Farm Acreage Shocks and Crop Prices: An SVAR Approach to Understanding the Impacts of Biofuels

Catherine Hausman, Maximilian Auffhammer, and Peter Berck*

Abstract

The last ten years have seen tremendous expansion in biofuels production, particularly in corn ethanol in the United States, at the same time that commodity prices (*e.g.*, corn) have experienced significant spikes. While supporters claim that biofuels are renewable and carbon-friendly, concerns have been raised about their impacts on land use and food prices. This paper analyzes how U.S. crop prices have responded to shocks in acreage supply; these shocks can be thought of as a shock to the residual supply of corn for food. Using a structural vector auto-regression framework, we examine shocks to a crop's own acreage and to total cropland. This allows us to estimate the effect of dedicating cropland or non-crop farmlands to biofuels feedstock production. A negative shock in own acreage leads to an increase in price for soybeans and corn. Our calculations show that increased corn ethanol production during the boom production year 2006/2007 explains approximately 27% of the experienced corn price rise.

*Department of Agricultural and Resource Economics, University of California, Berkeley. Corresponding author is Catherine Hausman (formerly Almirall), Department of Agricultural and Resource Economics, 207 Giannini Hall, University of California, Berkeley, CA 94720-3310, USA; Phone: 510-725-8860; Fax: 510-643-8911; E-Mail: calmirall@berkeley.edu. We are grateful to the editor and two anonymous reviewers whose suggestions have greatly improved the manuscript. Many thanks to Michael Jansson for helpful suggestions. Financial support provided by the Energy Biosciences Institute. All views are those of the authors alone and do not represent those of the EBI.

1 Introduction

Biofuels have been promoted as an alternative to petroleum products that bypasses some of the fundamental problems with the oil market: supporters claim that it is renewable (whereas conventional oil is exhaustible), produced in the U.S. (as opposed to regimes in some cases unfriendly to the U.S.), and carbon-friendly. As biofuels production has expanded, however, concerns have been raised about their direct and indirect impacts, particularly on land use and on food prices. The last ten years have seen tremendous expansion in biofuels production, particularly in corn ethanol in the United States, at the same time that commodity prices (*e.g.*, corn) have experienced significant spikes. In 2006, 2.1 billion bushels of corn went to ethanol production (approximately 15% of all corn production); in 2007 this rose to over 3 billion bushels (20% of corn production). From 2005 to 2006, corn prices rose by 47%, and from 2006 to 2007 they rose by another 28%. Not surprisingly, the "food versus fuel" debate has economists and policy makers asking how much of the increase in corn prices is due to the increased demand for corn from ethanol producers. On October 3, 2007, the BBC argued that "[i]t is one of the most hotly debated environmental topics of the year - whether the drive to produce alternative so-called green fuels will take food from the mouths of the hungry." As Roberts and Schlenker (2010) point out, the economic literature has yet to agree on the magnitude of these hypothesized effects.

While corn for ethanol has, to date, competed directly with the production of food commodities, scientists hope that future so-called "second generation" biofuels will use non-food crops and marginal lands (Heaton et al. 2008, Hill et al. 2006, and Robertson et al. 2008). The primary first-generation biofuel is corn ethanol, which uses a food crop and conventional sugar to ethanol fermentation to produce fuels used in transportation. Second-generation biofuels use non-food crops (*e.g.*, miscanthus or switchgrass) and a different technology, in which cellulosic plant material is converted into ethanol. This paper contributes to a growing literature analyzing the effect of corn ethanol on commodity and food prices. In this paper, we examine econometrically the effects of exogenous shocks in acreage supply on food crop prices, using 50 years of U.S. data. We analyze the effects of both shocks to a crop's own acreage and shocks to total cropland. Our approach allows us to calculate the effect of dedicating existing cropland to biofuels feedstock production as well as the effect of dedicating non-crop lands. We focus on shocks to acreage rather than shocks to crop yields, as there has been no evidence to date that biofuels production will change the yields of crops grown for food. Accordingly, the shock to acreage that we estimate is a useful summary statistic for shocks to the production of crops such as corn and soybeans.[1] These shocks to acreage can be thought of as a shock to the residual supply for corn for food; corn can be dedicated to food uses or fuel uses, so a shock to the demand

[1]We do allow the shock to acreage to feed through to shocks to yield, since farmers will re-optimize their yield decisions based on acreage decisions.

for fuel corn translates into a shock to the residual supply for food corn.

Our paper makes three main contributions to this literature. First, we rely on econometrically estimated coefficients using observed data for the U.S., an approach that steps away from the previous analyses that assumed supply and demand elasticities (*e.g.*, Sexton et al. (2008) and Rajagopal et al. (2007)). Second, we study the effects of removing either existing cropland or non-crop farmland on food crop prices. Third, in order to address the obvious endogeneity of supply and demand systems, we borrow econometric tools from the macroeconomic literature that leverage the timing of exogenous shocks.

The dynamic nature of agricultural production makes this question ideally suited for a structural vector auto-regression (SVAR) of the sort routinely used in the macroeconomic literature (*e.g.*, Christiano et al. (2005)). This framework allows us to leverage the timing of planting decisions versus harvest outcomes with a classic time-series methodology. We estimate a system of equations to explain the relationship between total cropland, corn and soybean acreage, and corn and soybean spot and futures prices. In particular, we use a factor-augmented structural vector auto-regression to allow for exogenous shocks to the entire system, including supply-side shocks, such as a spike in farm input prices, or a demand-side shock, such as increased foreign demand.

Time-series models such as SVARs are widely used in macroeconomic applications where variables are jointly determined, adjustment to long-run equilibrium is not instantaneous (which implies the importance of including lags in the model), and the underlying data generating process follows a specific timing mechanism. The above are all characteristics of an agricultural supply model, in which price and acreage are jointly determined and in which the effects of a shock can last several periods.[2] We show that a structural VAR can be used to leverage the sequencing inherent in agricultural supply. The resulting model is able to capture dynamics that may be missed with other models. Also, the system's dynamic nature allows us to estimate forecast error variance decompositions (FEVDs), which explain the percentage of variance that comes from specific shocks, and impulse response functions (IRFs), which trace out the effect of exogenous shocks across time.

Estimation results show that a reduction in corn area of one million acres (approximately 1% of U.S. corn acreage) leads to a corn price increase of \$0.04 per bushel (an approximately 1% increase). To put this in perspective, consider that from 2006 to 2007 (a boom production year), corn acreage dedicated to

[2]Previous work on dynamic agricultural systems beyond Nerlove's model (1956) is relatively limited. In the Nerlovian framework, farmers make production decisions (including acreage decisions) according to their expectations on crop prices and input prices. At planting time, they observe only last year's prices, planting-time spot prices, and futures prices for harvest-time delivery. Their decision can be modeled in a partial-adjustment framework, in which acreage at time t is a function of acreage at time $t-1$, plus futures prices and past spot prices. This equation appears in our model, but we generalize the framework by also modeling the movement of prices. Mushtaq and Dawson (2002) use a recursive vector auto-regression approach to investigate acreage response of various crops in Pakistan. They find that this approach is more appropriate than a Nerlovian partial-adjustment model, particularly in explaining adjustments to long-run equilibrium. However since their interest is in acreage response, they do not report the impact of shocks to acreage on prices.

ethanol increased by about 6 million acres in the United States. At the same time, corn prices rose by $0.88 per bushel. Thus our model finds that approximately 27% of the price increase was due to new ethanol production. For a negative shock in soybean acreage, we find a price increase of $0.23 per bushel, with a wider variation across alternative specifications. This much larger magnitude is partly explained by the fact that US soybean acreage has a much larger share of world production than does US corn acreage.

For the scenario in which one million acres is moved from non-crop farmland (*e.g.*, pasture and idle lands) to crop production, while holding corn or soybean acreage constant, we find a price decrease of $0.05 per bushel of corn and $0.08 per bushel of soybeans. The intuition for this result is that, while corn (or soybean) acreage is held constant, the production of substitute crops increases. Accordingly, demand for the crop falls. Historically, this has corresponded to increased production of other grains, but the intuition is consistent with a scenario of increased production of substitute biofuels feedstocks. Thus we find robust results that removing acreage from food production, in order to grow biofuels feedstocks, increases food crop prices. On the other hand, switching non-crop farmland to a substitute crop can lower food crop prices. As the U.S. and E.U. formulate biofuels policies, this empirical evidence on the costs for food production and commodity prices should be taken into account. This paper proceeds as follows: in section 2, we summarize the related literature. In section 3, we develop the econometric framework; in section 4 we present the data. Section 5 shows the main results, section 6 presents the robustness checks, and in section 7 we conclude.

2 Related Literature

A number of papers have simulated the impact of biofuels production on various economic outcomes, generally using supply and demand elasticities drawn from the literature. Banse and van Meijl (2008) use a global CGE model (modified from the GTAP model) to analyze the trade impacts of an EU Biofuels Directive. They find that cereals prices actually decline in the long run, but less than they would without the directive. This finding is a combination of an assumed inelastic demand and a high rate of productivity change. Rajagopal et al. (2007) use a stylized partial equilibrium model and find a 21% increase in corn price attributable to a $0.51 ethanol production tax credit in the US in 2006. A 2009 CBO study estimated that 10 to 15% of the food price increase from April 2007 to April 2008 was attributable to expanded ethanol production. The estimated impact on corn prices for the same period is higher: between 50 and 80 cents per bushel, or 28 to 47 percent of the total corn price increase. Rosegrant (2008) uses a partial equilibrium model from IFPRI and finds that biofuel demand accounted for 39% of the corn price increase from 2000 to 2007. Chakravorty et al. (2011) find in simulations that biofuels mandates drive land allocation changes rather than large food price increases. Two other oft-cited papers argue that biofuels policy has driven up corn prices but do not make direct calculations (Abbott, Hurt, and Tyner (2008) and Mitchell (2008)). Finally, Chakravorty, Hubert and Nostbakken (2009) provide an extensive literature review on the various models that have been applied to biofuels, including partial equilibrium agricultural models from FAPRI, IFPRI, and IIASA; general equilibrium models such as GTAP. Naylor et al. (2007) provide a useful summary of predicted crop price changes under the various biofuels scenarios found in the literature.

In related work, some papers analyze the welfare impacts in developing countries (*e.g.*, Runge and Senauer, 2007). Naylor et al. (2007) focus on the food security impacts of biofuels expansion. Roberts and Schlenker (2010) construct an elegant system of supply and demand for food calories, which they estimate using novel instrumental variable techniques based on weather shocks. The resulting price increases imply a large change in global consumer surplus.

Several papers have looked at some of the other possible explanations for recent crop prices. These explanations include export restrictions, growing food demand from developing countries, low investment and hence low productivity growth, weather shocks, crop diseases, depreciation of the U.S. dollar, increases in the price of crude oil, production cost increases, speculation in commodity markets, and the additional impact of low stocks (*e.g.*, Abbott, Hurt, and Tyner (2008), Chakravorty et al. (2011), Headey and Fan (2008), and Mitchell (2008)).

Finally, recent papers examine at the impact of biofuels policies on economic outcomes other than crop prices. Ando, Khanna, and Taheripour (2010) evaluate the impact of the Renewable Fuel Standard on the transportation sector; Khanna, Ando, and Taheripour (2008) evaluate the impact of ethanol production on greenhouse gas emissions and congestion externalities.

3 Econometric Framework

3.1 Structural Vector Auto-Regression: Scenario 1

We apply a structural vector auto-regression[3] to analyze what we call scenario 1, in which food acreage is removed and dedicated to a biofuels feedstock.[4] In this framework, we use a system of equations to explain the relationship between corn and soybean acreage, total cropland, and corn and soybean spot and futures price. We impose identification restrictions that take advantage of the timing of planting decisions in the United States. That is, agricultural producers set their acreage at planting time according to their expectation of harvest-time prices. In the classic Nerlovian framework, this implies that current acreage is

[3]In time series literature, "structural" vector auto-regressions refer to vector auto-regressions that allow for causal interpretations. This use of the word "structural" is different from that in the general econometrics literature; the system of equations need not explicitly model an optimization problem.

[4]Note that what matters for this model is not which feedstock is grown (*e.g.*, corn versus switchgrass) but where that feedstock is grown (land previously dedicated to food production versus non-crop farmland).

a function of past acreage, past prices and futures prices, and supply-side variables such as input prices, which are the only variables observable to farmers at planting time. Price at harvest is then a function of production (acreage times yield) and a number of demand-side market forces. Accordingly, the vector of variables of interest is

$$y_t \equiv \begin{bmatrix} corn\,futures_t \\ soy\,futures_t \\ supply\,variables_t \\ corn\,yield_t \\ corn\,acreage_t \\ total\,farmland_t \\ demand\,variables_t \\ corn\,harvest\,price_t \\ soy\,harvest\,price_t \end{bmatrix}$$

A generic vector auto-regression with exogenous variables x_t has the following structure

$$Ay_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_k y_{t-k} + Cx_t + B\varepsilon_t \tag{1}$$

where $\varepsilon_t \sim N(0, I_K)$ and $E(\varepsilon_s \varepsilon_t = 0), s \neq t$. We then impose restrictions according to our identifying assumptions. We can re-write the above equation as follows

$$y_t = A^{-1}A_1 y_{t-1} + A^{-1}A_2 y_{t-2} + ... + A^{-1}A_k y_{t-k} + A^{-1}Cx_t + u_t \tag{2}$$

where $u_t = A^{-1}B\varepsilon_t$, implying that u_t follows a white noise process. Our identifying assumptions are

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & 1 & 0 & 0 & 0 & 0 & 0 \\ a_{51} & a_{52} & a_{53} & 0 & 1 & 0 & 0 & 0 & 0 \\ a_{61} & a_{62} & a_{63} & 0 & 0 & 1 & 0 & 0 & 0 \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & 1 & 0 & 0 \\ a_{81} & a_{82} & a_{83} & a_{84} & a_{85} & a_{86} & a_{87} & 1 & 0 \\ a_{91} & a_{92} & a_{93} & a_{94} & a_{95} & a_{96} & a_{97} & 0 & 1 \end{bmatrix}$$

$$
B = \begin{bmatrix}
b_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & b_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & b_{33} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & b_{44} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & b_{55} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & b_{66} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & b_{77} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{88} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{99}
\end{bmatrix}
$$

Note that there are no restrictions on the lags (A_1through A_k) or on the coefficients on exogenous coefficients. The restrictions on A come from the timing of the agricultural production process in the United States. That is, corn and soybean futures (which are observed in March for harvest-time delivery) and supply-side variables (*e.g.*, input costs and loan rates) are generated before acreage decisions or fall prices have been observed; accordingly, they are functions of only lagged and exogenous variables. Total farmland, corn acreage, and corn yields are determined after futures prices have been observed but before fall prices are known. Thus the information set at time t for these variables consists of futures prices, supply-side variables, and lagged and exogenous variables. This is similar to the acreage function in a Nerlovian (partial-adjustment) framework. Demand-side variables (such as foreign production, affecting demand for US exports) that are determined in the summer are functions of futures prices, supply-side variables, US acreage and yields, and lagged and exogenous variables. Finally, harvest time prices are a function of that year's futures prices, supply-side variables, acreage and yield decisions, and demand-side variables. The structure of B imposes orthogonality of contemporary structural shocks. The exogenous variables x_t are constants and time trends. We focus on the case of one lag ($k = 1$), which we find is preferred according to Schwarz's Bayesian information criteria, but examine the robustness of our estimates to the inclusion of additional lags. The system is estimated via maximum likelihood.

3.2 Diffusion Indices

As mentioned above, the orthogonality conditions on the matrix B require that there be no omitted variables. Accordingly we control for spring-time supply-side variables (such as input prices and agricultural loan rates) and summer-time demand-side variables (such as US income and foreign production, which affects demand for US exports). However the curse of dimensionality prevents us from including all of these variables in the system; we would quickly run out of degrees of freedom. Thus we include diffusion indices (also known as principal components or factors) to control for these variables while avoiding the curse of dimensionality inherent in large VAR models.

Stock and Watson (2002) show that a large number of time series variables can be summarized with a few indices using principal components analysis. The end result is a linear combination of the original time series, with the linear coefficients chosen to incorporate as much of the variation in the original series as possible. This nonparametric approach begins with the objective function

$$(\hat{F}, \hat{\Lambda}) = argmin[(NT)^{-1} \sum_i \sum_t (x_{it} - \lambda_i F_t)^2] \tag{3}$$

where (F) are the factors and (Λ) the factor loadings. This is solved by setting $\hat{\Lambda}$ equal to the eigenvectors of $X'X$ corresponding to the largest eigenvalues. $\hat{F}$ is then found by setting $\hat{F} = X'\hat{\Lambda}/N$. This approach is applied separately to two sets of time series variables, one consisting of variables affecting crop supply and one of variables affecting crop demand. For each variable used in the indices we test for a unit root, take the first difference of the natural log, and standardize to mean zero and unit variance. Then we estimate a supply-side diffusion index and a demand-side diffusion index, which are both incorporated into the structural vector auto-regression. Thus we are able to control for international and domestic macroeconomic disturbances.

3.3 Robustness Checks

One potential concern with the above factor-augmented SVAR is its large size. Generally, smaller systems perform better in this framework than do larger systems. This concern obviously needs to be balanced with the potential of omitted variables bias. To address the concern, we also estimate a sparser model as a robustness check. This model contains only corn and soybean futures prices, corn acreage, farmland, and corn and soybean farmgate prices. Accordingly, the identifying restrictions on matrices A and B are:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & 1 & 0 & 0 & 0 \\ a_{41} & a_{42} & 0 & 1 & 0 & 0 \\ a_{51} & a_{52} & a_{53} & a_{54} & 1 & 0 \\ a_{61} & a_{62} & a_{63} & a_{64} & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{66} \end{bmatrix}$$

Additional robustness checks include using a log/log specification, allowing additional lags, varying the time period studied, and varying the factor indices for supply- and demand-side variables.

3.4 SVAR Framework for Scenario 2

In scenario 2, we consider the effect of growing a biofuels feedstock on acreage not previously dedicated to food-crop production. We hold own (corn or soybean) acreage constant, while decreasing non-crop farmland (thus increasing total cropland). In a robustness check, we show that results are fairly similar if the system is estimated with a positive shock to total cropland rather than a negative shock to non-crop farmland. Accordingly, the vector of variables of interest is

$$y_t \equiv \begin{bmatrix} corn\,futures_t \\ soy\,futures_t \\ supply\,variables_t \\ corn\,yield_t \\ corn\,acreage_t \\ non-crop\,farmland_t \\ demand\,variables_t \\ corn\,harvest\,price_t \\ soy\,harvest\,price_t \end{bmatrix}$$

The identifying restrictions on A and B are the same as for scenario 1. We again consider various robustness checks, including a sparser specification (ignoring yields, supply variables, and demand variables), additional lags, a log/log functional form, different supply- and demand-side indexes, and a varying time frame.

3.5 Forecast Error Variance Decomposition and Impulse Response Functions

The dynamics of the system imply that interpretation of either the reduced form or structural coefficients is difficult. Two tools for analyzing the coefficients are forecast error variance decompositions (FEVDs) and impulse response functions (IRFs). Forecast error variance decomposition tells us the percentage of the forecasting error for a variable due to a specific shock at a given horizon. Following Lutkepohl (1993), we define the FEVD at horizon h as $\omega_{jk,h} = \frac{\psi_{jk,0}^2+\ldots+\psi_{jk,h-1}^2}{MSE[\hat{y}_{j,t}(h)]}$, where $\psi_{mn,i}$ denotes the mn-th element of $(A^{-1}A_1)^i A^{-1}B$, and $MSE[\hat{y}_{j,t}(h)] = E(y_{j,t+h} - \hat{y}_{j,t}(h))^2 = \sum_{k=1}^{K}(\psi_{jk,0}^2 + \ldots + \psi_{jk,h-1}^2)$. Thus the FEVD at horizon h (for instance, $h = 2$) estimates the percentage of the total forecast error that comes from each orthogonalized structural shock.

The dynamic nature of the above system also allows us to estimate impulse response functions (IRFs), which trace out the effect of exogenous shocks on realizations of the random variables across time. Working from the VAR's moving average representation, we can write the structural impulse response function as follows: $y_t = \mu + \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}$. Thus the structural impulse response function traces each element of Ψ_i for each time period following a shock in period $i = 0$.

We report confidence intervals based on the delta method, as bootstrapped confidence intervals in our over-identified system require a great deal of computing power. Clearly this is an imperfect solution, given how the delta method can perform in a highly non-linear system. We do calculate confidence intervals for the main specifications using a parametric bootstrap method, and we find that the intervals are quite similar to those computed with the delta method. The bootstrapped intervals are slightly narrower but do not change the inference.

Finally, we can compute cumulative impulse response functions from the coefficients. The above structural impulse response functions give the dynamic path of each variable following a shock in period $i = 0$. A cumulative impulse response function gives the dynamic path of each variable as the shock is repeated in each period $i = 0, 1, 2 \ldots n$. and is given by $\Xi_n = \sum_{i=0}^{n} \Psi_i$.

4 Data

Data are obtained for US production of corn and soybeans from 1956 to 2007. Data on farmland, planted corn acreage, and planted soybean acreage (all measured in thousand acres) are obtained from the National Agricultural Statistics Service (NASS) at the USDA. Crop prices paid to farmers, in dollars per bushel, are also obtained from NASS. These are then deflated by the third-quarter GDP deflator, obtained from the Bureau of Economic Analysis (BEA), into 2007 dollars per bushel. Corn and soybean futures, available from

Datastream and the Wall Street Journal, are planting-time quotes for delivery at harvest time. For corn, they are the March 31 closing price for delivery in September. For soybeans, they are the March 31 closing price for delivery in November. The futures prices are deflated by the first-quarter GDP deflator (from the BEA). Corn and soybean yields, in bushels per planted acre,[5] are calculated from NASS production and acreage data. Data on total cropland (also measured in thousand acres) is obtained from NASS, but unfortunately it is only available to 2006. Accordingly, as a robustness check, we also use NASS data on acreage devoted to principal crops.[6]

All variables are examined for evidence of unit roots (table 1). We consider augmented Dickey-Fuller unit root tests and Phillips-Perron tests, with and without trends, for all variables. A unit root is rejected at the 5% level for corn acreage, corn and soy yields, and the supply and demand diffusion indices. We also perform a Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for each variable, with and without trends. Trend stationarity is rejected at the 5% level for corn and soy spot and futures, soy acreage, farmland, non-crop farmland, and soy yields.[7]

For the supply-side diffusion index, eleven variables are used. As these affect farmer planting decisions, they reflect only information available up until planting time (for corn and soybeans, first quarter data). Oil prices are given by West Texas Intermediate Oil Prices at the end of the first quarter, available from Global Financial Data (GFD). The national average loan rate for corn, soybeans, and wheat is available from the Commodity Research Bureau (CRB) from 1956 to 2003, and from the Economic Research Service (ERS) at the USDA for 2004 to 2007. Input prices, including automobiles, two USDA-calculated indices of producer prices paid, building materials, and farm wages are available from NASS. All prices are deflated by the BEA's GDP deflator. Note that fertilizer prices, an important input affecting crop profitability, are incorporated in the USDA-calculated indices of prices paid.

For the demand-side diffusion index, data should reflect information available up until (and including) harvest time. For corn and soybeans, this implies fourth quarter data. Oil prices are given by West Texas Intermediate Oil Prices, available from Global Financial Data (GFD). Third-quarter US Gross National Product is obtained from GFD. Oil prices and GNP are deflated by the US GDP deflator, also from GFD. Corn and soybean production from Southern Hemisphere countries (Argentina and Brazil), which harvest during the US summer months, are available from the Food and Agriculture Organization of the United Nations (FAO). The ideal diffusion index would also incorporate GDP from importing countries, but this is

[5] Results are extremely similar if we use yields from harvested, rather than planted, acres (figure 3).

[6] Principal crops include barley, beans, beets, corn, cotton, flax, hay, oats, peanuts, potatoes, rice, rye, sorghum, soy, sugarcane, tobacco, and wheat. For most crops, data on planted acreage is used; for beans, sugarcane, and tobacco only data on harvested acreage is available.

[7] We also test for Granger causality between all variables. As expected, lagged values of corn acreage and corn and soy spot prices help predict current values of corn futures prices; soy spot prices help predict soy futures prices; etc. Full results are available upon request.

not reliably available on a quarterly basis. Since the timing of innovations is important for the ordering of the VAR, incorporating GDP data updated yearly would be inappropriate.

The estimated indices are linear combinations of the logged, differenced, standardized variables. The first supply index accounts for 30% of the variation in all the series. The two USDA-computed input prices series and oil prices have the largest coefficients in this linear combination, thus the index loads primarily onto them. The second supply index accounts for 20% of the variation and loads primarily onto loan prices. The first demand index accounts for 24% of the variation in all the series and loads onto all six series. The second demand index accounts for 23% of the variation in all the series and loads primarily onto soy and corn production in Argentina.[8]

5 Results

We estimate the parameters using maximum likelihood,[9] then compute the percentage variance due to shocks (the Forecast Error Variance Decompositions) and the dynamic path of the variables following a shock (the Impulse Response Functions). The one-lag nine-equation SVAR gives the following structural FEVD for scenarios 1 and 2 with corn acreage (tables 2.1 and 2.2):[10]

[Tables 2.1 and 2.2: FEVD]

Thus shocks to corn acreage explain 7% of the one-step ahead forecast error in corn harvest prices, for scenario 1. The supply index contributes to 16% of the error, and unobserved variation to 51%. For scenario 2, shocks to non-crop farmland explain 4% of the one-step ahead forecast error in corn harvest prices. The majority of the variation comes from the supply index (21%) and unobserved variables (58%).[11]

The structural FEVDs for the soy scenarios are as follows:

[Tables 2.3 and 2.4: FEVD]

For scenario 1, shocks to soybean acreage contribute to slightly more (9%) of the own-crop harvest price than was the case for corn. The majority of the forecast error continues to come from variation in the supply index (24%) and unobserved variables (53%). The largest contributors to the forecast error are again the supply index and unobserved variables. Thus for both crops and both scenarios, shocks to acreage have historically contributed to a fairly small percentage of the one-step forecast error in spot prices.

Figure 1 shows selected structural impulse response functions estimated from the one-lag nine-equation

[8]Detailed descriptions of the estimation of these indices are available upon request.

[9]A table of estimated coefficients is in Appendix A.

[10]Standard errors are not shown because of space considerations. They are available upon request.

[11]FEVDs at longer horizons are available upon request. For the two-step forecast error in corn harvest prices, corn acreage still explains 7%. At two periods, shocks to non-crop farmland explain 11% of the forecast error. At two periods, shocks to soybean acreage contribute 7% of the forecast error in harvest price. For scenario 2, shocks to non-crop farmland contribute to 13% of the two-step soybean price forecast error. Thus for both crops and both scenarios, shocks to acreage have contributed to a fairly small percentage of the two-step forecast error in spot prices.

model for scenarios 1 and 2. In particular, we show the IRF graphs for the effect of negative corn (and soybean) acreage shocks to own price and the effect of negative non-crop farmland shocks to own corn and soybean prices.[12] For every 1 million acres of corn production removed, corn price increases in the first period by \$0.04 per bushel.[13] The effect lasts one additional period, and then falls back to zero. For every 1 million acres of soy production removed, soy price increases in the first period by \$0.23 per bushel. The effect is more persistent than it was for corn, although it does fall back towards zero. We expect the effect to be higher for soybeans than for corn, as US production of soybeans has historically had a much larger share of world production than has US corn.

For a negative one million acre shock to non-crop farmland (holding corn acreage constant), corn price falls. The initial effect is \$0.02 per bushel, peaking at \$0.05 per bushel. Eventually the effect returns to zero. Holding soybean acreage constant, the negative shock to non-crop farmland leads to a soybean price decrease of \$0.08 per bushel. What appears to be happening for both crops is that, holding own acreage constant, acreage of other crops is rising. Since grains are largely substitutable, this takes pressure off of the demand for the crop. Since own acreage was held constant, supply doesn't change, and the fall in demand lowers prices. This intuition could correspond historically to increases in, for instance, other food crops like wheat. There is no reason to expect the story not to hold for other biofuels feedstocks, *e.g.*, miscanthus or switchgrass.

Additionally, we compute cumulative impulse response functions (CIRFs). While the structural IRFs described above show the dynamic path of prices following a one-time acreage shock, a more likely biofuels scenario has a continual ramp-up of production. Figure 2 shows the CIRFs for the effect of repeated negative acreage shocks on crop prices. For repeated negative one million corn acreage shocks, corn prices initially rise \$0.04 per bushel, but then continue to rise, peaking at \$0.10 per bushel higher than they would have been, absent any acreage shocks. For repeated negative one million shocks to non-crop farmland, corn prices initially fall \$0.02 per bushel, but then continue to fall. For scenarios 1 and 2 with soybeans, the cumulative effect is also much larger than the one-time shock effect. Two caveats apply to the CIRF results. First, they inevitably have very large standard errors and the results after the first few periods must be interpreted with caution. Second, CIRFs won't capture changes in expectations. That is, after many periods a repeated "shock" could be incorporated into market expectations, altering the underlying data-generating process and no longer constituting a true "shock."

[12] Each SVAR has nine equations and therefore 81 estimated IRFs. These results are available upon request.

[13] All results reported are normalized to shocks of 1 million acres. The IRF as described in the methodology section yields estimates of a change in price following a shock of size ε in the acreage equation. This response can be rescaled by the corresponding element of matrix B to give a change in price following a shock of size 1 million acres.

6 Robustness Checks

As described in the modeling section, a number of robustness checks are considered. For instance, a far sparser SVAR (with only five equations) is considered for both crops and both scenarios. Results are quite similar (figure 3). A negative one million acre shock to corn production raises prices by \$0.07. For soybeans, the initial price increase is \$0.23 per bushel for a one million soy acre shock. Decreasing non-crop farmland (holding own acreage constant) lowers corn prices by \$0.06 and soy prices by \$0.11. Next, a log/log specification is estimated for both the main and the sparse models. For scenario 1 with corn, the results are quite similar. For scenario 2, the shape of the IRF is similar but shifted up. For scenario 1 with soy, the results are considerably dampened; however for scenario 2, the results are similar to the those in the linear specification.

[Figure 3: Robustness Checks]

The SVAR is also estimated with additional lags allowed in the system (figure 3). For scenario 1 with corn, the estimated IRF is quite similar for two lags. With three lags the initial effect is similar but the effect in later periods is unstable (and implausible). For scenario 1 with soy, the general effect is similar but (implausible) oscillations appear for both two and three lags. For scenario 2, the effect on corn prices is not robust to the inclusion of additional lags. The effect on soy prices is robust to two lags but not three. However the BIC-selected model is one lag for both scenarios and both crops. Moreover there is no theoretical reason to expect additional lags to be relevant.

Next the SVAR is estimated with different diffusion indexes (figure 3). The above results were for the inclusion of the primary diffusion indexes, which loaded mainly onto input prices (supply index) and fairly evenly across US GNP and southern hemisphere agricultural production. The model is also estimated with the secondary diffusion indexes, which loaded primarily onto loan prices (supply index) and Argentine corn and soy production. The results are very similar to those in the main specification (figure 3).

One robustness check uses an additional variable in each specification, to allow for crop rotation. Soy acreage is thus included in the corn specifications, and corn acreage in the soy specifications. Results are nearly identical (table 3). This is not surprising, since the determinants of soy acreage were already included in the corn equations and the determinants of corn acreage in the soy equations.

As described previously, a few other robustness checks use slightly different data. The main specification is estimated with yields per harvested (rather than planted) acres. Scenario 2 is estimated with a positive shock to total cropland, rather than a negative shock to non-crop farmland. Finally, data on principal crops (rather than total crops) is used in scenario 2, to allow for the inclusion of the year 2007. For all three checks, results are fairly similar to the main specification.

Finally, the SVAR is estimated with a varying time period. The same SVAR is estimated ten times, with 41 years included in each estimation (i.e., 1958 to 1998, 1959 to 1999, etc.). As can be seen from the estimated IRFs (figure 4), the results are quite robust to varying the time frame. For the two corn specifications, the estimation with smaller windows actually shows a larger response. This appears to be because the larger window includes 2007, an anomalous year, which is not included in the smaller time frames. Thus the inclusion of 2007, in which a number of supply and demand shocks hit commodity markets, may bias the results towards zero. Our scenario 1 results, which include 2007, could accordingly be interpreted as a lower bound.

[Figure 4: Rolling Time Frame]

7 Conclusion

Using a dynamic system of simultaneous equations, we explore the impacts on crop prices of changes in land use. We develop a structural vector autoregression model, allowing us to analyze impulse response functions and forecast error variance decompositions. These econometric tools, common to macroeconomic applications, provide elegant descriptions of the dynamics of the agricultural production process. We find significant and sustained increases in corn and soybean prices when crop acreage is removed. As described above, this is equivalent to a shock to the residual supply for food. For a reduction in corn area of one million acres, we estimate a corn price increase of $0.04 per bushel. The last year of our sample (which ends in 2007) saw by far the largest increase in corn acreage dedicated to ethanol - approximately 6 million acres. Real corn prices for this year paid to farmers rose by 88 cents. For 2007, our model therefore predicts that 27% percent of this price increase can be accounted for by the increased area planted to corn for ethanol. Extending our sample further is not currently feasible, yet we can provide a simple calculation for the 2007/08 year. USDA estimates indicate that area planted to corn for ethanol rose by 4.5 million acres and the price paid to farmers rose by $1.27. Our coefficient estimates based on data up to 2007 would indicate that 14% of this price increase can be attributed to ethanol. If we take a longer run perspective and average the shares over the years 2001-2008, our estimates explain 16% of the year to year fluctuation in prices due to changes in corn acreage planted to ethanol.

For a negative shock in soybean acreage, we find a price increase of $0.23 per bushel. This much larger magnitude is partly explained by the fact that US soybean acreage is a much larger share of world production than is US corn acreage. We also find significant and sustained decreases in crop prices when own acreage is held constant and total cropland is increased. A 1 million acre increase in US cropland leads to an approximately $0.06 to $0.11 decrease in corn and soybean prices. What sets our model apart from most is that the results extend to the production of biofuels besides corn ethanol. Any biofuel feedstock that is grown on land previously dedicated to corn will increase corn prices; this is crucial as the US considers the production of second- and third-generation biofuels.

A number of caveats should be mentioned. First, the magnitudes we see depend on the US share of world production. If this were to change substantially, we might expect a different multiplier. Second, the pathways for the observed responses are only hypothesized. For a scenario in which crop acreage is removed, it is intuitive that the crop's price will rise. Supply is constrained by the removal of acreage, a crucial input in the production process, and demand for the food crop has not changed. The economic rationale behind the second scenario is as follows: redirecting production of biofuels from food crops to second-generation crops will shift the demand for corn inwards, resulting in a drop in corn prices.

Our findings open up a number of possible extensions and future research projects. Scientists and policymakers have expressed hope that new biofuels feedstocks will be grown on land that does not compete with food crops, thus avoiding the effects of biofuels on food prices. Our evidence is suggestive that, if these additional crop lands released pressure from the corn market, corn prices could indeed decline. Verifying this, by analyzing the causal pathways at work, will be crucial as biofuels policies move forward.

References

[1] Abbott, P. C., C. Hurt, and W. E. Tyner. 2008. "What's Driving Food Prices?" Farm Foundation Issue Report.

[2] Ando, A. W., M. Khanna, and F. Taheripour. 2010. "Market and Social Welfare Effects of the Renewable Fuels Standard." *Handbook of Bioenergy Economics and Policy, Natural Resource Management and Policy* 33(4): 233-250.

[3] Banse, M., H. van Meijl, A. Tabeau, and G. Wolftjer. 2008. "Will EU Biofuel Policies Affect Global Agricultural Markets?" *European Review of Agricultural Economics*, 35(2): 117-141.

[4] BBC. 2007. "Will biofuel leave the poor hungry?" http://news.bbc.co.uk/2/hi/business/7026105.stm.

[5] Bureau of Economic Analysis. www.bea.gov.

[6] Chakravorty, U., M. Hubert, M. Moreaux, and L. Nostbakken. 2011. "Will Biofuel Mandates Raise Food Prices?" University of Albert, Department of Economics Working Paper No. 2011-01.

[7] Chakravorty, U., M. Hubert, and L. Nostbakken. 2009. "Fuel Versus Food." *Annual Review of Resource Economics* 1: 645-663.

[8] Christiano, L. J., Eichenbaum, M., and Evans, C. L. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113: 1-45.

[9] Commodity Research Bureau. 1956-2003. *Commodity Yearbook.*

[10] Congressional Budget Office. 2009. "The Impact of Ethanol Use on Food Prices and Greenhouse-Gas Emissions." Congress of the United States Congressional Budget Office Pub. No. 3155.

[11] Economic Research Service. www.ers.usda.gov.

[12] Food and Agricultural Organization of the United Nations. www.faostat.fao.org.

[13] Headey, D. and S. Fan. 2008. "Anatomy of a Crisis: The Causes and Consequences of Surging Food Prices." *Agricultural Economics* 39 Issue Supplement s1: 375-391.

[14] Hill, J., E. Nelson, D. Tilman, S. Polasky, and D. Tiffany. 2006. "Environmental, Economic, and Energetic Costs and Benefits of Biodiesel and Ethanol Biofuels." Proceedings of the National Acaddemy of Sciences 103(30): 11206−11210.

[15] Khanna, M., A. W. Ando, and F. Taheripour. 2008. "Welfare Effects and Unintended Consequences of Ethanol Subsidies." *Review of Agricultural Economics* 30(3): 411-421.

[16] Lutkepohl, H. 1993. *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

[17] Mitchell, D. 2008. "A Note on Rising Food Prices." World Bank Policy Research Working Paper No. 4682.

[18] Mushtaq, K. and Dawson, P. J. 2002. "Acreage Response in Pakistan: A Co-Integration Approach." *Agricultural Economics* 27: 111-121.

[19] National Agricultural Statistics Service. www.nass.usda.gov.

[20] Naylor, R. L., A. J. Liska, M. B. Burke, W. P. Falcon, J. C. Gaskell, S. D. Rozell, and K. G. Cassman. 2007. "The Ripple Effect: Biofuels, Food Security, and the Environment." *Environment* 49(9): 30-43.

[21] Nerlove, M. 1956. "Estimates of the Elasticities of Supply of Selected Agricultural Commodities." *Journal of Farm Economics* 38: 496-509.

[22] Rajagopal, D., Sexton, S. E., Roland-Holst, D., and Zilberman, D. 2007. "Challenge of Biofuel: Filling the Tank Without Emptying the Stomach?" *Environmental Research Letters* 2: 1-9.

[23] Roberts, M. J. and Schlenker, W. 2010. "Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate." NBER Working Paper No. 15921.

[24] Robertson, G. P., Dale, V. H., Doering, O. C., Hamburg, S. P., Melillo, J. M., Wander, M. M., Parton, W. J., Adler, P. R., Barney, J. N., Cruse, R. M., Duke, C. S., Fearnside, P. M., Follett, R. F., Gibbs, H. K., Goldemberg, J., Mladenoff, D. J., Ojima, D., Palmer, M. W., Sharpley, A., Wallace, L, Weathers, K. C., Wiens, J. A., and Wilhelm, W. W. 2008. "Sustainable Biofuels Redux." *Science* 322: 49-50.

[25] Rosegrant, M. W. 2008. "Biofuels and Grain Prices: Impacts and Policy Response." Testimony for the U.S. Senate Committee on Homeland Security and Governmental Affairs, May 7, 2008.

[26] Runge, C. F. and B. Senauer. 2007. "How Biofuels Could Starve the Poor." *Foreign Affaris* 86(3): 41-53.

[27] Sexton, S., Rajagopal, D., Zilberman, D., and Hochman, G. 2008. "Food Versus Fuel: How Biofuels Make Food More Costly and Gasoline Cheaper." *Agricultural and Resource Economics Update* 12: 1-6.

[28] Stock, J. H. and Watson, M. W. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97: 1167-1179.