Program on Dairy Markets and Policy Briefing Paper Series

Questions Farmers Should Ask When Reviewing Analyses of Dairy Program Proposals

Briefing Paper Number 12-01

January 2012

Andrew M. Novakovic*

Introduction

There are lots of ideas floating around about changes to dairy policy, and other agricultural policies as well. Some proposals are very different from existing programs. Their effects and outcomes are hard to predict. Some are tweaks and variations to earlier proposals. Some are tweaks to existing programs. The implications of small changes are usually easier to predict, but some levels of assuming and guessing are always involved in any policy analysis.

Typically, we all want to find out what analysts and modelers have to say about proposals, both the big ones and the little variations. The designers of a plan want confirmation that their idea will really work. Farmers and other people in dairy markets want to get a better feel for how a plan might affect them. Policymakers need to know how much a program would cost, as well as be convinced it really is a good idea.

In reading about, listening to, or thinking about analyses of new program proposals or even just tweaks, farmers and others ought to ask some basic questions before drawing conclusions about what a study says or means.

What kind of analysis is it?

We often hear comments like so-and-so's analysis says *this*, the XYZ study says *that*, Prof. Smith did a study that says *something else*. It is convenient to refer to all of these as "studies", but it is well to remember that there are all kinds of thinking and writing that might come under the heading of "study". Setting aside reports that are unabashed advocacy or sales pieces, the more analytical reports come in four basic flavors.

1. Conceptual or logical: the analysis is based on logical, economic analysis, typically drawing general conclusions that may paint a positive or negative picture without being

* *Andrew M. Novakovic is the E.V. Baker Professor of Agricultural Economics in the Charles H. Dyson School of Applied Economics and Management at Cornell University.*

very specific or precise. This can apply to a legal or sociological analysis as well as an economic one. *An example: increasing the minimum wage rate will reduce the number of jobs for entry-level or lower skilled laborers.*

2. Philosophical or values driven: the analysis is based not on objective, dispassionate observation or criteria; rather it explicitly is based on basic values about what needs to be done or how something needs to be done. It focuses on what "should be" rather than a more detached analysis of what "will be". This kind of analysis doesn't prove something won't work so much as it focuses attention on how something works. *An example: we should solve the deficit by cutting spending, not increasing taxes.*

3. Empirical – accounting or simple spreadsheets: the analysis gets into the numbers. It might look at historical data for patterns that justify a program or that suggest how a program might work – how often it might kick in, how big an effect it might have. It might look at accounting data to suggest possible impacts by different farm types. It might use relatively simple calculations, arithmetic, to describe added costs or benefits or suggest changes to trends. *An example: if we would have had margin insurance over the last 10 years instead of MILC, farmers would have received $X instead of $Y. Or, using the accounting records of 96 farms that participated in the Cornell Farm Business Summary project over the last 10 years, stabilizing average milk prices as has been predicted under the California plan would result in an adequate cash flow coverage in 9 of the last 10 years. In actuality, cash flow coverage was adequate in 6 of the last 10 years.*

4. Statistical – modeling: this is what researchers usually think of as a "study" of a proposed program. Not all studies use sophisticated statistical or computer models. Models differ dramatically in their detail and structure. Different mathematical or statistical frameworks can be used to build a model. *An example: based on our model of milk supply and dairy product demand, the average, annual US farm price of milk from 2012 to 2017 would be 74 cents lower if Plan B was chosen instead of Plan A.*

   Some modelers know a lot about technique but not so much about what they are modeling. Some know a lot about dairy markets or whatever they are modeling but aren't so talented in the mathematical part. Some aren't especially good at either aspect. It is the rare individual who excels at both. Teams that blend some of both backgrounds are probably the sweet spot for how to do policy analysis.

In assessing the results of any of these forms of research, we must assess the trustworthiness of the researcher(s). If we decide to believe the results, we are unavoidably required to trust the researcher. Few of us have the time or training to dig into modeling details and review the design of a model. At some point the reliability of a study is simply an extension of the credibility of the researcher; do we trust the person, team or company that did the analysis.

Whose Work Can Be Trusted?

Knowing when trust will be well placed is not straightforward. Three factors ultimately go into to that determination for most of us.

a. Is the researcher competent – does s/he know the industry and market, does s/he know how to do this kind of research

b. Is the researcher reliable – is there a track record to demonstrate the quality of previous work

c. Is the researcher conflicted – is there any reason to believe that the researcher brings a bias to the analysis that is likely to affect the outcome.

Assessing competence, reliability, or conflict is more art than science. One area in which I will specifically comment is conflict. In my work, I have learned that where a person is from or for whom they work is often assumed to lead to a bias that will overwhelm integrity. I believe that is far too harsh. It is hard from a distance to know the integrity of an analyst, but I would simply caution against jumping too quickly to conclusions. This cuts in both directions. An analyst employed in industry should not be assumed to always speak what his employer wants to hear. Likewise, an analyst from academia or a government or a not-for-profit organization should not automatically be assumed to be unbiased. This is where track-record – reliability - comes into play. In a similar vein, if a study seems to provide strong support for a position, we should not be too quick to conclude that the analyst is working for or trying to curry the support of the people who favor that position. For me, a good benchmark is whether the analyst, over time, occasionally has to tell a stakeholder or natural ally something they would rather not hear. If I, as a Cornell professor, always tell Northeasterners what they want to hear, then my credibility probably will be called into question no matter how honest my analyses are. If from time to time my honest analysis isn't what my local stakeholders were hoping to hear, my credibility, my integrity, is less likely to be called into question later. This is another aspect of having a track record.

Is the tool right for the task?

Especially with empirical or mathematical models, it is important to consider whether the tool is right for the task. A hammer may be a splendid hammer, but that doesn't make it the best way to take off a nut. Sometimes we are forced to choose between a really good hammer and a poor wrench. There is no general way to say which is the better choice, but when understanding and evaluating a result, it is necessary to know something about the tool.

Many models are national. Obviously those kinds of models can't directly tell us anything about what happens in regions, states, or individual farms. Many models are based on annual data and observations. They can't directly address dynamic effects that are monthly or seasonal. If the researcher uses a national model that logic tells us will have very different affects by region, then we may or may not get a very good idea about the aggregate, national effect but we certainly won't learn anything about regional or other more granular effects that could be very significant. When we look for solutions to something that is inherently daily, weekly, or monthly, like price volatility, it is unlikely that we will get the insights we want from an annual model.

In many cases, especially when time is of the essence, the researcher often must use a model that she knows isn't perfect simply because it is available and coming up with a new model just isn't feasible. In these cases, the researcher may try to compensate for the model's inherent limitations and infer, say, a regional or monthly implication from their national, annual model by using some logical or empirical side calculations. This is often simply as good as it is going to get, and we should not be overly critical of studies that stretch the capabilities of their models in this way, when the only other choice but doing nothing. But, users should understand that at that point the analysis becomes empirical or logical. That part is actually a separate analysis and of a different nature. Those results derive from something outside the statistical model

What is the scope of the analysis?

1. Representative or Synthetic farms – these are usually hypothetical models of a farm having a structure and composition that is internally consistent and believed to be typical of a group or class of farms often seen in the US or a region – e.g. a farm representative of a medium size, employing production technologies and management choices commonly used in the Upper Midwest, and having a financial structure typical of a mid-career farmer. These analyses can be really helpful in exploring the possible effects of different proposals or a proposal with different tweaks on one or a small number of types of farms.[1]
2. Sample or Survey farms – by this I mean farm data that derives from management surveys, accounting data, tax reports or the like. They are not hypothetical farms; they are actual farms. The downside is that these data are usually messy in some way. If data are used across several years, not all farms may participate in each year. The data may be very accurate, but they may be incomplete. Accounting data, for example, may not provide much information on production practices. Data from different sources, e.g., accounting data from this location or company and management survey data from another location or university, may sound similar but actually be so different as to make them impossible to compare, or inadvisable if not quite impossible. Net Farm Income in one data set may be a very different calculation from something with the same or similar name in another data set. These kinds of analyses are appealing because they use real farm data, and they can be helpful in understanding how one plan might have very different effects on different farms or what particular farm characteristic(s) seem to lead to different effects.[2]
3. Market models – most policy models are fundamentally market models. They try to represent the supply of and demand for farm milk. They don't say anything about impacts on a farm; rather they try to model the outcome on all farms, e.g, total production, average price, maybe even average gross margins or net farm income. As mentioned above, these models are most commonly national and annual. Another kind of detail relates to product sectors or the supply chain. If a model doesn't specifically

---

[1] *The representative farm models hosted by the Agricultural and Food Policy Center and Texas A&M University are a prime example of this type of analytical tool. http://www.afpc.tamu.edu/models/flipsim/ Another example comes from the International Farm Comparison Network located at the University of Kiel. http://www.ifcnnetwork.org/en/methods/dairyfarm/index.php*

[2] *A good example of this is the Agricultural Resource Management Survey (ARMS) data developed by the Economic Research Service and National Agricultural Statistics Service of the US Department of Agriculture. http://www.ers.usda.gov/Briefing/ARMS/ Farms are surveyed annually for a broad set of information and less frequently for detailed information for a certain farm type. From these data, accounting models of farms can be specified that are considered descriptive of the nation, or a state, or a region. Another example would be farm surveys that are done by various universities or other organizations, such as Farm Credit. These surveys are usually done to provide management information to clients, lenders or the like, but they provide a wealth of information that can be used for research purposes as well.*

model the cheese sector, or the wholesale level, or export markets, then it can't directly estimate the impacts on cheese sales or prices, wholesale prices vs. farm prices or retail prices, or export sales vs. domestic sales.[3]

How much of the results depend on assumptions or guessing?

Most policy proposals take us places we've not been before. How many people will sign up for a new margin program? How many people would get caught in a growth management program? Would retailers change their pricing strategies on fluid milk if there weren't monthly announcements of a Class I price?

An ideal model would internalize all of those thorny questions and make those decisions internal to the model. This is really hard to do and typically isn't feasible even with unlimited time and resources. Instead, the researcher has to make assumptions. This is reasonable, or at least unavoidable, but it is often not appreciated that the assumptions a researcher makes, usually in good faith, can drive the results. Indeed, I have often said that the assumptions are the results. If I assume that a program is unappealing and no one signs up for it, then of course it won't work. That is not a modeling result, it is the logical conclusion of my basic assumption. Of course, it can just as easily work the other way. The only way a modeler can deal with this is to offer analyses using different assumptions that give a sense of boundaries – high to low, strong to weak, big to small. This makes for more results to sift through and can be confusing, but it is a tacit admission that there are often many very important elements in how a new plan would work or how farmers or others would respond that the modeler simply cannot know or objectively predict.

When comparing results across studies, whether the studies seem to agree or disagree, it is really important to look for the important assumptions that the researchers inevitably had to make. Studies that seem to disagree may not actually disagree in how they see a plan would work so much as disagree on a critical assumption – for example, farmer's would sign up in droves or farmer's would not.

Is there a routine bias in assumptions? Are we more likely to be optimistic or pessimistic? I don't know for sure, but I think proponents of a plan have a built in tendency to be optimistic. Whether or not researchers have a predictable bias in this way, I don't know. I would observe that it has been rather typical lately that new agricultural programs have been created and justified on the basis of an overly optimistic assumption about farmer participation. This has been true for LGM-Dairy. It was true for the ACRE program introduced in the 2008 Farm Bill.

Some basic questions about assumptions and analyses

When practitioners look at analyses of new dairy proposals, they should ask some of these basic questions (as applicable):

1) What is assumed about participation rates

---

[3] *Numerous dairy market models have been developed over time. Some were created just for a specific project. Few survive as tools that are routinely maintained and available for use in short order. Two examples are the dairy model maintained by Prof. Scott Brown at the University of Missouri [http://web.missouri.edu/~browndo/] and the two major models maintained by Prof. Charles Nicholson at the California State University at San Luis Obispo [http://agb.calpoly.edu/content/about/faculty_staff/Nicholson].*

a) Participation in terms of Yes or No vs.

b) High, medium or low participation rates

2) What is assumed about participation levels

a) Minimal plan, maximal plan, or something in between (e.g., \$4, \$6, \$8)

b) Coverage (e.g., 10%, 50%, 75% of my milk)

c) Keep track of whether 50% participation means 50% farms or 50% of milk produced (which could be more or less than the milk from 50% of the farms, depending on which farms participate)

3) Are essential variables determined internally or as a scenario?

a) Milk prices are usually determined internally to a market model but imposed as a scenario (explicit assumption) in a farm model

b) Feed prices and prices of other inputs are usually based on a projection from another study or based on some historical trend or pattern. Obviously results from two studies can differ simply because one assumed "low" feed prices and the other assumed "high" feed prices. One needs to ask whether the differences in results are simply because of different assumptions about feed prices. In this case different "results" or conclusions from an analysis might disappear or diminish when the same assumption is used.

c) Other economic parameters may or may not be internal to a model. A different issue is whether they are directly included at all. A good example might be measures of economic growth or well-being that could influence demand, e.g. household income, unemployment, GDP growth and so on. Some models will have a variable like income in a demand equation. The level of income over time is likely to be based on some other model projection or a simple growth factor. In this case we can say that income growth effects demand, but, just as is true with feed costs, it is helpful to know if one analysis uses a rosy income scenario and another uses a pessimistic scenario. An altogether different issue arises when a model doesn't include the variable, say income, at all. In this case, when we ask what would the model predict if economic growth is stronger, the answer is that the model has no way of taking that into account. This may sound hard to believe, but all models have to define their boundaries. Even something as basic as income or economic growth may not be included in a model. Consider exchange rates, which impact how US prices look to someone in a foreign country or vice versa. This is an increasingly important variable, but its importance is new enough that few dairy models have figured out how to take it into account.

d) Is the analysis a simple projection or does it look at statistically expected values?

i) Most analyses involve straightforward projection. These analyses typically have a baseline and then look at changes in outcomes caused by changes in a program. If we change X, then our model predicts that in 2015 the price of milk would be this, production would be that, and so on. There is only one set of outcomes over a projection period for each scenario or set of policy actions. The only way you get a different outcome, say the price of all milk in 2015, is if you change something in the basic analysis. This could be something about the program, e.g., assuming higher participation rates; or it could be something about the market environment, e.g. higher feed prices.

This approach is the typical approach for policy analysis. These are usually studies that are done to see whether "the plan works" over a foreseeable, intermediate time frame.

ii) In some cases, analyses are "stochastic". This means that even when we assume basic parameters like a participation rate or income growth, we recognize that over time there are a bunch of factors we can't control; so there could be wide variety of outcomes. The price of milk isn't projected to be just one number in 2015, instead we say there is a 10% chance it will be $23 or higher, a 20% chance it will be $14 or lower, and so on. In this approach, the interaction of supply and demand does not determine a price outcome. Instead we say that prices are to some extent a question of probability. Typically, there will be a mathematical supply relationship that results in different production and sales outcomes caused by a randomly higher or lower milk price.

This approach is used by the Congressional Budget Office to see if "the plan saves money". When "scoring" a new proposal, CBO does not project a price of milk for 2015 so much as they say: given the range of prices that could occur in 2015 the probability of spending government money to implement the proposed program varies. If you combine the magnitude of a payout in a low price event with the probability of a low price event, along with all the other possible outcomes, you can calculate an expected payout.

This isn't the same as a discrete prediction, but you can't tell that just by looking at a table or chart of results, in only one number is printed for each year. This one number is a "most likely" or average but this isn't the same as a standard, one-number projection. In a stochastic analysis, the expected price might exceed an action trigger, this is it might be too high to trigger a government payment; yet, the analysis may concluded that there is a cost to the program. The cost derives from a probability that prices might be below the expected value.

When CBO establishes a Baseline, it provides base projections of supplies, demands, and prices for milk and dairy products. When it scores a policy proposal, it provides information on how much Plan A or B will cost the government compared to the Baseline. It does not provide estimates of prices and quantities, although those calculations stand behind the cost estimates.

All of this makes comparing studies using these different methods very tricky. It is good to keep in mind, at a minimum, that CBO uses only stochastic forecasting when they do a budget analysis or "score". Thus, their numbers will always look a little different from other policy studies simply because of that.

What is the scope or purpose of the analysis?

It is always wise to look at whether a study is a broad, general purpose analysis or a more narrowly targeted analysis.

The best example is the CBO scoring analysis. CBO is obliged to tell Congress what it thinks a new program will cost the federal government. Their modeling work is oriented to providing that result. It is not CBO's job to determine if a plan "works". They assume that the plan does what the Congress wants it to do. When CBO says a plan results in lower government cost then the current program, this might be because the new plan works better and more efficiently, or it may be that it doesn't work at all – it's just cheaper.

A model that is designed to see whether farmers can effectively use distiller's grains or another model that helps farmers decide whether to hedge milk prices both deal with issues of risk management and may have something useful to say with respect to the merits of a new proposal, but such models are obviously not able to provide complete insights about all the effects of a risk management program.

Additional Caveats and Cautions

Don't assume that a study report that leads with "it works" automatically means the plan can be afforded or that it works equally well for everyone.

Don't assume that a study report that leads with "it saves money" automatically means that the plan also achieves the market or firm objectives.

Don't get hung up on comparing results across studies unless you know that they made the same assumptions about program design and economic situation. Instead, look at the baseline in each study and how the new program achieves different results from the baseline. Model results will be internally consistent. If a model says Plan A will result in lower prices, that general result should hold true whether the researcher uses a high feed price scenario, or a low participation scenario, or whatever. There can be exceptions to this rule but typically we can be confident about relationships like "lower" even when we might not be so sure about $2 lower. When we try to compare one study's $2 lower with another study's $4 lower, then we need to figure out if there was something simple, like feed price assumptions, that explain the difference or whether it is something related to how the program is estimated to work.

When a report says that Plan A "saves money" or results in a higher price, remember to check what this change is relative to. Typically the something that is higher or lower is so relative to the projected base, not last year or a current number. In this way, a new proposal may actually cost more money than was spent historically but be calculated to "save money" because the government would have spent even more if it stayed with the existing plan. Similarly a plan that results in a "higher price" may actually have prices that are lower than today's price but are higher than the expected future price without the plan.

Don't assume you will get all the benefit and someone else will bear all the cost, or vice versa. Too often we brush off program effects by saying that it only affects big guys, or cheese guys, or highly leveraged guys. That may be true, but it may not be as true as we think.

Look at the big picture. Plan A may have lower prices than Plan B but that doesn't necessarily mean that net farm income, return on equity, total income or the like are moving in the same direction. Maybe you are actually better off in the long haul with the plan that has lower prices because there is an offsetting factor.

Look at the big picture. Don't look at an investment penalty in the short term without comparing it to an investment benefit in the longer term.

Look at the big picture. Something may be a great deal for farmers and be quite the opposite for consumers. If so, it may be that the economic benefit of the plan is short-lived, never mind the political repercussions. Thirty dollar milk sounds terrific to a dairy farmer, but if it eventually results in a big decline in demand then it may not be so terrific after all.

Final Words

Studies and analyses of new program proposals play a very valuable role in helping us decide whether or not we like a proposal or think it will be helpful. But, be careful in looking at the results of several different analyses. You might make the right decision for the wrong reasons, or you might make the wrong decision.

If you are a numbers person, jump in, but pay close attention to how those numbers are generated. If you aren't a person who likes to dive into a table full of numbers, don't just go with your gut. Listen to people you trust who do delve into the numbers. Honest analyses should get you thinking. If an analysis doesn't seem to agree with your current thinking, don't write it off as wrong or biased until you have challenged your own thinking a bit.

As the saying goes, a certain amount of skepticism is healthy. I don't know if unbridled pessimism, or optimism, is healthy, but we probably shouldn't assume an analysis is good just because we like what it says, or bad because we don't.

A good study doesn't give us all the right answers so much as it helps develop the right questions. A study that comes up with answers we didn't expect shouldn't be cast aside as a bad study until we have challenged ourselves to think hard about whether we might have overlooked something or been too optimistic (or pessimistic) in our original thinking about a plan.