirical model (i.e., total land endowment $\bar{A}_i$, lagged area $\ln(A_{it-1})$, exogenous shifters such as temperature and precipitation, real exchange rates $\ln(RER_{i,us,t-1})$, trade costs τ_{it}, the expansion terms from (15), $\theta_{iit}X_{it}$ and $\sum_{j\neq i}^{m} \theta_{ijt}X_{jt}$, and the ω indexes).

The null hypothesis is that geography does not matter. In other words, the integrated world market is a good depiction of how farmers in different parts of the world make decisions about areas to plant. The alternative view is that geographic patterns of trade influence planting decisions in individual countries. Formally:

$$\begin{aligned} H_0 &: [\alpha_1, \alpha_2, \alpha_3] = 0 \\ H_A &: [\alpha_1, \alpha_2, \alpha_3] \neq 0 \end{aligned} \tag{18}$$

5 Empirical Implementation

The validity of the hypothesis test proposed above depends on the extent that the covariates in the regression model (17) are controlling for factors that could otherwise obscure the meaning of the parameters of interest. Moreover, meaningful estimates can only be obtained if the data offers enough variability. These two issues, along with the compromises needed to take the regression to the data, are discussed next.

5.1 Covariates

It is important to ensure that the reference price is uncorrelated with the residuals η_{it} so we can be reasonably confident that the parameter estimates are capturing the intended effects. The first potential concern is that the US price is influenced by the production decisions of other countries and thus its presence as an explanatory variable will induce simultaneous equations bias. This potential source of simultaneity is controlled for by using the lagged US price, which cannot arguably be influenced by current production decisions.

A second potential issue is that many of the variables that affect land use decisions also affect the US price for periods that can extend over several years — if these covariates are not controlled for, they will be captured by the disturbances, which will then be correlated with the error term, leading to biased parameter estimates. Fortunately, the theoretical framework offers strong guidance on which covariates should be included. In particular, the inclusion of exogenous factors such as weather and changes in demand ensure that climatic and economic shocks that can affect the U.S. and its competitors alike (e.g. El Niño or the Asian financial crisis in the mid 1990s) are included on the right hand side of (17). Weather is controlled for by including country-specific annual mean temperature and precipitation during a globally defined growing season (See Appendix B). Changes in demand are considered by including the GDP of both exporters and importers. Following the Armington specification in Equation 15, the GDPs are weighted by revenue (own GDP) and bilateral market (partners' GDP) shares.

Another source of bias is the presence of a lagged dependent variable on the right hand side of (17). In (17), the dependent variable A_{it} is a function of $\eta_{it} = \mu_i + \varepsilon_{it}$, where μ_i is a country specific fixed effect. As μ_i does not change over time, it turns out that the lagged area term A_{it-1} is also a function of μ_i, thus inducing correlation between A_{it-1} and η_{it}. Using a fixed effect estimator (as we do) alleviates this problem by explicitly controlling for μ_i. However, the underlying within-transformation is based on the differences in A_{it-1} and ε_{it} from their respective means. In calculating the mean value $\varepsilon_{i.}$, ε_{it-1} is used, and thus, by construction, A_{it-1} is correlated with $\varepsilon_{i.}$ (Nickell, 1981). The bias induced by correlation between A_{it-1} and $\varepsilon_{i.}$ decreases with the time dimension of the dataset. Our panel consists of 38 countries with observations ranging from 7 to 25 years; thus, it is sensible to assume the time dimension of the panel to be fixed. To deal with the bias, several suggestions have been offered (Baltagi, 2008, p.147-148). The method of choice in this paper is a bootstrap correction due to Everaert and Pozzi (2007). This is because, with the small sample in both the time and cross-sectional dimensions, the bootstrap correction satisfactorily approximates analytical corrections and, as shown in the Monte Carlo experiments of Everaert and Pozzi (2007), is more efficient than commonly used GMM estimators.

Another issue that deserves a comment is the data on trade barriers from the price transmission equation (6). Although these are not observable at the required level of detail regarding product, country coverage, and time span needed here, they are already captured by the indexes ω. This can be seen by writing the price that country i charges in market j as the product of the market price in i times a bilateral trade cost τ_{ij} which equals one plus (or minus) and ad-valorem trade cost, so $P_{ij} = P_i\tau_{ij}$. τ_{ij} can be thought of as a comprehensive measure of tariffs, sanitary barriers, shipping costs, etc. This allows rewriting ω_j as:

$$\begin{aligned} \omega_j &= \sum_{j\neq(i,US)}^{n-2} \theta_{ij}\delta_{us,j} \\ &= \sum_{j\neq(i,US)}^{n-2} \frac{P_{ij}X_{ij}}{P_iX_i}\frac{P_{us,j}X_{us,j}}{P_{us}X_{us}} \\ &= \sum_{j\neq(i,US)}^{n-2} \frac{\tau_{ij}X_{ij}}{X_i}\frac{\tau_{us,j}X_{us,j}}{X_{us}} \end{aligned} \tag{19}$$

which shows that the ω_js capture the variability of trade costs across partners, and although the subscript t is omitted to avoid clutter, over time. By the same logic, the indexes ω_i and ω_{us} capture the barriers faced by the U.S. when entering country i and the barriers faced by country i when entering the U.S. The time variation of the indexes is a potential source of endogeneity (current market shares may influence land decisions); thus, in the estimation, one-year lags of 3 year moving averages are used.

A final aspect of the estimable version of (17) is the inclusion of country-specific fixed effects which capture other omitted sources of variation. So, even if they are not constant over time, it can be plausibly assumed that they vary systematically across countries. This is another strategy with a long lineage in the agricultural economics literature used to deal with the lack of availability of data on management practices (Griliches, 1957; Mundlak, 1961), input prices and the prices of competing crops (Lyons and Thompson, 1981). In this case, they are designed to capture cross-country differences in agricultural policies.

Summarizing the discussion thus far, the equation to be estimated takes the following form:

$$\begin{aligned}
\ln(A_{it}) = {} & \mu_i + \gamma_0 \ln(A_{i,t-1}) + \beta \ln(P_{us,t-1}) \\
& + \alpha_1[\omega_{i,t-1}^{3-yr-ma} \ln(P_{us,t-1})] + \alpha_2[\omega_{us,t-1}^{3-yr-ma} \ln(P_{us,t-1})] + \alpha_3[\omega_{j,t-1}^{3-yr-ma} \ln(P_{us,t-1})] \\
& + \gamma_1 \ln(TMP_{it}) + \gamma_2 \ln(PRE_{it}) + \gamma_3 \theta_{ii,t-1} \ln GDP_{i,t-1} + \gamma_4 \sum_{j \neq i}^{m} \theta_{ij,t-1} \ln GDP_{j,t-1} \qquad (20) \\
& + \gamma_5 \ln RER_{i,us,t-1} + \gamma_6 \omega_{i,t-1}^{3-yr-ma} + \gamma_7 \omega_{us,t-1}^{3-yr-ma} + \gamma_8 \omega_{j,t-1}^{3-yr-ma} \\
& + \varepsilon_{it}
\end{aligned}$$

where TMP_{it} and PRE_{it} are temperature and precipitation and GDPs are used as proxies of the expansion in demand X_i and X_j. The rest of the variables, the parameters to be estimated, and the properties of the error term, have already been discussed.

5.2 Data

Table 1 presents summary statistics for all the variables used in the regressions. The dependent variable is the sum of the annual area harvested for barley, oats, maize, and sorghum for 36 countries for the period 1975-2002, sourced from FAO (2009). The representativeness of this sample in terms of area covered and international trade has already been discussed in Section 2.

Key variables for this study are the competition indexes based on the market shares of exporters in different importing countries. From the modeling framework, two types of shares are needed. When a given country is seen as an exporter (and thus indexed by i) the relevant share is θ_{ijt}: the share of country i's exports of coarse grains to country j in country i's total coarse grains output at time t — we will refer to this share as a sales share. When the same country is seen as an importer (and thus indexed by j), the needed share is δ_{ijt}, the share of coarse grains originating in country i in the total consumption of coarse grains in country j at time t — a budget share.

In both cases, domestic sales are considered; thus, the first step is to obtain the shares θ_{iit} and δ_{jjt}. Details about the procedures and sources used to build these shares and the competition indexes discussed below are in Appendix B. Figure (7) plots the budget shares allocated to domestic θ_{ii} (horizontal axis) against the domestic sales share δ_{jj}. In the upper left corner is Argentina, whose domestic needs are completely satisfied by domestic output, but for which, at the same time, the domestic market is just a small fraction of total sales, leaving large amounts of excess supplies. Other countries with large productions and excess supplies are Thailand, France, South Africa, and Hungary. In the opposite corner at the bottom right is Japan. This country has a very large demand for coarse grains relative to production, and its limited production goes entirely to its feed industry. As we move toward the upper right corner, countries become self-sufficient and have fewer excess supplies to export, as in the case in India.

The market shares just described are used to calculate the competition indexes between a given country i with the U.S. in third markets (ω_{jt}), in country i's own market (ω_{it}), and in the US market (ω_{kt}). These variables, bounded by $[0, 1]$ (they are the product of two different shares.), are shown in the three panels of Figure 7, which shows, for each country, a box-plot describing the temporal distribution of the competition index. As will be discussed shortly, this variability is central in the econometric identification of the price effects. Starting with the leftmost panel, from bottom to top, Argentina shows the largest median value (thick black bar inside the box) of the index of competition with the U.S. in third markets. Moreover, going from the minimum value (vertical bar at the end of the left whisker) to the maximum value (vertical bar at the end of the right whisker), it can be seen that the intensity of competition varies through time over a range spanning almost 20 percentage points. The distribution of Argentina's index is skewed toward the right, as evidenced by a mass of observations between the median and the third quartile (outer vertical side of the right box) larger than the mass to the left between the median and the first quartile (outer vertical side of the left box).

The leftmost panel of Figure 7 shows that most countries do not compete significantly with the U.S. in third markets. The largest competitors (judging by the median values) are Argentina, Thailand, Zimbabwe, Australia, France, and Germany. The second panel gives similar information, but this time the interest is in competition with US imports in each country. Ranking the countries using the medians, Japan's coarse grains sector faces the highest competition, and in contrast with the previous panel, the U.S. plays an active role in several of the individual markets. Finally, the third panel (rightmost), shows competition with the U.S. in the US market. Only for Canada, Argentina, and Chile, does this seem to be important, although the value of the indexes is low.

The price transmission parameter β in Equation 20 is identified solely on the basis of temporal variation in the US price. The US price is a quantity weighted index of the price of coarse grains in the most important US markets (see details in Appendix B). The index is normalized to unity in 1975. As can be seen in Table 1, the US price index has a mean value that coincides with its mean of 0.95 (However, this price index does not seem to be normally distributed as evidenced by the fact that 75% of the observations are below 1 — See Q3 in Table 1 — which indicates that the distribution of prices is skewed toward the left.) During the period 1975-2002, the US price index has ranged from 0.63 to 1.46, suggesting that price spikes are more acute than price troughs and also that the US price index varies significantly, providing a good basis for identifying the US price elasticity of harvested area, β.

However, to estimate the parameters α_i $(i = 1..3)$, the competition indexes ω are included by themselves and combined with the US price index. Because the indexes ω vary both over time (as does the US price index) and across countries, it is informative to investigate whether the interaction term retains enough variability to be able to identify the parameters α. To this end, in Table 2, the column "Source of variation" decomposes each interaction into its three components. For instance, the variance of the interaction $\ln(P_{ust}) \times \omega_{it}$ is decomposed in the variances due to year, country, and the interaction *per se*. The only case in which temporal variation has no explanatory power is the interaction that captures the price elasticity of area due to competition within the US market. In all cases, cross-country differences are a significant explanatory variable of the behavior of the ωs, but they are disproportionately so in the case of interactions in third markets. Finally, the "Interaction" component shows the variance not explained by time or cross country variations. This portion ranges from 33% in the case of ω_j to 52% in the case of ω_k to 55% in the case of ω_i. This suggests that the interaction terms have enough variability to properly estimate the parameters α.

The construction of the rest of the variables shown in Table 1, temperature (TMP), precipitation (PRE), and real exchange rates (RER), is detailed in Appendix B. This last variable has a number of missing values due to the incomplete series on nominal exchange rates. Excluding the observations with the missing RER, the dataset employed in the estimation of Equation (20) is a panel with 35 cross-sectional units (countries) and a time dimension that varies from 28 years (31 countries) to 5 years (2 countries)[6]

6 Results

The null hypothesis proposed at the end of the Section "Conceptual Framework" (p. 18), is that global cropping patterns are independent of bilateral trade flows. To test this hypothesis, the restrictions $\alpha_1 = \alpha_2 = \alpha_3 = 0$ are imposed on Eq. 20. Then, both the restricted and unrestricted models are estimated using least squares with country-specific dummy variables (LSDV estimation), and a Lagrange Multiplier (LM) test is used to compare them.

The estimates of the restricted model are shown in the first column of Table 3 under the heading IWM (for integrated world market). For each variable, the three values are stacked. The first is the LSDV parameter estimate. The second is a traditional standard error calculated assuming that the regression residuals have a constant variance (within and across countries) and are serially uncorrelated. To verify whether these two assumptions hold, the bottom panel of Table 3 shows the p-values of a Breusch-Pagan test for heteroskedasticity and a test for serial autocorrelation due to Wooldridge (2001, p.275). In general, homoskedasticity is rejected at the 1% significance level, while the null hypothesis of non-autocorrelation cannot be rejected (with the exceptions of models A and D, which will be discussed shortly). In principle, it would be natural to fix the standard errors to be robust to the presence of heteroskedasticity. However, Stock and Watson (2008) indicate that in panels with fixed T and large N, the conventional white estimator is inconsistent, so they recommend using errors that are robust to both heteroskedasticity and serial autocorrelation. Thus,

[6] The number of included years and countries are as follows (years, countries): (5,2), (8,1), (10,2), (11,1), (12,1), (16,1),(17,1), (18,2), (20,1), (23,2), (28,31).

despite the weak evidence of autocorrelation, the standard errors calculated using the corrections suggested by these authors are shown for each parameter underneath the traditional standard errors.

Interestingly, in many cases, the robust errors tend to overstate the degree of significance of the parameter estimates relative to their traditional counterparts. This is particularly worrying in the case of the unrestricted model A, for which the robust errors of $\hat{\alpha}_2$ and $\hat{\alpha}_3$ are less than half the traditional ones. In the case of $\hat{\alpha}_2$ the robust standard errors make the coefficient appear statistically significant. It is preferable to err on the conservative side of less significance, so the rest of this discussion is based on the traditional standard errors.

Returning to the IWM regression shown in Table 3, the included regressors explain around 66% of the temporal cross-country variation in harvested areas (see adjusted R^2 in the middle panel of Table 3). The coefficient of the lagged area term ($\hat{\gamma}_0$) is large and highly significant, evidencing the persistence of harvested areas over time. The coefficient on temperature ($\hat{\gamma}_1$) is negative and large, but the precision of the estimates is low. The coefficient on precipitation ($\hat{\gamma}_2$) is quite small and insignificant. The terms associated with country size — own-GDP weighted by output consumed at home ($\hat{\gamma}_3$) and the revenue-share-weighted average of the GDPs of the trading partners ($\hat{\gamma}_4$) — are significant and negative. A possible explanation for this is that the income variables are capturing the downward trend of global harvested areas shown in Figure 7 (p. 38) and briefly discussed in the Section "Background on the Global Coarse Grains Sector." In line with expectations, a depreciation of the real exchange rate ($\hat{\gamma}_5$) has a positive effect on output (doubling the RER leads to approx. 1% increase in harvested area). The elasticity of area with respect to the US price, $\hat{\beta}$ is statistically significant and plausibly sized. With a value of 0.04, this elasticity implies that if the US price were to double, areas harvested everywhere else would increase by 4.4%.

The next column shows the unrestricted model labeled "A." The two models are similar, with two important exceptions. First, the coefficient on the real exchange rate lost its significance due to a slight reduction in the size of the estimate. More importantly, the coefficient on the US price, $\hat{\beta}$, is no longer significant. However, the interaction term between degree of competition overseas (ω_j) and the US price is highly significant. This suggests that a price increase in the U.S. will affect areas harvested by the proportion $0.011 \times \omega_{j,t-1}^{3-yr-ma}$, and thus, the size of the actual elasticity is contingent on the importance of output of country i exposed to markets where the U.S. is an important supplier.

To investigate whether model A explains the temporal cross-country variability of areas harvested better than the IWM model, the residuals of the restricted model are regressed on all the covariates of model A. The sample size and the R^2 of this regression are used to compute an LM statistic that in turn is distributed Chi-squared (Wooldridge, 2002, p. 186). The null hypothesis is that the excluded parameters, α_1, α_2, α_3, and the parameters on the ω terms, are uncorrelated with the residuals of the restricted model, for if they are, the restricted parameter estimates are inconsistent. The lowest panel of table 3 (row "LM test") shows the probability (p-values) of rejecting this null hypothesis given that the excluded restrictions are indeed zero. This probability is 3.7% in the comparison of models A and IWM, so the null hypothesis of the LM test is rejected at the 5% level of significance. In other words, the inclusion of the interaction terms represents a model improvement relative to the IWM model. Table 3 also shows that an asymptotically equivalent Likelihood ratio test (row "LR test") supports the finding of the LM test.

Before further exploring what is driving the apparent superiority of model A over the IWM model, recall that the discussion of covariates (see Section 2.5 "Empirical Implementation") indicated that the presence of a lagged dependent variable in the right hand side of the regression causes the so called Nickell (1981) bias. In order to assess how severe this bias is in the current setting, a bootstrap correction due to Everaert and Pozzi (2007) is used (see details in Appendix C). The results for the restricted (IWM) and unrestricted (A) models are shown in Table 4. Three general conclusions can be drawn from these results. First, the LSDV parameter $\hat{\gamma}_0$ is severely biased downward in both models. Second, the rest of the parameter estimates are virtually identical. In particular, the price elasticities of area $\hat{\beta}$ (in both models) and $\hat{\alpha}_3$ (in model A) are for all practical matters equal. Third, the pattern of significance based on the bootstrap sample confirms the pattern of significance observed using traditional standard errors (as opposed to the robust ones), thus supporting the earlier choice of using them as the basis of inference.

The most consequential finding of the bootstrap correction relates to the estimation of long run elasticities, $(1-\gamma_0)^{-1}\beta$ for the IWM model or its counterpart for the Armington model, which will be evidently biased if the LSDV estimates are employed in their calculation. Pesaran and Zhao (1998) point out that simply using the bootstrap corrected parameter from Table 4 will not solve the problem due to the nonlinearities involved in the long run elasticity. They compare four different methods, including a bootstrap correction and conclude that "none of the estimators seem to be effective when the coefficient of the lagged dependent variable is around 0.8." Due to this limitation, this paper will not further pursue the calculation of long run elasticities. Moreover, due to the similarity between the bootstrap-corrected and the LSDV estimates, the rest of the discussion will continue focusing on the latter.

In Table 3, a puzzling feature of model A is that, although the addition of the interaction terms is an improvement over the restricted model,the exclusion restrictions — with parameters $\hat{\gamma}_7$,$\hat{\gamma}_8$, $\hat{\alpha}_1$, and $\hat{\alpha}_2$ — are statistically zero. To investigate whether they are zero because of interactions with the other price terms or simply because they do not add any additional explanatory power to the restricted model, columns B, C, and D report regressions keeping one price interaction at a time. Just as before, the lowest portion of Table 3 shows LM and LR tests for the pairwise comparison of the IWM model against these models. This exercise gives evidence that neither competition in country i's own market (model B including ω_i and its interaction with the US price) nor competition with the U.S. (model C including ω_{us} interacted with the US price) improve the IWM formulation (with a p-value of 0.085, the LM test rejects the IWM model only marginally in the case of model B). In contrast, model D, which retains competition in third markets reiterates the significance of the interaction term and also rejects, albeit in a weaker fashion, the IWM model.

The fact that $\hat{\beta}$ is statistically insignificant in model A prompts the question of whether this is because of multicollinearity with the interaction terms. To investigate this possibility, the lower panel of Table 3 shows the variance inflation factors (VIF) of $\hat{\beta}$ for all the estimated models. The VIF indicates the proportion by which the variance of $\hat{\beta}$ is inflated due to the addition of the exclusion restrictions. In model IWM, the VIF is 1.016, indicating that there are no other variables collinear with the US price term. When all the exclusion restrictions are added, the VIF jumps to 1.275, and as noted before, $\hat{\beta}$ loses its significance. This is a relatively low VIF. If for example the standard error of $\hat{\beta}$ in model A were reduced by 30%, the parameter estimate would still not be significant. In model B, the VIF reduces to 1.169, and $\hat{\beta}$ regains its significance; however, in model D, the VIF is lower (1.09), and $\hat{\beta}$ is insignificant. This seems to reinforce the notion that, in these regressions, higher VIFs are disconnected from the significance of $\hat{\beta}$.

6.1 Implications

Despite the lack of significance of a large number of terms in model A, this model is kept as the preferred regression as it is the best specified model in light of the discussion in the Sections "Conceptual Framework" and "Empirical Implementation". To appreciate the differences in the elasticities implied by each model, Figure 7 plots 95% confidence intervals (CI) of the elasticities from model A using only the statistically significant part, $\hat{\alpha}_3$, evaluated at the mean value of the competition indexes w_j. These are contrasted with the elasticities from the IWM model with point estimate $\hat{\beta}$ given by the black dashed line and a 95% CI represented by the grey band. Note that model A gives just a handful of elasticities that are larger than IWM's $\hat{\beta}$. In particular, Argentina and Thailand appear to be the most responsive countries. Also, there is a group of countries, from China to Australia, for which IWM's $\hat{\beta}$ seems to fairly represent the elasticities that take into account competition with the U.S. The rest of the countries appear to have much lower elasticities under model A than under the IWM model. Indeed, a non-trivial number of countries, from Sri Lanka downward, appear disconnected from price events in the U.S. This is of course a natural consequence of having the area elasticities depend on the intensity of competition with the U.S. overseas and is in line with the simulation results in the Section 2.3 "Trade Assumptions and Global Land Use Predictions," particularly those that tie the intensity of supply response to the degree of competition in third markets.

From the elasticities just discussed, it is clear that the two pricing mechanisms will result in different area responses. To understand the implications of these differences, the IWM model and model A are used to predict the changes in harvested area following a 15% increase in the US price index of coarse grains, while keeping all the variables constant at 1992 levels. Unlike the discussion on the price elasticities, these predictions use all the parameter estimates of model A regardless of their statistical significance. This implies that the predicted standard errors of model A are much larger than those of model IWM, a fact that does not change the main message of the following discussion.

The percentage change in areas in each country is plotted in the left panel of Figure 7. Under the IWM model, the percentage changes are constant because the IWM elasticity $\hat{\beta}$ is assumed to be constant across countries. Under model A, countries with a larger exposure of their exported outputs to competition with the U.S. will have larger area responses. According to the model predictions, and under the conditions that prevailed in 1992, Argentina, Australia, and South Africa will have the largest supply responses. The right panel of Figure 7 shows how the percentage changes translate into area. With differential elasticities, the incorporated areas no longer depend on country size, and thus countries such as Argentina, Australia, and South Africa appear to incorporate more area under model A than under the IWM model. But for China and India, the opposite is true.

The total hectarage predicted by the IWM model is around 840 thousand hectares (2,075 thousand acres) worldwide. Meanwhile, for the same shock, model A predicts approx. 789 thousand hectares (1,950 thousand acres). Using average yield estimates of coarse grains (FAO, 2009), the total tonnage predicted by the IWM model is 2.23 million MT (88.18 million bushels - corn equivalent), while the total tonnage predicted by the Armington model is 2.48 million MT (97.63 million bushels - corn equivalent). In other words, the model with Armington-type elements of competition predicts that roughly the same quantity of coarse grains needs to be produced in less area simply because the countries that compete more heavily with the U.S. have better yields.

6.2 Influential Countries

The results discussed thus far suggest that only a reduced number of countries compete with the U.S. The question of whether it is possible that these countries are driving the results is tackled next. Table 5 shows the parameters of interest obtained — $\hat{\beta}$ from the IMW and A models and α_3 from model — estimating the restricted and unrestricted model, dropping one country at a time. The first column of the table shows the country dropped. Starting with Argentina, when it is dropped from the sample, the elasticity of the IWM model is reduced from 0.044 (in Table 3) to 0.037. The significance of the parameter estimate, however, is mostly unchanged. Indeed, it does not matter which country is dropped from the sample; the elasticity $\hat{\beta}$ of the IWM model remains statistically significant (p-values < 0.05) at roughly 0.04.

Next, the elasticities from the Armington-D model, $\hat{\beta}$ and $\hat{\alpha}_3$ are presented along with their significance level. In general, it does not matter which country is dropped; $\hat{\beta}$ loses its statistical significance when the interaction term involving competition with the U.S. is added to the regression. Likewise, the elasticity estimated using the interaction term, α_3 remains around 0.011 and statistically significant in all the cases (p-values < 0.03) except when Argentina is dropped that the p-value increases to 8%, this is not a surprising result given that Argentina is the country that competes most fiercely with the U.S. in third markets (recall Figure 7 discussed in the Section "Empirical Implementation").

Finally, the p-values obtained from an LM test identical to the one described at the beginning of this section shows that, as before, except for the case where Argentina is dropped, there is around a 5% probability of rejecting the null hypothesis that the exclusion restrictions associated with the competition indexes (α_1- α_3 and the parameters γ attached to the ω terms) are zero given that this was true. When Argentina is dropped, the p-value is 29%, thus, the null cannot be rejected: this is an expected result given the large degree of competition between Argentina and the U.S.

7 Conclusions

As the indirect land use effects of biofuels expansion transcend from the scientific arena to the specifics of biofuels policy design, the need to accurately predict where production will occur in the world, how much land will be needed, and what type of land will be brought into production becomes more important. Sophisticated large-scale economic models have been used for this task; however, the results are surrounded by important uncertainties in some of the key assumptions shaping the global distribution of production. More importantly, these models have very different predictions about where iLUC changes will occur. This paper tackles one of these assumptions, specifically the one governing the international trade exchanges of coarse grains, and investigates which of the two competing views — an integrated world market vs. products differentiated by places of origin — has better grounding in the historical data.

The study uses a standard profit maximization framework to derive a reduced form land equation that controls for a rich set of factors influencing production. This land equation is the base for nesting two pricing mechanisms characterizing the trade assumptions: one-world-price for the integrated world market and prices combined with competition indexes for the differentiated products. The focus is on the changes in the harvested areas of coarse grains triggered by changes in the US price of these products. This is justified because the U.S. is a large player in the market for coarse grains. Moreover, the U.S. is home to ambitious biofuels objectives and it is in this country where the indirect land use impact of biofuels policies have been more closely linked to biofuels regulation. Thus, the competition indexes combine the share of a country's output in a given market with the importance of the U.S. as a supplier of that market. The study employs data for a cross section of countries representative of global production and trade covering the period 1975-2002. This relatively long time period has had important changes in agricultural technology and trade policy. However, controlling for trends (via incomes) and trade costs (via competition indexes) helps to capture these changes in a framework that ensures a clean identification of the area elasticities to changes in the US price.

The main finding is that the integrated world market is rejected in favor of the Armington model. The dominance of the differentiated products model has important consequences for the global distribution of land response following a shock in the US price. Simple predictions using the parameter estimates of the two competing models show that by allowing competition in third markets, the distribution of land conversion is to some extent independent of current land under production. Moreover, the results highlight that by taking competition into account, the total estimates of land converted can be significantly different from those that assume an integrated world market due to differences in productivity across countries. The size and significance of the area elasticities to changes in the US price are robust to the correction of the bias generated by the presence of a lagged dependent variable and to variations in the countries included in the sample.

References

Abler, D.G., and D. Pick. 1993. "NAFTA, Agriculture, and the Environment in Mexico." *American Journal of Agricultural Economics* 75:794–798.

ACE. 2009. "American Coalition for Ethanol statement on House agreement calling for scientific review before EPA imposes international land use effects on biofuels." www.ethanol.org.

Allen, E., and H. Lutman. 2009. "ERS USDA Briefing Room - Corn: Trade." http://www.ers.usda.gov/Briefing/corn/trade.htm, USDA Economic Research Service briefing room.

Armington, P.S. 1969. "A Theory of Demand for Products Distinguished by Place of Production." *Staff Papers - International Monetary Fund* 16:159–178.

Baltagi, B.H. 2008. *Econometric Analysis of Panel Data*, 4th ed. New York: John Wiley.

Dimaranan, B. 2006. *Global Trade, Assistance, and Production: The GTAP 6 Data Base*. West Lafayette, IN: Center for Global Trade Analysis, Purdue University.

Everaert, G., and L. Pozzi. 2007. "Bootstrap-based bias correction for dynamic panels." *Journal of Economic Dynamics and Control* 31:1160–1184.

Fabiosa, J.F., J.C. Beghin, F. Dong, A. Elobeid, S. Tokgoz, and T.E. Yu. 2008. "Land Allocation Effects of the Global Ethanol Surge: Predictions from the International FAPRI Model." Working Paper No. 08005, Iowa State University, Department of Economics, Ames, IA.

FAO. 2009. "FAOSTAT." http://faostat.fao.org/.

Gehlhar, M. 2005. "Chapter 15.B: Reconciling Merchandise Trade Data." In *GTAP 6 Data Base Documentation*. GTAP Center, pp. 15–B–1–11.

Griliches, Z. 1957. "Specification Bias in Estimates of Production Functions." *Journal of Farm Economics* 39:8–20.

Hertel, T., D. Hummels, M. Ivanic, and R. Keeney. 2007. "How confident can we be of CGE-based assessments of Free Trade Agreements?" *Economic Modelling* 24:611–635.

Hertel, T.W. 1999. *Global trade analysis*. Cambridge University Press.

Heston, A., R. Summers, and B. Aten. 2006. "Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania." http://pwt.econ.upenn.edu/php_site/pwt_index.php.

Keeney, R., and T. Hertel. 2009. "The Indirect Land Use Impacts of U.S. Biofuel Policies: The Importance of Acreage, Yield, and Bilateral Trade Responses." *American Journal of Agricultural Economics* Forthcoming.

Lobell, D., and C. Field. 2007. "Global scale climatecrop yield relationships and the impacts of recent warming." *Environmental Research Letters* 2.

Lyons, D., and R. Thompson. 1981. "The Effect of Distortions in Relative Prices on Corn Productivity and Exports: A Cross-Country Study." *Journal of Rural Development* 4:83:102.

McCarl, B.A., and U.A. Schneider. 2000. "U.S. Agriculture's Role in a Greenhouse Gas Emission Mitigation World: An Economic Perspective." *Review of Agricultural Economics* 22:134–159.

Mitchell, T.D., and P.D. Jones. 2005. "An improved method of constructing a database of monthly climate observations and associated high resolution grids." *International Journal of Climatology*, pp. 693–712.

Monfreda, C., N. Ramankutty, and J.A. Foley. 2008. "Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000." *Global Biogeochemical Cycles*, Mar., pp. 1:19.

Mundlak, Y. 1961. "Empirical Production Function Free of Management Bias." *Journal of Farm Economics* 43:44–56.

Mundlak, Y., and D.F. Larson. 1992. "On the Transmission of World Agricultural Prices." *The World Bank Economic Review* 6:399–422.

Nickell, S. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49:1417–1426.

Pesaran, M.H., and Z. Zhao. 1998. "Bias Reduction in Estimating Long-run Relationships from Dynamic Heterogenous Panels." In *Analysis of Panels and Limited Dependent Variables: A Volume in Honour of G.S. Maddala*. Cambridge: Cambridge University Press, pp. 297–321.

Russo, C., R. Green, and R.E. Howitt. 2008. "Estimation of Supply and Demand Elasticities of California Commodities." Working Paper No. 08-001, Giannini Foundation, Department of Agricultural and Resource Economics, University of California, Davis, Davis, CA.

Searchinger, T., R. Heimlich, R.A. Houghton, F. Dong, A. Elobeid, J. Fabiosa, S. Tokgoz, D. Hayes, and T. Yu. 2008. "Use of U.S. Croplands for Biofuels Increases Greenhouse Gases Through Emissions from Land-Use Change." *Science* 319:1238–1240.

Shumway, C.R., R.D. Pope, and E.K. Nash. 1984. "Allocatable Fixed Inputs and Jointness in Agricultural Production: Implications for Economic Modeling." *American Journal of Agricultural Economics* 66:72.

Stock, J.H., and M.W. Watson. 2008. "Heteroskedasticity-Robust Standard Errors for Fixed Effects Panel Data Regression." *Econometrica* 76:155–174.

The World Bank. 2009. "World Development Indicators Online." url-http://go.worldbank.org/6HAYAHG8H0.

USDA. 2009. "ERS/USDA Data - Feed Grains Database." http://www.ers.usda.gov/data/feedgrains/.

Wang, M., and Z. Haq. 2008. "Ethanol's Effects on Greenhouse Gas Emissions (E-letter to Searchinger et al.)." http://www.sciencemag.org/cgi/eletters/319/5867/1238#10977.

Wooldridge, J.M. 2001. *Econometric Analysis of Cross Section and Panel Data*, 1st ed. The MIT Press.

—. 2002. *Introductory Econometrics: A Modern Approach*, 2nd ed. South-Western College Pub.

Table 1: Descriptive Statistics

	mean	Std.Dev.	min	Q1	median	Q3	max
Area (Mill Ha)	4.10	7.41	0.03	0.50	1.49	4.30	43.88
Temperature (C)	19.22	5.47	5.70	16.39	20.00	23.17	28.05
Precipitation (ml)	93.21	78.84	0.00	29.56	71.42	141.37	420.72
GDP (Billion 2000 US$)	274.06	640.44	2.12	27.89	87.86	256.08	4680.00
US Price (Index)	0.95	0.16	0.63	0.84	0.95	1.00	1.46
ω_i	0.05	0.12	0.00	0.00	0.00	0.03	0.86
ω_{us}	0.00	0.01	0.00	0.00	0.00	0.00	0.12
ω_j	0.01	0.04	0.00	0.00	0.00	0.00	0.41

Note: Q1 and Q3 are the first and third quartile of the distribution of each variable. The terms ω are competition indexes between any country i in the sample and the U.S. in country i's own market, the US market, or in third markets (j).

Table 2: ANOVA of the Interaction Between the US Price and the Competitiveness Indexes

Source of variation	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
Year ($P_{us,t} \times \omega_{it}$)	27.00	2961.92	109.70	2.50	0.00
Country ($P_{us,t} \times \omega_{it}$)	35.00	35895.29	1025.58	23.39	0.00
Interaction ($P_{us,t} \times \omega_{it}$)	945.00	41440.70	43.85		
Year ($P_{us,t} \times \omega_{kt}$)	27.00	5.47	0.20	1.27	0.16
Country ($P_{us,t} \times \omega_{kt}$)	35.00	118.74	3.39	21.29	0.00
Interaction ($P_{us,t} \times \omega_{kt}$)	945.00	150.59	0.16		
Year ($P_{us,t} \times \omega_{jt}$)	27.00	364.08	13.48	2.55	0.00
Country ($P_{us,t} \times \omega_{jt}$)	35.00	10702.92	305.80	57.92	0.00
Interaction ($P_{us,t} \times \omega_{jt}$))	945.00	4989.23	5.28		

Note: The ANOVA decomposes the variability of each interaction term into its time (Year) and cross-section (Country) dimensions. The residual (Interaction) is the remainder not explained by the two included dimensions. The ANOVA table shows for each of these components the degrees of freedom (Df), the explained sum of squared residuals (Sum.Sq), the mean of the Sum.Sq, and an F-test that indicates whether the mean of the different groups in Year and Country are statistically different.

Table 3: Regression Results, Diagnosis, and Comparisons

	IWM	A	B	C	D
$\ln(A_{i,t-1}$, $\hat{\gamma}_0$	0.787	0.783	0.783	0.789	0.784
	(0.021)***	(0.021)***	(0.021)***	(0.021)***	(0.021)***
	(0.040)***	(0.040)***	(0.039)***	(0.040)***	(0.041)***
$\ln(TMP_{it})$, $\hat{\gamma}_1$	-0.111	-0.091	-0.087	-0.111	-0.116
	(0.083)	(0.083)	(0.083)	(0.083)	(0.083)
	(0.065)*	(0.066)	(0.065)	(0.065)*	(0.066)*
$\ln(PRE_{it})$, $\hat{\gamma}_2$	0.015	0.016	0.015	0.015	0.016
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
	(0.011)	(0.010)	(0.011)	(0.011)	(0.010)
$\theta_{ii,t-1} \ln GDP_{i,t-1}$, $\hat{\gamma}_3$	-0.025	-0.025	-0.026	-0.025	-0.024
	(0.010)**	(0.012)**	(0.010)**	(0.010)**	(0.012)**
	(0.009)***	(0.013)*	(0.009)***	(0.009)***	(0.013)*
$\sum_{j\neq i}^{m} \theta_{ij,t-1} \ln GDP_{j,t-1}$, $\hat{\gamma}_4$	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)**	(0.001)***	(0.001)***	(0.001)**	(0.001)**
	(0.001)*	(0.001)**	(0.001)**	(0.001)*	(0.001)*
$\ln RER_{i,us,t-1}$, $\hat{\gamma}_5$	0.009	0.008	0.007	0.009	0.010
	(0.005)*	(0.005)	(0.005)	(0.005)*	(0.005)*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
$\ln(P_{us,t-1})$, $\hat{\beta}$	0.044	0.021	0.048	0.035	0.026
	(0.020)**	(0.026)	(0.024)**	(0.021)*	(0.022)
	(0.024)*	(0.021)	(0.023)**	(0.024)	(0.024)
$\omega_{i,t-1}^{3-yr-ma}$, $\hat{\gamma}_6$		0.001	0.001		
		(0.001)**	(0.001)*		
		(0.000)***	(0.000)***		
$\omega_{us,t-1}^{3-yr-ma}$, $\hat{\gamma}_7$		0.006		0.007	
		(0.010)		(0.010)	
		(0.006)		(0.006)	
$\omega_{j,t-1}^{3-yr-ma}$, $\hat{\gamma}_8$		0.001			0.001
		(0.002)			(0.002)
		(0.001)			(0.001)
$\omega_{i,t-1}^{3-yr-ma} \ln(P_{us,t-1})$, $\hat{\alpha}_1$		-0.001	-0.001		
		(0.002)	(0.002)		
		(0.002)	(0.002)		
$\omega_{us,t-1}^{3-yr-ma} \ln(P_{us,t-1})$, α_2		0.050		0.057	
		(0.035)		(0.035)	
		(0.016)***		(0.012)***	
$\omega_{j,t-1}^{3-yr-ma} \ln(P_{us,t-1})$, $\hat{\alpha}_3$		0.011			0.012
		(0.005)**			(0.005)**
		(0.002)***			(0.002)***
N	894.000	894.000	894.000	894.000	894.000
SSR	8.501	8.373	8.454	8.468	8.444
Adj R^2	0.656	0.658	0.657	0.656	0.657
Log likelihood	812.488	819.250	814.958	814.240	815.482
VIF of $\hat{\beta}$	1.016	1.275	1.169	1.045	1.090
Heteroskedasticity (p-values)	0.000	0.002	0.000	0.000	0.000
Autocorrelation (p-values)	0.129	0.084	0.131	0.104	0.097
LM test (p-values)	0.000	0.037	0.085	0.174	0.051
LR test (p-values)	0.000	0.035	0.085	0.173	0.050

Note: The dependent variable is the log of annual area harvested in each country. The number of years included for each country vary between 7 up to a maximum of 28 (1975-2002). Standard errors, traditional and robust to heteroskedasticity and serial autocorrelation of general form (Stock and Watson, 2008) are underneath each parameter estimate. The middle panel presents some measures of goodness of fit. The lowest panel is a diagnosis of the regression. First, the variance inflation factor (VIF) of the US price coefficient is presented. Then, the last four rows show the probability (p-values) of not rejecting the following null hypotheses: 1) the LSDV residuals are homoskedastic (Breusch-Pagan test for heteroskedasticity); 2) the LSDV residuals are serially uncorrelated (test for short panels by Wooldridge (2001, p.275)); and 3) the omitted terms in the "Base" model but present in models A to D add no explanatory power to the former equation. These tests are a Lagrange Multiplier (LM) distributed Chi-square..
*** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.10$

Table 4: Bootstrap Corrected Parameters

Variable	IWM		A	
$\ln(A_{i,t-1})$, $\hat{\gamma}_0$	0.864	***	0.859	***
$\ln(TMP_{it})$, $\hat{\gamma}_1$	-0.101		-0.072	
$\ln(PRE_{it})$, $\hat{\gamma}_2$	0.016	*	0.017	*
$\theta_{ii,t-1} \ln GDP_{i,t-1}$, $\hat{\gamma}_3$	-0.018	*	-0.018	
$\sum_{j \neq i}^{m} \theta_{ij,t-1} \ln GDP_{j,t-1}$, $\hat{\gamma}_4$	-0.001	**	-0.001	**
$\ln RER_{i,us,t-1}$, $\hat{\gamma}_5$	0.006		0.005	
$\ln(P_{us,t-1})$, $\hat{\beta}$	0.043	**	0.019	
$\omega_{i,t-1}^{3-yr-ma}$, $\hat{\gamma}_6$			0.001	*
$\omega_{us,t-1}^{3-yr-ma}$, $\hat{\gamma}_7$			0.008	
$\omega_{j,t-1}^{3-yr-ma}$, $\hat{\gamma}_8$			0.001	
$\omega_{i,t-1}^{3-yr-ma} \ln(P_{us,t-1})$, $\hat{\alpha}_1$			-0.000	
$\omega_{us,t-1}^{3-yr-ma} \ln(P_{us,t-1})$, $\hat{\alpha}_2$			0.052	
$\omega_{j,t-1}^{3-yr-ma} \ln(P_{us,t-1})$, $\hat{\alpha}_3$			0.011	**

Note: Parameter estimates corrected using bootstrap algorithm by Everaert and Pozzi (2007). Significance levels based on confidence intervals using procedure described in Appendix C. *** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.10$

Table 5: Parameter Estimates and Hypotheses Tests Dropping one Country at a Time

Country	$\hat{\beta}$ IWM	p-value	$\hat{\beta}$ Arm-D	p-value.1	α_3 Arm-D	p-value.2	LM p-value
Argentina	0.04	0.04	0.03	0.12	0.01	0.07	0.29
Australia	0.04	0.02	0.03	0.11	0.01	0.01	0.03
Austria	0.04	0.02	0.02	0.14	0.01	0.01	0.05
Bulgaria	0.05	0.01	0.03	0.11	0.01	0.01	0.06
Brazil	0.05	0.01	0.03	0.09	0.01	0.01	0.06
Canada	0.04	0.03	0.02	0.16	0.01	0.01	0.05
Switzerland	0.04	0.02	0.02	0.14	0.01	0.01	0.05
Chile	0.04	0.04	0.02	0.23	0.01	0.01	0.04
China	0.04	0.02	0.03	0.13	0.01	0.01	0.05
Colombia	0.05	0.01	0.03	0.06	0.01	0.01	0.06
Spain	0.05	0.02	0.03	0.12	0.01	0.01	0.05
France	0.05	0.02	0.03	0.12	0.01	0.01	0.04
Greece	0.04	0.02	0.03	0.13	0.01	0.01	0.05
Hungary	0.05	0.01	0.03	0.10	0.01	0.01	0.06
Indonesia	0.04	0.02	0.02	0.14	0.01	0.01	0.05
India	0.04	0.02	0.03	0.13	0.01	0.01	0.05
Italia	0.05	0.02	0.03	0.12	0.01	0.01	0.05
Japan	0.05	0.01	0.03	0.09	0.01	0.01	0.04
Sri Lanka	0.05	0.01	0.03	0.09	0.01	0.01	0.05
Morocco	0.04	0.03	0.02	0.18	0.01	0.01	0.04
Madagascar	0.05	0.01	0.03	0.12	0.01	0.01	0.05
Mexico	0.04	0.02	0.02	0.16	0.01	0.01	0.05
New Zealand	0.04	0.03	0.02	0.16	0.01	0.02	0.08
Peru	0.05	0.01	0.03	0.09	0.01	0.01	0.06
Philippines	0.05	0.02	0.03	0.12	0.01	0.01	0.05
Poland	0.04	0.02	0.03	0.13	0.01	0.01	0.05
Portugal	0.04	0.02	0.03	0.13	0.01	0.01	0.04
Thailand	0.04	0.04	0.02	0.15	0.01	0.02	0.11
Turkey	0.04	0.02	0.02	0.14	0.01	0.01	0.05
Uganda	0.05	0.01	0.03	0.11	0.01	0.01	0.06
Uruguay	0.04	0.02	0.02	0.14	0.01	0.01	0.04
Venezuela	0.06	0.00	0.04	0.03	0.01	0.01	0.07
Vietnam	0.05	0.01	0.03	0.11	0.01	0.01	0.05
South Africa	0.05	0.01	0.03	0.11	0.01	0.01	0.06
Zambia	0.04	0.02	0.02	0.17	0.01	0.01	0.04
Zimbabwe	0.04	0.02	0.02	0.13	0.01	0.01	0.04

Note:The parameter estimates shown in the columns, as well as the LM comparing the restricted (IWM) and unrestricted (A) models are estimates n times, where n is the number of countries. Each time, one country is excluded. The exclude country is shown in the rows.

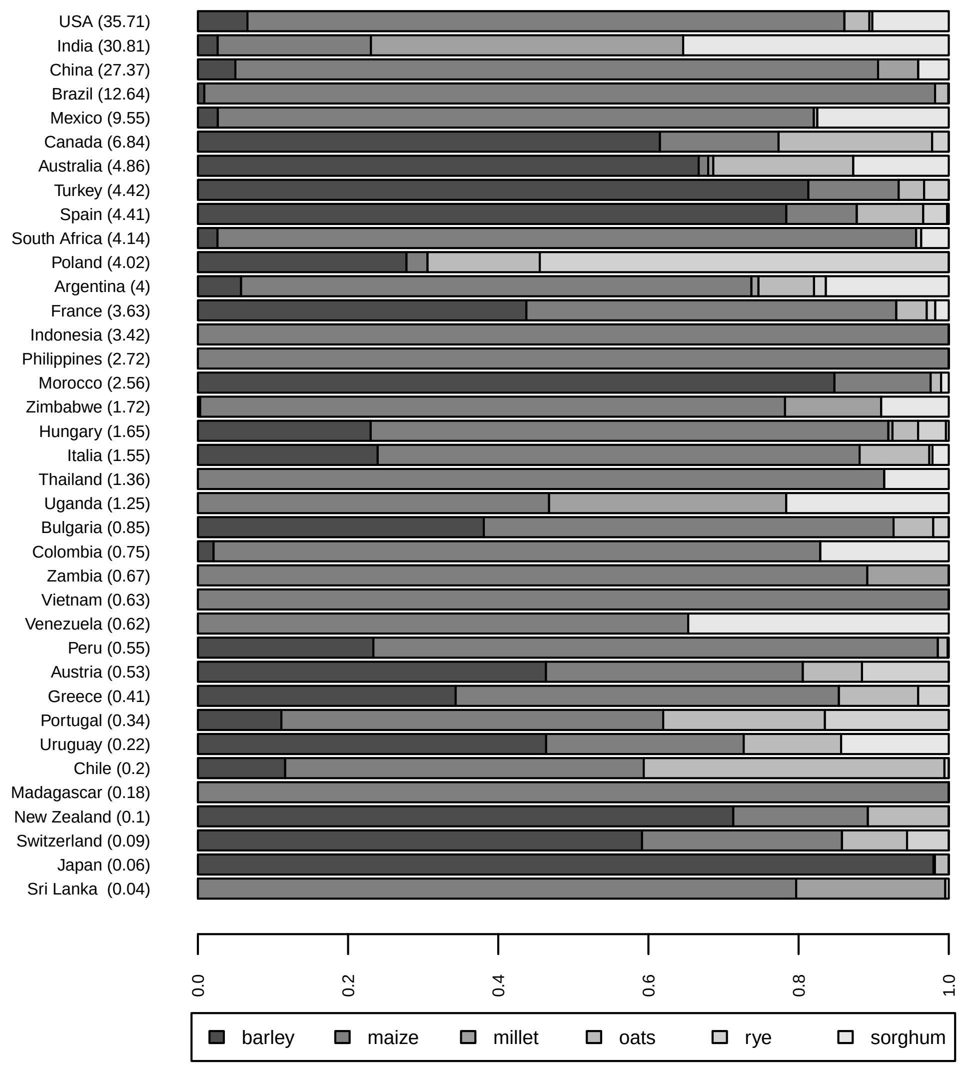

Figure 1: Composition of the Area Devoted to Coarse Grains in Selected Countries.
Note: The percentages are based on harvested areas of individual crops averaged over 1992:2002. Countries are ranked according to their total harvested area averaged during the same period. The total area, in parentheses is in million hectares. Source: FAO.

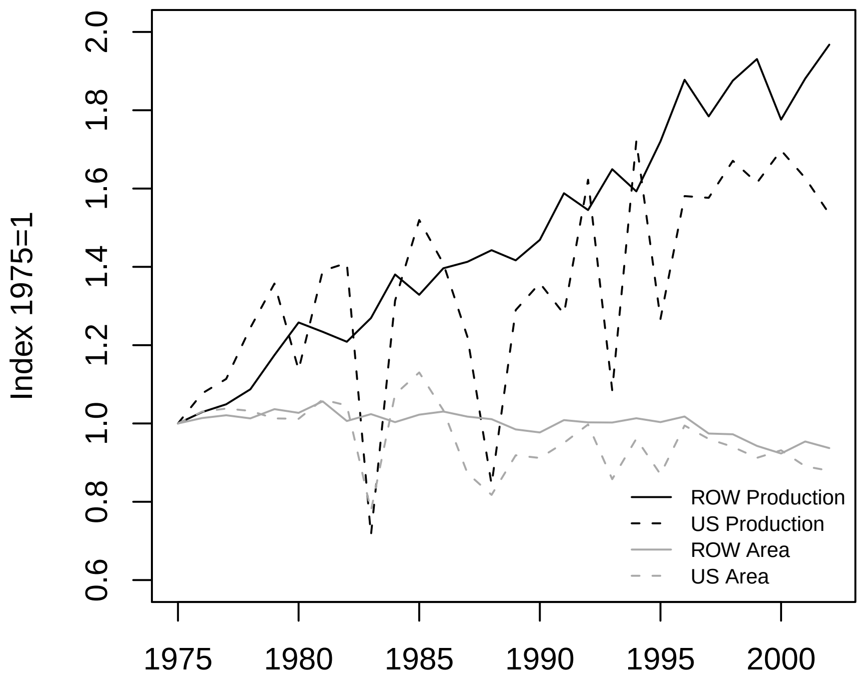

Figure 2: Production and Harvested Area of Coarse Grains in the U.S. and in the Rest of the World (ROW).
Note: The four series are relative to 1975. Source: Allen and Lutman (2009).

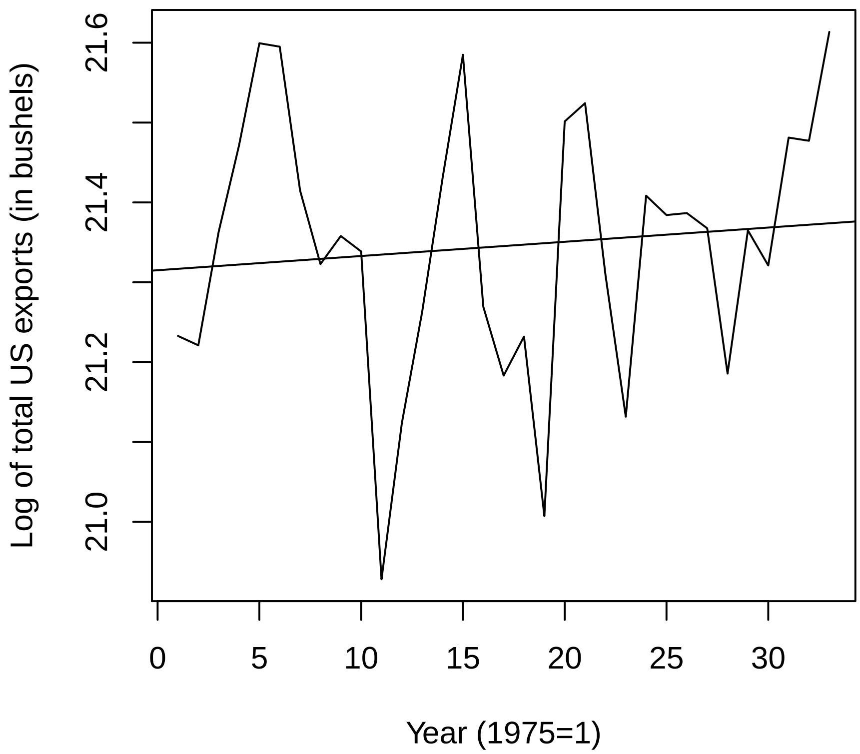

Figure 3: U.S. Exports of Coarse Grains are Volatile in the Short Run but Relatively Stable Over Time.
Note: The straight line in the plot is obtained by regressing the total log of exports on a trend variable. The obtained equation is (standard errors in parentheses): $log(exports_t) = 21.315(0.06) + 0.002T(0.003)$. Source: Allen and Lutman (2009).

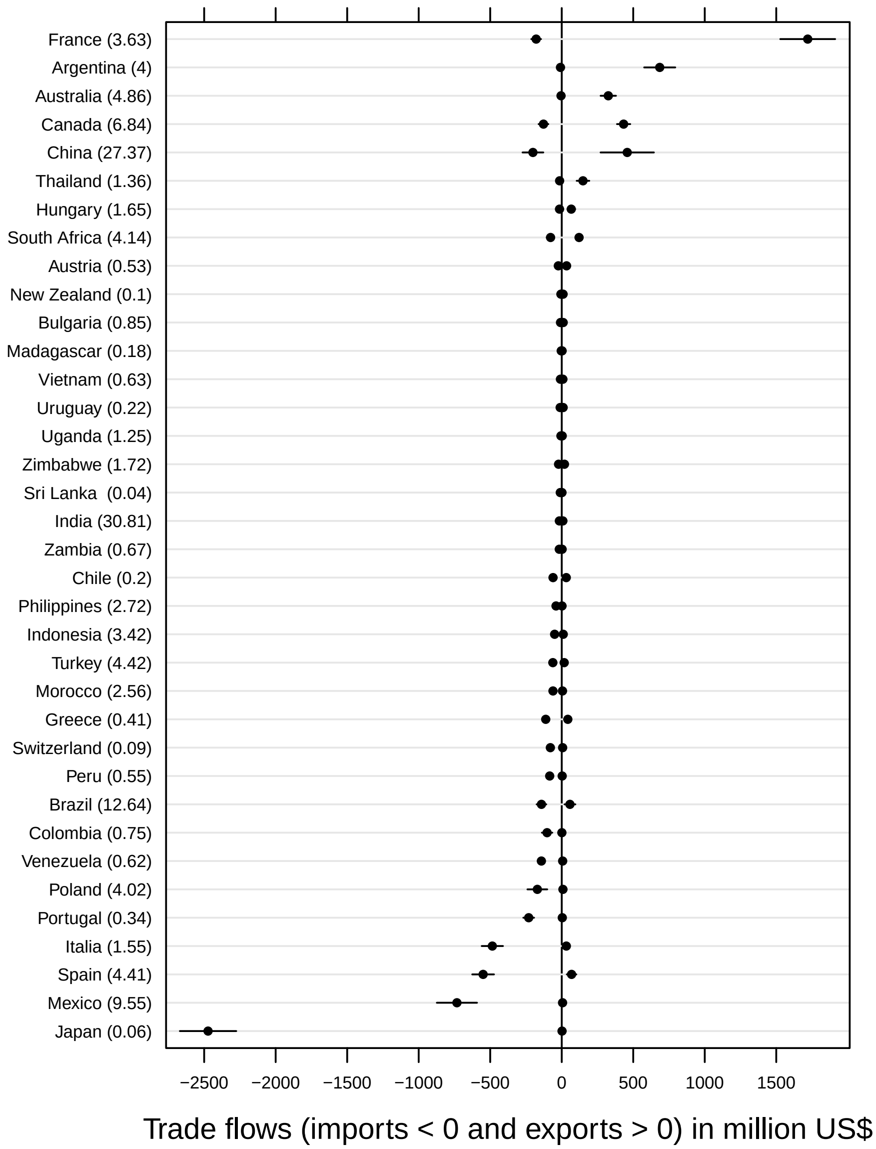

Figure 4: Imports, Exports, and Area of Selected Countries.
Note: The midpoints are averages 1975-2002 while the segments extend to the lower and upper 95% confidence interval of the annual distribution of trade flows. Countries are ordered by their net trading position (exports - imports) using the averages 1975-2002. The trade data is in millions U.S.$. Sources: GTAP V6 inGehlhar (2005).

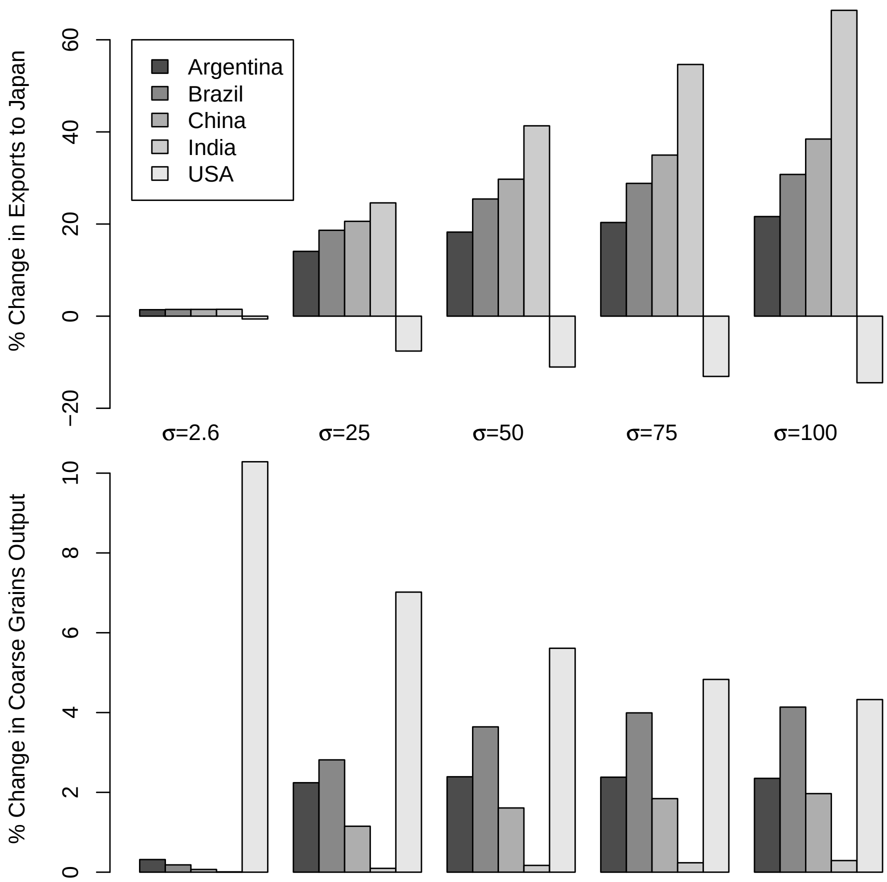

Figure 5: Supply Response and the Elasticity of Substitution.

Note: The upper panel shows, for selected countries, the percentage change in exports to Japan following a 15% increase in the U.S. industrial demand of U.S. coarse grains under different values of the elasticity of substitution between import sources, σ_M. The experiments were performed using a partial equilibrium closure (income, factor prices, and output in the non-coarse grains sector are held fixed) in the standard GTAP model (Hertel et al., 2007) using version 6 of the database (Dimaranan, 2006). The lower panel shows, for the same experiment, the percentage changes in output.

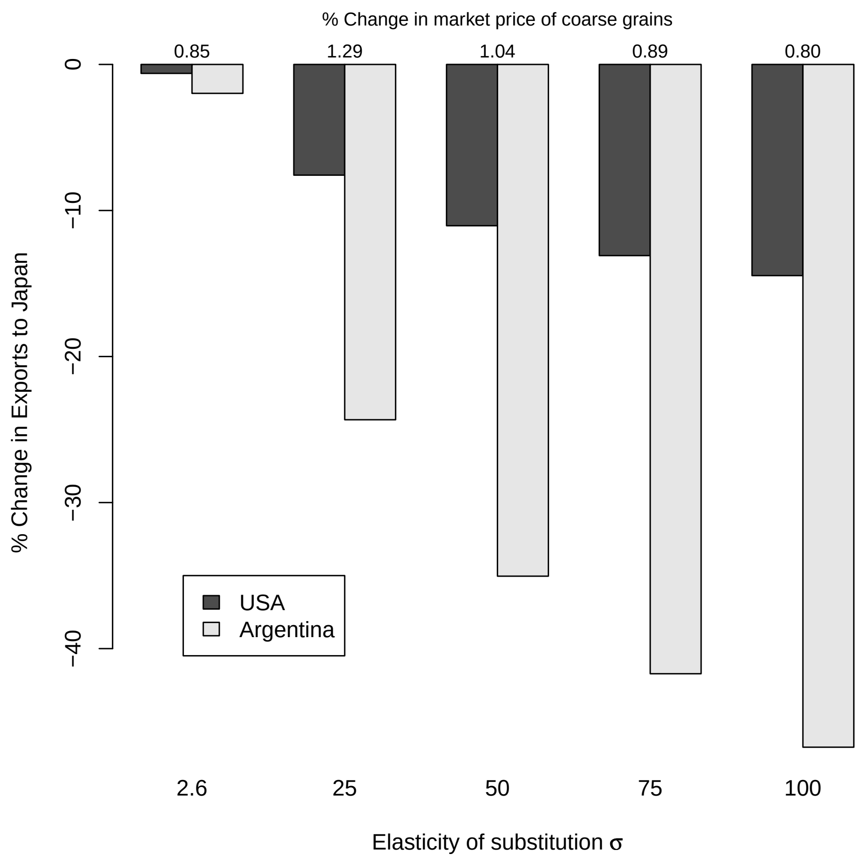

Figure 6: Market Shares and the Elasticity of Substitution.

Note: For the same change in prices, the exports of Argentina to Japan suffer a larger reduction than those of the U.S. due to the fact that the U.S. has a much larger market share (74%) of the Japanese market than Argentina (approx. 4%).

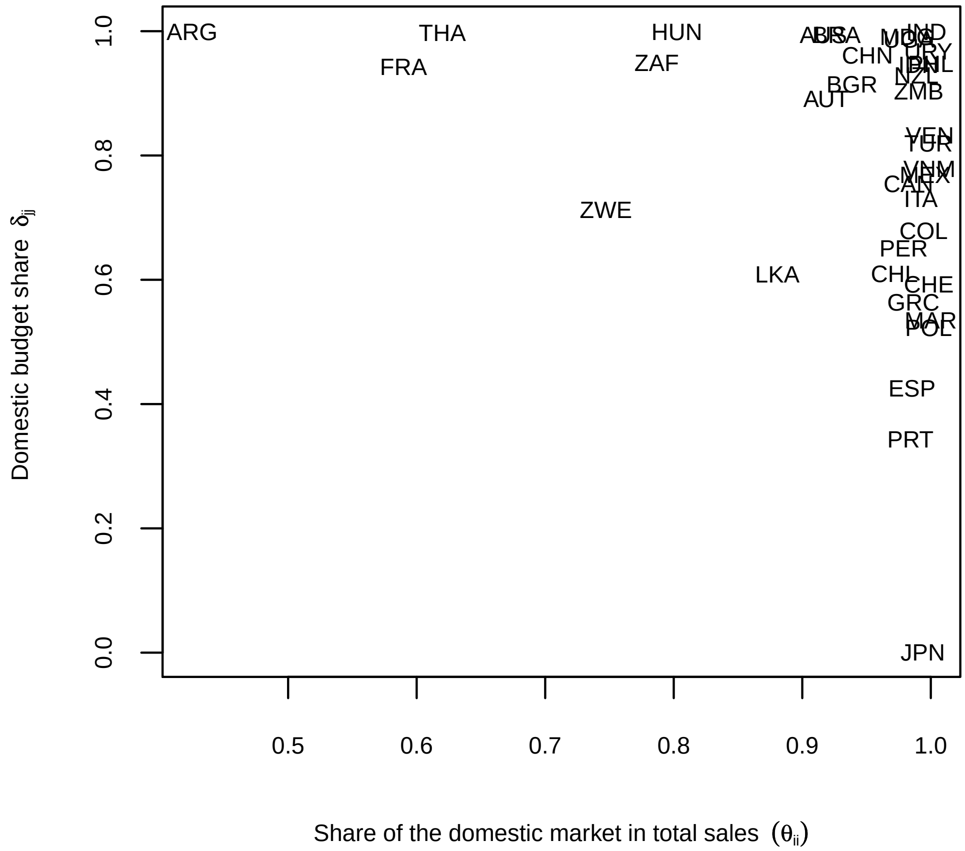

Figure 7: The Ability of Domestic Production to Supply Domestic Markets.
Note: The shares are averages over the period 1975-2002.

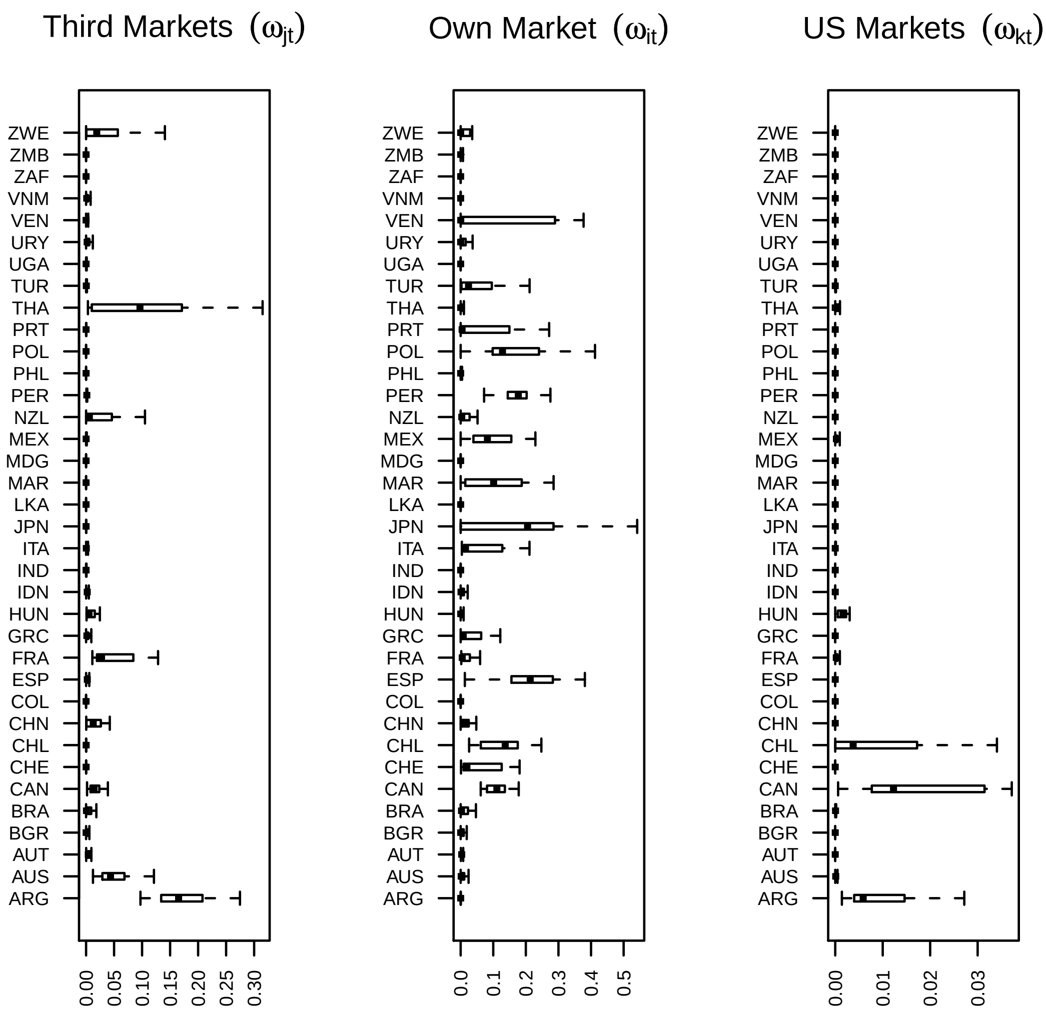

Figure 8: Competition of the Sample Countries with the U.S. in Third Markets, Their Own Markets, and US Markets.

Note: The box-plots show the distribution of the indexes over 1975-2002. The units of the indexes are between zero and unity.

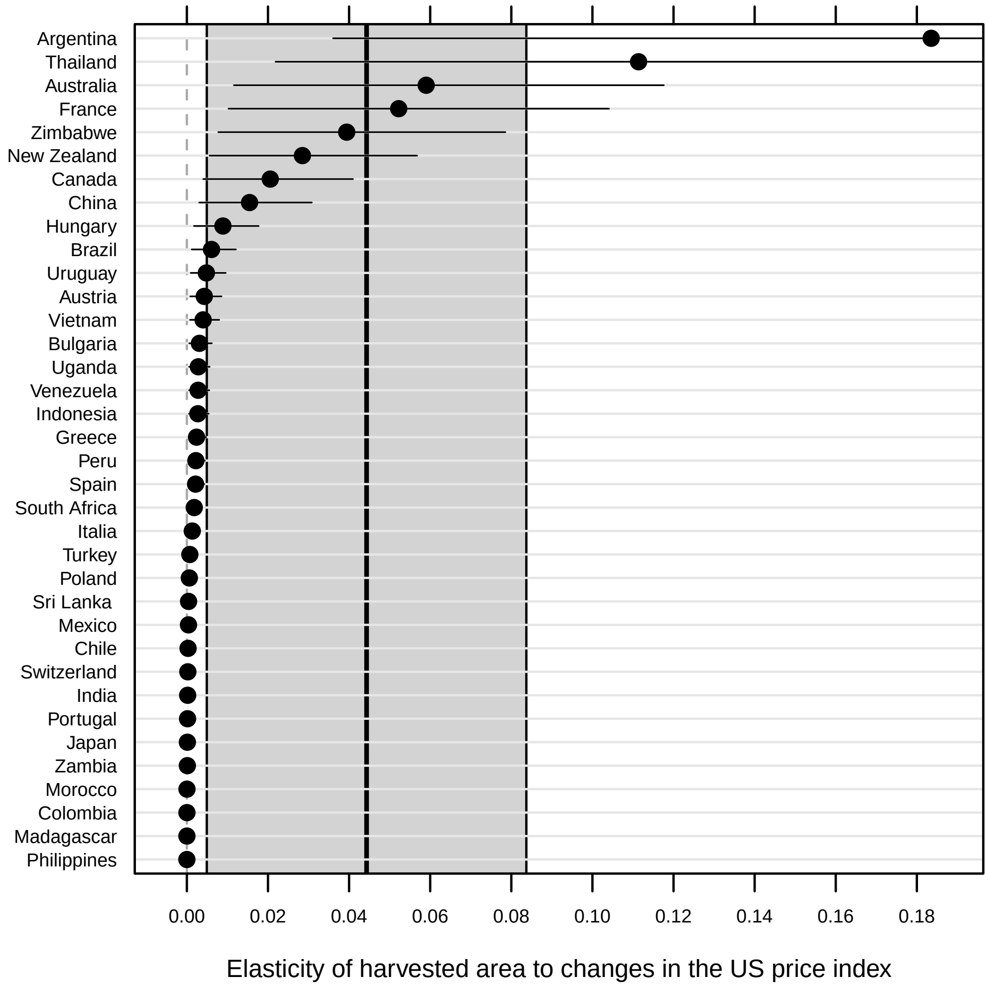

Figure 9: Short Run Elasticities of Area Harvested to Changes in the US Price of Coarse Grains. Note: The mean values of the competition index with the U.S. are used to construct the confidence intervals (thick dots are the point estimates, while the segments indicate the 95% confidence interval). The dashed black line is the point estimate β from the IWM model, while the gray band represents its 95% CI. Source: Table3.

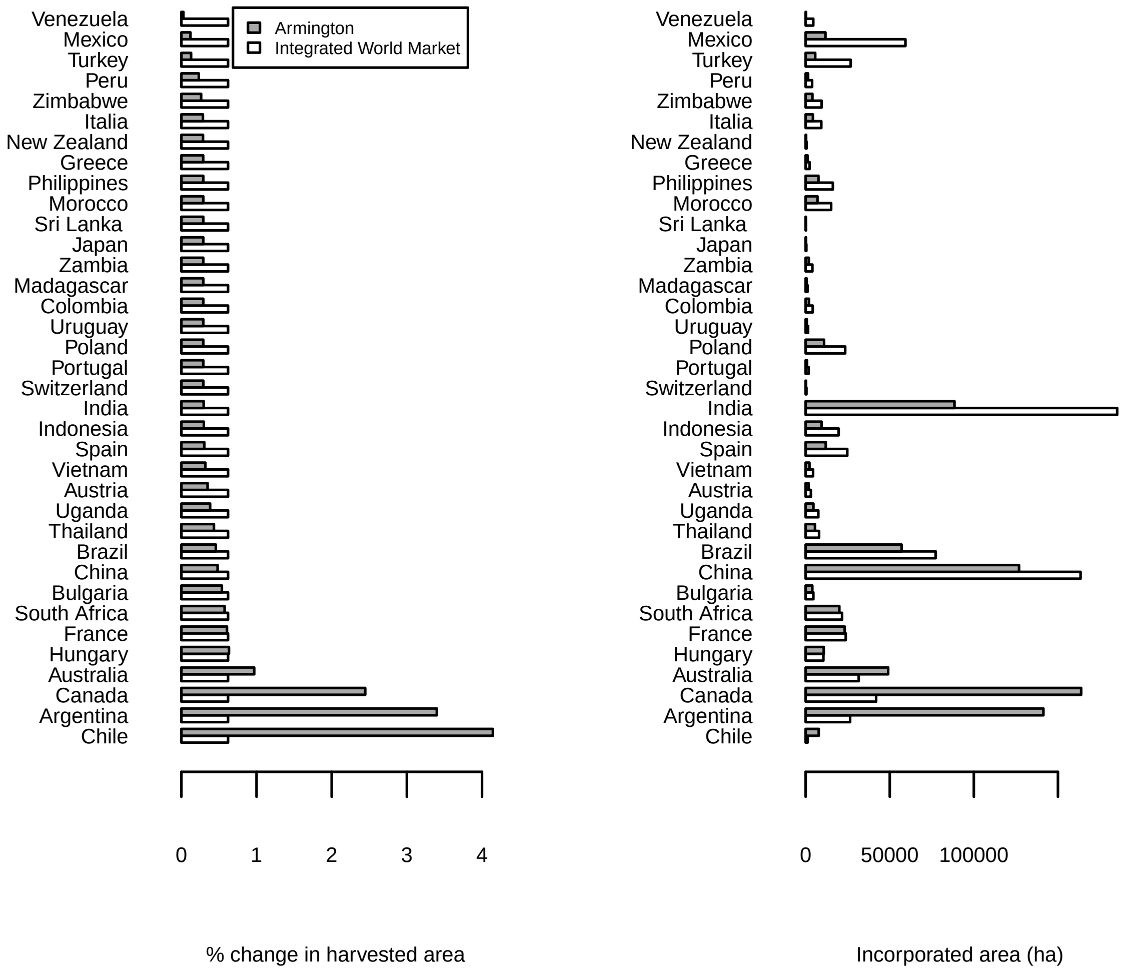

Figure 10: Pricing Mechanisms and Area Response.

Note: In the left panel, the values shown are the difference (in percentage terms) between harvested area predictions with and without shocking the US price index by 15%, keeping all the other variables constant at 1992 levels. The right panel translates the percentage changes in hectares. Source: Table 3.

Appendices

A The Armington specification of trade flows

Start with some notation. Exporting countries are indexed by i, markets are indexed by j. There are n countries trading among themselves. Armington (1969) assumes separable utilities for different groups of goods. Goods are assumed to be differentiated by place of origin. Given changes in relative prices, consumers substitute among origins according to an elasticity of substitution assumed to be invariant to the share of each good in total consumption and equal across countries. This allows aggregating total consumption of coarse grains in market j (X_j) using a constant elasticity of substitution (CES) function:

$$X_j = \left(\sum_i^n B_{ij} X_{ij}^{\rho} \right)^{\frac{1}{\rho}}, \quad \rho = 1 - \frac{1}{\sigma} \tag{A-1}$$

where B_{ij} is a constant bilateral preference weight, σ is the elasticity of substitution and X_{ij} are market j's imports from each origin i, including itself; thus, the summation above includes all n countries.

The price of X_j is an aggregate price of the prices P_{ij} attached to each X_{ij}. This aggregate price is consistent with the optimality of the choices implied by the utility maximization. Following Armington (1969, p.165), the price P_j of the aggregate X_j must fulfill the following condition:

$$P_j = P_{ij} \left(\frac{\partial X_j}{\partial X_{ij}} \right)^{-1}, \forall i. \tag{A-2}$$

To see the implications of this condition, solve for P_{ij} in (A-2) and multiply both sides by X_{ij}:

$$P_{ij} X_{ij} = P_j \frac{\partial X_j}{\partial X_{ij}} X_{ij} \tag{A-3}$$

Summing (A-3) over all the exporters i, and employing Euler's theorem[7] coupled with the fact that

[7] Given $y = f(x, z)$, if $f(\theta x, \theta z) = \theta^n f(x, z)$, then $\frac{\partial f(.)}{\partial x} x + \frac{\partial f(.)}{\partial z} z = ny$

X_j in (A-1) is homogeneous of degree 1, the following result is obtained:

$$\sum_{i}^{n} P_{ij} X_{ij} = P_j \sum_{i}^{n} \frac{\partial X_j}{\partial X_{ij}} X_{ij} = P_j X_j \tag{A-4}$$

The expression just derived has two important implications. First, these prices define the budget constraint, if consumers allocated a portion E_j of their total income to the consumption of coarse grains, $E_j \leq P_j X j = \sum_i^n P_{ij} X_{ij}$. Second, solving for P_j in (A-4), and linearizing it in percentage changes (indicated by variable names in lowercase) yields:

$$p_j = \sum_{i}^{n} \delta_{ij} p_{ij}, \quad \delta_{ij} = \frac{P_{ij} X_{ij}}{P_j X_j} \tag{A-5}$$

In other words, the percentage change in the price of the CES aggregate (A-1), p_j, is the weighted sum of the changes in the prices charged by all the suppliers of j, including j itself, using as weights the value shares of the consumption of each good in the total consumption of coarse grains in j.

Maximization of (A-1) subject to the money constraint implied by the budgetary allocation to coarse grains (E_j) yields bilateral demands of the form:

$$X_{ij} = \frac{B_{ij}^{\sigma} E_j P_{ij}^{-\sigma}}{\sum_i^n B_{ij}^{\sigma} P_{ij}^{1-\sigma}} \tag{A-6}$$

Substituting these *Marshallian or uncompensated* demands in the CES aggregate (A-1) yields the indirect utility function. Solving the indirect utility function for E_j yields the expenditure function:

$$E(\mathbf{P_{ij}}, X_j) = \left(\sum_{i} B_{ij}^{\sigma} P_{ij}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} X_j \tag{A-7}$$

Where the first term on the right hand side is the CES price index P_j. Substituting $E(\mathbf{P_{ij}}, X_j)$ for E_j in (A-6) and using the definition of the CES price index yields the *Hicksian or compensated* CES demands used in the text:

$$X_{ij} = B_{ij}^{\sigma} X_j \left(\frac{P_{ij}}{P_j} \right)^{-\sigma} \tag{A-8}$$

B Data description

B.1 Climate variables

The climate variables are calculated aggregating the climate data of the Climate Research Unit (Mitchell and Jones, 2005) from the grid cell level (30 min resolution) up to the national level using as weights the production value of each grid cell calculated using Monfreda, Ramankutty, and Foley (2008)'s productivity data and FAO (2009) prices. These growing seasons are the contiguous months within the growing seasons for the major growing regions (Lobell and Field, 2007). These are July-Aug. (maize), May-Aug. (barley) and Aug. (sorghum). Besides these crops, millet, oats, and rye are included to better match the GTAP trade category "gro." Also for the aggregated "gro," with growing seasons differing across crops, the monthly climate value of each grid-cell is weighted by the fraction of the production value corresponding to the crops grown in the grid-cell during that month.

B.2 Shares and competition indexes

From the modeling framework, two types of shares are needed. When a given country is seen as an exporter (which is indexed by i) the relevant share is θ_{ijt}, the share of country i's exports of coarse grains to country j in country i's total coarse grains output at time t. When the same country is seen as an importer (and thus indexed by j), the needed share is δ_{ijt}, the share of coarse grains originating in country i in the total consumption of coarse grains in country j at time t. In both cases, domestic sales are considered; thus, the first step is to obtain the shares θ_{iit} and δ_{jjt} given by:

$$\theta_{iit} = \frac{P_{iit}X_{iit}}{P_{it}Y_{it}}, \quad \delta_{jjt} = \frac{P_{jjt}X_{jjt}}{P_{jt}X_{jt}} \tag{A-9}$$

Obtaining data on prices or value terms for total output, consumption, and domestic sales proved difficult without compromising the country and time coverage of the sample. Fortunately, the equality of the price terms in the numerators and denominators of the shares in (A-9) allows the obtaining of shares using quantity values, which are readily available from FAO (2009). A small difficulty is that it would not be appropriate to add quantities of corn, sorghum, etc., to get the aggregated coarse grains: the empirical choice was to calculate the own-country shares using data on maize, which for most countries, comprises the bulk of produced and traded coarse grains.

This implies that the *bilateral* share θ_{ijt} can be decomposed in the following way:

$$\theta_{ijt} = \frac{P_{ijt}X_{ijt}}{P_{it}X_{it}} \equiv \frac{\sum_{j\neq i} P_{ijt}X_{ijt}}{P_{it}X_{it}} \frac{P_{ijt}X_{ijt}}{\sum_{j\neq i} P_{ijt}X_{ijt}} \equiv (1-\theta_{iit})V_{ijt} \tag{A-10}$$

where:

$$(1-\theta_{iit}) = \frac{\sum_{j\neq i} P_{ijt}X_{ijt}}{P_{it}X_{it}} \tag{A-11}$$

is the proportion of country i's output that goes to foreign markets, and

$$V_{ijt} = \frac{P_{ijt}X_{ijt}}{\sum_{j\neq i} P_{ijt}X_{ijt}} \tag{A-12}$$

is the share of country i's exports to country j in country i's total coarse grains exports. The data on V_{ijt} comes from the GTAP database, Version 6 (Gehlhar, 2005).

Likewise, δ_{ijt} can be decomposed and rewritten as follows:

$$\delta_{ijt} = \frac{P_{ijt}X_{ijt}}{P_{jt}X_{jt}} \equiv \frac{\sum_{i\neq j} P_{ijt}X_{ijt}}{P_{jt}X_{jt}} \frac{P_{ijt}X_{ijt}}{\sum_{i\neq j} P_{ijt}X_{ijt}} \equiv (1-\delta_{jjt})V_{ijt} \tag{A-13}$$

where V_{ijt} was defined before, and $(1-\delta_{jjt})$ is the proportion of country j's consumption that comes from overseas.

The shares θ_{ijt} and δ_{ijt} are used to obtain the competition indexes:

$$\begin{aligned} \omega_j &= \textstyle\sum_{j\neq(i,k)}^{n-2} \theta_{ij}\delta_{kj} \\ \omega_k &= \theta_{ik}\delta_{kk} \\ \omega_i &= \theta_{ii}\delta_{ki} \end{aligned} \tag{A-14}$$

which stand for competition between country i and country k in country j's market, in k's market and in country i's market.

B.3 US price index of coarse grains

The US price index uses export quantity weights to combine price information obtained from USDA (2009) on the following markets:

- Corn, No. 2, Yellow at Gulf ports, Louisiana Gulf, barge delivered IN $ per bushel
- Oats, No. 2 heavy white at Minneapolis IN $ per bushel
- Barley, No. 2 or better, feed at Market: Duluth IN $ per bushel
- Sorghum, No. 2, Yellow at Gulf ports, Louisiana Gulf, barge delivered IN $ per cwt

The period covered is 1975-2002.

B.4 Macroeconomic variables

The data on GDPs are in 2000 US$ and come from The World Bank (2009). The real exchange rates (RER_{it}) denominated in dollars and based on the year 2000 are constructed using consumer price indexes The World Bank (2009), and nominal exchange rates from the Penn World Tables (Heston, Summers, and Aten, 2006) using the formula:

$$RER_{it} = \frac{NER_{it}}{NER_{i,2000}} \frac{CPI_{us,t}}{CPI_{i,t}} \tag{A-15}$$

C Bootstrap based bias correction for dynamic panels

This section is taken from Everaert and Pozzi (2007) and is included here only as a reference. Given a model of the form:

$$y_t = \gamma y_{t-1} + \beta x_t + D\eta + \varepsilon = \mathbf{W}\delta + D\eta + \varepsilon \tag{B-1}$$

for $i = 1, ..., N$ (cross-sectional dimension) and $t = 2, ..., T$ (time dimension) and where y_t is the dependent variable, x_t is an explanatory variable, η are fixed effects, and ε_i an error term, the idea is to get a correction of the biased vector parameter δ:

$$\hat{\delta} = (\mathbf{W}'\mathbf{A}\mathbf{W})^{-1}\mathbf{A}y \tag{B-2}$$

where $\mathbf{A}$ (...) is an indempotent transformation matrix that eliminates the individual effects η. We know that $\hat{\delta}$ is a biased estimator fpr δ, i.e.

$$E(\hat{\delta}) = \int_{-\infty}^{\infty} z f_{\hat{\delta}}(z) dz \neq \delta, \tag{B-3}$$

From (B-3) it is clear that $\bar{\delta}$ is an unbiased estimator for δ if it satisfies:

$$\hat{\delta} = \lim_{J \to \infty} \frac{1}{J} \sum_{j=1}^{J} \widehat{\delta}_j^*(\bar{\delta}) \tag{B-4}$$

For practical purposes, a bias-corrected estimate for δ can be obtained by searching over the parameter space until a vector of parameters $\hat{\delta}$ is found that satisfies (B-4). This search is implemented through an iterative bootstrap procedure which simulates the distribution of the LSDV estimator when sampling from (2) with some vector of known parameters, say $\tilde{\delta}$. This bootstrap procedure is described in the following steps:

1. Estimate the individual effect $\tilde{\eta} = (T-1)^{-1} D'(y - W\tilde{\delta})$ and the residuals $\tilde{\varepsilon} = y - W\bar{\sigma} - D\tilde{\eta}$.

2. Choose the number of bootstrap samples B, and proceed as follows in bootstrap sample j, with $j = 1, ..., B$:

a. Obtain a bootstrap sample $\tilde{\varepsilon}^b$ from rescaled estimated residuals $\tilde{\varepsilon}^r$ (see MacKinnon, 2002).

b. Calculate a bootstrap sample $y^b = W^b\tilde{\delta} + D\tilde{\eta} + \tilde{\varepsilon}^b$ where $W^b = (y_{-1}^b, x)$ and with initialization $y_{i1}^b = y_{i1}$. So we condition on the initial values $y_i 1$ and on x.

c. Obtain the LSDV estimator $\tilde{\delta}_j^b = (\tilde{\gamma}_j^b, \tilde{\beta}_j^b)$.

3. Calculate the mean of the LSDV estimator $\tilde{\delta}_j^b$ over the B bootstrap samples as $\tilde{\delta}^b = B^{-1} \sum_{j=1}^{B} \tilde{\delta}_j^b(\tilde{\delta})$.

The mean of the bootstrap distribution $\tilde{\delta}^b$ can now be used to evaluate $\tilde{\delta}$ as an estimator for δ. From 6 we know that for $\tilde{\delta}$ to be an unbiased estimator $\bar{\delta}$ for δ, the mean $\tilde{\delta}^b$ of the bootstrap distribution $\tilde{\delta}$ should equal the original biased LSDV estimates $\hat{\delta}$, i.e. $\omega = \hat{\delta} - \tilde{\delta}^b = 0$. In order to find a parameter vector $\tilde{\delta}$ that satisfies this condition, we iterate over the bootstrap procedure outlined in steps 1-3 and evaluate $\tilde{\delta}_{(k)}$, in each iteration k, as an estimator for δ by calculating $\omega_{(k)} = \hat{\delta} - \tilde{\delta}^b_{(k)}$. If $\omega_{(k)} = 0$, $\tilde{\delta}_{(k)}$ is taken to be the unbiased estimate $\bar{\delta}$ for δ. If $\omega_{(k)} \neq 0$, $\tilde{\delta}_{(k)}$ is updated as $\tilde{\delta}_{(k+1)} = \tilde{\delta}_{(k)} + \omega_{(k)}$ and we iterate over the bootstrap procedure until this condition is satisfied. As the biased LSDV estimator $\bar{\delta}$ can be thought of as being our first guess for the vector of population parameters δ, we initialize the algorithm by setting $\tilde{\delta}_{(1)} = \hat{\delta}$.