The Sources of Measured Agricultural Productivity Growth[1]

Robert G. Chambers[2 3]

June 15, 2011

[1]Preliminary version for purposes of discussion. Please do not cite without the author's permission.

[2]Professor, University of Maryland, College Park, 20742

[3]I would like to thank Simone Pieralli for performing the calculations that are reported here and both he and Ariel Ortiz-Bobeia for conversations that have contributed greatly to this paper.

Almost half a century ago, Griliches (1963) presented his pioneering enquiry into the sources of US agricultural productivity growth. This study was the culmination of a series of studies by Griliches and his mentor (Griliches, 1957, 1960, 1963; Schultz, 1947, 1956, 1958) on the factors explaining a startling empirical observation by Barton and Cooper (1948):

> Output per unit of all inputs has shown an upward trend since World War I, as a result of a remarkable stability of total inputs and a steady upward trend in the volume of farm output....These considerations are extremely important in analyzing the changes in economic conditions of agriculture over the last quarter century.

Schultz (1956) defined an *ideal input-output formula* as "...one where *output over input*, excluding of course, changes in their quality, stayed at or close to one", and he attributed this definition in a footnote to, the then graduate student, Griliches. Seemingly they saw the goal of the productivity analyst as eliminating the residual between input and output growth that Abramovitz (1956) had recently called a "measure of our ignorance". For US agriculture, Griliches (1963) suggested that making adjustments for changes in the quality of inputs (particularly, labor and capital) and correcting for scale economies, which his empirical analysis suggested then persisted in US agriculture, would achieve this goal.

Those quality corrections to inputs have long since been incorporated in official United States Department of Agriculture (USDA) TFP calcualtions. Moreover, after that time US agriculture underwent a dramatic transformation characterized by a remarkable concentration of its traditional "small family farms" into fewer and fewer commercially viable operations. Even so, as Figure 1 attests, Barton and Cooper 's (1948) observation is still valid. The Schultz-Griliches *ideal* input-output relationship has not been achieved, even though US agriculture has now witnessed almost a century of marked stability in aggregate input use. The residual remains

The longer the residual persists, the more productivity analysts try to explain it. By far the most common approach is to regress some measure of agricultural productivity, typically multi-factor or total factor productivity, on potentially causal factors such as public expenditures on agricultural research and development.[1] Because most computations of total factor productivity rely explicitly upon the economic theory of index numbers, this analysis often attempts to explain productivity change under the maintained hypothesis of efficient and economically rational behavior.[2]

[1] An excellent summary as well as guide to many of the econometric, philosophical, and practical issues involved in this particular mehtod of research evaluation can be found in Alston, Norton, and Pardey (1995).

[2] Another approach is induce rates of technical change from estimated cost or profit functions for representative

I have no problem, *per se,* with the assumption of either efficiency or rationality. In fact, my prior belief is that farmers are usually highly rational and very technically efficient. Instead, I question whether the observed behavior of our empirical TFP measures jibes with both rational behavior and the physical realities of farming.

The kernel of my concern is nicely captured by Figure 2, which depicts rates of TFP change for US agriculture for the last six decades. As I have pointed out elsewhere (Chambers, 2008), calculated TFP changes for US agriculture are highly variable exhibiting both significantly positive and significantly negative rates of change. This phenomenon is most marked starting in the early 1980s.[3]

In one sense, what's happening is obvious and, perhaps, even trivial. Clearly, agriculture is highly variable. Farmers make many production decisions before a host of potentially important and variable factors, such as weather, are known. Thus, the trivial answer: It's the weather. And I, for one, would not disagree that an important component of this variability is weather induced. In fact, I would go further. Because input use is so stable in US agriculture and weather is not, I would argue that *more than anything else*, *Figure 2 represents a weather index expressed in output terms.* The trouble, of course, is that Figure 2 is not supposed to be a weather index. Rather, by the economic theory of index numbers, Figure 2 should depict changes in the locations of production frontiers.

It is precisely such situations that have led many researchers to insist upon a stochastic frontier approach to measuring productivity. But for me, that approach also can seem a tad too facile in the following sense. Clearly agriculture and other industries are affected by stochastic factors or in a real-valued world by random variables. No argument. But the real question is how to represent those stochastic factors.

In decision theory, those stochastic factors are represented, following Savage (1954) and Arrow and Debreu, by first defining a "state space", where each element of that state space corresponds to a complete description of the world under all possible conditions. Then acts are defined as maps

agents and then use secondary regression analysis to decompose the factors contributing to the evolution of those rates. Of course, this also implicitly assumes rational behavior.

[3]This phenomenon is not peculiar to US agriculture (Chambers, 2008) or even to agriculture in general, as any glimpse at sectoral TFP growth rates for the United States and other developed nations will reveal. But that said, my concern is with agricultural technologies, and in the remainder of the paper I will stick to them even though the principles that I elicit can clearly be applied to TFP measurement in other sectors, such as finance, with exhibit high degrees of variability.

from that state space to some outcome space. When that outcome space is the set of reals, those maps are random variables. The domain of technologies and preferences in this setting are random variables. Or put another way, the technology is defined over random variables, but it is not itself random.[4] Once the state's realization is known everything is deterministic. Where things become perhaps too facile for me is when those random variables are defined in econometrically convenient terms without a proper accounting of how they relate back to the underlying state space, and the physical reality we attach to that state space.

That's the simple task I set myself. Try to be more careful about incorporating information about the state space into representations of stochastic technologies. If done properly, it might allow us to parse measured productivity growth into measures of frontier shifts due to changes in technical knowledge and shifts associated with realizations of "good" or "bad" states of Nature. Thus, in a sense, I see the goal of the paper as similar to Griliches and Schultz: try to explain as much of the residual, be it called technical change or efficiency change, as possible in terms of the physical technology. Or, in more econometric terms, filter the noise process of some of its explainable heterogeneity. And I perceive this as a step that should be taken before one tries to explain whatever residual finally remains.

So, in what follows, I start by looking at an aggregate agricultural data set,[5] drawn from the 48 continental US states covering the period 1960 through 2004. Using data envelopment analysis measures, I follow Färe, Grosskopf, Norris, and Zhang (1994) first to construct measures of intertemporal productivity growth for the meta-technology facing these 48 states over time, and second to decompose those measures of productivity growth into a 'technical change' component and an 'efficiency change' component. To keep the presentation of empirical results manageable, the empirical discussion focuses on two representative, but distinctly different, US states: California and Iowa.

When that is done, it turns out that a large component of agricultural productivity change (particularly in Iowa) is attributed to 'efficiency change'. Because no overarching theory of inefficiency yet holds sway, that observation begs an explanation. One way to explain it is to regress efficiency scores on explanatory variables. My interpretation is that it is further evidence that something is missing from the model. I find it hard to credit that producers who are on the frontier at one point

[4] I won't let this fact stop from me talking about stochastic and nonstochastic technologies in what follows. In this case, it seems the linguistic convenience of the abuse of terminology outweighs any linguistic imprecision that results.

[5] These data were generously supplied by V. Eldon Ball and consist of state-level information on aggregate output, capital, land, labor, and materials in agriculture.

in time suddenly forget where the frontier is or are so surprised by technical innovations that they routinely lag behind it.

The analysis that forms the core of the paper is next. I attempt to incorporate, in as simple a manner as possible, the Arrow-Debreu notion of a state-contingent technology into an empirical representation of the agricultural meta-technology and to use that representation to decompose measured productivity growth into three components: efficiency change, technical change, and, for lack of a better term, heterogeneity, which in our case is attributable to weather.[6] Those measures are then compared to the results obtained from the more traditional Färe et al. (1994) analysis.

This is a simple-minded paper. True enough. There are virtually no mathematics. The goal is not to develop more sophisticated means of examining existing data. Instead, it's to use simple methods from standard efficiency analysis, albeit viewed from a different theoretical perspective, to learn more from the data that we already possess.

1 Nonstochastic Productivity Measures

Our starting point is perhaps the simplest possible notion of a technology. There is a single output, and production possibilities at time t are governed by

$$T(t) = \{(x, y) : y \leq f(x, t)\},$$

where $x \in \mathbb{R}_+^N$ denotes inputs controlled by the producer, $y \in \mathbb{R}_+$ denotes output, t now indexes the state of knowledge available at time t, and $f(x, t)$ represents the production function.

The associated productivity index for observations (x^0, y^0, t^0) and (x^1, y^1, t^1) is defined, following Caves, Christensen, and Diewert (1982) and many others, as the geometric average of two Malmquist productivity indices:

$$P^{t^0,t^1}(y^0, x^0; y^1, x^1) := \left(\frac{y^0}{f(x^0, t^1)} \frac{f(x^1, t^1)}{y^1} \times \frac{y^0}{f(x^0, t^0)} \frac{f(x^1, t^0)}{y^1} \right)^{\frac{1}{2}}.$$

Following Färe et al. (1994), I decompose this measure into two components:

$$P^{t^0,t^1}(y^0, x^0; y^1, x^1) = E^{t^0,t^1}(y^0, x^0; y^1, x^1)\, T^{t^0,t^1}(y^0, x^0; y^1, x^1),$$

where

$$E^{t^0,t^1}(y^0, x^0; y^1, x^1) := \frac{y^0}{f(x^0, t^0)} \frac{g(x^1, t^1)}{y^1}$$

[6] In a related paper, I show with a co-author how similar methods can be applied when the source of the heterogeneity is not weather but driven by other environmental factors (Chambers and Kafkalas, 2011).

is the 'efficiency' component of the productivity index and

$$T^{t^0,t^1}\left(y^0,x^0;y^1,x^1\right) := \left(\frac{f\left(x^0,t^0\right)}{f\left(x^0,t^1\right)}\frac{f\left(x^1,t^0\right)}{f\left(x^1,t^1\right)}\right)^{\frac{1}{2}}$$

is the technical-innovation or technical change component of the productivity index.

As already mentioned, our data set consists of a panel of observations on aggregate output and inputs for 48 US states over the 1960 to 2004 period. Let $\left(y^{kt},x^{kt}\right)$ represent the input-output vector at time $t=1,2,...,45$ for state $k=1,2,...,48$. The standard DEA approximation to $T\left(t\right)$ with free disposal of inputs and outputs, constant returns to scale, and no technical regress is

$$T^D\left(t\right) = \left\{\begin{array}{c}(y,x) : y \leq \sum_{k=1}^{48}\sum_{v=1}^{t}\lambda_{kv}y^{kv}, x \geq \sum_{k=1}^{48}\sum_{v=1}^{t}\lambda_{kv}x^{kv}, \\ \lambda_{kv} \geq 0, v=1,...,t\end{array}\right\}.$$

The corresponding approximation to the production function is

$$f^D\left(x,t\right) = \max\left\{\begin{array}{c}\sum_{k=1}^{48}\sum_{v=1}^{t}\lambda_{kv}y^{kv} : x \geq \sum_{k=1}^{48}\sum_{v=1}^{t}\lambda_{kv}x^{kv}, \\ \lambda_{kv} \geq 0, v=1,...,t\end{array}\right\}.$$

1.1 Empirical Results for State-level Intertemporal Productivity Indices

The empirical focus, throughout the paper, is on intertemporal productivity measures for California and Iowa. Figure 3 presents information on aggregate agricultural output and input for California, while Figure 4 presents the same information for Iowa.[7]

Several patterns are noticeable. In 1960, Iowa and California were roughly the same when compared in terms of aggregate agricultural inputs and outputs. But by 2004, California's aggregate output was almost 80% higher than Iowas. And while California's aggregate input use had grown, Iowas had fallen after peaking in the mid 1970s.

Figure 5 presents year-to-year intertemporal productivity indices computed for both states on the basis of the DEA approximation to the meta-technology. It is apparent, as also evidenced by Figure 3, that California experienced almost continual productivity growth, with only a few instances of productivity setbacks.[8] Iowa has a more checkered history, particularly since the 1980s. Its productivity growth often exceeds that of California, sometimes by as much as 10%, but it also has far more, and more drastic, productivity set backs than California. Despite these differences,

[7] In all cases indexes are defined taking Alabama 1996 as 1.

[8] For the form in which the indices are calculated, numbers less than one signal productivity growth (less input per unit of output) while numbers greater than one (more input per unit of output) imply productivity setbacks.

the long-run patterns are moderately close. For California, the average of the intertemporal productivity indices is approximately .97, while for Iowa, it is slightly over .98. On average, California grew faster by about 1% per annum. But, as is visually apparent, Iowa's productivity exhibited far more variability. For example, the coefficient of variation for its intertemporal productivity index at .11 was about 4 times as large as that for California.

A pat explanation is to hand: California's climate is more moderate than Iowas. And, as such, one expects more variable traditional measures of productivity growth. That is, in fact, one of the two reasons why I chose these two states to focus upon. The other is that both California and Iowa are traditionally thought of as being among the most "efficient" of the agricultural states in the United States. They have very different forms of agriculture, but one routinely expects them to be among the industry leaders.

Figure 6 presents Iowa's computed year-to-year intertemporal productivity indices and year-to-year technical change indices graphed against the same axis for 1980 forward. Over that period, Iowa experienced fairly steady technical change on the order of about 1 to 3% per year with the exception of 1995 to 2000, during which there was almost no technical progress. Iowa's intertemporal productivity index over the same period oscillated between extremes of slightly over 1.2 and slightly below .8 suggesting productivity changes of over 20%. The explanation for that oscillation, in terms of the current decomposition, are "efficiency changes".

One interpretation of efficiency change is in terms of catching up to a continually expanding meta-frontier that grows with technical innovation and progress (Färe et al., 1994). For example, if in period t^0, a particular state is operating inside its technical frontier, then one component of $E^{t^0,t^1}\left(y^0, x^0; y^1, x^1\right)$,

$$\frac{y^0}{f\left(x^0, t^0\right)},$$

is less than one. If it moves to the frontier in period t^1, that is catches up to it, the remaining component of the efficiency index

$$\frac{f\left(x^1, t^1\right)}{y^1},$$

equals one. Consequently, the intertemporal efficiency index is less than one. On the other hand if the country is on the frontier in period 0 and then "falls behind" as the frontier shifts out in period 1, the intertemporal efficiency index is greater than one. In what follows, we shall refer to situations where the intertemporal efficiency index is less than one as *catching up* and situations where it is greater than one as *falling or lagging behind.* Apparently Iowa has been doing a lot of lagging behind and catching up since 1980.

Another possibility is that there is some form of heterogeneity that we are not capturing with the DEA hull. More intuitively, a residual exists, which we call efficiency change, that the intersection of the data and the current model cannot explain. One might be tempted to dismiss this as a typical example of DEA's shortcomings, but, as I have emphasized earlier a similar pattern emerges in USDA TFP calculations.

Turning to California, Figure 7 suggests a pattern that is quite different than in Figure 6. Now there is closer agreement between the calculated intertemporal productivity indices and the calculated technical change indices. This, of course, signals that the residual efficiency index is at or close to 1 throughout these 24 years.

2 Stochastic Technologies and Productivity Measurement[9]

The problem this section tackles is to incorporate the physical reality that farming takes places under conditions of uncertainty into our model. The approach taken is that taken by Arrow and Debreu, and much later by John Quiggin and myself (Chambers and Quiggin, 2000). Uncertainty is represented by a set of states, Ω, from which a neutral player, 'Nature', makes a draw. Ω provides a comprehensive and mutually exclusive description of the possible states of the world to which the producer is exposed, but which are beyond his or her control. Random variables are defined as (measurable) maps from the set of states, Ω, to the reals. Random variables can thus be thought as vectors $\tilde{f} \in \mathbb{R}^{\Omega}$ where

$$\tilde{f} = \{f(s) : s \in \Omega\},$$

and $f(s)$ denotes the realized value (*ex post value)* of the random variable if 'Nature' chooses s.

The production technology involves using multiple inputs to produce a single stochastic output.[10] That stochastic output is represented by the random variable $\tilde{z} \in \mathbb{R}_{+}^{\Omega}$. At time t, the technology in a convenient abuse of notation is represented by a set

$$T(t) = \{(\tilde{z}, x) : x \text{ can produce } \tilde{z} \text{ at time } t\},$$

where $\tilde{z} \in \mathbb{R}_{+}^{\Omega}$ denotes the stochastic output, and $x \in \mathbb{R}_{+}^{N}$ denotes the inputs that the producer chooses.[11] The interpretation of the technology is as follows. Before the producer knows Nature's

[9] Readers wishing more detail on the basics of this approach to specifying stochastic production technologies can refer to Chambers and Quiggin (2000).

[10] We stick to the single output case here for consistency sake. The basic concepts easily extend to the multiple-output case.

[11] $T(t)$ as defined for the nonstochastic technology is a subset of $\mathbb{R}^{N+1}$. Here it is a subset of $\mathbb{R}^{\Omega} \times \mathbb{R}^{N}$.

draw $s \in \Omega$, he or she picks $(\tilde{z}, x) \in \mathbb{R}_+^{\Omega} \times \mathbb{R}_+^N$ from within T. Then Nature makes her draw. If the draw is s, the realized state is $s \in \Omega$, and the realized (*ex post*) output is $z(s)$, while if $s' \neq s$ is drawn, the *ex post* output is $z(s')$.

Making this representation both empirically operational and compatible with what has gone before requires some further assumptions that impinge in important ways on the producer's freedom of choice in choosing his or her input-stochastic output mix. Chambers and Quiggin (2000) discuss these issues at length. For our purposes, it suffices to assume that we can represent the technology by the set

$$T^{\Omega}(t) = \{(\tilde{z}, x) : z(s) \leq g(x,s,t), s \in \Omega\},$$

where g is a real valued function. This representation corresponds to the *state-contingent production function* axiomatically studied by Chambers and Quiggin (2000).

A state-contingent production function has a number of advantages for applied work. Most importantly, it permits the use of *ex post* observations on output to construct an empirical approximation to the technology using DEA methods. Second are its intuitive advantages. It corresponds in a reasonable fashion to *stochastifying* the technology used earlier. And it is by far the most common empirical representation of stochastic technologies once sufficient structure is placed upon Ω. But, it has shortcomings. Chambers and Quiggin (2000, 2007) discuss these at length, and O'Donnell, Chambers, and Quiggin (2009) have shown that if the true technology is not of this form, empirical representations based upon it can be seriously biased in approximating the frontier and measuring efficiency.

In our terminology, $\tilde{z}$ is a random variable because it belongs to $\mathbb{R}_+^{\Omega}$. Its maximal realized values are defined, for given x and state of the technology, t, by $g(x, s, t)$ which maps Nature's draw from Ω into a real number that represents an upper bound on the producer's choice for output in that state of Nature. That upper bound function, as a map from Ω to the reals, defines a random variable. As with our nonstochastic technology, the producer can choose to be technically inefficient in any state of Nature.

It is important to note, however, that even though $\tilde{z}$ is random, we have not and will not attribute any probability distribution to it. All our results are free from any assumption on an associated probability measure. So, in this sense, this representation of the technology is different from the more commonly encountered stochastic frontiers frequently encountered in efficiency and productivity analysis. One can, however, arrive at the stochastic frontier model by associating s with an econometric error term, and then assuming a probability distribution for that error term.

That is not the approach that I use. Instead, I follow O'Donnell and Griffiths (2006) and several others and identify Ω with physically observable outcomes that are beyond the control of the producer, and that for the most part only become known to the producer after his or her input decisions are made.[12] More on that shortly.

Now, I introduce a productivity measure that accommodates the presence of s and Ω. With yet another abuse of notation, that index is defined:

$$P^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right) := \left(\frac{z^0}{g\left(x^0,s^1,t^1\right)}\frac{g\left(x^1,s^1,t^1\right)}{z^1}\frac{z^0}{g\left(x^0,s^0,t^0\right)}\frac{g\left(x^1,s^0,t^0\right)}{z^1}\right)^{\frac{1}{2}}, \tag{1}$$

where z^k corresponds to the observed output for observation k. As before, the productivity index is the geometric average of the ratio of two Malmquist indices.

This index can be decomposed into three components:

$$P^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right) = E^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)\Omega^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)T^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right),$$

where

$$E^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right) := \frac{g\left(x^1,s^1,t^1\right)}{z^1}\frac{z^0}{g\left(x^0,s^0,t^0\right)},$$

$$\Omega^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right) := \left(\frac{g\left(x^0,s^0,t^1\right)}{g\left(x^0,s^1,t^1\right)}\frac{g\left(x^1,s^0,t^0\right)}{g\left(x^1,s^1,t^0\right)}\right)^{\frac{1}{2}},$$

and

$$T^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right) := \left(\frac{g\left(x^0,s^0,t^0\right)}{g\left(x^0,s^0,t^1\right)}\frac{g\left(x^1,s^1,t^0\right)}{g\left(x^1,s^1,t^1\right)}\right)^{\frac{1}{2}}.$$

As before, $E^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)$ and $T^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)$ represent an efficiency index and a technical-innovation index, respectively. The new term, $\Omega^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)$, is the geometric average of two separate indices of the effect that different draws by Nature, s^0 and s^1, have on maximal feasible output. One component holds the input bundle constant at x^0 while evaluating

[12] Another reason for the choice of the stochastic production function specification of the technology is that it makes the timing of the observation on Nature's draw from Ω less critical in technical efficiency analysis. For example, suppose that the producer were allowed to observe Nature's draw before choosing his input and output. If the technology assumes the form in this paper, which Chambers and Quiggin (2000) dub 'output cubical', the interpretation would then be that the producer faces a state-specific production function which governs his technical choices. Timing of the observation of Nature's draw is, of course, critical to the input and output choices a producer makes. But because we are only examining technical possibilities in this paper and not optimal economic choices, whether the producer knows the draw or not is not essential.

the technology at the state of knowledge indexed by t^1 and then evaluates the frontier shift caused by different realizations of Nature's draw. The other component holds the input bundle at x^1 and the state of knowledge at t^0 to evaluate the frontier shift.[13]

With an index and the corresponding decompositions defined, the next step is an empirical representation comparable to the empirical representation used earlier. Formally, I need a comprehensive description of all possible states of the world that can be implemented empirically. That's not possible practically. Therefore, a compromise is necessary, and of needs I must settle for something that describes production conditions relevant to the producer and that are beyond his control.[14]

For this application, I take Ω to be a subset of $\mathbb{R}^2_+$ that is associated with observations on two climatic variables: one representing solar radiation and the other moisture. These data, which are also state level, are drawn from Schlenker and Roberts (2008) and correspond to degree days between 8^o and 30^o Celsius and inches of precipitation between the months of March and August. As our specification indicates, these observations are incorporated directly into the definition of the frontier as though they were 'inputs'. What kind of inputs they are is another matter. In particular, it is apparent that precipitation does not satisfy global free disposability.

At first blush, it might seem that the degree day measure might satisfy free disposability. In fact, in the existing economic literature on climate change from which these variables are drawn, degree days between 8^o and 30^o are referred to as 'beneficial' degree days while degree days defined for temperature ranges outside that span (for example, above 34^o C) are often viewed as detrimental (Schlenker, Haneman, and Fisher, 2005; Deschênes and Greenstone, 2007).

Things may not be so simple.[15] First, the notion of degree days, which originated in the biological literature, was originally intended for other purposes. Physical scientists found that the

---

[13] This decomposition is, in fact, arbitrary. It was chosen to afford an easy comparison with the Färe et al. (1994) decomposition that was used earlier. Therefore, $\Omega^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)$ is most properly thought of as a weather residual. To see the arbitrariness, note that one could just have easily taken

$$\Omega^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)=\left[\frac{g\left(x^0,s^0,t^0\right)}{g\left(x^0,s^1,t^0\right)}\frac{g\left(x^1,s^0,t^1\right)}{g\left(x^1,s^1,t^1\right)}\right]^{\frac{1}{2}},$$

and induced a balancing technical change index that is different from $T^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)$. The problem of defining 'path independent' decompositions is treated in Henderson and Russell (2005).

[14] By beyond control, it is not meant that the producer cannot prepare for different realizations, for example, by building greenhouses or installing irrigation systems. Rather, it means that he or she cannot affect Nature's choice.

[15] This section reflects lessons learnt in discussions with Ariel Ortiz-Bobeia and from reading Ortiz-Bobeia (2011).

rate at which a plant progresses through its stages of development is approximately linear in a measure of thermal time. This is the relationship captured by the notion of a degree day. So while the degree-day measure is related to the rate at which the plant progresses through these stages, it is not necessarily related to the resulting harvested mass, which is our traditional output measure (Ortiz-Bobeia, 2011). Moreover, degree days here correspond to the March to August period. For much of the United States, March to August does cover most of the growing period, but not for all of the United States.

The DEA approximation to $T^{\Omega}(t)$ under the assumption of weak disposability of s is

$$\hat{T}^{\Omega}(t) = \left\{ \begin{array}{c} (z,x,s) : z \leq \sum_{k=1}^{48} \sum_{v=1}^{t} \lambda_{kv} z^{kv}, x \geq \sum_{k=1}^{48} \sum_{v=1}^{t} \lambda_{kv} x^{kv}, s = \sum_{k=1}^{48} \sum_{v=1}^{t} \lambda_{kv} s^{kv}, \\ \lambda_{kv} \geq 0 \end{array} \right\},$$

and the approximation to the state-contingent production function is

$$\hat{g}(x,s,t) = \max \left\{ \begin{array}{c} \sum_{k=1}^{48} \sum_{v=1}^{t} \lambda_{kv} z^{kv}, x \geq \sum_{k=1}^{48} \sum_{v=1}^{t} \lambda_{kv} x^{kv}, s = \sum_{k=1}^{48} \sum_{v=1}^{t} \lambda_{kv} s^{kv}, \\ \lambda_{kv} \geq 0 \end{array} \right\}.$$

Some differences between this representation of a stochastic technology and more familiar specifications of stochastic production functions are worth noting. First is the absence of any assumption on an underlying probability measure. That's not to deny that one exists, but rather a specific assumption is superfluous. Second, the randomness in the technology is driven by Nature's draw, which in the current case is two-dimensional, rather than one dimensional. Imagine trying to approximate $g(x,s,t)$ as follows

$$g(x,s,t) = m(x,t,\varepsilon)$$

where ε is a one-dimensional random variable intended to capture the effect of Nature's draw. Specifically, because ε is a random variable, it must be a map or function from Ω to the reals. A more suggestive notation is thus

$$g(x,s,t) = m(x,t,\varepsilon(s)).$$

This may or may not be plausible. *A priori*, it is difficult to judge. But it does imply that the various components of s (which is a vector) are separable from (x,t) in the technology. In our case, for example, it implies that precipitation is separable from other forms of moisture delivery, which may or may not be incorporated in x. And unless the elements of Ω interact in a 'nice' fashion, ε could exhibit significant heterogeneity across observations that might be difficult to capture realistically using traditionally convenient econometric specifications.

2.1 Empirical Approximations to and Decompositions of State-level Intertemporal Productivity Indices for T^{Ω}

Figures 8 and 9 present intertemporal productivity indices for California and Iowa, respectively. In each case, three separate indices are reported. One repeats the intertemporal productivity index reported earlier that ignored the weather data. The other two indices correspond to the DEA approximations to (1) calculated under different disposability assumptions.

For both California and Iowa, there appears to be relatively close general, but certainly not complete, agreement between the calculated indices. For both states, there are years (1997 for California, 1993 for Iowa) where there is significant disagreement between the nonstochastic approximation to the productivity index and the indices calculated for the stochastic representations of the technology. And for both California and Iowa, the stochastic productivity indices show some strikingly large productivity increases. Moreover, the productivity gains are more marked (perhaps unbelievably so) for the weakly disposable version of the technology than for the freely disposable version of the technology. For example, on average, the nonstochastic productivity index for California shows productivity growth of about 3% while the stochastic measures suggest growth on the order of 6 and 8%. For Iowa, the nonstochastic indicator suggests productivity growth on average of about 2% per annum, while the stochastic measures suggest average growth of about 3 and 6% annually.

Looking more closely, there are instances where the nonstochastic productivity index suggests that productivity fell, but one or both of the stochastic productivity indices suggest that productivity had actually risen. So, for example, in 1996 the nonstochastic productivity index calculated for California suggests a productivity decline of about 3 to 4%, while both stochastic productivity indices indicate productivity growth on the order of 2%. In 1999, the difference is even more marked. The nonstochastic productivity index indicates a productivity decline of about 3 to 4%, while the weakly disposable measure indicates growth on the order of 7% and the freely disposable measure on the order 8%.

Similarly, for 2002 the nonstochastic productivity measure shows a decline in Iowa's productivity on the order of 9% while the stochastic indices show growth on the order of 10 to 11%. The pattern reverses in 2003, the nonstochastic productivity index suggests growth on the order of almost 20% while the stochastic measures indicate much more modulated growth on the order of 2%. But then the pattern reverses again in the final year.

Figure 10 presents intertemporal efficiency indices for California for the nonstochastic representation of the technology and for the weakly disposable stochastic approximation. As noted earlier, California routinely exhibits intertemporal efficiency indices at, or very close to one, for the nonstochastic representation of the technology. That tendency is dramatically increased by the inclusion of climatic variables. For example, using the nonstochastic approximation, between 1961 and 1980 California had an intertemporal efficiency index that differed from unity 15 times. According to that measure, California routinely lagged behind the frontier and caught up to it only to repeat the cycle. Incorporating the climatic variables into our analysis eliminated all of these departures from unity. Similarly from 1981 forward, California's computed intertemporal efficiency index for the nonstochastic technology was different from unity 11 times, while that number fell to 6 when climatic variables were incorporated.

The intertemporal efficiency index for $T^{\Omega}(t)$ also suggests that California's instances of lagging behind the frontier and then catching back up to it were less dramatic than the nonstochastic efficiency index suggests. Overall before the climatic variables were included in the analysis, California's average intertemporal efficiency index was unity through 3 decimal places, and its coefficient of variation was approximately .02. After the climatic variables were included, the average intertemporal efficiency index was unity out through 5 decimal places, and its coefficient of variation was approximately. 01. In fact, each time California lagged behind the meta-frontier (1983, 1988, and 1996), it immediately caught up to the meta-frontier in the next year.

A similar, though less dramatic, pattern emerges in Figure 11 where I present the corresponding intertemporal efficiency indices for Iowa. As with California, including climatic variables reduces the number of times that the intertemporal efficiency index departs from unity. Both measures give Iowa an intertemporal efficiency index of unity between 1961 and 1968. But after 1968, the nonstochastic representation of the technology *never* yields an intertemporal efficiency index of one. In other words, if Iowa was ever on the productive frontier and thus could be judged as a true leader in technical innovation during that period, it didn't stay there for long. Instead, according to that measure, it routinely lagged behind the frontier trying to catch up to innovations made elsewhere.

The version including climatic variables, however, has an intertemporal efficiency index of unity 21 times after 1968. So almost 60% of the lagging and catching up behavior has been eliminated. The stochastic intertemporal efficiency index suggests that throughout the 1960s and the early 1970s, Iowa, like California, was helping to set the meta-frontier for the agricultural technology. In other words, it was a technical leader. Its first instance of lagging behind the frontier occurred in 1974, and after that it took a few years to catch up to the frontier. That first instance of lagging behind the frontier corresponds to the first-oil shock. By 1979, Iowa was back on the frontier and stayed there until 1983 when it again started to lag behind the meta-frontier. Recall, however, that in January 1983 President Reagan had announced the first payment-in-kind (PIK) program that involved producers relying intensively on USDA payment programs idling production capacity. Iowa caught back up to the frontier in 1985 and stayed there until it lagged behind in the late 1980s only to catch up in the 1990s.

This pattern suggests a different story about technical innovation and productivity growth for Iowa than do the nonstochastic numbers. Iowa now looks more like California. Over large periods of time, it seems to be a leader in pushing the meta-frontier for agriculture outward. When exposed to shocks, such as the first oil shock and Reagan's PIK program, it lagged behind the frontier briefly before it moves back to a leadership role.[16]

In a sense this is the end of the story. Or rather, now that we have seen these numbers we know how the story should end. When weather was ignored, the Färe et al. (1994) decomposition suggested that Iowa, in particular, was doing a lot of "catching up to" and "falling away from" the productive frontier. California, too, was doing some catching up to and falling way from the productive frontier.

For California, incorporating climatic variables has eliminated almost all of California's lagging behind and catching up activity. Instead, California seems to be on the productive frontier pretty much all of the time. Because California is typically regarded as one of the leading agricultural states and has been for a very long time, this is not unexpected.[17] For Iowa the story is similar, but not quite so dramatic. Still, almost 60% of that activity has been eliminated by the incorporation of climatic variables into the analysis.

Therefore, where before the productivity and technical change indices were very divergent for Iowa and somewhat less so for California, we should now expect to see much less divergence for California and somewhat less divergence for Iowa. And as Figures 12 and 13 illustrate, that is what

[16] When the preceding analysis is replicated for the version of the technology satisfying free disposability of s, not surprisingly, the number of instances where either California or Iowa exhibits and efficiency index of one. For California, instead of 6 instances of an efficiency index differing from unity, there are 8. And for Iowa, the number of departures from an index of one is reduced by slightly over 30 % instead of 60%.

[17] For example, Schultz's (1947) analysis suggests that Pacific Coast agriculture had the highest output per farm worker as far back as 1929.

happens. For California, the technical change and productivity indices now coincide almost exactly for much of the sample period. For Iowa, the match is less perfect, but still noticeable.

2.2 A Closer Look at $\Omega^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)$

In the introduction, I suggested that measured TFP growth for US agriculture more properly reflects a weather index than an actual productivity index. So far, not much has been said about $\Omega^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)$. Partly, that's because it is actually a residual that is identified by the measures already discussed. In this sense, it contains no information about the technology that has not been examined by other means. But sometimes, it helps to look at matters from different angles.

Figures 14 and 15 presents the empirical results or $\Omega^{s^0,s^1,t^0,t^1}\left(z^0,x^0;z^1,x^1\right)$, graphed against the same axis, for three states, California, Iowa, and Maryland. I include Maryland for several reasons. First, unlike the other two, it is an eastern state. Thus, one expects it to experience somewhat different weather patterns than the other two. Second, and more importantly, Maryland's agricultural industry is heavily dominated by broiler production, which is largely undertaken indoors under closely controlled environmental conditions.[18]

Figure 14 compares California and Maryland, and Figure 15 compares California and Iowa. Each of these series represents a state-level weather index in the sense that each captures year-to-year variation in productive capacity that is associated purely with variations in s. But they are not weather measures in the same sense that either components of s are. Instead, they measure s indirectly through its impact on productive ability. In fact, it is not too much of an intuitive stretch to suggest that these measures, which are in fact statistics, portray some of what we might characterize as noise in productivity measurement.

The most noticeable feature of Figure 14 is the tendency of Maryland's weather index to be at or very close to one throughout the period. This does not mean that Maryland does not have variable weather patterns. It does. Rather, it means that those weather patterns have relatively little impact on state-level agricultural productivity, and hence there is relatively little noise here. Given the concentration of poultry production in Maryland, this is not surprising. When compared to Maryland, California's weather index appears much more variable even though California is noted for the moderation of its climate. Again, this reflects the fact that weather variation in California plays more of a role in affecting California's agricultural productivity than in Maryland.

[18] It also happens to be the state where I grew up, where I currently live, and whose University pays my salary.

Turning to Figure 15 reveals that the variability in Iowa's weather index dwarfs that of California. A large part of this is concentrated on the extreme outcome for 1993 in Iowa (reflected in an extreme productivity outcome). But even ignoring this apparent outlier, the patterns in variability are distinct.[19]

3 Conclusion

The main lesson learnt is that including our measures of s into a state-contingent production function changes DEA based measures of productivity change and efficiency change for California and Iowa.[20] That's hardly surprising. Incorporating new variables into efficiency and productivity analysis routinely changes results. So, in and of itself, that's not really saying much. What impresses me is that the changes tend to confirm what I know to be conventional wisdom about those two states in a fashion that also accords with the basic nature of agricultural production. Agricultural producers operate in a stochastic environment. I do not believe that it is inherently more stochastic (if any sense can be attributed to that notion) than other industries, but I do know that our productivity and efficiency models routinely deny its very nature. And the evidence that this matters is palpable to anyone who bothers to look at routinely computed TFP numbers. Thus, if we are to take those numbers seriously, perhaps we should take the modeling process more seriously in this regard.

Venturing beyond that is dangerous. The analysis here is preliminary and is intended to be a start towards a more thorough analysis of how the Arrow-Debreu framework can be integrated in a meaningful fashion into agricultural productivity measurement and accounting. While the start seems encouraging, there are reasons for caution. As usual, attempting to solve one problem introduces new problems, and here that is manifested in some very large productivity estimates that appear to be driven by outlier effects. There are a number of possible causes to this ranging from the obvious possibility of misspecification to the use of poor measures of Nature's actual draws. As already noted, the measures included here were originally developed for purposes other than those to which they have been put. These and other matters need to be resolved before anything approaching a definitive statement can be made on whether the inclusion of weather variables truly

[19] Replicating Figure 15 for the stochastic technology consistent with free disposability still leads to weather indices for the states that are distinctly different across states. However, for that version of the technology, California's index exhibits the most variability.

[20] In fact, this is broadly true for all 48 states in the sample.

makes a difference in productivity analysis. But what is clearly true is that it does make a difference in the numbers we currently calculate, and that's a start.

4 References

Alston, J. M., G. W. Norton, and P. Pardey. *Science Under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*. Ithaca and London: Cornell University Press, 1995.

Arrow, Kenneth. "Le Role Des Valeurs Boursiers Pour La Repartition La Meilleur Des Risques." *Cahiers Du Seminair D'Economie.*Paris: CNRS, 1953.

Barton, G., and M. R. Cooper. "Relation of Agricultural Production to Inputs." *Review of Economics and Statistics* 30 (1948): 117-26.

Chambers, R. G. "Stochastic Productivity Measurement." *Journal of Productivity Analysis* 30 (2008): 107-20.

Debreu, G. *The Theory of Value*. New Haven, CT: Yale University Press, 1959.

Deschenes, O., and M. Greenstone. "The Economic Impacts of Climate Change: Evidence From Agricultural Output and Random Fluctuations in Weather." *American Economic Review* 97 (2007): 354-85.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang. "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries." *American Economic Review* 84 (1994): 66-83.

Griliches, Z. "The Sources of Measured Productivity Growth: United States Agriculture, 1940-60." *Journal of Political Economy* 71 (1963): 331-46.

———. "Specification Biases in Estimates of Production Functions." *Journal of Farm Economics* 39 (1957): 8-20.

Griliches, Zvi. "The Demand for a Durable Input: Farm Tractors in the United States, 1921-1957." *The Demand for Durable Goods.* ed. A. C. Harberger, 181-207. Chicago: University of Chicago Press, 1960.

Henderson, D. J., and R. R. Russell. "Human Capital and Convergence: A Production-Frontier Approach." *International Economic Review* 46 (2005): 1166-205.

O'Donnell, C, R. G. Chambers, and J. Quiggin. "Efficiency Analysis in the Presence of Uncertainty." *Journal of Productivity Analysis* forthcoming (2009).

Ortiz-Bobeia, A. "Towards a More Transparent Econometric Approach of Assessing Climate Change Impacts on Agriculture." University of Maryland, 2011.

Savage, L. J. *Foundations of Statistics*. New York: Wiley, 1954.

Schlenker, W., W. M. Hanemann, and Fisher. "Will US Agriculture Really Benefit Fom Global

Warming? Accounting for Irrigation in the Hedonic Approach." *American Economic Review* 95 (2005): 395-406.

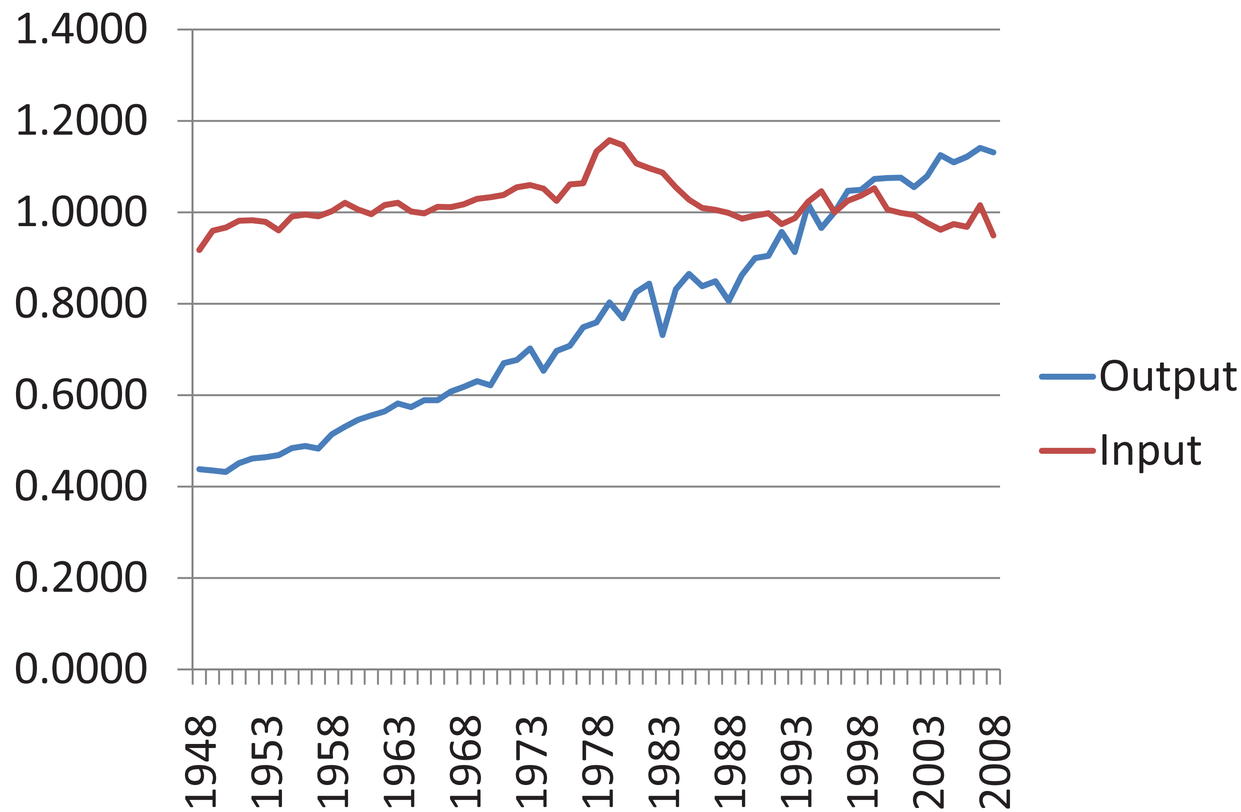
Schlenker, W., and M. Roberts. "Estimating the Impact of Climate Change on Crop Yields: The Importance of Nonlinear Temperature Effects." NBER Working Paper 13799, 2008.

Schultz, T. W. "How Efficient Is US Agriculture?" *Journal of Farm Economics* 39 (1947): 644-58.

Schultz, T. W. "Output-Input Relationships Revisited." *Journal of Farm Economics* 40 (1958): 924-32.

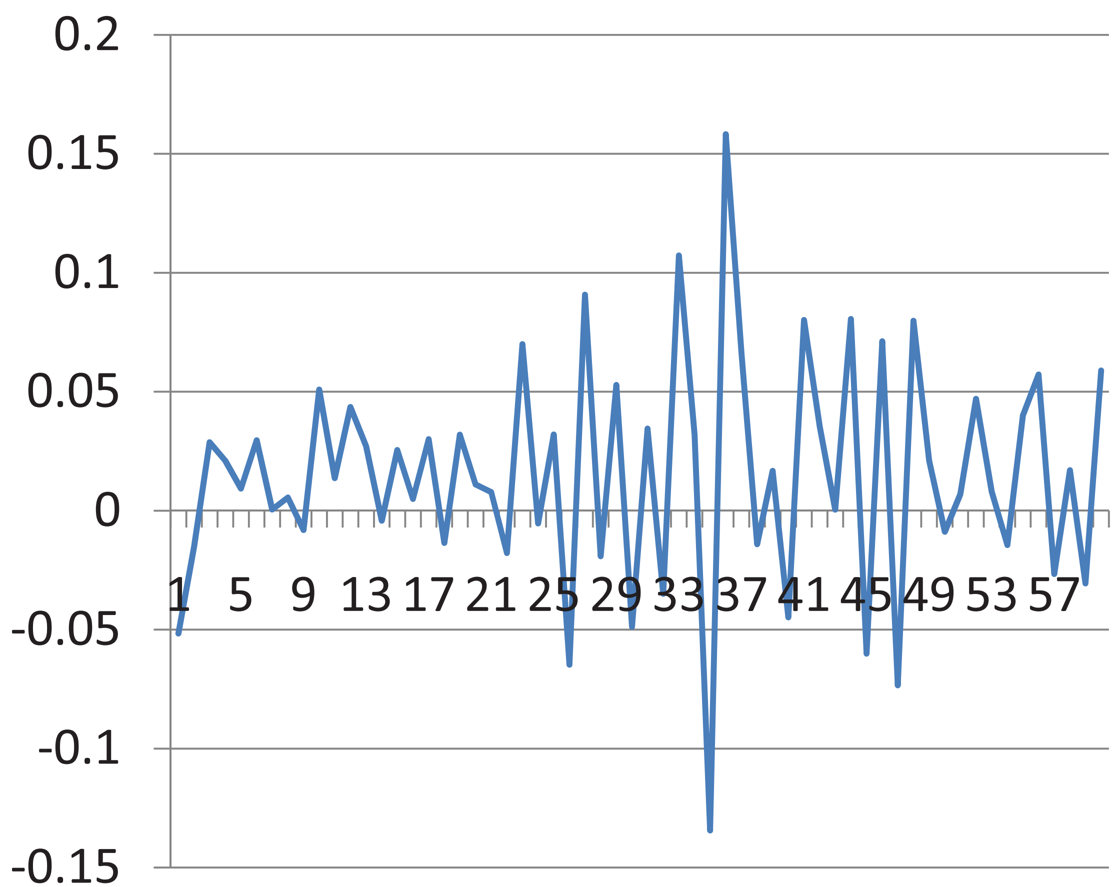
Schultz, T. W. "Reflection on Agricultural Production, Output, and Supply." *Journal of Farm Economics* 48 (1956): 748-62.

Figure 1: Aggregate US Agricultural Input and Output (1948-2008)



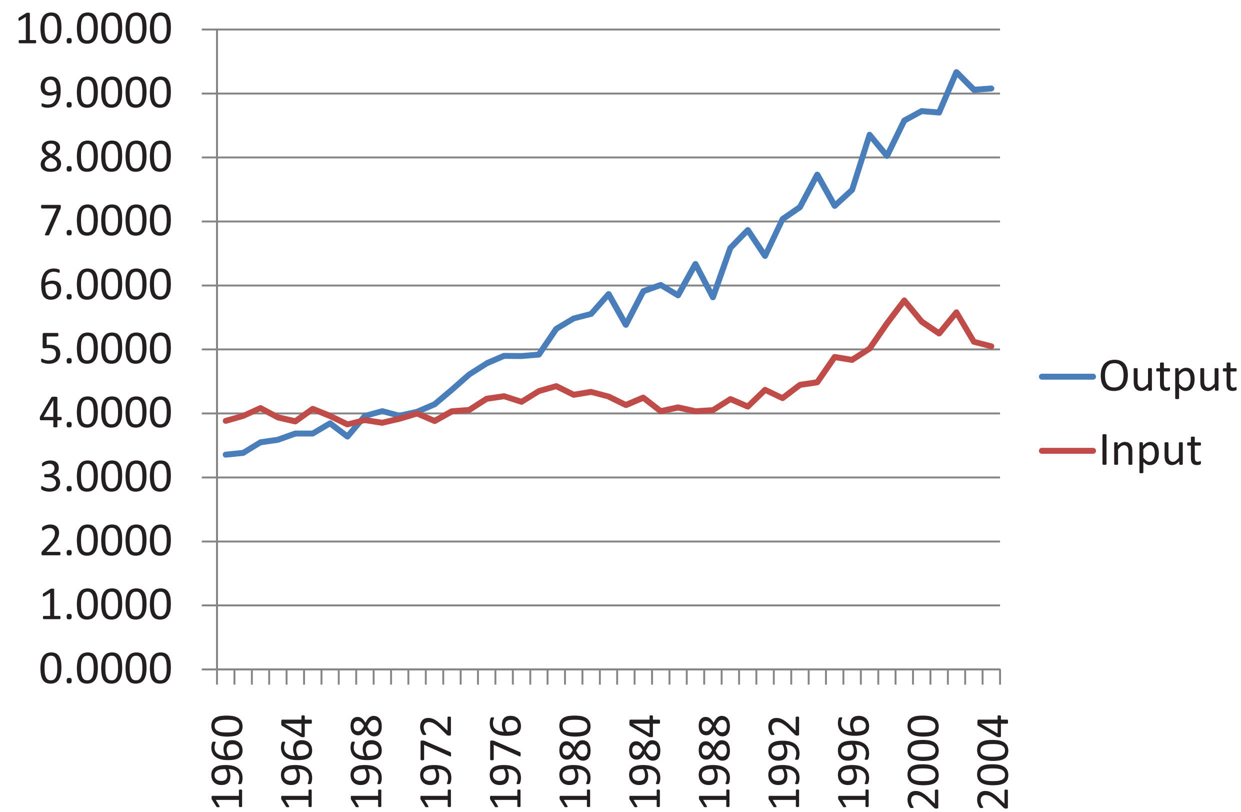
Source: Economic Research Service, United States Department of Agriculture

Figure 2: US Agriculture TFP Change (1949-2008)



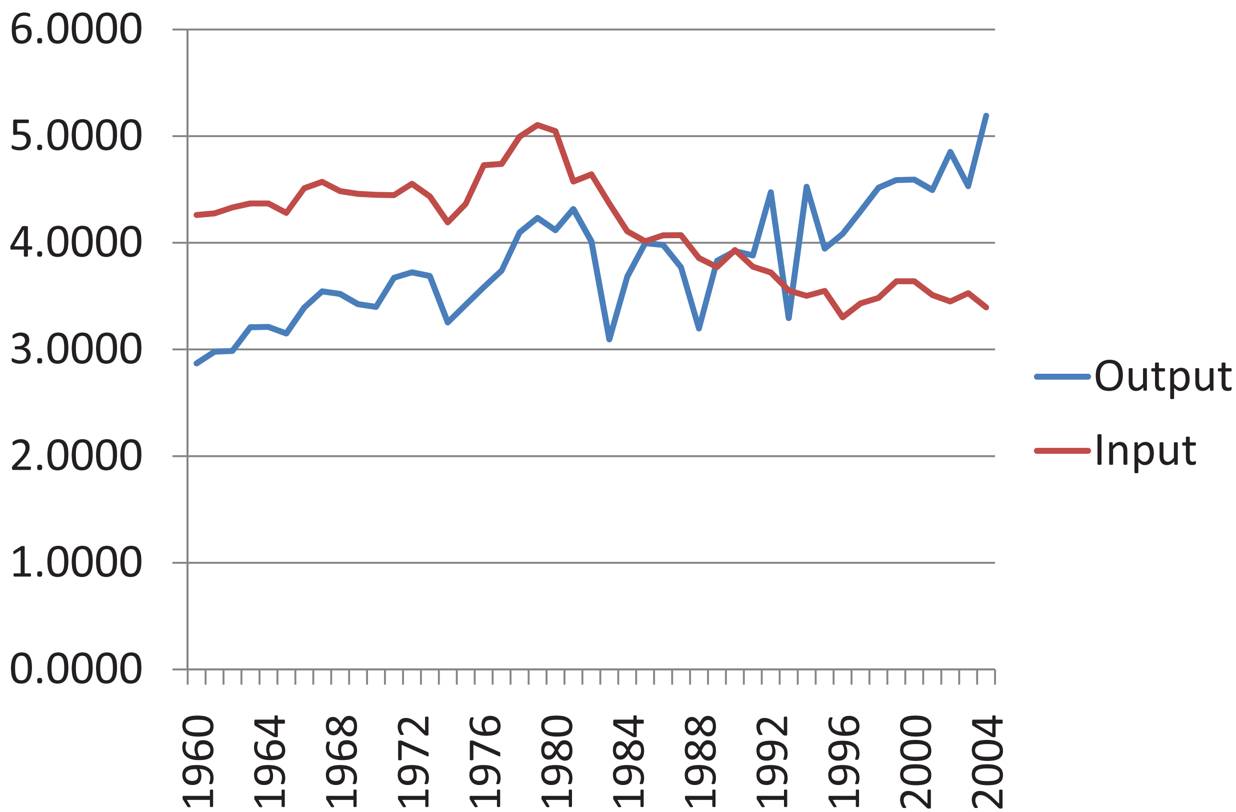
Source: Computed from ERS/USDA official statistics

Figure 3: CA Aggregate Output and Input (1960-2004)



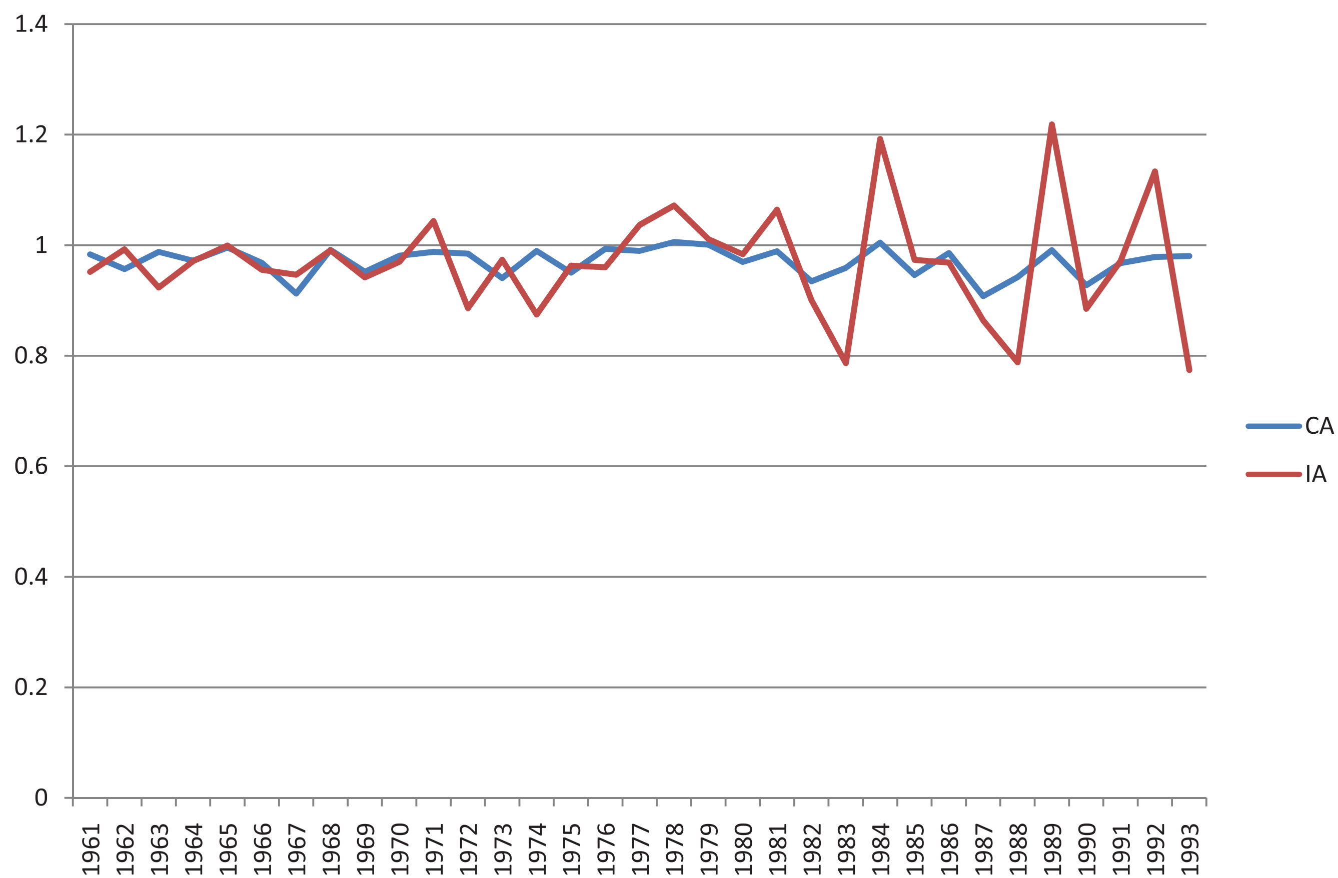
Source: ERS/USDA

Figure 4: IA Aggregate Output and Input (1960-2004)



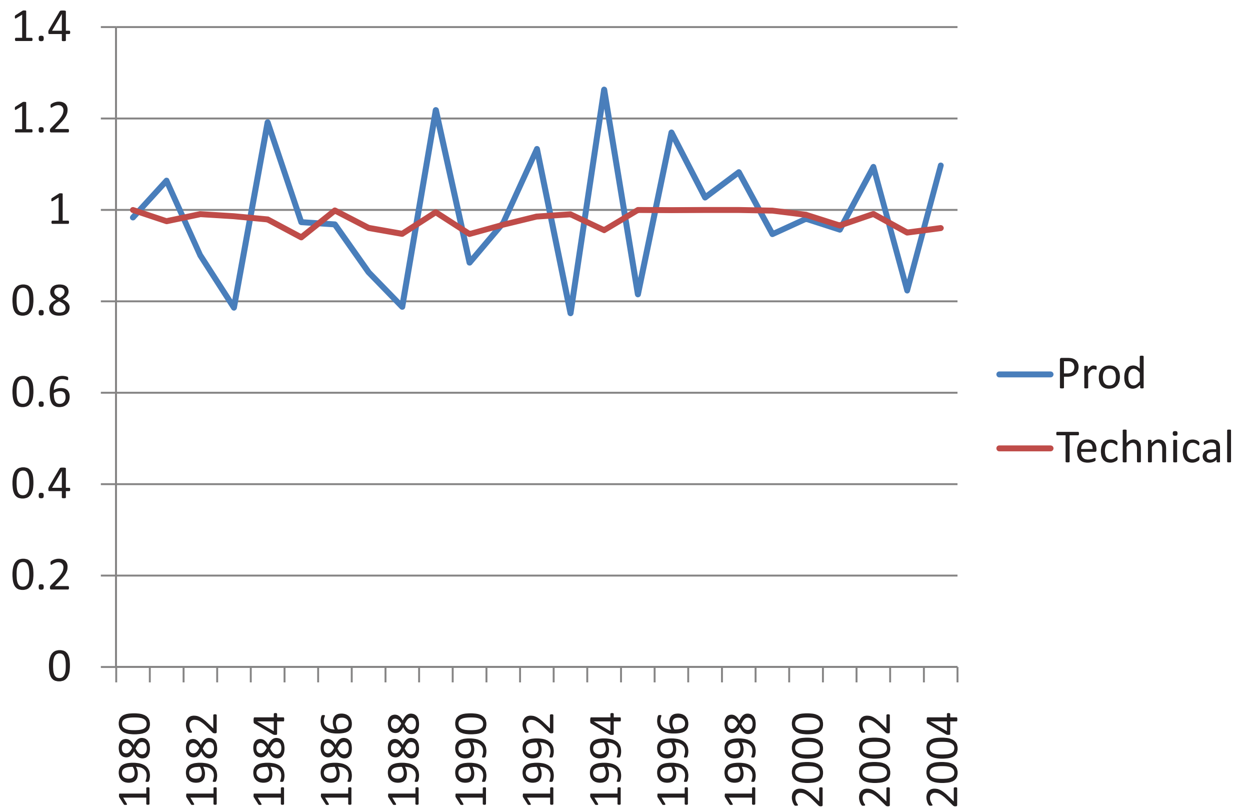
Source: ERS/USDA

Figure 5: CA and IA Intertemporal Productivity Indices (1961-2004)



Source: Computed

Figure 6: IA Intertemporal Productivity and Technical Change Indices (1980-2004)



Source: Computed

Figure 7: CA Intertemporal Productivity and Technical Change Indexes (1980-2004)

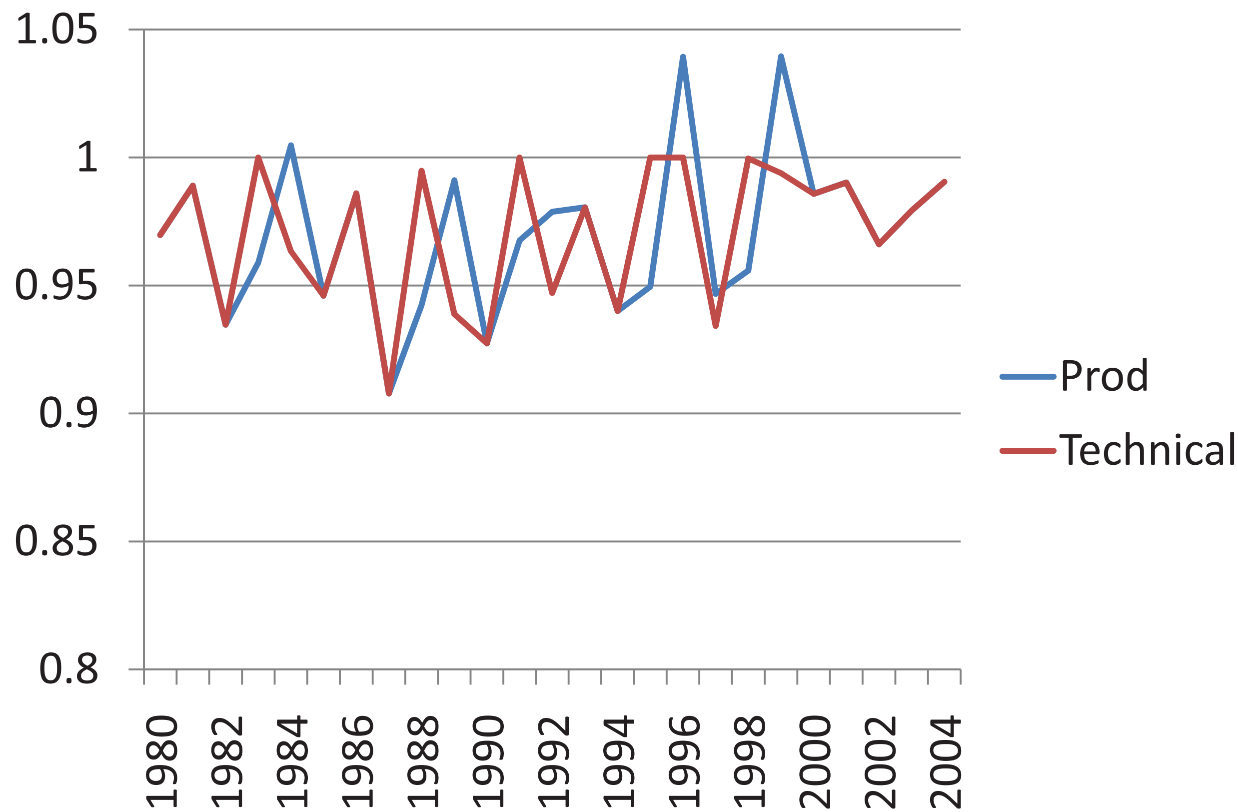


Figure 8: CA Intertemporal Productivity with and without Weather

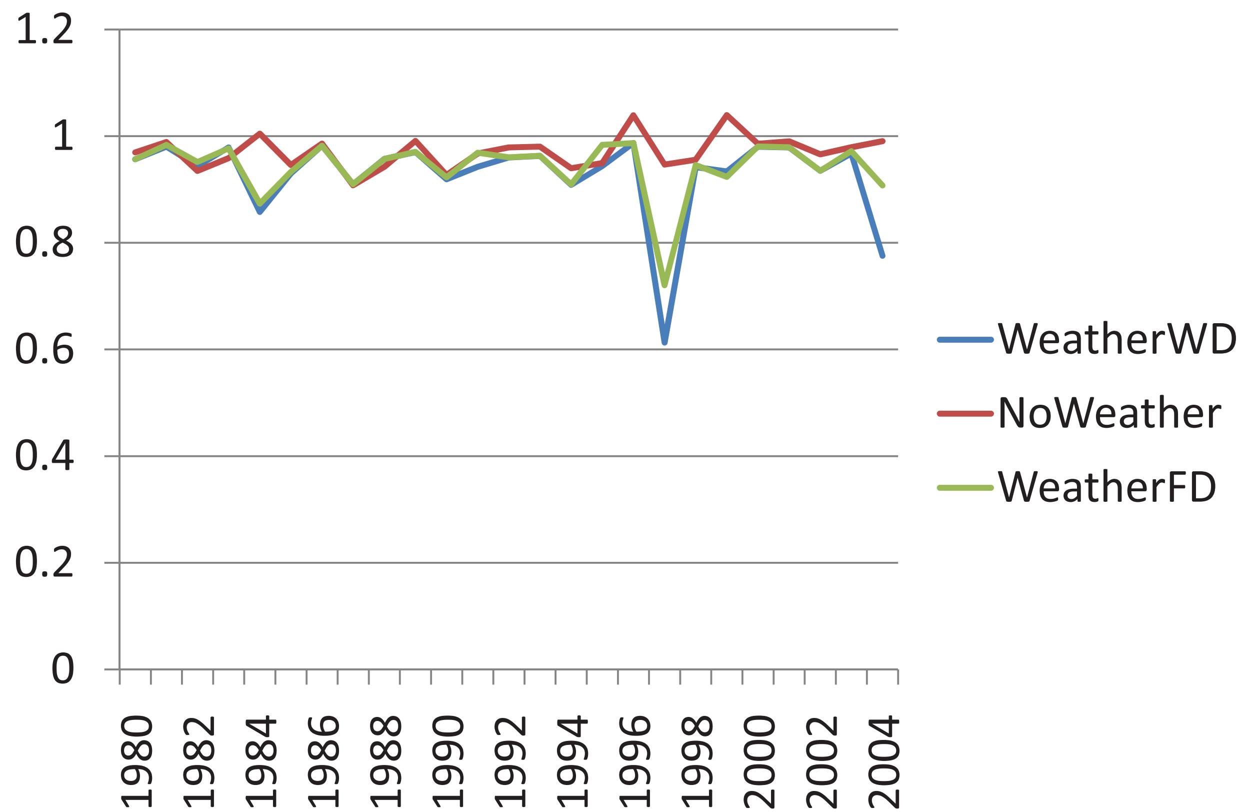


Figure 9: IA Intertemporal Productivity with and without Weather

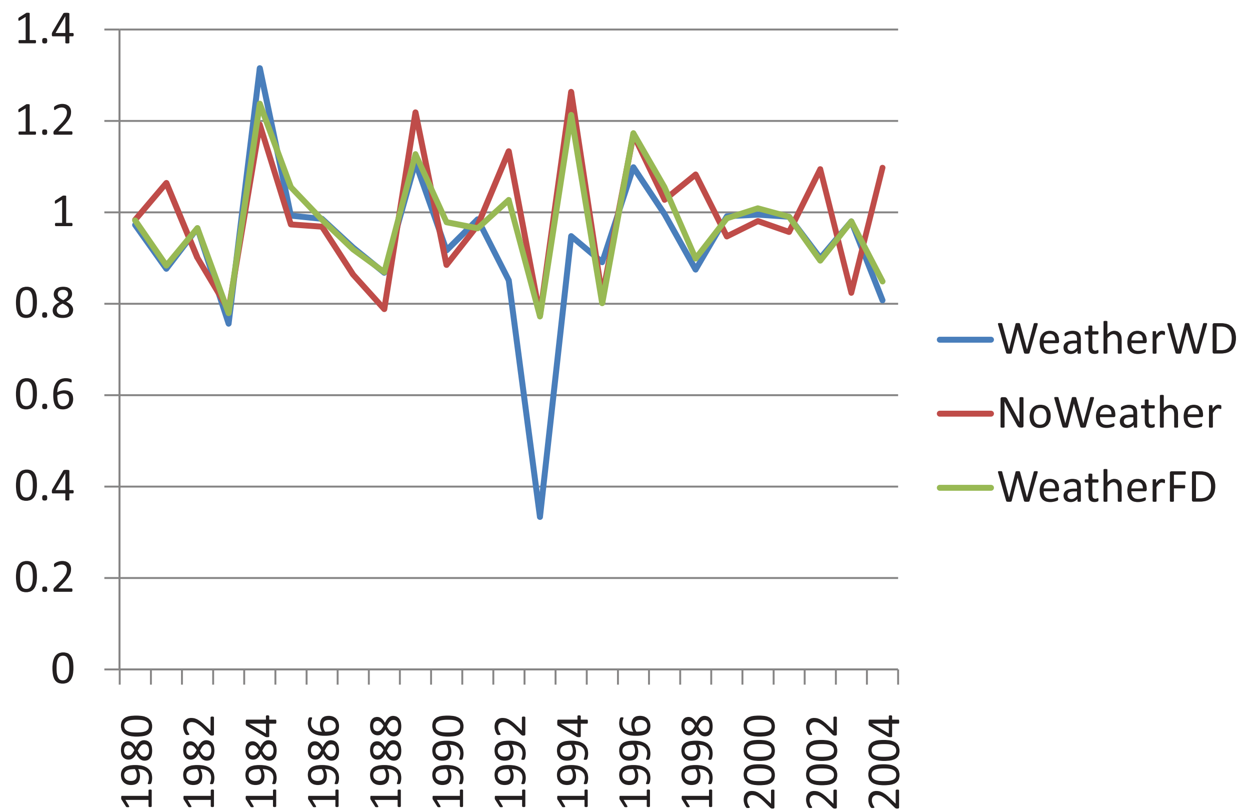


Figure 10: CA Intertemporal Efficiency Index with and without Climatic Variables

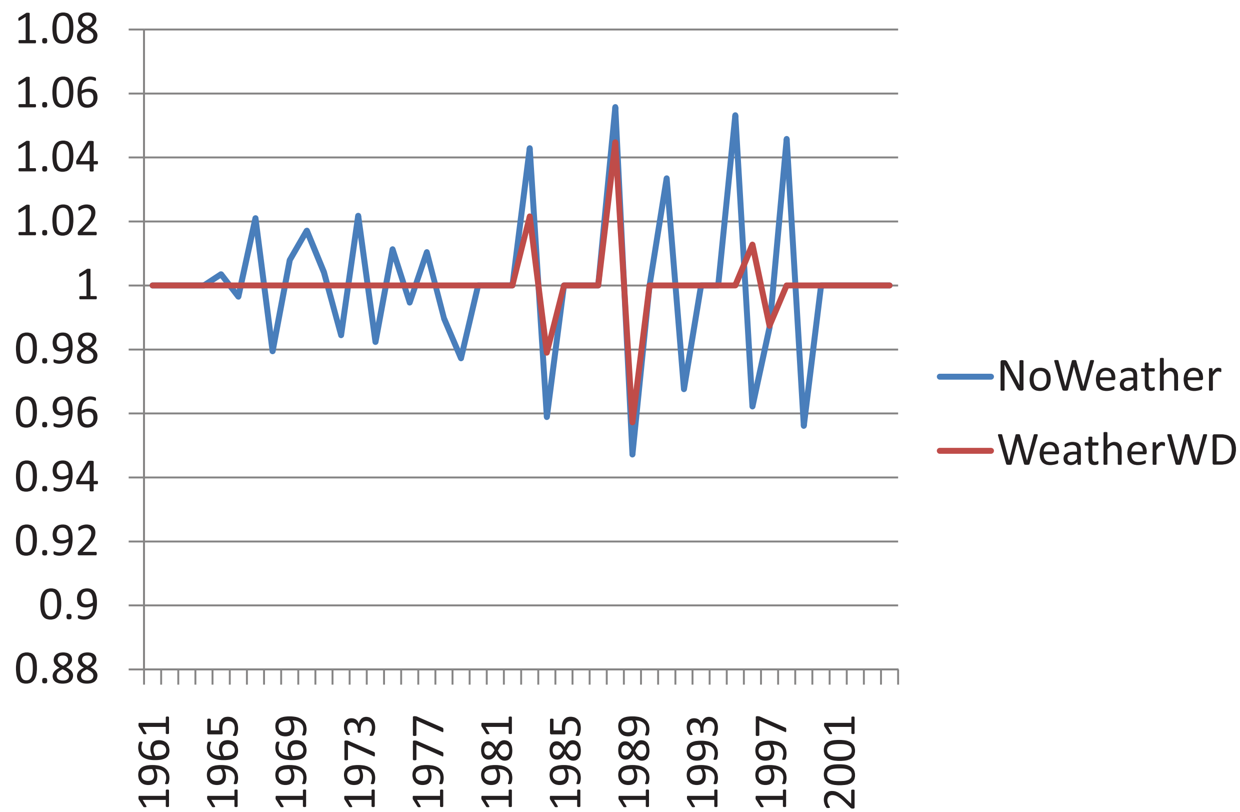


Figure 11: IA Intertemporal Efficiency Index with and without Climatic Variables

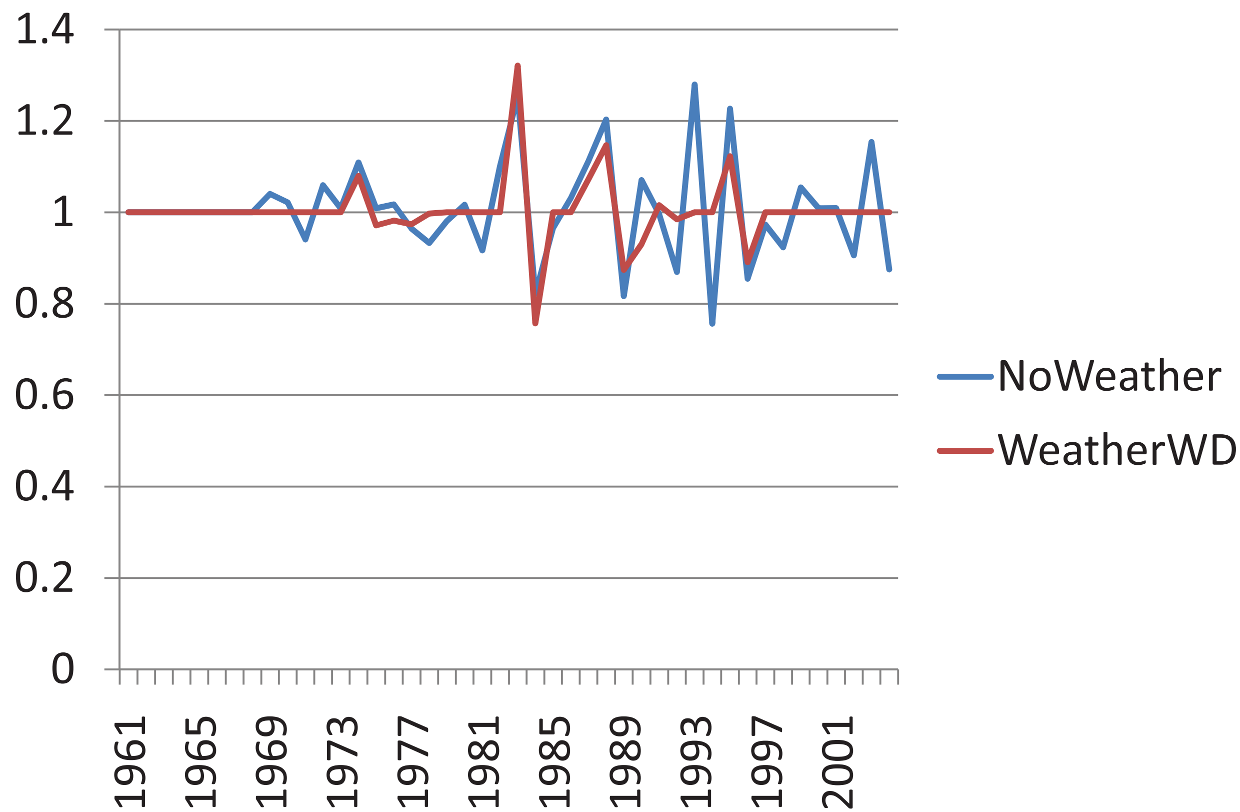


Figure 12: CA Intertemporal Productivity and Technical Change Indices with Climatic Variables

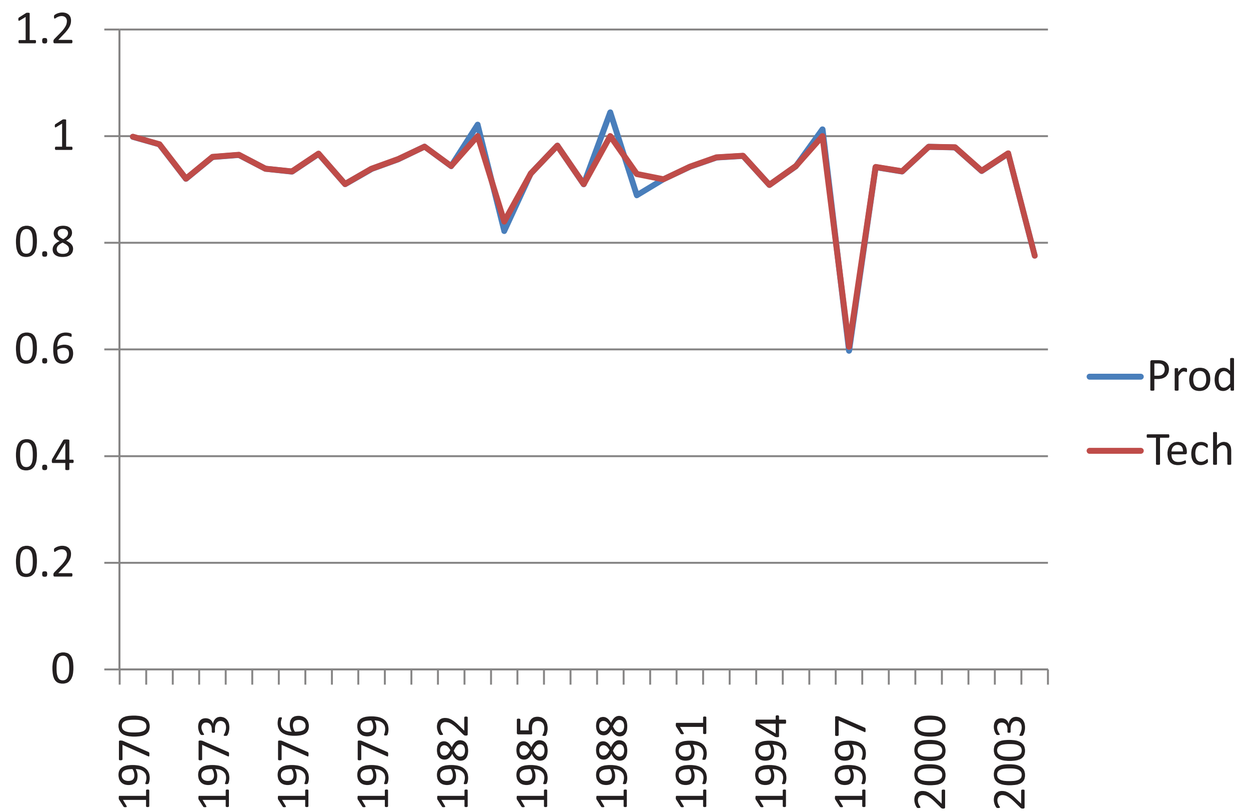


Figure 13: IA Intertemporal Productivity and Technical Change with Climatic Variables

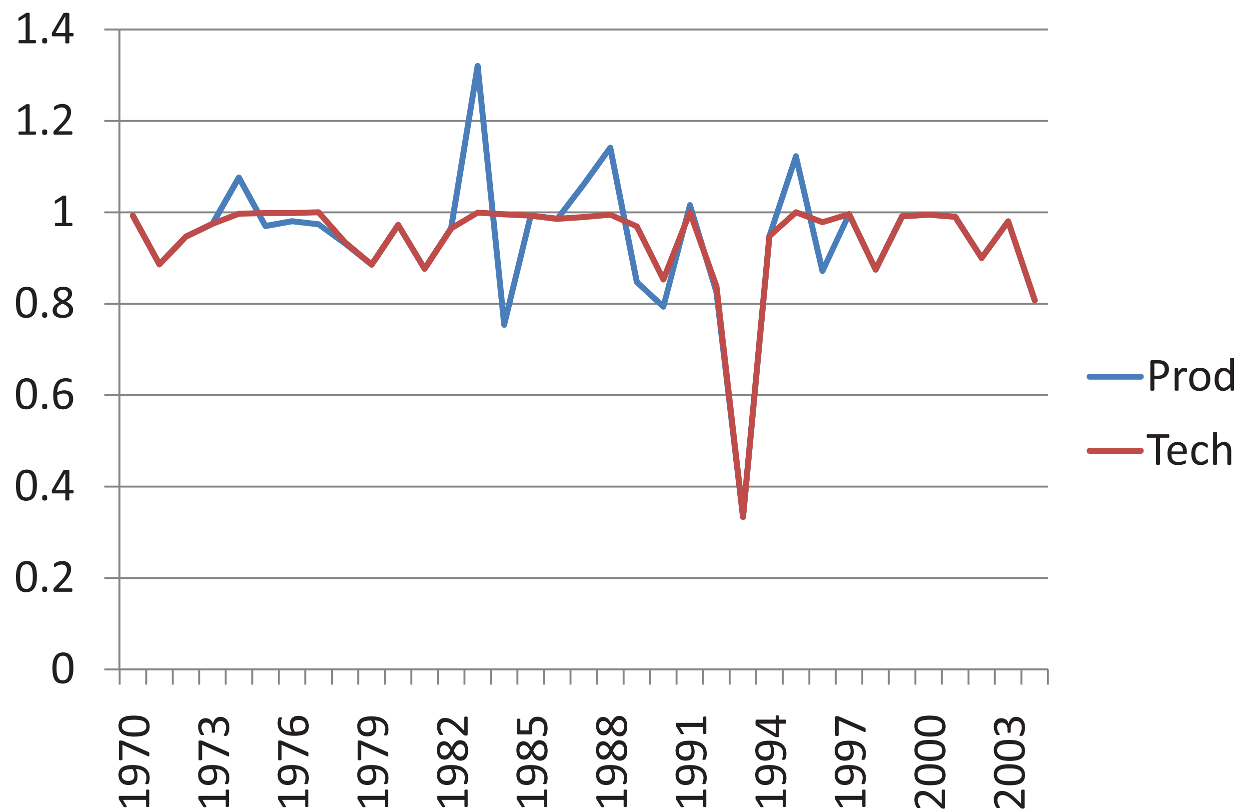


Figure 14: California and MD Omega Indices

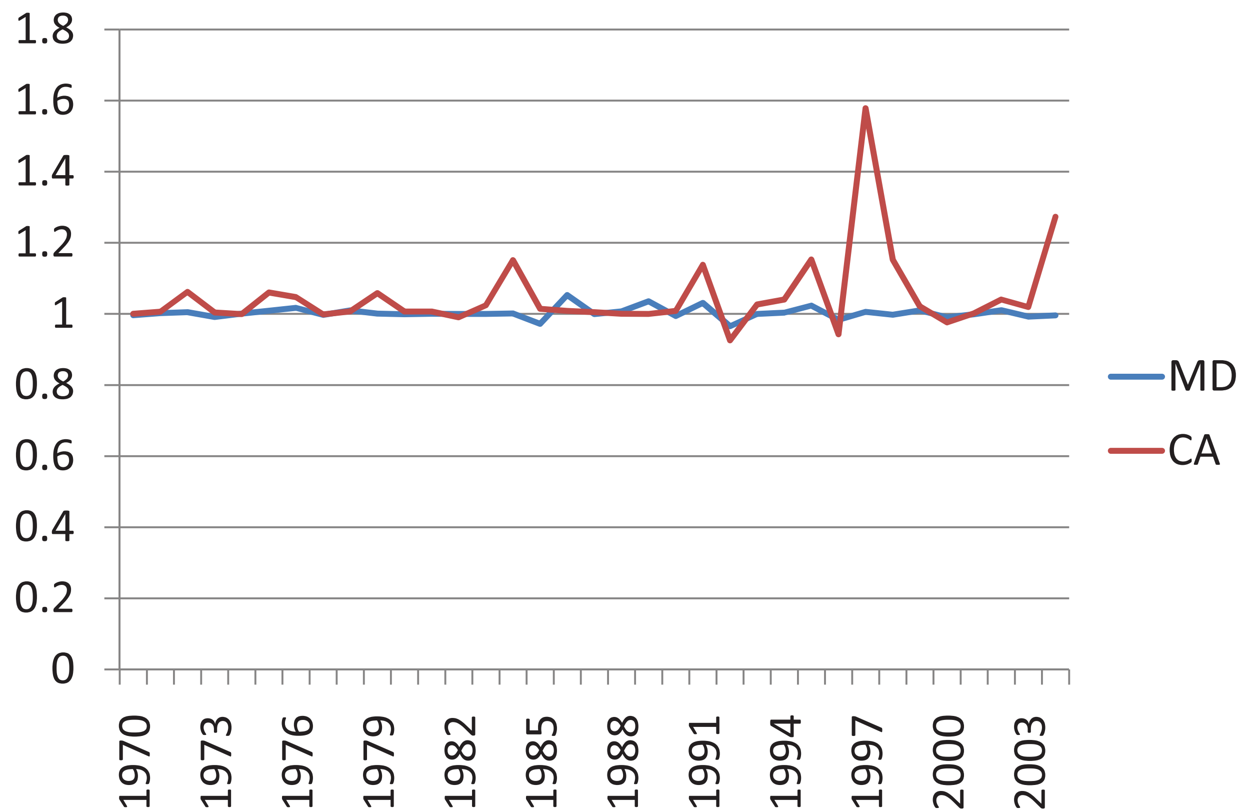


Figure 15: IA and CA Omega Indices