What Is the Cost of a Renewable Energy–Based Approach to Greenhouse Gas Mitigation?
Anthony Oliver and Madhu Khanna
Land Economics 93:3, August 2017

Appendix A: Description of the Model

1. The Biofuel and Environmental Policy Analysis Model with Electricity Sector (BEPAM-E)

BEPAM-E is a nonlinear, dynamic, multi-sector, price-endogenous, open-economy, partial equilibrium, mathematical programming model that simulates U.S. agricultural, transportation fuel, and electric power sectors including international trade with the rest of the world (ROW). Market equilibrium is found by maximizing the sum of consumers' and producers' surpluses in the agricultural, transportation fuel, and electric power sectors subject to various material balance constraints and technological constraints underlying commodity production and consumption in a dynamic framework. BEPAM considers regional supply of crop and biofuel feedstocks at Crop Reporting District (CRD) level where crop production costs, yields, and resource endowments are specified for each CRD and each crop. The CRD is also used as a spatial unit to model electricity generation from existing power plants by fuel type, while generation from new electricity capacity is considered at Electricity Market Regions (EMR).[1] The model endogenously determines the agriculture and transportation sector variables of food consumption, gasoline, diesel, and biofuel consumption, imports of gasoline and sugarcane ethanol, mix of biofuels, and regional land allocation among different food, feed and fuel crops and livestock activities over a given time horizon. It also endogenously determine electricity sector variables such as generation by energy type (coal, natural gas, oil, wind, co-fired biomass, dedicated biomass, and co-product), regional electricity consumption, inter-regional electricity transmission, and GHG emissions.

A brief overview and major components of the model are described below, followed by a mathematical representation of the model objective function and individual constraints.

1.1 Description of Spatial Units

The sector model considers the 48 contiguous United Sates where production and consumption of the outputs and inputs at different levels of spatial aggregation as is described here. Crop Reporting District, State, and Electricity Market Region. There are 306 CRDs included in the model. These CRDs are designated by the USDA and are made up of a number of counties that are contiguous and within the same state. The 48 contiguous United States and the District of Columbia included. There are 20 electricity market regions that are aggregation of group of adjacent states approximating the National Energy Reliability Corporation (NERC) designations.[2]

1.2 Transportation Fuel Sector

The demand for transportation fuels is driven by the demand for vehicle kilometers

[1] Exisiting electricity generation capacity is aggregated to the CRD by taking the sum of capacity and the generation-weighted average heat-rate of all power plants located in a given CRD for each type of energy source. The definition of EMRs is described in the spatial units section.
[2] The EMRs used here do not directly correspond to NERC regions, as they are all based on state boundaries.

traveled (VKT) that are produced by blending liquid fossil fuels and biofuels. BEPAM includes linear demand curves for VKT with four types of vehicles, including conventional gasoline (CV), flex fuel (FFV), gasoline-hybrid (HV), and diesel vehicles (DV). We also include the demand for kilometers traveled with plug-in and battery electric vehicles (EV), but in the current version of the model its consumption is determined exogenously. In producing VKT, we consider the difference in the energy content of alternative fuels, fuel economy of each type of vehicle, and technological limits that liquid fossil fuels and biofuels can be blended in particular types of vehicles. Therefore, in addition to a minimum ethanol blend for gasoline blend vehicles to meet the oxygenate additive requirement, we impose the blend limits as specified by Energy Information Administration (EIA 2010a)for each of these four types of vehicles due to their technological constraints in blending biofuels with fossil fuels. We exogenously shift demand curves for VKT with each type of vehicle over time as projected by the Annual Energy Outlook (EIA 2010b) to capture the demand due to the growth in income and population.

We include linear supply curves for domestic gasoline and for gasoline supply from the ROW. The excess supply of gasoline to the U.S. is determined by the difference between gasoline demand and supply in the ROW. We also include a linear supply curve for diesel that is assumed to be produced domestically.

The biofuel sector includes several first and second generation biofuels. First generation biofuels include domestically produced corn ethanol and imported sugarcane ethanol, biodiesel produced from soybean oil, DDGS-derived corn oil and waste grease. As second generation biofuels, we include cellulosic ethanol and biodiesel derived from cellulosic biomass such as crop or forest residues and dedicated energy crops (miscanthus and switchgrass). Biomass can be converted to either cellulosic ethanol blended with gasoline or biomass-to-liquids (BTL) blended with diesel using the Fischer-Tropsch process. While the feedstock costs of producing biofuels are determined endogenously in the agricultural sector, processing costs of biofuels are assumed to decline with cumulative production using the experience curve approach (de Wit et al. 2010).

1.3 Agricultural Sector

The agricultural sector component includes fifteen conventional crops, eight livestock products, two perennial bioenergy crops, crop residues from the production of corn and wheat, forest residues, and co-products from the production of corn ethanol and soybean oil. In the crop and livestock markets, primary crop and livestock commodities are consumed either domestically or traded with the ROW (exported or imported). Primary crop commodities can also be processed or directly fed to various animal categories. Domestic and export demands and import supplies are incorporated by assuming linear price-responsive demand/supply functions. The commodity demand functions and export demand functions for tradable row crops and processed commodities are shifted upward over time at exogenously specified rates.

The model incorporates spatial heterogeneity in crop and livestock production, where the production costs and yields of individual crop/livestock activities and resource endowments are specified differently for each region based on crop/livestock budgets reported by various agricultural experiment stations and the National Agricultural Statistics Service (NASS) database. Crops can be produced using alternative tillage, rotation, and irrigation practices. As spatial decision units we consider 295 CRDs in 41 U.S. states in five major regions.[3] Crop yields increase over time at exogenously given rates based on econometrically estimated trends and price responsiveness of crop yields in the U.S. (Chen et al. 2011). Following (Hertel, Tyner, and Birur 2010), we assume that marginal lands have a crop productivity that is two thirds of the average cropland productivity. Yields of the bioenergy crops are assumed to be the same on marginal lands and regular croplands, but they vary regionally. In the absence of observed yield data for energy crops in most regions, we use the simulated biomass yields obtained from MISCANMOD (Jain et al. 2010).

The model includes five types of land in each CRD, namely cropland, idle land, cropland pasture, pasture land and forestland pasture. Cropland availability in each CRD is assumed to change in response to crop prices. Idle land and cropland pasture are assumed to move in and out of cropland and idle state. They can also be converted to the production of energy crops at a conversion cost assumed to be equal to the returns from producing the least profitable annual crop. We fix pasture land and forestland pasture at 2007 levels, while the land enrolled in the Conservation Reserve Program (CRP) is kept at levels authorized by the Farm Bill of 2008. Due to concerns about the impact of monocultures of perennial grasses on biodiversity or sub-surface water flows, we impose a 25% limit on the amount of land in a CRD that can be converted to perennial grasses.

In the livestock sector, we consider chicken, turkey, lamb, beef, pork, wool, dairy and eggs. Livestock production requires nutrition in terms of protein and calories that are provided by feed crops and byproducts of crop processing, such as soymeal and Distiller's Dried Grains with Solubles (DDGS, a byproduct of corn ethanol). The demand for feed from the production of livestock products is then determined by the number of livestock animals and their nutrition requirements through the least cost feed mix approach. We impose upper bounds for the share of DDGS in total feed consumption for each livestock category as in (Babcock, Barr, and Carriquiry 2010). We model the supply of chicken, turkey, lamb, pork, wool, diary and eggs at the national level, and restrict their supply to be a weighted average of the historical supplies obtained from the NASS database. The beef and dairy cattle activities are modeled at CRD level because they require grazing land and compete for land with crop production in each region. Thus, the supply of beef is restricted by the number of cattle which in turn depends on the amount of grazing land available at regional level.

[3] Western region includes Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington and Wyoming; Plains includes Nebraska, North Dakota, Oklahoma, South Dakota, Texas and Kansas; Midwest includes Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio and Wisconsin; South includes Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi and South Carolina; Atlantic includes Kentucky, Maryland, New Jersey, New York, North Carolina, Pennsylvania, Tennessee, Virginia, and West Virginia

1.4 Electricity Sector

The electricity sector component of the model considers electricity producers and consumers in the 48 contiguous states on an annual basis. The electricity sector is modeled regionally, where existing power plants by fuel type are modeled at the CRD-level, and new electricity generation and electricity demand are considered at the EMR-level. Electricity demand functions for each EMR are found by aggregating state level, sector specific (residential, commercial, and industrial) demand functions. Electricity is supplied from existing capacity and from new capacity that is endogenously determined.

Existing power plants are aggregated to the CRD level by energy type which consists of coal, natural gas, oil, hydroelectric, nuclear, geothermal, waste, solar, biomass, wind, and other. This power plant capacity is regionally heterogenous in the energy type, nameplate capacity, and conversion efficiency. The decision can also be made to expand generation capacity that uses natural gas, wind, or biomass energy. The expansion of biomass capacity can be accomplished by co-firing biomass at an existing coal power plant, generating it at a new dedicated bio-power plant, or generating net co-product electricity from cellulosic ethanol refining. Biomass must be transported from where it is produced to a coal power plant to be co-fired and incurs a transportation cost for this based on the distance between production and use. Wind capacity can be expanded at the EMR-level and its cost depend upward sloping wind supply curves that are based on the quantity of wind resources available at a given level of cost.

The price of Fossil fuels which are supplied to power plants are based on supply curves. Natural gas supply is specified at the national level, which gives a national wholesale price for natural gas based on the quantity consumed.[4] Power plants which consume natural gas must pay EMR-specific transmission and distribution cost in addition to the wholesale price. Coal and fuel oil prices are specified at the state level and are fixed on an annual basis.

Electricity generated at a power plant must be transmitted to end-use consumers through transmission infrastructure. All electricity generation incurs a transmission cost and percentage loss before it can be consumed. The model also allows for electricity generated in a given electricity market region (EMR) to be transmitted to an adjacent market region to meet demand in that region subject to a transmission capacity constraint.

1.5 Mathematical Representation

We describe the algebraic model below. For convenience, we use lower case symbols to denote the exogenous parameters and upper case symbols to represent endogenously determined variables. We first describe the objective function then the constraints; the description of constraints follows Chen et al. (2014).

[4] Natural gas demand from sectors other than electricity increases exogenously according to AEO projections (EIA 2012); and is held constant across scenarios.

1.5.1 **Objective function**

The objective function is the sum of discounted consumers' and producers' surpluses obtained from production, consumption and international trade in the agricultural, transportation fuel, and electric power sectors over a planning horizon of T and the terminal values of standing perennial grasses in year T. The algebraic expression is given in (1):

$$\begin{aligned}
Max: \sum_{0}^{T} e^{-rt} \Bigg\{ & \sum_{z} \left[\int_{0}^{DEM_{t,z}} f^{z}(.)d(.) + \int_{0}^{EXP_{t,z}} f^{z}(.)d(.) - \int_{0}^{IMP_{t,z}} f^{z}(.)d(.) \right] \\
& + \int_{0}^{MIL_{t,v}} f^{v}(.)d(.) + \int_{0}^{MDSL_{t}} f^{mdsl}(.)d(.) + \int_{0}^{G_{t,row}} f^{g}(.)d(.) \\
& + \sum_{er} \int_{0}^{QE_{er,t}} f^{er}\left(QE_{er,t}\right)d(.) - \sum_{r,et} ec_{et} QR_{et,r,t} - \sum_{s,nt} lc_{nt} QM_{er,nt,t} \\
& - \int_{0}^{NG_{t}} g(NG_{t})d(.) - \sum_{er,age} np_{er,t} NGM_{er,age,t} \\
& - \sum_{ff,r(er)} fp_{ff,er,t} FF_{ff,r,t} - \sum_{er} \int_{0}^{QW_{er,t}} h^{er}(.)d(.) - pc * BE_{t} \\
& - cf \sum_{r} CF_{r,t} - \sum_{r_1,r_2} tr_{r_1,r_2} BT_{r_1,r_2,t} - t_e GHG_t - \sum_{r,q} rc_{r,q} ACR_{t,r,q} \\
& - \sum_{r,p} pc_{r,p} ACR_{t,r,p} - \sum_{r,q} rs_{r,q} ACR_{t,r,q} - \sum_{r,p} cc_{r} \Delta ACR_{t,r,p} \\
& - \sum_{fr,r} \int_{0}^{FR_{fr,r,t}} f^{fr,r}(.)d(.) - \sum_{k} lc_{k} LIV_{t,k} - \sum_{i} sc_{i} PRO_{t,i} \\
& - \int_{0}^{GAS_{t,O}} f^{O}(.)d(.) - \int_{0}^{DSL_{t}} f^{dsl}(.)d(.) - ec_{c} ETH_{t,c} - ex_{b} ETH_{t,b} \\
& - pc_{b} BTL_{t} - \sum_{j} pc_{j} OILDSL_{t,j} \Bigg\} + e^{-rT} \sum_{r,p} \left(v_{r,p} - w_{r}\right) ACR_{T,r,p}
\end{aligned} \tag{1}$$

The first integral term in the first line of (1) represents the areas under the demand functions from which consumers' surplus is derived. Each integral is associated with a crop, livestock, or processed commodity for which a domestic market demand is considered: $DEM_{t,z}$ denotes the endogenous domestic demand variable in year t; z =(i,j,k) denotes the index set for crop commodities (i), processed products from crops (j), and livestock commodities (k); $f^{z}(.)$ denotes the inverse demand function for the associated commodity; and d(.) denotes the integration variable. The next two integral terms account for the areas under the inverse demand functions for exports, $EXP_{t,z}$ and the areas under the import supply functions, $IMP_{t,z}$ (such as sugar and sugarcane ethanol).

The first integral term in the second line of (1) represents the area under the inverse demand function for kilometers traveled with alternative vehicle choices (v) (denoted by $MIL_{t,v}$). The second integral term denotes the area under the inverse demand function for VKT with diesel vehicles (denoted by $MDSL_t$). The third integral term in the same line represents the area under the inverse demand function for gasoline consumed by the ROW (denoted by $G_{t,row}$). The demand functions for crop and livestock products, kilometers traveled and ROW gasoline consumption are all characterized by linear demand functions in the current version, but other functional forms, such as constant elasticity demand functions, can be incorporated without difficulty.

The first integral term in the third line of (1) pertains to the electricity sector and represents the area under the inverse demand curves for electricity. The choice variable $QE_{er,t}$ is the quantity of electricity consumed in EMR *er* in year *t* from all energy sources. $f^{er}(.)$ is the inverse demand function for electricity in *er*. The second term in this line is a summation that represents the variable operations and maintenance (O&M) cost for electricity generation from existing power plants and co-firing. The parameter ec_{et} is the O&M cost for electricity generated from technology type *et* and the variable $QR_{et,r,t}$ represents the electricity generated from technology type *et* in CRD *r* and year *t*. The third term in this line represents levelized cost of electricity generation from new power plants. The parameter lc_{nt} is the levelized cost of generating electricity from a new power plant with technology type *nt* (*nt* is a subset of *et*) and the choice variable $QM_{er,nt,t}$ is the quantity of electricity generated in EMR *er* from a new power plant with technology type *nt* in year *t*.

The first integral term in the fourth line of (1) represents the area under the national supply curve for natural gas. The endogenous variable NG_t is the quantity of natural gas used to produce electricity for all EMRs in year *t* and the function $g(.)$ represents the supply curve for natural gas. The second term in this line represents the regional transmission and distribution costs of natural gas for electric power region EMR. The parameter $np_{er,t}$ is the transmission and distribution cost per unit of natural gas by EMR er in year t and the endogenous variable $NGM_{er,age,t}$ is the quantity of natural gas used in EMR er for power plants with type $age \in \{exisitng, new\}$ in year *t*.

The first summation term in the fifth line of (1) represents the fuel cost of power plants that consume coal or fuel oil to generate electricity. The parameter $fp_{ff,er,t}$ denotes the fuel cost for fuel type $ff \in \{coal, fuel\ oil\}$, in EMR *er*, year *t* and the variable $FF_{ff,r,t}$ is the quantity of fuel type *ff* used for electricity generation in CRD *r*, year *t*. The second term in the line represents the area under the supply curve for electricity generated from new wind capacity. The choice variable $QW_{er,t}$ is the quantity of electricity generated from new wind capacity in EMR *er* in year *t*. The functions $h^{er}(.)$ are the supply curves for new wind generation in EMR *er*. The last term in this line represents the processing costs of cellulosic biomass for power generation at either a dedicated biomass power plant or co-fired at a coal power plant.[5] The parameter *pc* is the

[5] This cost does not apply to electricity generated as a co-product at cellulosic ethanol refineries.

unit processing cost and the variable BE_t is the quantity of biomass used to generate electricity in year *t*.

The first term in the sixth line of (1) represents the cost of retrofitting a coal power plant for co-firing. The parameter *cf* is the cost of retrofitting a MW of coal plant capacity for co-firing and the variable $CF_{r,t}$ is the amount of capacity converted for co-firing in region *r*. The second term in this line represents the transportation cost of biomass for electricity generation. The parameter tr_{r_1,r_2} is the per-metric ton cost of transporting biomass from CRD r_1 to CRD r_2 and the variable $BT_{r_1,r_2,t}$ is the quantity of biomass transported from CRD r_1 to CRD r_2 in year t. The third term in this line represents the externality costs due to greenhouse gas emissions (GHG_t) with an exogenously fixed carbon tax of t_e.

The last term in the sixth line and the seventh line of (1) include the production costs of row crops, perennial crops, and collecting crop and forest residues for biofuel and electricity production. The land allocated to row crops and perennial crops (acreage) in region *r* and year *t* is denoted by $ACR_{t,r,q}$ and $ACR_{t,r,p}$, respectively. The production of row crops may use one of various production practices, denoted by *q,* which differ by crop rotation, tillage, and irrigation. Parameters denoting the production costs of row and perennial crops are $rc_{r,q}$ and $pc_{r,p}$. Leontief production functions are assumed for both row crops and perennial crops production. The third term in the seventh line represents the cost of collecting crop residues, such as corn stover and wheat straw, and we use $rs_{r,q}$ to denote the collection costs of crop residues. The last term in the seventh line denotes the costs of converting marginal lands (such as idle land and crop pasture land) for producing perennial crops, which include costs of land clearing, wind rowing and any necessary activities for seedbed preparation. The amount of marginal lands converted for perennial grasses are denoted by $\Delta ACR_{t,r,p}$ and cc_r represents the conversion cost per unit of marginal land.

The first term in the eighth line of (1) represents the area underneath the regional-specific supply functions of forest residues and pulpwood ($FR_{ft,r,t}$), where *ft* is a set consisting of forest residues and pulpwood. The second term of the eighth line includes the costs associated with livestock activities, where lc_k denotes the cost per unit of livestock category *k* (again employing Leontief production functions), which is assumed to be the same across all regions. Variable $LIV_{t,k}$ represents the quantity of livestock *k* processed in year *t*. The last term represents the total cost of converting primary crops (corn, soybeans, sugarcane and sugar beet) to secondary (processed) commodities (oils, soymeal, refined sugar, and HFCS). The amount of processed primary crop *i* in year *t* is denoted by $PRO_{t,i}$, and sc_i denotes the processing cost per unit of *i*.

All the terms in the ninth line and the first two terms of the tenth line of (1) consists of costs accruing to the transportation fuel sector. The first integral represents the area under the supply functions for gasoline from domestic producers and the ROW, $GAS_{t,O}$ (where index *o* denotes the source of gasoline supply), whose consumption and price are to be determined endogenously. The second integral represents the area under the supply functions for petroleum diesel from domestic producers, DSL_t. The next two terms represent the processing costs of corn and cellulosic ethanol in the refinery, namely $ETH_{t,c}$ and $ETH_{t,b}$, followed by two terms representing the processing costs of biodiesel produced from biomass (denoted by BTL_t) and vegetable oils (denoted by $OILDSL_{t,j}$). Finally, the last term in (1) reflects the value of the remaining economic life of standing perennial grasses beyond the planning period *T*, denoted by $v_{r,p}$, net of the return from the most profitable cropping alternative in region *r*, denoted by w_r. The latter is used to account for the opportunity cost of land.

Gasoline kilometers ($MIL_{t,v}$) are produced by blending gasoline and ethanol, which are assumed to be perfect substitutes in generating kilometers travelled with per liter of ethanol providing two-thirds energy of per liter of gasoline. The amount of kilometers generated for each type of gasoline-based vehicles by the use of gasoline and ethanol is formulated in equation (2) below, where $E_{t,v}$ and $G_{t,v}$ represent the consumption of ethanol and gasoline for each type of vehicle and $\gamma_{t,v}$ is the fuel efficiency (kilometers per liter):

$$MIL_{t,v} = \gamma_{t,v}\left[\frac{2}{3}E_{t,v} + G_{t,v}\right] \;\; \forall t,v \tag{2}$$

We impose a minimum ethanol blending rate for gasoline vehicles in order to meet the oxygenate additive requirement, and also impose an upper limit for blending ethanol with gasoline for each type of vehicle. The latter is based on the technological characteristics of each vehicle type. Equation (3) below shows the gasoline demand and supply balance for the U.S. and the ROW markets:

$$\sum_v G_{t,v} + G_{t,row} = \sum_O GAS_{t,O} \;\; \forall t. \tag{3}$$

Equation (4) below establishes the balance for the ethanol demand and supply.

$$\sum_v E_{t,v} = ETH_{t,c} + ETH_{t,b} + IMP_{t,sugarcane\ ethanol} \;\; \forall t \tag{4}$$

Similarly, we derive the production function of diesel kilometers that are produced by blending petroleum diesel and biodiesel. We incorporate the difference in energy contents between biodiesel produced from alternative sources and petroleum diesel, and use $\kappa_j < 1$ to represent this relationship. The amount of diesel kilometers generated is formulated in equation (5) below:

$$MDSL_t = r_t\left[\kappa_b BTL_t + \sum_j \kappa_j OILDSL_{t,j} + DSL_t\right] \;\; \forall t \tag{5}$$

The regional material balance equations link the production and usage of primary crops, as shown in constraint (6) for primary crop product *i* produced and marketed by region *r*:

$$MKT_{t,r,i} + \{CE_{t,r}\}_{i=corn} \le \sum_q y_{r,q,i} ACR_{t,r,q} \;\; \forall t,r,i \tag{6}$$

where $MKT_{t,r,i}$ denotes the marketed amount of primary crop product *i* and $y_{r,q,i}$ is the yield of product *i* per unit of the land allocated to crop production activity *q* in region *r*. For corn, $MKT_{t,r,i}$ includes non-ethanol uses and $CE_{t,r}$ is the amount of corn converted to ethanol producers (which appears only in the balance constraint for corn).

The amount of primary crop commodity *i* available in the market (excluding the corn used for ethanol) comes from the domestic supply of that commodity from all regions ($MKT_{t,r,i}$). This total amount is either consumed domestically ($DEM_{t,i}$), exported ($EXP_{t,i}$), processed to secondary commodities ($PRO_{t,i}$), or used for livestock feed ($FED_{t,i}$). This is expressed in constraint (7) below:

$$DEM_{t,i} + PRO_{t,i} + FED_{t,i} + EXP_{t,i} \leq \sum_{r} MKT_{t,r,i} \;\; \forall t, i \tag{7}$$

Similar to (7), a balance equation is specified for each processed commodity. Like primary commodities, processed commodities can also be consumed domestically, exported, or fed to animals, as shown in constraint (8) below:

$$\begin{aligned} &DEM_{t,j} + FED_{t,j} + EXP_{t,j} + \frac{OILDSL_{t,j}}{a_j} \\ &\leq \sum_{i} v_{i,j} PRO_{t,i} + \left\{ \sum_{r} v_{i,j} CE_{t,r} \right\}_{j=ddg, i=corn} \qquad \forall t, j \end{aligned} \tag{8}$$

where a_j is the conversion rate of vegetable oil *j* to biodiesel and $v_{i,j}$ denotes the conversion rate of primary product *i* to processed product *j*. A particularly important component of the model that links the crop and fuel sectors sector is the conversion of corn and cellulosic biomass to ethanol. During the conversion of corn to ethanol, a secondary commodity *DDGS*, is produced as a byproduct. The amount of DDGS produced is proportional to the amount of corn used for ethanol, $CE_{t,r}$ through a fixed conversion rate $v_{\text{corn,ddg}}$, and it can either be fed to livestock as a substitute for corn or exported.

The relations between biofuels production and crop production activities are expressed below:

$$ETH_{t,c} = \alpha \sum_{r} CE_{t,r} \;\; \forall t \tag{9}$$

$$ETH_{t,b} = \beta_e \sum_{r} BM_{t,r,eth} \;\; \forall t \tag{10}$$

$$BTL_t = \beta_{dsl} \sum_{r} BM_{t,r,dsl} \;\; \forall t \tag{11}$$

where α, β_e and β_{dsl} denote the amounts of biofuels produced from per unit of corn or cellulosic feedstocks. $BM_{t,r,eth}$ and $BM_{t,r,dsl}$ represent the quantity of biomass used for cellulosic ethanol and BTL production in region *r* and year *t*, respectively.

Biomass can be utilized for the production of liquid fuels and electricity. The amount of biomass used for ethanol production, biodiesel production, and electricity generation at each CRD cannot exceed the total biomass production at that CRD, which is expressed in the accounting inequality below:

$$\begin{aligned} &BM_{t,r,eth} + BM_{t,r,dsl} + BM_{t,r,elec} \\ &\leq \sum_{p} by_{r,p} AC_{t,r,p} + \sum_{q} ry_{r,q} AC_{t,r,q} + \sum_{fr} FR_{fr,r,t} \;\; \forall t, r \end{aligned} \tag{12}$$

where $by_{r,p}$ and $ry_{r,q}$ are the biomass and crop residue yields in region *r* for respective perennial and crop production activities. The endogenous variable $BM_{t,r,elec}$ is the quantity of biomass utilized for electricity generation in region *r* and year *t*.

Land is the only primary production factor considered in the model. In each region, the total amount of land used for all agricultural production activities cannot exceed the available land ($al_{t,r}$), which is specified separately for irrigated and non-irrigated land categories. Due to the steady increase in biofuels consumption assumed in the model, the demand for agricultural land is expected to increase through the conversion of marginal lands (not currently utilized) to cropland. The extent of conversion is assumed to depend on variations in crop prices over time. Therefore, in the model we determine the agricultural land supply 'endogenously'. Specifically, for a given year *t* in the planning horizon *T*, we solve the model assuming a fixed regional land availability for each year of the 10-year production planning period ahead considered in that run. From the resulting multi-year equilibrium solution, we take the first-year values of the endogenous commodity prices and use them to construct a composite commodity price index (CPI). Based on the CPI generated thereby we adjust the land availability for the subsequent run (which considers another 10-year planning period starting with year *t*+1 in a rolling horizon framework). The land constraint is shown in (13).

$$\sum_q ACR_{t,r,q} + \sum_p ACR_{t,r,p} \leq al_{t,r} \quad \forall t, r \tag{13}$$

To prevent unrealistic changes and extreme specialization in land use, which may be particularly serious at regional level, we restrict farmer's planting decisions to a convex combination (weighted average) of historically observed acreage patterns ($h_{r,ht,i}$) where subscript *ht* stands for the observed time periods prior to the base year. Historical land uses may be valid when simulating farmer's planting decisions under 'normal' conditions. However, they may be too restrictive for future land uses given the increasing demand for biofuels which may lead to dramatically different land use patterns that are likely to occur in the future to produce the required bioenergy crops. To address this issue we introduce 'hypothetical' acreage patterns ($h'_{r,n,i}$) for each row crop and each region (Chen and Onal 2012). To generate hypothetical acreage patterns (crop mixes), we first use the historical data on prices and acreages of row crops in each region to estimate the acreage elasticities for each row crop with respect to its own price and cross-price changes while controlling other factors, such as social- economic changes and time trend. Then we estimate a number of hypothetical acreages (crop mixes) using these price elasticities and considering a systematically varied set of crop prices. The resulting set of actual and hypothetical crop mixes are used in constraint (14) to limit the flexibility in planting decisions, where $\theta_{i,q}$ represents the share of row crop *i* in production activity q and $W_{t,r,*}$ represents the weight assigned to historical or hypothetical crop mixes. The latter are defined as variables to be endogenously determined by the model.

$$\sum_q \theta_{i,q} ACR_{t,r,q} = \sum_{ht} h_{r,ht,i} W_{t,r,ht} + \sum_n h'_{r,n,i} W_{t,r,n} \quad \forall t, r, i \tag{14}$$

The sum of the endogenous weights assigned to individual mixes must be less than or equal to 1 (convexity requirement), as shown in equation (15).

$$\sum_{ht} W_{t,r,ht} + \sum_{n} W_{t,r,n} \leq 1 \;\; \forall t, r \tag{15}$$

A similar set of crop mix constraints is also introduced for irrigated crops, which we do not show here, using only the historically observed irrigated land use patterns (no hypothetical mixes for irrigated crops).

Since miscanthus is a non-native grass species, a large scale miscanthus production may have unforeseen impacts on biodiversity and water quality. To prevent extreme specialization in the production of perennial grasses in some regions, we restrict the land allocated to perennial grasses (miscanthus and switchgrass) not to exceed 25% of total land availability in each region ($al_{t,r}$). The constraint is shown in (16).

$$\sum_{p} ACR_{t,r,p} \leq 0.25 al_{t,r} \; \forall t, r \tag{16}$$

In the livestock sector, we define production activity variables (number of animals) at national level for each category of livestock except the beef and dairy cattle. Cattle production is given special emphasis in the model for two reasons. First, cattle requires grazing land, therefore competes with crop production activities on total land in each region. Second, besides requirements of feed crops directly fed to different types of livestock, DDGS is also used as a feed item that may substitute other livestock feed crops. The regional cattle production activities are aggregated in equation (17) to obtain the total cattle activity at national level:

$$LIV_{t,cattle} = \sum_{r} CTL_{t,r} \;\; \forall t \tag{17}$$

where $CTL_{t,r}$ is the number of cattle stock in region *r* and year *t*. Cattle supply is constrained by the grazing land availability. Therefore, for each region we specify the grazing rates and the supply of grazing land, $GL_{t,r,g}$, where *g* denotes the type of grazing land (namely pasture land, forest land and cropland that can be used for grazing -such as wheat and oats). The latter is updated over time (for each region and land type) based on the expected livestock product prices and feed price index (similar to the way agricultural land is updated). Constraint (18) below relates the usage of grazing land and cattle activity in each region:

$$CTL_{t,r} \leq \sum_{g} \frac{GL_{t,r,g}}{ga_{r,g}} \;\; \forall t, r \tag{18}$$

where $ga_{r,g}$ denotes the amount of grazing land required per unit of cattle.

Equations (19) and (20) establish the balances between nutrition needs of livestock activities, in terms of protein and calories, and the amounts of nutrients provided by primary feed crops (grains) and byproducts of crops processing (i.e., soymeal and DDGS):

$$nr_{k,nu} LIV_{t,k} = \sum_{i} nc_{i,nu} F_{t,i,k} + \sum_{j} nc_{j,nu} F_{t,j,k} \;\; \forall t, k \tag{19}$$

$$FED_{t,z} = \sum_{k} F_{t,z,k} \;\; \forall t, k \; and \; \forall z = i, j \; used \; for \; feed \tag{20}$$

where $nc_{z,nu}$ denotes the nutrition content per unit of feed item *z*, and $nr_{k,nu}$ and $F_{t,i(j),k}$ are the required amount of nutrient *nu* per unit livestock and the amount of feed item i or j used by livestock category k, respectively.

To avoid unrealistic changes in feed mixes, we impose historical feed mixes used by all livestock categories. Constraints (21) and (22) constrain the consumption of feed to be within a convex combination of historical feed uses.

$$FED_{t,z} = \sum_{ht} hf_{z,ht} WF_{t,ht} \tag{21}$$

$$\sum_{ht} WF_{t,ht} \leq 1 \tag{22}$$

where parameter $hf_{z,ht}$ denotes historical observed usage of feed crops, and $WF_{t,ht}$ represents the weight assigned to historical feed crop mixes. We also impose appropriate upper bounds for DDGS to restrict the consumption of DDGS in total feed consumption for each type of livestock.

Livestock commodities can be consumed domestically or exported. The total supply of each livestock commodity is then related to the respective livestock production activity through a fixed yield coefficient, denoted by ly_k. Constraint (23) establishes this relationship:

$$DEM_{t,k} + EXP_{t,k} \leq ly_k LIV_{t,k} \;\; \forall t, k \tag{23}$$

Constrain (24) pertains directly to the electricity sector where it describes how electricity production by source is related to consumption. The quantity of electricity consumed is equal to the sum of all electricity produced from all sources considered (indexed by *at*): wind, coal, natural gas, fuel oil, geothermal, biomass, and other sources, including generation at new power plants, plus electricity transmitted from other regions er_1 $\left(\sum_{er_1,at} IT_{er_1,er,at,t}\right)$, minus electricity transmitted to other regions er_2 $(\sum_{er_2,at} IT_{er,er_2,at,t})$. We also consider electricity generated from cellulosic ethanol production as a co-product, which is denoted by $CP_{r,t}$.

$$\begin{aligned} QE_{er,t} = QW_{er,t} + \sum_{r(er),et} \left(QR_{et,r,t}\right) + \sum_{r(er)} \left(CP_{r,t}\right) + \sum_{nt} QM_{er,nt,t} \\ + \sum_{er_1,at} IT_{er_1,er,at,t} - \sum_{er_2,at} IT_{er,er_2,at,t} \;\; \forall er, t\,. \end{aligned} \tag{24}$$

Electricity is generated or produced according to production functions with a fixed input-output relationship (Leontief production functions). The generation of electricity from existing power plants is described by constraint (25)

$$QR_{et,r,t} \leq hr_{et,r} FF_{ff(et),r,t} \;\; \forall et, r, t \tag{25}$$

The amount of electricity generated from existing power plants in CRD *r*, year *t*, from technology *et* is a function of the thermal efficiency of the power plant (heat-rate), represented by $hr_{et,r}$, and the input of fossil fuel of type *ff*. This production applies to existing power plants that use fossil fuels as inputs, and does not apply to existing power plants that do not use fossil fuels as an input (e.g. geothermal, wind, solar, and nuclear).

The co-firing of biomass at existing coal power plants is described by constraint (26), where the coal-energy equivalent quantity of biomass transported to CRD *r*, $BC_{r,t}$, is converted to electricity using regional-specific coal power plant's conversion efficiency (denoted by $hr_{coal,r}$).

$$QR_{\text{bio},r,t} \leq hr_{coal,r} BC_{r,t} \quad \forall r,t \tag{26}$$

As shown in (27), we restrict the percentage of coal-fired biomass at an existing power plant to be less than ten percent of the amount of coal used at that power plant.[6]

$$0.1 FF_{\text{coal},r,t} \geq BC_{r,t} \quad \forall r,t \tag{27}$$

Generation at existing power plants is subject to a capacity constraint. Constraint (28) describes that generation of electricity from technology *et*, in CRD *r*, year *t* cannot be greater than the nameplate capacity $cap_{et,r}$ times the capacity factor $cf_{et,r}$, converted from megawatts to megawatt-hours by the parameter *mw*.

$$QR_{et,r,t} \leq mw * cf_{et,r} cap_{et,r} \quad \forall r,t, and\ et \neq coal \tag{28}$$

For coal power plants, generation from coal and biomass sources cannot be greater than the nameplate capacity times the capacity factor, converted to megawatt-hours, which is represented by constraint (29).

$$QR_{coal,r,t} + QR_{bio,r,t} \leq mw * cf_{\text{coal},r} cap_{\text{coal},r} \quad \forall r,t \tag{29}$$

Moreover, electricity generated from biomass co-firing cannot be greater than the sum of the capacity already retrofitted, ccf_r, and cumulative converted capacity by year *t*.

$$QR_{bio,r,t} \leq mw * cf_{\text{coal},r} * \left(ccf_r + \sum_{t=1}^{t} \left(CF_{r,t}\right)\right) \quad \forall r,t \tag{30}$$

Electricity generation occurring at new natural gas power plants is described by constraint (31). Generation occurs at the EMR level with a uniform thermal efficiency.

$$QM_{er,ng,t} \leq nr_{ng} NGM_{er,new,t} \quad \forall er,t \tag{31}$$

where nr_{ng} is the heat-rate of new natural gas power plants.

The natural gas balance equation relates natural gas used by the electricity and other sectors to the national quantity supplied:

$$NG_t = ngo_t + \sum_{er,age} NGM_{er,age,t} \quad \forall t \tag{32}$$

where ngo_t is the exogenous demand for natural gas from other sectors.

Constraint (33) is the balance constraint for electricity generated from new dedicated bio-power plants. Quantity of electricity generated in EMR *er*, year *t*, $QM_{er,\text{bio},t}$, equals the amount of biomass utilized by all bio-power plants within that EMR multiplied by a thermal efficiency:

$$QM_{er,\text{bio},t} \leq \sum_{r(er)} nr_{\text{bio}} BD_{r,t} \quad \forall er,t \tag{33}$$

where nr_{bio} is the heat-rate for new dedicated biomass power plants and $BD_{r,t}$ is the quantity of biomass used by dedicated biomass power plant for electricity production in CRD *r* and year *t*.

$$CP_{r,t} \leq ecp * BM_{t,r,eth} \quad \forall r,t, \tag{34}$$

Constraint (34) states that electricity produced as a co-product from cellulosic ethanol production should be no more than the amount of biomass used to produce cellulosic ethanol

---

[6] Co-firing constraints at other percentages are examined in the sensitivity analyses.

multiplied by a conversion coefficient *ecp* that represents amount of electricity generated per ton of biomass.

The amount of electricity generated from new power plants of other energy types: solar, geothermal, hydroelectric, and nuclear energy (indexed by *ot*) is constrained to be no more than the projections for these technologies from the AEO 2011, which is described by constraint (35):

$$QM_{er,ot,t} \leq aeo_{er,ot,t} \tag{35}$$

where $aeo_{er,ot,t}$ is a parameter that is an adjusted projection of electricity generated from technology type $ot \in \{hydro, solar, geothermal, nuclear\}$ in EMR er and year *t*.[7]

Constraint (36) states that the amount of biomass utilized to generate electricity cannot exceed the amount of biomass transported from all CRDs:

$$\sum_{r_1} BT_{r_1,r_2,t} \geq BD_{r_2,t} + BC_{r_2,t} \quad \forall r_2, t \tag{36}$$

$$\sum_{r_2} BT_{r_1,r_2,t} \leq BM_{r_1,t,elec} \quad \forall r_1, t \tag{37}$$

Constraint (37) relates the total amount of biomass transported from CRD r_1 to all CRDs to the amount of biomass used for electricity generation in that CRD. Biomass produced for the electricity sector, $BM_{r,t,elec}$ must be less than the amount of biomass produced minus that used for the production of liquid fuels for the transportation sector (see constraint (12)).

Constraint (38) below is the material balance constraint for biomass used for electricity generation.

$$BE_t = \sum_r \left(BD_{r,c} + BC_{r,c}\right) \ \forall \ t \tag{38}$$

Constraint (39) states that inter-regional transmission of electricity is constrained to only be allowed between adjacent regions and to not exceed historical levels.

$$\sum_{at} IT_{er_1,er_2,at,t} \leq tcap_{er_1,er_2} \quad \forall er_1, er_2, t \tag{39}$$

where the variable $IT_{er_1,er_2,at,t}$ is the quantity of electricity transmitted from er_1 to er_2 generated from technology type at in year *t.* The parameter $tcap_{er_1,er_2}$ denotes the observed quantity of electricity transmitted from er_1 to er_2 in the past.[8]

The amount of electricity imported from all other EMRs for a given EMR er_1 is:

$$IMP_{er_1,t} = \sum_{er_2,at} IT_{er_1,er_2,at,t} \quad \forall er_1, t \tag{40}$$

The amount of electricity that EMR er_1 exports to all other EMRs ($EXP_{er_1,t}$) is:

$$EXP_{er_1,t} = \sum_{er_2,at} IT_{er_1,er_2,at,t} \quad \forall er_1, t \tag{41}$$

[7] The AEO projection for these sources is adjusted by differencing the baseline quantity as of 2007 with the projected for each succeeding year and creating $aeo_{er,ot,t}$ from it (EIA 2010a).

[8] Inter-regional transmission constraints are based on historical data as is described in the Data section.

We assume that once wind and dedicated biomass capacity is built it will be fully utilized over its lifetime. Constraints (42) and (43) ensure that generation from these sources cannot decrease relative to the previous year.

$$QM_{er,\text{bio},t} - QM_{er,\text{bio},t-1} \geq 0 \qquad \forall er, t > 1 \tag{42}$$

$$QM_{er,\text{wind},t} - QM_{er,\text{wind},t-1} \geq 0 \qquad \forall er, t > 1 \tag{43}$$

Equation (44) accounts for total domestic GHG emissions generated from U.S. agricultural, transportation and electricity sectors. The first line of (44) includes GHG emissions from the production of conventional crops, perennial energy crops, and crop residues collection, while the second and third lines consist of emissions from gasoline and petroleum diesel consumption and biofuel conversions. We use δ_i to denote unit lifecycle GHG intensity of transportation fuels/crop production. The last line of (44) accounts for emissions from the production of electricity. We use γ_{ff} to denote life cycle GHG emissions from all power plant that burn fossil fuels.

$$\begin{aligned} GHG_t = & \sum_{r,q} (\delta_{r,q} + \delta_{r,q,s}) ACR_{t,r,q} + \sum_{r,p} \delta_{r,p} ACR_{t,r,p} + \delta_g \sum_v G_{t,v} \\ & + \delta_d DSL_t + \delta_{btl} BTL_t + \sum_j \delta_j OILDSL_{t,j} + \delta_c ETH_{t,c} \\ & + \delta_b ETH_{t,b} + \delta_s IMP_{t,sugarcane\ ethanol} \\ & + \sum_{ff,r} \gamma_{ff} QR_{ff,r,t} + \sum_{ff,er} \gamma_{ff} QM_{ff,er,t} \quad \forall t \end{aligned} \tag{44}$$

The constraint for the regional RPS policy is modeled as a constraint on the proportion of electricity generated from renewable sources in a region to the total electricity consumed in that region, thus allowing for inter-regional trading of electricity generated from renewables:[9]

$$rps_{er,t} QE_{er,t} \leq \sum_{renewable,r(er)} QR_{renewable,r,t} + QM_{er,renewable,t} \; \forall er, t \tag{45}$$

The parameter $rps_{er,t}$ is the percentage of renewables required in region er in year t. The right hand side of (45) is the sum of electricity generated from renewable sources, where $renewable$ is an index for all renewable energy and is a subset of et.

Constraint (46) establishes the consumption mandate specified by the Renewable Fuel Standard (RFS). The RFS is implemented as a blend mandate where biofuel (including ethanol and biodiesel) is blending with conventional fuel at exogenously given rates θ_t:

$$\sum_v E_{t,v} + 1.5 BTL_t + 1.5 \sum_j OILDSL_{t,j} \geq \theta_t \left(\sum_v G_{t,v} + DSL_t \right) \forall t. \tag{46}$$

Lastly, we require all choice variables to take non-negative values.

[9] An adjustment is made to the qualifying renewable electricity generation for EMRs that have states where some existing renewable energy does qualify in achieve the RPS.