TECHNICAL APPENDICES

A Framework for Assessing the Economic Impacts of Agricultural Equipment Emission Reduction Strategies on the Agricultural Economy in the San Joaquin Valley

May 1, 2018

Michael McCullough
Lynn Hamilton
Agribusiness Department
California Polytechnic State University

Duncan MacEwan
Jay E. Noel
Richard E. Howitt
ERA Economics, LLC

CARB Contract #13-331

Table of Contents

List of Figures

List of Tables

Introduction

The following technical appendices outline the suite of Excel tools created to evaluate additional regulatory policy in the San Joaquin Valley. The tools are to be used in conjunction with each other and show the estimated effects of an additional regulatory policy or incentive program at the farm level, agricultural and related industries, and the broader San Joaquin Valley economy.

The suite of tools developed includes: (i) 22 stochastic farm budget models, (ii) a calibrated regional agricultural production optimization model (SWAP), and (iii) a regional input-output (or "multiplier") model (IMPLAN). While the study was limited in its primary data gathering approach to 22-case studies and supplemental secondary data, all three types of models have been made flexible in their inputs so they could be easily updated. The stochastic farm models can easily be updated with current price, yield, interest rates, and production costs. The regional models are flexible in the agricultural market conditions and water supply conditions, and with limited work could include additional case studies.

A technical overview of the stochastic farm models appears in Appendix A. This is followed by a technical overview of the SWAP model developed for this analysis in Appendix B, and a technical overview of IMPLAN in Appendix C. There is a brief overview of the tools themselves and a detail of the survey instruments used to generate the scholastic farm models.

Appendix A: Farm Simulation Models Technical Overview

Farm Simulation Models and Data

The farm models developed for this study are simulation models. Simulation models are used extensively in farm level analyzes[1]. Simulation models are numerical or mathematical representations of a real-world process. One use of a simulation model is to evaluate potential economic impacts of changes in market conditions or policy changes on firm profitability before those changes occur.

The farm simulation models in this study are representations of farm income statements. The income statement has two basic components: 1) farm revenue and 2) the cash costs of production. The simulation variable of interest is returns to land and management or total farm revenue less total cash costs of production[2]. The returns to land and management variable is simulated 500 times to statistically represent all possible combinations of returns to land and management. A cautionary note is needed at this point. The case study simulation models are

---

[1] The Agricultural and Food Policy Center (AFPC) at Texas A&M University provides a robust set of publications that document the use of simulation modeling for analyzing the impacts of policy, market, and technological changes on farm profitability. https://www.afpc.tamu.edu/pubs/.

[2] Returns to land and management is like the financial measure of earnings before interest on fixed assets, taxes, depreciation, and amortization (EBITA). It is a proxy for a firm operating profitability as well as a proxy for cash flow.

useful in assisting understanding the economic consequences of a regulatory policy change and can facilitate better decision-making but they should not be used to predict point estimates.

There were twenty-two farm simulation models developed from the farm case study interviews. That data obtained in those case study interviews included crop prices, yields, in some interviews total revenue, costs of production, and regulatory costs. The question addressed by these farm simulation models is what changes in returns to land and management would occur given changes in total farm income and regulatory costs. Changes in total farm income are driven by changes in crop prices and yields. Crop prices and yields can vary significantly over time. That variability needs to be taken into consideration as regulatory costs change over time. For example, if a specific crop has a high price and yield, then the economic impact of a regulatory cost change could be relatively small as compared to that crop experiencing low price and yield. Thus, differing combinations of crop price and yields will result in differing economic impacts of a regulatory cost change.

The price and yield variability and thus the total revenue variability in the farm simulation models is accounted for by estimating probability distributions. The data to estimate the probability functions came from different sources. Table 1 shows the source of the data for each of the case study crops. There were differences in the amount of price, yield, and revenue data provided by the case study growers. The case study interviews attempted to obtain 9 years of price and yield data or, absent price and yield data, total revenue data over the period 2004 – 2012. Thirteen growers provided 5- 9 years of price and yield data, two growers provided 9 years of revenue data, six growers did not provide any price, yield or revenue data, and one grower provided 9 years of yield data but no price data. The data to estimate the price and yield probability functions for 5 of the six growers that did not provide that data was obtained from the California Agricultural Statistical Service (CASS). It was not possible to obtain price or yield data for one of the citrus case studies, and price data for one tomato case study was obtained from the California Tomato Growers Association.

Table 1. Sources of Farm Simulation Model Data

Model	Cost of Production	Stochastic			Deterministic	
		Price	Yield	Total Revenue	Price	Yield
Large Tree Nut 1	Grower	Grower	Grower	N.A.	N.A.	
Large Tree Nut 2	Grower	Grower	Grower	N.A.	N.A.	
Medium Tree Nut	Grower	CASS	CASS	N.A.	N.A.	
Small Tree Nut	Grower	CASS	CASS	N.A.	N.A.	
Large Silage 1	Grower	Grower	Grower	N.A.	N.A.	
Large Silage 2	Grower	Grower	Grower	N.A.	N.A.	
Medium Silage 1	Grower	Grower	Grower	N.A.	N.A.	
Medium Silage 2	Grower	Grower	Grower	N.A.	N.A.	
Medium Silage 3	Grower	Grower	Grower	N.A.	N.A.	
Large Cotton	Grower	Grower	Grower	N.A.	N.A.	
Medium Cotton	Grower	Grower	Grower	N.A.	N.A.	
Large Grape 1	Grower	Grower	Grower	N.A.	N.A.	
Large Grape 2	Grower	CASS	CASS	N.A.	N.A.	
Medium Grape	Grower	Grower	Grower	N.A.	N.A.	
Small Grape	Grower	CASS	CASS	N.A.	N.A.	
Large Citrus	Grower	N.A.	N.A.	N.A.	Yes	Yes
Medium Citrus	Grower	N.A.	N.A.	Grower	N.A.	
Small Citrus	Grower	CASS	CASS		N.A.	
Medium Stone Fruit	Grower			Grower	N.A.	
Small Stone Fruit	Grower	Grower	Grower	N.A.		
Large Processing Tomato	Grower	Grower	Grower	N.A.		
Medium Processing Tomato	Grower	CA Tomato Growers Assn	Grower	N.A.	N.A.	

Twenty-one of the case study simulation models are stochastic simulation models and one is a deterministic farm model. The stochastic case study farm models are characterized by treating price and yield or total revenue, if price and yield data was not available, as random variables. One of the citrus farm simulation models is a deterministic simulation model as neither price, yield, nor total revenue data was available. The citrus farm price and yield are fixed at 2012 figures. A basic assumption of stochastic farm simulation models is that future prices and yields or total revenue variability will mimic past price and yield or total revenue variability.

As discussed earlier, historical variability in crop prices, yields, and total revenue need to be considered in determining the impact of a policy change on returns to land and management. Allowing price and yield to be random results in total revenue being random. Thus, the impact of a regulatory cost or change in a regulatory cost can be evaluated under random total revenues. The randomness of the total revenue will be reflected in a random return to land and management.

As mentioned above, the simulation models have two components, total farm revenue and the crop cost of production. The base year for the crop cost of production was 2012. The crop cost of production data was collected during the case study interviews. All the cash costs of production are used in the simulation models. They include cultural costs, harvest costs, operating interest cost and cash overhead costs, see Table 4 of the main report.

Other than two case study crops, medium citrus and medium stone fruit, empirical probability functions[3] are estimated for crop price and yield. The medium citrus and medium stone fruit models have empirical probability functions estimated for total revenue.

Basically, the simulation models have a stochastic total revenue component and a static cost of production budget with two exceptions. Operating interest cost can be adjusted by changing the interest rate in some of the models and harvest cost can be adjusted as the stochastic yield changes as the simulation is run. This is discussed in more detail later in this technical overview.

The intent of the case study farm simulation models is to allow for the evaluation of changes in the returns to land and management due to differing regulatory policies. The models will provide information on the probability distribution of the returns to land and management before and after changes in regulatory costs and provide for a visual or graphical representation of those changes. The stochastic case study models simulation results are not intended or expected to be predicative, rather the results can provide a range of the relative impact of proposed policies on returns to land and management.

Composition of the Case Study Farm Simulation Models

There are eight worksheets in each stochastic case study farm model workbook. Five sheets are protected from analyst view. The information and data on those sheets is both hidden and protected with each sheet protected by a unique password. The projected sheets contain confidential producer information. That protection was at the request of the California Air Resources Board staff[4] and representatives from the agricultural community. The five protected sheets are: Detail Cost of Production, Regulatory Cost, Price and Yield Stochastics, or Total Income Stochastics, Validation, and Analysis Cost of Production. The unprotected sheets are the Analysis, Sim Results, and Stoplight worksheets and will be discussed in greater detail.

The following will provide an overview of the worksheets and a general example of the information each contains. These worksheet sheets are not related to any given case study

[3] The empirical distribution was selected because there were nine or less price and yield observations which is too few observations to estimate an alternate statistically valid continuous distribution (e.g. normal distribution). The shape of the empirical probability function is defined by the price/yield/total revenue data. The function's input data (prices/yields/total revenues) are discrete (yearly data) but the estimation of the function allows for a continuous probability density function by interpolation between the discrete data points.

[4] The confidentially protection is stated in the contract between Cal Poly, ERA Economics, and the California Air Resources Board.

model. The numbers shown on them, if any, are made up to illustrate what is contained on any given sheet. The worksheet descriptions are as follows:

Worksheet 1: The Detailed Cost of Production worksheet has the cultural costs (chemical costs, labor costs, irrigation costs, machinery costs and other costs relative the production of the crop), harvest cost, interest on operating capital cost, and cash overhead costs of production.

Worksheet 2: The regulatory cost sheet has detailed cost information on education/training for regulatory compliance, air quality regulations, water quality regulations, pesticide use regulations, and other regulatory costs that are specific to individual case study crop farms. Data on this sheet is used on the Analysis sheet.

Worksheet 3: The Price and Yield Stochastics sheet contains the case study price, yield or total revenue time series. This sheet is where the price/yield/total revenue empirical probability functions are estimated.

Worksheet 4: The validation worksheet has several purposes. The first is to simulate price/yield/total revenue series from the estimated empirical probability functions. The simulated price/yield/total revenue series were then statistically tested to determine whether the means of the historical and simulated time series were statistically equal and whether the variances of the historical and simulated time series were statistically equal. The statistical test results for all the case study farm models were the simulated means and historical means were statistically equal and the same was true for the variances.

Worksheet 5: The Analysis Cost of Production worksheet is a consolidated version of the Cost of Production worksheet. It is put in the case study models for ease of use in transferring consolidated crop cost of production data to the analysis worksheet. The Analysis Cost of Production worksheet differs slightly between case study crop models. That is due to case study growers using different cost of production accounting methods. For example, harvest costs can be a flat charge which can occur in a custom harvesting situation or are dependent on yield. The Analysis Cost of Production worksheet accounts for those differences.

Worksheet 6: The Analysis worksheet is the key worksheet in case study stochastic farm simulation models. Table 2 shows a representative Analysis worksheet. It contains the basic data to determine returns to land and management under different regulatory cost scenarios.

Table 2. An example of an Analysis worksheet.

Price	
Yield	
Total Revenue	
Cultural Costs	
Harvest Cost	
Interest on operating capital	
Cash Overhead Costs	
Total Cost Per Acre	
Air Quality Regulatory Cost	
Other Regulatory	
ARB Policy	
Net Farm Income with Regulatory Costs Included	$0
Net Farm Income with Regulatory Costs Excluded	$0

Price Switch	
Set Price	
Yield Switch	
Set Yield	
Interest Rate Switch	
Interest Rate	
ARB Policy Switch	
Set Cost	
Percentage Cost	

There are two data input areas in the Analysis worksheet. The default entries for revenue section of the Analysis worksheet are stochastic price and stochastic yield. The stochastic price and stochastic yield are multiplied to get the stochastic total revenue. The stochastic price and yield are in terms of per acre.

Total Cultural Cost, Harvest Cost, Operating Interest Cost, and Cash Overhead Cost come from the Analysis Cost of Production worksheet. All costs are on a per-acre basis. Harvest cost can be either fixed or variable. Whether harvest cost is fixed or variable depends on the crop and grower. Typically, a variable harvest cost varies with yield. An example would be boxed citrus where labor is payed based upon how many boxes they fill in a day. A fixed harvest cost does not change with yield. An example would be a custom harvest where a fixed rate harvest cost is paid to a harvest contractor.

Operating Interest can also be fixed or variable in the farm simulation models. An interest on operating capital cost was fixed if a case study grower self-financed and indicated a set interest on operating capital cost or a case study grower did not include an interest on operating cost. If there was an interest on operating cost figure that was determined by a given loan percentage, it was considered variable and used to calculate the interest on operating capital cost. The interest rate for calculating the operating interest cost is in the Detailed Cost of Production worksheet.

Cash overhead is taken directly from the cost of production data provided by the case study growers. That data is on the Detailed Cost of Production worksheet and Analysis Cost of Production worksheet.

The regulatory costs are divided into an air quality regulatory cost and other regulatory costs. The separation was at the request of California Air Resources Board personnel. These costs are on the regulatory cost sheets.

The ARB policy cell is where any additional regulatory cost may be inputted. There are currently two ways to calculate an ARB policy. There is an IF statement in the ARB policy cell. The ARB policy cell can be activated by placing a 1 in the cell below the ARB Policy Switch cell. The analyst can then select to use a percentage of cost (e.g. 3%) in the cell below the percentage cost cell. The IF statement in the ARB policy cell can be changed to be any cost figure the analyst chooses to use. The default for the ARB Policy is have a blank in the cell under the ARB Policy cell and to place a set cost in the Set Cost field.

The variable of interest in the simulation models is returns to land and management. The returns to land and management is also a stochastic variable since it is determined in part by the stochastic total revenue variable. Two net returns to land and management variables are in the farm simulation models. The first is returns to land and management with regulatory cost included and the second is net returns to land and management with regulatory costs excluded. This was done so a comparison of the economic impact of regulatory costs could be evaluated.

The second set of input data available in the farm simulation models is user determined. That is, an analyst can input the following data to replace the default data that is currently used in the farm simulation models.

There are switches for replacing the stochastic price/yield/total revenue stochastic variables, for using a different interest rate for the calculation of the interest on operating cost variable, and ARB policy switches that can be used to place a value in the ARB Policy cell. If a number greater than zero is placed in a switch box and then a value is place in the set box, the number in the set box will replace the stochastic number in respectively the price or yield cells. Likewise, a similar approach can be used to set a different interest rate on interest on operating capital cost and a revised interest on operating capital expense will replace the current one. The ARB policy cost can be set by placing a number in the set cost cell or by placing a number greater than zero in the switch box and a percentage number in the percentage cost cell.

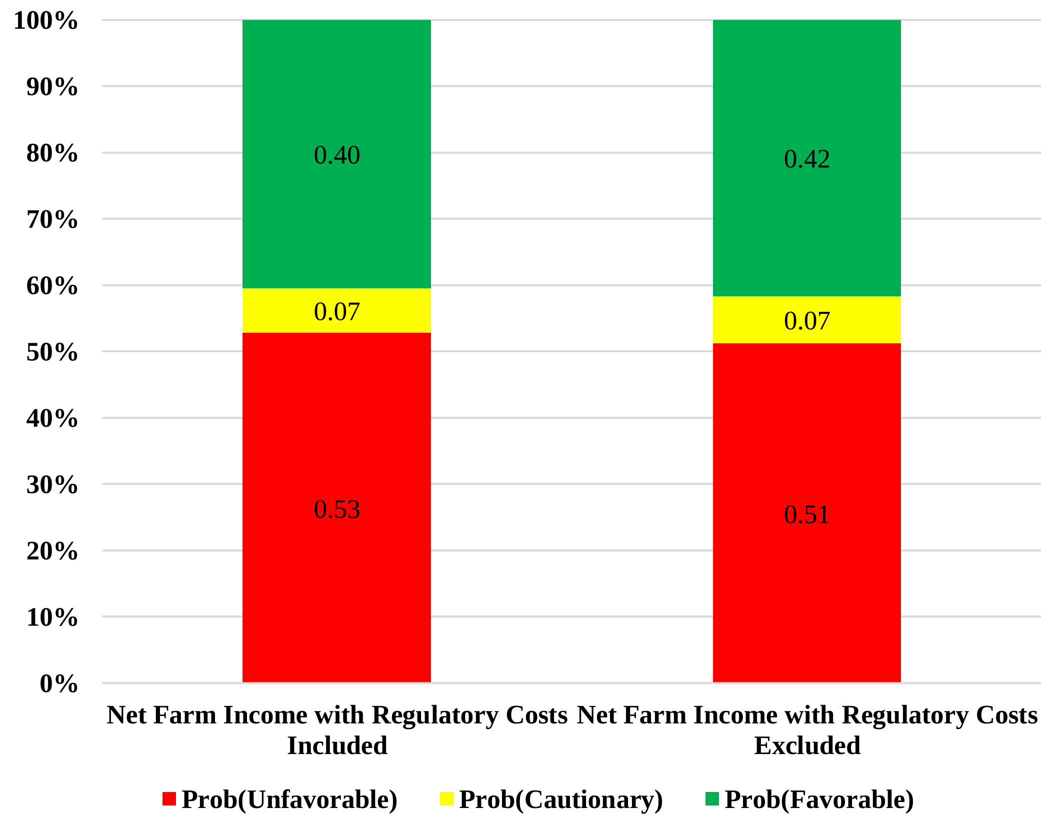
It should be noted that other costs such as cultural cost or cash overhead costs can also be changed by simply replacing them with chosen costs.

Worksheets 7 and 8 in the case study stochastic farm models are outputs from the simulation of the model and a graphical representation of the simulation results. The SimResults worksheet contains the simulated returns to land and management data. There were 500 simulation iterations run for both returns to land and management with and without regulatory costs.

The last worksheet is the Stoplight worksheet. The Stoplight worksheet is one way that simulated returns to land and management can be presented. It is basically a single column bar chart. The simulation tool used in development of the case study simulation models is SIMITAR. SIMITAR has a stoplight tab that can be used to present the simulation results in a single column bar chart.

Figure 1 shows a Stoplight chart for the probability that returns to land and management will be less than -$160/acre, between -$160.00 and $0.00, or greater than $0.00.

Figure 1. StopLight Chart for the Probability of the Return to Land and Management will be Less Than -$160.00, between -$160.00 and $0.00 Greater Than $0.00



The breakpoints on the bar chart can be chosen when preforming the farm model simulation.

The chart shows that returns to land and management with regulatory cost included has 40% probability of being greater than $0.00/acre, a 7% probability of being between $0.00/acre and negative $160/acre and 53% probability of being lower than -$160/acre. This is a convenient way to present the results of the simulation but the results can be placed in other graphical forms or plotted as box plots, probability density function plots, cumulative distribution function plots, and probability plots. The choice of presentation can be based on whom the results are being presented to or other considerations.

Every attempt has been made to make the case study farm simulation models as flexible as possible given the imposed data confidentially restrictions. Although the farm case study model workbook has five password protected worksheets the price, yield, interest rate variables can be easily adjusted. The ARB policy variable can also be adjusted and other policy instruments can be introduced to the analysis worksheet. The simulations are straight forward and the results of the simulations can be plotted as probability distributions, and different types of graphical presentations. The case study stochastic farm models should provide the analyst significant flexibility in addressing “what if” questions relative to the on-farm emission policies.

Appendix A SWAP Model Technical Overview

The Statewide Agricultural Production (SWAP) model is developed for this study to evaluate the regional changes in agricultural production resulting from changes in regulatory costs. SWAP is a regional economic optimization model that is developed and maintained by ERA Economics and covers over 97 percent of irrigated agriculture in California. The model has been peer reviewed and the underlying calibration framework has been used by economists in various contexts for over 30 years (Howitt et al. 2012). This section provides an overview of the development history of the model, as well as a description of the modifications made to the SWAP model specifically for this research contract. In addition, the technical details of the model are presented. The following sections of the report describe the technical and data adjustments to the model that were completed for this project, and a concluding section offers some thoughts on future work.

The following changes to the calibration code and data are made for this project:

1. Crop groups are split into small, medium, and large operations, for applicable crop groups. Data from the 22 case studies are combined with supplemental data from USDA, UC Cooperative Extension, published studies, Agricultural Census, and other sources to establish these new production budgets, and the spatial location of the crops produced.

2. All of the underlying crop cost-of-production data in the SWAP model are updated to reflect the crop groupings based upon the case study interviews and regional disaggregation of the San Joaquin Valley. In addition, water supply and elasticity estimates are updated to the most recent data available.

3. The SWAP model calibration is updated to include the small and large farms. This involves updating the Positive Mathematical Programming calibration routine to include the new crop groups. In addition, the SWAP model includes downward-sloping linear crop demands. The code is updated to calculate the aggregate (small and large farm) supply and aggregate demand for each crop, since this determines the market clearing price. In other words, whether a crop was produced on a small or large farm does not alter the market to which it is sold.

4. The model is additionally specified to operate on future-condition time-scale (2030) which reflects future conditions and regulatory costs. This includes updating crop demand shifts (driven by income and population growth), real production costs, real energy cost, and perennial (orchard) retirement.

5. Finally, the model is specified to account for drought conditions. California Department of Water Resources (DWR) are combined with federal project and local project supply information to generate a "drought condition" baseline which can be included in the regulatory impact assessment.

SWAP Model Development History

The SWAP model is an economic optimization model of California's agricultural economy. It is grounded in economic theory and based on a well-established calibration method known as Positive Mathematical Programming (PMP). The interested reader can find details about the initial calibration approach in Howitt (1995) and can find a useful summary of the evolution of the method in Mérel and Howitt (2014). With respect to California, the SWAP model is the most recent in a series of similar models that have been developed and improved over the last 30 years in California. Initial models include a San Joaquin Valley Drainage model and the California Regional Agricultural Production Model. These models were largely developed by Richard Howitt at UC Davis and his graduate students. The transition from an academic model to one that is acceptable for public (and regulatory) policy analysis is a key step in the development of SWAP and its predecessor models.

The first model to gain wide-spread acceptance by California water managers and state and federal agencies was the Central Valley Production Model (CVPM). CVPM was developed by Richard Howitt in coordination with Steve Hatchett of CH2M and Farhad Farnam of the California Department of Water Resources, along with a team of DWR and U.S. Bureau of Reclamation staff and other consultants. CVPM development took place in the late 1980's and early 1990's, and it was ultimately applied to the Central Valley Project Improvement Act to evaluate the effect of changes in State Water Project and Central Valley Project to agricultural users. The major change to bring CVPM from an academic model to a policy-evaluation tool was: (i) rigorous data update, (ii) calibration update, (iii) peer-review process, and (iv) linking the model to state and federal surface and groundwater models. To this day, these are the key features that differentiate the models applied for academic studies and those used for policy analysis.

CVPM was the central model of California water and agriculture between 1994 and 2009. The defining feature of CVPM was that it had linear marginal PMP cost functions and a limited production technology. In particular, the production technology included only a water – irrigation efficiency isoquant, and did not include a full production function. The SWAP model was originally developed by Richard Howitt as an academic model to correct these two inconsistencies. In particular, the PMP cost function is exponential, and therefore always positive, avoiding the situation where the linear PMP cost function intercept can be negative. Second, the SWAP model includes a full Constant Elasticity of Substitution (CES) production technology, allowing for a richer specification of intensive margin adjustments. The SWAP model was adopted by the DWR during the California Water Plan Update 2009 on a project with Richard Howitt and Duncan MacEwan. Duncan MacEwan worked with Steve Hatchett to benchmark the CVPM and SWAP models, and concurrently, the model regions were expanded from 21 regions in the Central Valley to the current 27. The academic version of the SWAP model still includes 21 regions. Following the adoption of the SWAP model by DWR, Duncan MacEwan and Steve Hatchett worked with U.S. Bureau of Reclamation under the Sites Reservoir feasibility study to update the data, code, and calibration. Since 2009 the SWAP model has been the model for California water and policy analysis for agriculture. In 2013, Richard

Howitt and Duncan MacEwan led a data update to develop the current version of the model, known as version 6.1, which continues to be developed and maintained by ERA Economics. A new version of the model will be released in 2017, with a comprehensive update to the model code, data, and structure.

The SWAP model developed for this ARB regulatory analysis is based off the SWAP v6.1 model code and data. However, it represents a wholesale update to the calibration and code, as highlighted in the previous section, and discussed in more detail in the following sections.

SWAP Model Technical Overview

The SWAP model is a regional agricultural production and economic optimization model that simulates the decisions of farmers across 97 percent of agricultural land in California. The model assumes that growers maximize profits (revenue minus cost) by choosing total input use (e.g., total crop acres) and input use intensity (e.g., applied water per acre) subject to market, resource, and technical constraints. Growers are assumed to face competitive markets, where no one grower can influence crop prices, but an aggregate change in production can affect crop price. This competitive market is simulated by maximizing the sum of consumer and producer surplus.

The SWAP model allows for flexibility in production technology and input substitution, and it has been extended to allow for a range of analyses, including interregional water transfers and climate change effects. The SWAP model estimates the changes in agricultural production using a simulation/optimization framework based on the principle of Positive Mathematical Programming (PMP) (Howitt 1995). The model takes land allocation, input use, crop prices, yields, and costs as input and estimates how agricultural production will respond to changes in water supply, prices, costs, or other policy shocks. The benefit (or cost) of changes in regulations or other policies can be determined from the change it produces in the net value of agricultural production relative to a base (e.g. without regulation alternative) condition.

This version of the SWAP model was designed to be data-driven in order to easily represent different analytical circumstances without changing the model code. For example, the model can be linked to water supply models if ARB staff are interested in the impact of other regulations that may affect agriculture jointly with air emissions.

SWAP Model Theory

The SWAP model self-calibrates using a three-step procedure based on Positive Mathematical Programming (PMP) (Howitt 1995) and the assumption that farmers behave as profit-maximizing agents. In a traditional optimization model, profit-maximizing farmers would simply allocate all land, up until resource constraints become binding, to the most valuable crop(s). In other words, a traditional model would have a tendency for overspecialization in production activities relative to what is observed empirically. PMP incorporates information on the marginal production conditions that farmers face, allowing the model to exactly replicate a base year of observed input use and output. Marginal conditions may include inter-temporal effects of crop rotation, proximity to processing facilities, management skills, farm-level effects such as risk and input smoothing, and heterogeneity in soil and other physical capital. In the SWAP model, PMP is used to translate these unobservable marginal conditions, in addition to observed average conditions, into a cost function.

Unobserved marginal production conditions are incorporated into the SWAP v6.1 model through increasing land costs. Additional land into production is of lower quality and, as such, requires higher production costs, captured with an exponential “PMP” cost function. The PMP cost function is both region and crop specific, reflecting differences in production across crops and heterogeneity across regions. Functions are calibrated using information from acreage response elasticities and shadow values of calibration and resource constraints. The information is incorporated in such a way that the average cost data (known data) are unaffected.

PMP is fundamentally a three-step procedure for model calibration that assumes farmers optimize input use for maximization of profits. In the first step a linear profit-maximization program is solved. In addition to basic resource availability and non-negativity constraints, a set of calibration constraints is added to restrict land use to observed values. In the second step, the dual (shadow) values from the calibration and resource constraints are used to derive the parameters for the exponential PMP cost function and Constant Elasticity of Substitution (CES) production function. In the third step, the calibrated CES and PMP cost function are combined into a full profit maximization program. The exponential PMP cost function captures the marginal decisions of farmers through the increasing cost of bringing additional land into production (e.g. through decreasing quality). Other input costs, (supplies, land, and labor) enter linearly into the objective function in both the first and third step. Calibrating production models using PMP has been reviewed extensively in the peer-reviewed literature, Merel and Howitt (2014) provide an excellent overview of the literature.

The SWAP model, and calibration by PMP, is a complicated process thus sequential testing is very useful for model validation, diagnosing problems, and debugging the model. At each stage in the SWAP model there is a corresponding model check. In other words, the calibration procedure has particular emphasis on the sequential calibration process and a parallel set of diagnostic tests to check model performance. Diagnostic tests are discussed in Howitt et al. (2012).

While the base levels of input use and output are calibrated against recent data, the responsiveness of farmers to changes in expected prices and costs is based on econometric estimates based on a 30-year data set of observed grower response to these factors. The SWAP model includes a series of sub-routines to forecast crop demand, real energy prices, pumping head, and technological innovation.

SWAP Data

SWAP model v6.1 data, before additional information added under this ARB study discussed below, include land use, crop prices, yields, input costs, water costs, use, and availability, and relevant elasticity estimates. The version 6.1 of the model calibrates to land use data for 2010. We note that the calibration base year is different from the simulation base year. 2010 is the calibration base year because it is the most recent and consistent dataset covering all of

California agriculture and water on the spatial scale required by the SWAP model. The base year for a policy impact (regulatory) analysis is defined by the level of development.

We briefly summarize the core data in the SWAP model in this section, and go into additional detail for the adjustments made in the ARB model in the following section. All prices, costs, land use, and market conditions, used to calibrate the model are for 2010, and then indexed to current dollars. Having calibrated the model using the 2012 data, the model is then used for policy simulations under future conditions, as described in a subsequent section. Table 3 summarizes the components of the model calibration data.

Table 3. Primary Model Inputs

Input	Description
Land Use	Cropping patterns for the 2010 base year, by region and crop group.
Crop Prices	Farm-gate price received for the crops produced.
Crop Yields	Average yield for each crop.
Interest Rates	All interest rates for short-term production loans, and long-term capital recovery are normalized to 2010 levels, equal to 4.75% and 5.75%, respectively.
Land Costs	The land rental rate or capital recovery cost of the land, per acre.
Other Supply Costs	Other supply costs include variable material and chemical inputs to crop production, such as fertilizers and pesticides.
Labor Costs	Labor costs include on-farm, contractor, and custom operator labor inputs into to crop production.
Surface Water Costs	Surface water costs per acre-foot for each region are a weighted average of the rates paid by growers in major water districts in the region.
Groundwater Costs	Groundwater pumping costs are calculate per acre-foot and include the amortized fixed cost of the well and pump, an annual maintenance cost, and variable energy pumping cost per foot of dynamic head.
Irrigation Water	Crop applied water requirements, in acre-feet per acre, are used to establish the average applied irrigation water for each crop and region.
Elasticities	Demand and supply elasticities are estimated from a 30-year time-series of cropping data and used to establish the statewide market for crops produced.

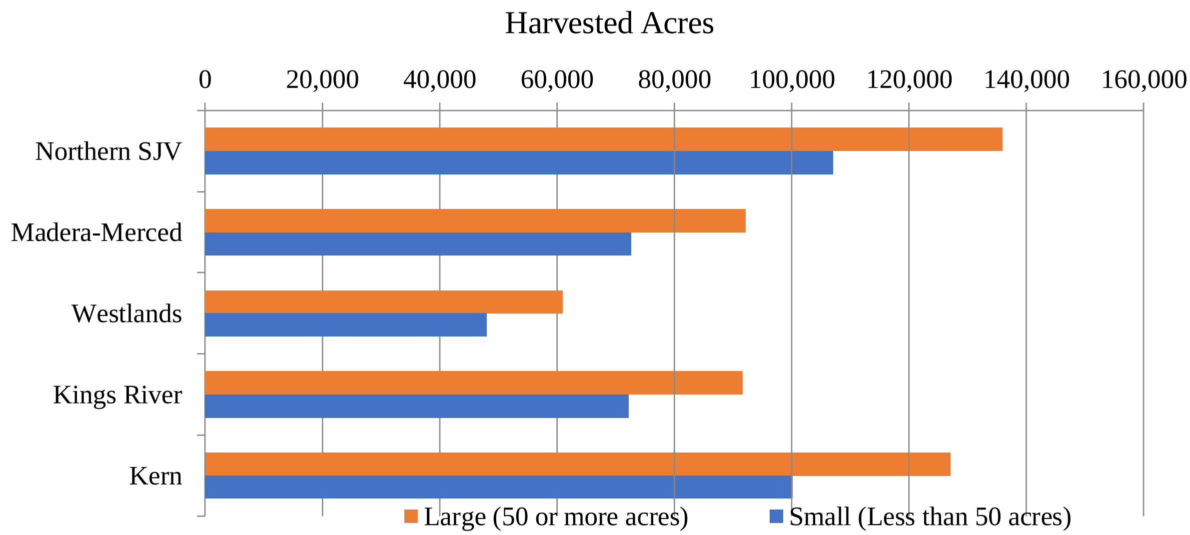
Air Emission Regulation SWAP Model Development

There are 5 central changes to the SWAP model, discussed in the introduction of this appendix, undertaken by ERA to generate the model used in this study.

The first task was to split crops into small and large operations, for applicable crop groups. The Agricultural Census is the only source of information on the distribution of farm size. However, these data are generated from surveys and are only available at the California-level of detail. As such, it was necessary to create a disaggregation approach and apply that to the individual SWAP model regions. ERA compiled additional data on the total number of farms in each county from the Agricultural Commissioners and Agricultural Census.

A three-step process is used to generate small and large farm distribution by SWAP model region. First, county farm data are used to calculate the proportion of farms, by crop type, in each county. Concurrently, the Agricultural Census data (statewide) are used to calculate the proportion of small and large farms, by crop. These two data sets are then used to calculate the small and large farm distribution, by county and crop. Finally, a GIS-based tool developed by ERA to map county-level data to the individual SWAP model regions is used to calculate the small and large farm distribution for each SWAP model region. Figure 2 illustrates an example for almonds in the San Joaquin Valley.

Figure 2. Almond and Pistachio Crop Group, Small and Large Farm Distribution



Other crops show a similar pattern, however, not all crops are split by small and large farm size. Some crops are only produced at a single scale. For example, cotton operations are typically larger farms. As such, this crop is not split by small and large farms. These crops are simply reported as average. Table 4 summarizes the breakdown of crops by farm size in the updated SWAP model.

Table 4. SWAP Model Crop Groups and Farm Size Distribution, San Joaquin Valley

Crop	Size	Acres	Crop	Size	Acres
Alfalfa	Small	115,138	Onions and Garlic	Large	40,947
Alfalfa	Large	375,136	Other Deciduous	Small	202,176
Almonds and Pistachios	Small	399,968	Other Deciduous	Large	104,163
Almonds and Pistachios	Large	507,560	Other Field	Average	508,240
Corn	Small	133,064	Misc. Vegetables	Small	51,039
Corn	Large	436,136	Misc. Vegetables	Large	111,752
Cotton	Average	283,000	Pasture	Average	113,382
Cucurbits	Small	32,306	Potatoes	Average	14,000
Cucurbits	Large	19,217	Processing Tomatoes	Small	20,259
Dry Beans	Small	21,050	Processing Tomatoes	Large	178,469
Dry Beans	Large	28,441	Rice	Average	9,978
Fresh Tomatoes	Small	15,306	Safflower	Average	10,314
Fresh Tomatoes	Large	10,652	Sugar Beets	Average	600
Grain	Small	54,963	Subtropical	Average	69,898
Grain	Large	198,439	Vines	Small	288,617
Onions and Garlic	Small	13,917	Vines	Large	160,724

The next phase of the analysis involved a comprehensive update of the SWAP model crop budget and calibration data. All of the 22 case studies were integrated into the SWAP model input data. This involved parsing-out the data compiled by Cal Poly and integrating into the SWAP model crop budget input categories. This task also involved an important update to the crop budgets, including regulatory costs. Prior to this study, these costs were not included in the SWAP model. The ERA team went through each budget line-by-line to determine which regulatory costs were missing from the model. For example, many of the baseline crop budgets in the SWAP model include a cost for water quality coalitions. As such, this cost would be compared to the Cal Poly case study budget, and excluded if it was already included. If the default budgets in SWAP showed a different number than the case study budgets, ERA defaulted to the case study budgets. The regulatory compliance costs are summarized in the main text of this report.

The regulatory cost per acre is added to the appropriate input category. For example, dust control costs are added to the land cost category, and other operational regulatory costs are added to the other supply cost category. In this way, the cash cost of the inputs are specified to be consistent with the crop budgets, case studies, and supplemental regulatory cost data.

The SWAP model calibration is updated to include small and large farms. This involves two steps: (i) updating the Positive Mathematical Programming calibration routine to include the new crop groups, and (ii) updating the set definitions to properly calculate aggregate supply and demand. First, a series of mapping matrices are added to the model input code to map small and large farms into a universal set for the aggregate crop group. The PMP calibration routine proceeds with the small and large farms split out individually. That is, there is a crop-specific

PMP-cost function calibrated for small and large crops. Formally, for each small land large crop, the total land cost function is defined as

$$TC(x_{land}) = \delta e^{\gamma x_{land}} \tag{1}$$

where δ and γ are the intercept and the elasticity parameter for the exponential land (x_{land}) area response function, respectively. These parameters are obtained from a regression of the calibration shadow values on the observed quantities, restricted by the first order conditions, and elasticity of supply for each crop group from previous studies. For clarity, consider a single PMP cost function within a single region for a specific crop, defined as

$$MC = \frac{\partial TC}{\partial x_{land}} = \delta\gamma e^{\gamma x_{land}}, \tag{2}$$

where marginal cost equals cash cost plus marginal opportunity cost. The acreage supply elasticity, η, is

$$\eta = \frac{\partial x_{land}}{\partial TC}\frac{TC}{x_{land}}, \tag{3}$$

where

$$\frac{\partial x_{land}}{\partial TC} = \frac{1}{\delta\gamma e^{\gamma x_{land}}}. \tag{4}$$

Simplifying and noting that the logarithmic version of the equation is linear,

$$\ln(\eta\delta\gamma x_{land}) + \gamma x_{land} = \ln(R). \tag{5}$$

Thus, there are two conditions, Equations (2) and (5) which must be satisfied at the calibrated (observed) base level of land use. The former is the PMP condition and holds with equality, the later is the elasticity condition and is fit by least-squares. This is calculated for small and large crop groups individually.

Following the calibration, the regional demand elasticities are calibrated. However, since the market clearing price is determined by the aggregate supply and demand for each crop, regardless of small or large farm production, the fully-calibrated model must be aggregated back into the small and large crop group categories. The implicit assumption is that small and large producers are serving the same market. The matrices generated before the PMP calibration are used to map small and large farm production into total production, and a single downward-sloping linear demand function is calibrated for each crop. Formally, for each crop group, regardless of small or large size, a linear inverse-demand function with two parameters is used, for crop *i* in region *g* as follows:

$$p_i = \xi \alpha_i^1 - \alpha_i^2 \left(\sum_g \sum_j y_{gij} \right). \tag{6}$$

The crop price is p_i and parameters α_i^1 and α_i^2 represent the intercept and slope of the crop-specific inverse demand curve, respectively. The parameter ξ is a potential parallel shift in demand due to exogenous factors (e.g. income and population shifts) and y_{gij} is the total production by region and crop.

The end result is a model that calibrates to small and large farms individually, and thus is able to estimate the different supply response by each farm size in response to new regulations. However, the model is theoretically consistent and there is a single market-clearing price for each crop. The demand functions are calibrated using aggregate elasticity estimates, and shift linearly over time in response to changing income and population.

The final task is to develop the policy simulation model. This involves adjusting the calibrated CES program to account for factors that change over time. These generally fall into one of two categories: (i) factors that change independent of policy, and (ii) factors that change with the policy of interest. In this case, the latter is the incremental change in regulatory cost. All other adjustments fall under the former definition.

Factors that change independent of the regulatory policy are driven by the time-horizon (or level-of-development) of the model. The model developed for this ARB study operates under current conditions (2016) and future conditions (2030). Future conditions include demand shifts driven by increasing real income and changes in population, increasing real production costs, real energy cost, and perennial (orchard) retirement. Income and population shifts are based on data from the California Department of Finance. Increases in real energy costs are based on the Department of Water Resources Forward Price Projections, and affect the energy cost to lift and pressurize water for irrigation. Orchard retirement is calculated based on the average life (typically 25 – 30 years) of a representative orchard and the distribution of existing orchard ages.

The model is specified to account for drought conditions. California Department of Water Resources (DWR) data are combined with federal project and local project supply information to generate a “drought condition” baseline which can be included in the regulatory impact assessment. The drought conditions specify a surface water shortage to each region in the San Joaquin Valley based on the average surface water deliveries from all sources to that region under drought conditions. Additional groundwater pumping to offset the surface water shortage is allowed up to the maximum installed well capacity of each region. The costs of surface water shortage and additional groundwater energy pumping costs are included in the assessment.

Finally, the model is specified to include the change in regulatory cost for each crop. With endogenous prices, the objective function is to maximize the sum of producer and consumer surplus, as defined in Equation (8). Following the variable definitions specified previously, the first term of the objective function is the consumer surplus given the linear demand functions. The second term captures the region-specific gross revenue associated from deviations in regional prices from the base prices (these are denoted regional marketing costs). The third term is the region and crop specific PMP land costs. These include both the direct costs of land reported in the base data and the marginal costs inferred from the shadow values on the resource and calibration constraints. The fourth term accounts for the labor and other supply costs across all regions and crops. The fifth term of the objective function is the sum of irrigation water costs by region, crop, and water source. This term is written separately to emphasize that SWAP includes water costs that vary by source. Finally, the sixth term is the additional regulatory cost per acre by crop and region, multiplied by the total number of acres in each region. As shown, this is an explicit additional cost per acre, over and above the regulatory costs specified in the crop production budgets.

$$
\begin{aligned}
\underset{x_{gij},wat_{gw}}{Max}\ PS+CS = & \sum_i \left(\xi \alpha_i^1 \left(\sum_g \sum_j y_{gij} \right) + \frac{1}{2} \alpha_i^2 \left(\sum_g \sum_j y_{gij} \right)^2 \right) \\
& + \sum_g \sum_i \left(rm_{gi} \left(\sum_j y_{gij} \right) \right) \\
& - \sum_g \sum_i \left(\delta_{gi} \exp\left(\gamma_{gi} x_{gi,land} \right) \right) \\
& - \sum_g \sum_i \left(\omega_{gi,\mathrm{supply}} x_{gi,\mathrm{supply}} + \omega_{gi,labor} x_{gi,labor} \right) \\
& - \sum_g \sum_w \left(\varpi_{gw} wat_{gw} \right) \\
& - \sum_g \sum_i \left(RC_{gi} x_{gi,\mathrm{land}} \right).
\end{aligned} \tag{7}
$$

The fully calibrated model is used to simulate the response of the agricultural industry to changes in regulatory costs, holding other factors constant. The model is consistent with the underlying data, profit maximizing production conditions, and economic theory. It is a robust framework that evaluates the incremental cost of additional regulations, as required in an economic impact analysis.

SWAP Model Output Parameters

The SWAP model includes a full CES production technology. As such, intensive and extensive margin adjustments are simulated within the model and can be output. The focus of this study is to develop an Excel-based tool which summarizes the output of the SWAP model. The parameters included in the output include:

1. Water use, by source and region

2. Groundwater pumping and cost, by region
3. Crop acreage, by crop and region
4. Fallowing, by region
5. Gross farm-gate revenues, by crop and region

The output from the SWAP model is exported using the GDX-xls linkage. This exports the model parameters into an excel database, and this database is then used as input to the IMPLAN model and Analysis Tool, described in the following technical appendices.

Assumptions and Limitations

The SWAP model is an optimization model that makes the best (most profitable) adjustments to regulations, water supply, and other conditions. Constraints can be imposed to simulate restrictions on how much adjustment is possible or how fast the adjustment can realistically occur. Nevertheless, an optimization model can tend to over-adjust and minimize costs associated with detrimental changes or, similarly, maximize benefits associated with positive changes.

SWAP does not explicitly account for the dynamic nature of agricultural production; it provides a point in time comparison between two conditions. This is consistent with the way most economic and environmental impact analysis is conducted, but it can obscure sometimes important adjustment costs.

SWAP also does not explicitly incorporate risk or risk preferences (e.g., risk aversion) into its objective function. Risk and variability are handled in two ways. First, the calibration procedure for SWAP is designed to reproduce observed crop mix. To the extent that crop mix incorporates risk spreading and risk aversion through diversification, the starting, calibrated SWAP base condition will also. Second, variability in water delivery, prices, yields, or other parameters can be evaluated by running the model over a sequence of conditions or over a set of conditions that characterize a distribution, such as a set of water year types.

Appendix B IMPLAN Model Technical Overview

This study utilizes the IMPLAN Version 3.1 software package, in conjunction with the 2013 R3 IMPLAN data file for San Joaquin Valley counties in California. The IMPLAN software functions as an input-output economic model, which estimates the effects of assumed exogenous changes in final demand within the specified geographic region, in this case, the San Joaquin Valley. The model leverages a robust data set of national and regional economic accounts[5] that document purchasing relationships between industries through multiple rounds of spending. The software also incorporates institutional demand and inter-institutional transfers, which reflect purchases made by households and government agencies. The IMPLAN input-output model contains detailed industry account data for 536 unique industries, which can be cross-referenced to the corresponding 6-digit NAICS code industries, as well as households and assorted institutions.

The output values typically generated through IMPLAN input-output analysis include the direct, indirect, and induced economic impacts. These are most commonly expressed in terms of output, but may also be expressed in terms of employment, labor income, or one of the four components of value added (employee compensation, proprietor income, other property type income, or tax on production and imports). By definition, the direct impacts are associated with the initial dollars spent within the study area. Amounts paid to entities located outside of the study area are excluded from the analysis because these dollars do not circulate within the local economy and thus cannot be counted among the local economic impacts. The indirect impacts are estimated as industries that receive direct spending for the purchase of materials and services necessary for production. Induced impacts result from household consumption made possible by wages paid to workers and by income generated to proprietors and institutions. When combined, this cycle of spending represents an economic multiplier effect, by which a direct increase in exogenous demand, as input from the economic primary production model, results in a total economic effect that is greater than the initial direct impact.

The IMPLAN model is used to evaluate changes in economic activity in ancillary industries. Changes in economic activity are expressed in terms of the following.

1. **Output value**. The gross sales value of an industry. In crop production, for example, this measure is equal to the price of the crop multiplied by the total production.
2. **Value added**. The net contribution of an industry to the San Joaquin region economy. It is equivalent to the commonly-cited national measure of economic activity known as Gross Domestic Product (or GDP).
3. **Employment**: The number of full time equivalent jobs in a sector.

---

[5] The IMPLAN model is developed and maintained by MIG, Inc. Additional details about the IMPLAN model and underlying calibration data can be found at, http://implan.com/.

Each measure of economic activity can be decomposed into direct, indirect, and induced components. This is sometimes referred to as the multiplier effect. The components are defined as follows.

1. **Direct.** The economic effects of direct sales activity by an individual agricultural sector. For example, the farm-gate revenues from crop production.
2. **Indirect.** The economic effects of intermediate input purchases by the sector. For example, irrigation supply purchases for crop production.
3. **Induced.** The economic effects of spending by employees in all other industries. For example, farm workers purchase housing and food in the San Joaquin region.

The multiplier effect can be expressed in terms of indirect and induced, together or individually, and may include or exclude direct effects. The SWAP model output is a direct input into the IMPLAN model in the form of gross sales (farm-gate revenue) value. ERA Economics developed a mapping between SWAP model categories and the corresponding IMPLAN model categories based on the NAICS code of each industry.

Limitations

ERA Economics completed a recent analysis that found that some of the default data in IMPLAN may not be fully representative of economic activity associated with agriculture. For example, data on the ratio of value added to output is based on national benchmark data provided by the Bureau of Economic Analysis (BEA). While IMPLAN utilizes detailed county-level data to define the output values associated with agricultural commodities, the model estimates total value added based on national benchmark data provided by the BEA. The breakdown of value added is similarly derived, since the BEA reports things like proprietor income only at the farm level, with no detail provided that would help to distinguish between commodities. While the IMPLAN model offers a detailed spending profile for intermediate expenditures, the list of inputs purchased is held constant across commodities. Likewise, purchasing coefficients are held constant in proportion to the intermediate expenditure coefficient. While the model reflects the best information that is available on a national basis, it is possible to leverage the robust systematic methods associated the IMPLAN input-output software, while also improving the accuracy of the resulting multipliers by modifying it using alternative data sources that better reflect the characteristics of producers and related industries.

In short, the IMPLAN data is based on the best information available, and there is no better option available. However, future analyses can use the method developed by ERA Economics to refine the industry expenditures and resulting multipliers associated with agricultural industries in the San Joaquin Valley.

Appendix C Analysis Tool Technical Overview

The Analysis Tool is a bit of a misnomer as it includes a single tool for the SWAP model and IMPLAN output, and a series of workbooks (stochastic farm budgets) for the case studies. The reason for this is because the farm budgets are included as inputs to the SWAP model. As such, in order for the analyst to evaluate both farm-level changes in the probability of making a profit on any representative farm and regional changes in production plus links to related industries, it is necessary to provide two separate tools. In general, we refer to the Analysis Tool as the fundamental output of the analysis which includes the case study stochastic farm budgets.

Stochastic Farm Budgets

During the developmental phase of the project, the project team met a number of times with ARB staff and representatives of the agricultural community. During these meetings adjustments to the level of data exposure and the user interface of the economic models was discussed. It was agreed upon that in the interest of the confidentiality of the growers, only aggregate output would be viewable to ARB staff and the general public. Namely, the stochastic farm models would report cultural costs in aggregate (herbicide/fungicide/insecticide, irrigation costs, machinery costs), harvest costs as they are commonly tied to yield, and any other specific costs necessary for analysis. A line item for total regulatory costs is provided and a few interfacing cells for use in analysis are provided. The output generated consists of Total Revenue less Total Operating Cash Costs (before ARB policy is implemented) and Total Revenue less Adjusted Total Operating Cash Costs (after the policy is implemented).

Analysis Tool User Guide

The user selects the conditions to observe the potential impacts of a regulatory change on the yellow control tab of the Analysis Tool. The control tab includes the specifications for a baseline and alternative scenario. The user selects the year (level of development) and whether it is an average or dry year, which apply to both the baseline and alternative scenarios. The year of observation can be 2016, which shows the immediate impacts of an increase in regulatory costs, or 2030, which shows the early long run impacts of an increase in regulation. The baseline scenario always shows the projected outcome when regulatory costs are unchanged for all farms. The alternative scenario shows the simulated outcomes given a rise in regulatory costs for agricultural producers. The user selects the dollar increase per acre in regulatory costs for the alternative scenario, and whether all farms, only small farms, or only large and medium farms are affected by the increase in regulatory costs. Regulatory costs can increase by 3, 6, 15, 20, or 30 dollars per acre. These options are summarized on the orange definitions tab.

Once the user has selected the conditions for observation, the main findings are reported on the summary output tab. This tab graphs the level and percentage change in acreage, revenue, and acre-feet of applied water attributable to the increase in regulatory costs.

The blue tabs in the analysis tool show the impacts of the proposed regulatory change in more detail. Each tab has three tables. The first table is the most useful for policy analysis and it shows the level and percentage changes attributed to the change in regulatory costs. For example, on the blue Acreage tab, the first table shows the change in acreage of each crop and farm size attributable to the change in regulatory costs. Scrolling to the right, the user will find several useful graphs depicting the level and percentage changes by crop and farm size, and by region. Scrolling down, the second table shows the expected acreage of each crop on small and large farms by region in the alternative scenario, that is, the scenario with higher regulatory costs as set on the Control tab. Scrolling further down, the third table shows the acreage of each crop and farm size by region in the baseline scenario where there is no increase in regulatory costs for any farms.

The second blue tab is the Revenues tab, and it shows the expected change in farm revenues for large and small farms of each crop by region. The third blue tab, the Applied Water tab, shows the expected change in acre-feet of applied water, the fourth tab, Water by Region, shows the expected change in surface water and groundwater applied in each region, and the fifth blue tab, Groundwater Cost, shows the expected change in the groundwater costs as demands for water change in response to changes in acreage and regulatory cost changes.

Changes in regulatory costs impact agriculture directly by changing the costs of production, and they impact other sectors of the economy indirectly through changes in demand for agricultural inputs, employment, and income and expenditures in the economy. The analysis tool uses IMPLAN multipliers to measure the direct, indirect, and induced impacts of changes in agricultural production on the local economy.

The first green tab shows the change in direct, indirect, and induced economy-wide impacts of a rise in regulatory costs using IMPLAN. The first green tab is the Output tab, and the first table shows the change in direct, indirect, and induced impacts throughout the local economy. The change in direct impacts reflect the change in agricultural output attributed to the change in regulatory costs. The change in indirect impacts reflects the change in output from agricultural suppliers as a result of the increase in regulatory costs. When production costs rise and agricultural output declines, farms purchase fewer supplies. Consequently, economy-wide output decreases by more than the decrease in agricultural output alone. Finally, when output declines in any sector, incomes decrease and there are fewer expenditures in the economy at large. This is reflected in the change in induced impacts. The sum of the change in direct, indirect, and induced impacts is the total economy-wide change in output expected from a change in regulatory costs. The second and third tables on this tab show the direct, indirect, and induced economy-wide impacts of agricultural production in the alternative and baseline scenarios, respectively.

The second green tab, the IMPLAN-Employment tab, shows the expected impacts on employment in the local economy. The first table shows the change in employment attributed to the rise in regulatory costs. Direct impacts reflect changes in employment in the agricultural industry. Indirect impacts reflect changes in employment in industries that supply inputs to the agricultural industry, and induced impacts reflect changes in employment caused by changes in income and consumer spending throughout the local economy. Finally, the third green tab, the IMPLAN-Value-Added tab shows the expected impacts on value-added. That is, the impact on local industries’ output net the value of its inputs.