Should Farmers Follow the Recommendations of Market Advisory Services? A Hierarchical Bayesian Approach to Estimation of Expected Performance

by

Silvina M. Cabrini, Scott H. Irwin, and Darrel L. Good[1]

July 2009

[1] Silvina M. Cabrini is a former Graduate Research Assistant for the AgMAS Project in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign. Scott H. Irwin is the Laurence J. Norton Chair of Agricultural Marketing in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign. Darrel L. Good is a Professor in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign.
This material is based upon work supported by the Cooperative State Research, Education and Extension Service, U.S. Department of Agriculture, under project Nos. 98-EXCA-3-0606 and 00-52101-9626. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture. Additional funding support from the Aurene T. Norton Trust is gratefully acknowledged.

Contact author: Silvina M. Cabrini. (scabrini@pergamino.inta.gov.ar). Instituto Nacional de Tecnología Agropecuaria. Estación Experimental Pergamino. Área Economía. 54 – 2477 – 439005. (Argentina).

Should Farmers Follow the Recommendations of Market Advisory Services? A Hierarchical Bayesian Approach to Estimation of Expected Performance

Abstract

This article employs a Bayesian hierarchical approach to estimate individual expected performance of market advisory services in corn, soybeans, and wheat. This estimation procedure is a conservative approach compared to traditional estimation, since it reduces estimation error in the expected gains from following top-performing advisory programs. Two versions of the model are estimated. The first combines information across the entire sample, while the second includes skeptical beliefs based on the efficient market hypothesis. For a skeptical decision maker who is willing to follow an advisory program only if it is expected to increase price received by more than 1%, one of the corn programs, most of the soybean programs, and a few wheat programs may be valuable marketing alternatives. A skeptical decision maker who is willing to follow an advisory service only if it is expected to increase price received by more than 5% would prefer to adopt a strategy that mimics the market benchmark rather than follow any advisory program in each of the three crops.

Key words: Bayesian estimation, corn, market advisory service, pricing, performance, soybeans, wheat

Should Farmers Follow the Recommendations of Market Advisory Services? A Hierarchical Bayesian Approach to Estimation of Expected Performance

Agricultural market advisory services are popular with U.S. farmers (Patrick, Musser, and Eckman 1998; Isengildina et al. 2006). For a subscription fee, these firms provide market analysis and pricing advice to farmers. In particular, they make recommendations on how to market crops using various instruments, including cash sales, forward, futures, and options contracts. Marketing recommendations are specific, indicating the portion of a crop that should be marketed, the marketing tool, and the timing of transactions. Market advisory services conduct market research and employ fundamental and/or technical analysis to identify profitable marketing alternatives (Isengildina et al. 2006).

Several studies have evaluated the performance of market advisory services (e.g., Gehrt and Good 1993; Kastens and Schroeder 1996; Irwin, Good, and Martines-Filho 2006). These studies analyze whether advisory services are valuable to farmers by comparing the price obtained when following the marketing recommendations of services and external benchmarks. For example, Irwin, Good, and Martines-Filho (2006) find positive performance for advisory programs when compared to the average price offered by the market for corn and soybeans, with price differences around 3¢/bu. for corn and 15¢/bu. for soybeans.[1] Batts, Irwin, and Good (2009) report negative performance for wheat, with average net advisory prices falling below market benchmarks by about 9¢/bu.

Previous studies on the pricing performance of advisory services provide thorough evaluations of advisory services as a *group*. However, pooled performance estimates do not provide answers to other interesting questions such as: Is there any advisory program that can outperform benchmarks by, say, more than 5%? What is the expected performance of the top-ranked program? What is the expected performance of the program a farmer currently follows? Farmers and other market participants are likely to be interested in these questions, which require estimation of expected pricing performance for *individual* advisory programs. Performance of individual advisory services is especially interesting since there is a wide range of pricing performance across programs. Irwin, Good, and Martines-Filho (2006) and Batts, Irwin, and Good (2009) report that the difference between minimum and maximum net advisory prices in 2001 was $0.87/bu. for corn, $0.93/bu. for soybeans and $0.94/bu. for wheat, which represents a substantial proportion of average market prices for that crop year, $2.00/bu. for corn, $5.34/bu. for soybeans and $2.59/bu. for wheat. Moreover, for corn and soybean advisory programs, there is some evidence of performance persistence at the extremes of advisory service performance rankings, which suggests that a subset of services may have more attractive expected performance than other services. Furthermore, Cabrini, Irwin, and Good (2007) find that more active corn and soybean market advisory programs (i.e., a tendency to make large bets on price movements) have better pricing performance compared to less active programs.

The estimation of individual expected pricing performance for advisory programs is a challenge because of the small number of observations available. In the most complete evaluations to date (Irwin et al. 2006; Batts, Irwin, and Good 2009) a maximum of 10 time-series observations is available for any individual advisory program. The simplest procedure to estimate the expected performance for advisory programs is to average past performance observations for each program. In estimation problems like this, individual sample averages tend to over-fit the data and unusually high and low expected performance values may appear due to estimation error. This implies that the gains from following top performing programs are likely to be overestimated when traditional estimates are considered (Jorion 1986; Michaud 1989; Grauer 1995; Kosowski, Naik, and Teo 2007).

A model that combines information from advisory programs as a group is a more appealing estimation procedure than traditional estimation approaches. Bayesian hierarchical models have this characteristic. A hierarchical model produces estimators that are weighted averages of separate and pooled estimates. This type of estimator is called a *shrinkage* estimator, since individual estimates are shrunk to common values. Shrinkage estimation is a more conservative approach compared to traditional estimation, since it reduces estimation error in the expected gains from following top-performing advisory programs. Several studies in the finance literature have employed shrinkage estimators to compute expected stock investment returns, and results indicate that these estimators outperform traditional sample estimates, in particular when the sample size is small (Jorion 1986; Grauer 1995).

The purpose of this article is to evaluate whether farmers can increase price received by following the marketing recommendations of individual market advisory programs in corn, soybeans and wheat. Hence, the focus of this article is estimation of expected performance for individual programs. A Bayesian hierarchical approach is employed to estimate individual expected performance for advisory services tracked by the AgMAS project over the 1995 to 2004 crop years. Two versions of the model are estimated. The first version is based exclusively on the sample data and the second version includes skeptical beliefs about the performance of advisory services. The posterior distribution of individual expected performance is computed by simulation. Bayesian point estimates and 90% probability intervals are employed to identify subsets of programs that represent attractive marketing alternatives for farmers.

Data and Non-Bayesian Estimates

Data for this article are obtained from AgMAS performance evaluations (Irwin et al. 2006; Batts, Irwin, and Good 2009). The AgMAS project tracked agricultural market advisory services and computed corn, soybean, and wheat net advisory prices for the 1995 to 2004 crop years. The net advisory price is the price received by a farmer in a specified geographic area (e.g., central Illinois) who markets the crop according to an advisory program's recommendations, net of storage and brokerage costs. All programs that were tracked for two or more crop years are considered here.[2] AgMAS performance evaluations for wheat are applied to two classes of wheat: soft red winter wheat and hard red winter wheat. Since advisory program performance is very similar for both types of wheat, this article considers only the performance of market advisors for hard red winter wheat. From this point forward, whenever the term "wheat" is used, it is meant to represent hard red winter wheat. A complete list of programs tracked by AgMAS is presented in table 1.

The primary measure of advisory program performance is the difference between the price received by a farmer who markets grain following a program's recommendations and a given benchmark price:

(1) $$y_{jt} = NAP_{jt} - BP_t$$

where NAP_{jt} is the net advisory price for program *j* in crop year *t* and BP_t is the benchmark price in crop year *t*. The market benchmark is computed as the average price over the two-year marketing window defined for each crop: i) September 1st of the year before harvest through August 31st of the year after harvest for corn and soybeans, and ii) June 1st of the year before harvest through May 31st of the year after harvest for wheat. AgMAS performance evaluations also consider alternative benchmarks.[3] Since performance results are not overly sensitive to the benchmark considered, this article presents a complete set of results only for the 24-month market benchmark. The results computed using other benchmarks are mentioned briefly in the results section and are available from the authors upon request.

The traditional estimator for an advisory program's performance is simply the individual sample average:

(2) $$\bar{y}_j = \frac{1}{T_j}\sum_{t=1}^{T_j} y_{jt}$$

where T_j is the number of past performance observations available for program j. Figure 1 shows traditional point estimates and 90% confidence intervals for the expected performance of advisory programs against the 24-month market benchmark in corn, soybeans, and wheat. The values in figure 1 are obtained by estimating expected performance separately for each advisory program. Point estimates are sample averages (equation 2) and confidence intervals are computed using the standard errors of these averages.[4] Panel A shows expected performance estimates for corn. Point estimates range from 30¢/bu. above the market benchmark to 19¢/bu. below. Slightly more than half of the programs have positive expected pricing performance, but only three have an expected price that is significantly greater than the benchmark price $(\alpha = 0.10)$. In addition, three programs have significantly negative performance in corn. Panel B shows expected performance estimates for soybeans. Point estimates range from 70¢/bu. above the benchmark to 29¢/bu. below. Nearly 80% of the programs have positive point estimates for expected performance and eight have significantly positive performance $(\alpha = 0.10)$. None of the programs has significantly negative performance in soybeans. Finally, panel C shows expected performance estimates for wheat. Point estimates range from 55¢/bu. above the benchmark to 54¢/bu. below. Half of the programs have positive point estimates for expected performance and only one has significantly positive performance $(\alpha = 0.10)$. One of the programs has significantly negative performance in wheat.

An alternative approach to estimation of expected pricing performance is to assume there is not enough data to estimate individual performance, and therefore, information is pooled to obtain one estimate of expected performance for the group of advisory programs. The precision-weighted pooled estimate is:

$$(3) \qquad \hat{y}^{pool} = \frac{\sum_{j=1}^{N} \frac{1}{\hat{\sigma}^2_{\bar{y}_j}} \bar{y}_j}{\sum_{j=1}^{N} \frac{1}{\hat{\sigma}^2_{\bar{y}_j}}}$$

where N is the number of advisory programs considered and $\hat{\sigma}^2_{\bar{y}_j}$ is the variance estimator of $\bar{y}_j$. Note that this pooled estimator is different from the simple average of individual expected performance across programs. The pooled estimator is a weighted-average, where the weights are the inverse of the squared standard error of each estimate. A simple average would be reasonable under the assumption that individual estimates have the same error, or that the standard deviation of performance and the number of observations are the same for all programs. However, as illustrated in figure 1, the data employed in this study suggest that standard deviation of performance is different across programs, and hence, a weighted pooled estimate is more appropriate. For farmers deciding whether to follow a market advisor or simple spreading of crop sales, this pooled estimator provides more accurate information compared to a simple average since it assigns more weight to the performance data from programs that are more stable and with more years of tracking history.

Under the assumption that all advisory programs have the same expected performance, pooled estimates fully describe expected performance for the group of advisory programs considered. The last dot in each of the panels in figure 1 is the pooled estimate for expected performance. This value is 0.5¢/bu. for corn, 8.5¢/bu for soybeans, and 2.5¢/bu. for wheat. Based on these estimates, risk-neutral farmers would be willing to follow an advisory program rather than applying a naïve strategy of spreading sales along the marketing window for the three crops. However, the associated 90% confidence intervals are [-0.6¢/bu., 1.7¢/bu.], [5.8¢/bu., 11.1¢/bu.], and [-1.2¢/bu. to 6.2¢/bu.]; therefore, expected performance for advisory programs as a group is significantly positive only in the soybean market.[5]

Bayesian Hierarchical Model

Separate individual or pooled estimates imply relatively extreme assumptions about the expected performance of advisory programs. On one hand, separate estimates imply that expected performance is completely independent across programs. On the other hand, pooled estimates imply that all programs have the same expected performance. Market advisors operate in the same markets and have access to the same public information so performance of the different programs should not be completely unrelated. However, advisors may have different skills in analyzing information or have access to different sources of private information, with the implication that it may be unreasonable to assume that all advisors have identical expected performance. A situation between these two alternatives seems more reasonable and can be implemented by Bayesian hierarchical models (Gelman et al. 2003). A hierarchical model based on the normal distribution produces *shrinkage* estimators that are weighted averages of individual and pooled estimates. For example, the shrinkage estimator for the expected performance of advisory program *j* is a weighted average of the individual sample mean $\left(\bar{y}_j\right)$ and the Bayesian pooled estimate $\left(\hat{\mu}^{pool}\right)$:[6]

(4) $$\hat{\theta}_j^{shrink} = \left(1-w\right)\overline{y}_j + w\hat{\mu}^{pool}.$$

The coefficient *w* is defined as the *shrinkage intensity* since it indicates how much individual estimates are shrunk towards pooled values.

There are two levels of parameters in the normal hierarchical model for market advisory service performance. The expected performance for each advisory program, $\theta = \left(\theta_1,...,\theta_N\right)$, is in the lower level and hyperparameters, $\left(\mu,\tau\right)$, that combine information for all programs in the sample, are in the higher level. The general structure of a Bayesian hierarchical model includes a prior distribution for the parameters, $p\left(\theta\right)$, that can be decomposed into a conditional prior given the hyperparameters, $p\left(\theta|\mu,\tau\right)$, and the prior for the hyperparameters, $p\left(\mu,\tau\right)$, called the hyper-prior:

(5) $$p\left(\theta\right) = p\left(\mu,\tau\right)p\left(\theta|\mu,\tau\right).$$

Then the related joint posterior distribution can be expressed as:

(6) $$p\left(\mu,\tau,\theta|y\right) \propto p\left(\mu,\tau,\theta\right)p\left(y|\mu,\tau,\theta\right) = p\left(\mu,\tau,\theta\right)p\left(y|\theta\right)$$

where y is the sample information (data distribution). The last equality holds because the hyperparameters affect $p\left(y\right)$ only through the parameters θ. The key characteristic of this model is that individual performance parameters share a common prior. This prior distribution is not subjective nor is it based on information that precedes data collection; instead it is constructed from the entire sample. The prior is based on the notion that performance across programs is not independent, and in this context, not only are data on price performance of a particular program helpful in estimating the expected performance for that program, but also information from the rest of the programs contribute to the estimation. A description of the main points of the model employed in this article follows.

Performance for program *j* is assumed to be normally distributed with mean θ_j and variance v_j^2. The simplifying assumption that variances are known is made such that:[7]

(7) $$y_{jt} \left| \theta_j \sim N\left(\theta_j, v_j^2\right)\right. . \qquad \textit{distribution of } y_{jt}$$

As noted earlier, individual expected performance estimates share the same prior. Specifically, this prior is a normal distribution with mean μ and variance τ:

(8) $$\theta_j \left| \mu, \tau \sim N\left(\mu, \tau\right)\right. \qquad \textit{prior distribution of } \theta_j$$

where the parameter τ defines the prior uncertainty and, as explained below, determines the shrinkage intensity. Combining the sample likelihood derived from equation (7) with the prior distribution (equation 8), the posterior distribution of θ_j conditional on μ and τ is obtained:

(9) $$\theta_j \left| \mu, \tau, y \sim N\left(\hat{\theta}_j, V_j\right)\right. \qquad \textit{conditional posterior distribution of } \theta_j$$

$$\text{where } \hat{\theta}_j = \frac{\frac{1}{\sigma_{\bar{y}_j}^2}\bar{y}_j + \frac{1}{\tau^2}\mu}{\frac{1}{\sigma_{\bar{y}_j}^2} + \frac{1}{\tau^2}},\ V_j = \frac{1}{\frac{1}{\sigma_{\bar{y}_j}^2} + \frac{1}{\tau^2}}, \text{ and } \sigma_{\bar{y}_j}^2 = \frac{v_j^2}{T_i}.$$

The above equation shows that the posterior distribution for each program's expected performance is also normal with a mean equal to the weighted average of the sample mean for that program and the mean of the prior distribution. Note that the point estimate for θ_j is the shrinkage estimator, $\hat{\theta}_j^{shrink}$, presented in equation (4), and that the greater the variance of the sample mean $\sigma^2_{\bar{y}_j}$ the more the individual estimate is shrunk towards μ.[8] Finally, the greater the prior uncertainty, measured by τ^2, the lower the shrinkage intensity.

Up to this point, the posterior distribution of expected performance was defined in terms of the hyperpameters μ and τ. A full Bayesian treatment of hierarchical models includes the definition of a prior distribution for the hyperparameters. Following Gelman et al. (2003), an uninformative prior is employed here for $\hat{\mu}$. The use of this uninformative prior in hierarchical models is reasonable since the entire sample is employed to estimate μ and the total number of observations is large enough to justify relying only on the sample for estimation of this parameter. The posterior distribution of μ conditional on τ is also normal with a mean equal to the Bayesian precision pooled estimate ($\hat{\mu}$):[9]

(10) $$\mu|\tau, y \sim N\left(\hat{\mu}, V_\mu\right)$$ *conditional posterior distribution of* μ

where $\hat{\mu} = \dfrac{\sum_{j=1}^{N} \frac{1}{\sigma^2_{\bar{y}_j} + \tau^2} \bar{y}_j}{\sum_{j=1}^{N} \frac{1}{\sigma^2_{\bar{y}_j} + \tau^2}}$ and $V_\mu = \left[\sum_{j=1}^{N} \frac{1}{\sigma^2_{\bar{y}_j} + \tau^2}\right]^{-1}$

Finally, the posterior distribution of τ is:

(11) $$p\left(\tau|y\right) \propto p\left(\tau\right) V_\mu^{1/2} \prod_{j=1}^{N} \left(\sigma^2_{\bar{y}_j} + \tau^2\right)^{-1/2} \exp\left(-\frac{\left(\bar{y}_j - \hat{\mu}\right)^2}{2\left(\sigma^2_{\bar{y}_j} + \tau^2\right)}\right)$$ *posterior distribution of* τ

An uninformative uniform prior distribution for τ is also assumed. According to Gelman (2006), this type of distribution performs well when the number of groups (advisory programs), is greater than two or three, as is the case in this article. Note that the distribution of τ depends on the dispersion of y_{jt} within programs $\left(\sigma^2_{\bar{y}_j}\right)$ and the dispersion of $\bar{y}_j$ across programs

$\left[\left(\bar{y}_j - \hat{\mu}\right)^2\right]$. Small values of τ will be more likely and the optimal shrinking intensity will be higher the higher the variability of y_{jt} within programs and the lower the dispersion of $\bar{y}_j$ across programs. Also, the number of observations has an effect on the shrinkage intensity. Separate estimates for programs with fewer time-series observations, which are less reliable, have higher values of $\sigma^2_{\bar{y}_j}$ and will be shrunk more towards pooled estimates.

Skeptical Beliefs

Some farm decision-makers may be skeptical about the ability of advisory services to outperform the market, and therefore, unwilling to base expected performance estimates exclusively on past performance observations, as in the basic hierarchical model outlined in the previous section. Farmers could be strongly influenced by the efficient market hypothesis and believe that it is difficult if not impossible to enhance income based on the marketing recommendations of advisory services (Brorsen and Anderson 1994; Zulauf and Irwin 1998; Tomek and Peterson 2005). These views can be incorporated in the Bayesian hierarchical model by adding a prior whose parameters imply that all services have an expected performance close to zero. A similar problem has been considered in the finance literature for the performance of mutual fund managers (Baks, Metrick, and Wachter 2001; Busse and Irvine 2006; Kosowski, Naik, and Teo 2007).

In this article a normal distribution is employed as a prior for skeptical beliefs (equation 12). While there are other ways to model skeptical beliefs, the normal distribution is chosen here to simplify the computation of the posterior distribution:

(12) $$\theta_j \sim N\left(\mu_0, \tau_0^2\right) \qquad \textit{prior distribution of } \theta_j \textit{ under skeptical view}$$

The mean (μ_0) and variance $\left(\tau_0^2\right)$ of the skeptical prior depend on the strength of the skeptical beliefs. The stronger the skeptical beliefs, the closer μ_0 is to zero and the smaller the values of τ_0^2. At one extreme, when both μ_0 and τ_0^2 equal zero it is implied that the decision maker is absolutely sure that advisory programs will obtain an average price equal to the market benchmark price and no data set would change this view. On the other hand, small positive values of μ_0 and large values of τ_0^2 imply that skeptical beliefs are not as strong.

The parameters of the prior distribution for a given view can be obtained from the answers to the following simple questions about advisory services performance:

<u>Question 1</u>: What is the most likely difference between the price obtained by following the recommendations of advisory programs and the market benchmark price?

<u>Question 2</u>: What is the probability that an advisory program outperforms the market benchmark, on average, by more than 5%?

Different from the first version of the Bayesian model, the skeptical prior is based on expectations on the magnitude of advisory program performance that precede the data collection. The answers to these two questions should not depend on the data set employed in the estimation. The mean of the prior is set equal to the answer to the first question and the variance is determined based on the mean and the answer to the second question. In particular, the skeptical view employed in this article corresponds to a decision maker who believes that expected performance is most likely to be zero but there is a 1% probability that expected performance is more than 5% of the market benchmark price. The computation of prior parameters based on the answer to the questions is described in the Appendix. This approach is similar to the one employed by Baks, Metrick, and Wachter (2001), where a prior distribution for mutual funds managers' ability to beat the market is obtained from the answer to similar questions. [10]

The posterior distribution of expected performance from the Bayesian hierarchal model with skeptical beliefs θ_j^{skep} combines the posterior distribution of the hierarchical model with the skeptical prior to obtain the following posterior distribution:

(13)
$$\theta_j^{skep} \left| \mu, \tau, \mu_0, \tau_0, y \sim N\left(\hat{\theta}_j^{skep}, V_j^{2\ skep}\right)\right.$$

$$\text{where } \hat{\theta}_j^{skep} = \frac{\frac{1}{V_j}\hat{\theta}_j + \frac{1}{\tau_0^2}\mu_0}{\frac{1}{V_j} + \frac{1}{\tau_0^2}} \text{ and } V_j^{2\ skep} = \frac{1}{\frac{1}{V_j} + \frac{1}{\tau_0^2}}$$

Simulation Procedures

In the hierarchical model, it is assumed that observed data (pricing performance observations) are normally distributed in the population with a different mean for each group (program), and the group means (average performance for each program) are also normality distributed. To assess whether these normality assumptions are supported by the data, the Jarque-Bera normality test was applied to the performance observations for each program, as well as the distribution of average performance across programs. For corn, normality was not rejected at the 5% significance level for any of the 35 programs, while normality for the distribution of average performance across programs was rejected. For soybeans, normality was not rejected for any of the 34 programs or average performance across programs. For wheat, normality was rejected for 2 out of 28 programs and for average performance. Overall, the small number of departures indicate the normal hierarchical model is a reasonable estimation alternative for the problem being evaluated in the current article.

The computation of the posterior distribution of expected performance is accomplished via simulation. First, the sample information is used to compute the posterior distribution of τ, $p(\tau|y)$ (Equation 11). Figure 2 presents the posterior distribution for the three crops. As shown in panel A, the most likely value of τ is about 3¢/bu. for corn. The most likely values for soybeans and wheat are around 10¢/bu. (panels B and C of figure 2). Recall that τ is the measure of uncertainty for the common prior distribution. A value of $\tau = 0$ indicates no differences in expected performance across programs with all separate estimates being completely shrunk toward the pooled estimator. Note that the figures show that $\tau = 0$ is highly unlikely for advisory performance. Larger values of τ indicate that larger differences in expected performance across programs are more likely. Therefore, the distributions of τ imply that expected performance estimates have higher shrinkage intensity in corn than in soybeans and wheat.

Based on the probability distributions plotted in figure 2, the cumulative distribution of τ is computed. This cumulative density function is employed to simulate τ using the inverse cumulative density function method. The following step is to simulate μ by drawing from its conditional posterior normal distribution $p(\mu|\tau, y)$ (equation 10), given the simulated values for τ. The final step is to simulate θ_j by sampling from the relevant conditional posterior normal distribution $p(\theta_j|\mu, \tau, y)$ (equation 9) given the simulated values for τ and μ.

In the second version of the Bayesian model, skeptical beliefs are added and individual estimates are shrunk towards zero. The degree of shrinkage towards zero depends on the prior uncertainty (τ_0^2 in equation 12) and the variance of individual parameters (V_j in equation 9). Recall that the subjective prior employed in this article corresponds to a decision maker who believes that expected performance is most likely to be zero with a 1% probability that the expected difference between the advisory revenue and the benchmark is more than 5% of the benchmark. The average market benchmark prices are \$2.28/bu., \$5.88/bu., and \$3.02/bu. for corn, soybeans, and wheat, respectively. Therefore, this statement implies a mean of zero $\left(\mu_0 = 0\right)$ and a standard deviation , τ_0^2 , of 5¢/bu., 13¢/bu., and 6¢/bu. for the three crops, respectively. Similar to the Bayesian model without skeptical beliefs, the shrinkage intensity towards zero varies across programs depending on the uncertainty in individual estimates. The computation of the posterior distribution of θ_j^{skep} under skeptical beliefs is again simulated from the normal distribution given the values of the parameters ($\mu_0, \tau_0, \theta_j, V_j$). A complete explanation of the steps followed in the model estimation is included in the Appendix.

Results and Discussion

Tables 2 through 4 present individual expected performance point estimates based on traditional and Bayesian models for corn, soybeans and wheat, respectively. Programs are ordered from highest to lowest based on traditional performance estimates (third column). The Bayesian point estimates are the median of the posterior distribution of the parameters. Note that expected performance estimates for most programs are strongly shrunk towards pooled values (the Bayesian pooled performance estimates, $\hat{\mu}$, in the model for the entire sample are 0.5¢/bu., 13¢/bu. and 2¢/bu for corn, soybeans and wheat, respectively) with a shrinkage intensity generally higher for corn programs compared to soybean and wheat programs. For instance, *AgResource* (program #9) has expected performance of 30¢/bu., 70¢/bu., and 9¢/bu. under traditional estimation for corn, soybeans, and wheat, respectively, and corresponding Bayesian estimates of 2¢/bu., 27¢/bu., and 3¢/bu. The shrinkage intensity for this program (*w* in equation

4) is 97% for corn, 76% for soybeans, and 90% for wheat. In some cases traditional and Bayesian hierarchical estimates are quite similar. For example, AgLine by Doane-cash only (#7) has an expected performance of 14¢/bu. in soybeans under traditional estimation and 13¢/bu. under Bayesian estimation, with a shrinkage intensity of 50%. Recall that the lower the precision of the individual estimate, the greater the shrinkage intensity towards the pooled values. Therefore, programs with wide confidence intervals for the separate estimates (figure 1) have higher shrinkage intensity.

Tables 2 through 4 also present expected performance point estimates for the Bayesian model under skeptical beliefs. It is evident that the point estimates with and without skeptical beliefs are similar for corn, with differences being smaller than 1¢/bu. in all cases. In other words, estimation results do not change much when skeptical views are added to the model. This occurs because the performance of corn advisory programs as a group matches the skeptical prior distribution reasonably well. That is, on average, expected performance of advisory programs is close to zero and therefore the skeptical prior is similar to the prior without skeptical beliefs. The effect of adding skeptical beliefs is stronger in the wheat model and the strongest in the soybeans model. This is the case because performance of advisory programs as a group in soybeans is superior compared to performance in the corn and wheat markets.

Note that there are some differences in the ordering of programs under the different estimation procedures. This is due to differences in shrinkage intensity across programs. For example, *AgriVisor (aggressive cash)* (#13) is ranked 4th according to the traditional estimation in corn and 1st for the hierarchical model with and without skeptical beliefs. Not surprisingly then, a decision maker will choose different programs depending on the beliefs that he/she is willing to incorporate in the model.

Figure 3 provides a graphical representation of the point estimates for expected performance against the market benchmark under the different estimation procedures. The dots above the lower gray line are programs expected to outperform the market benchmark by more than 1% (2¢/bu. for corn, 6¢/bu. for soybeans, and 3¢/bu. for wheat). The dots above the higher gray line are programs expected to outperform the market benchmark by more than 5% (11¢/bu. for corn, 29¢/bu. for soybeans, and 15¢/bu. for wheat). This figure succinctly illustrates shrinkage effects in the Bayesian hierarchical models.

As shown in panel A of figure 3, traditional expected performance point estimates indicate that 18 out of the 35 corn programs have positive performance, 12 programs are expected to outperform the benchmark by more than 1%, and 4 programs by more than 5%. In contrast Bayesian estimates indicate only one program (#13) is expected to outperform the benchmark by more than 1% and none are expected to outperform the benchmark by more than 5%.

Panel B in figure 3 shows that performance of advisory programs in soybeans generally is superior to performance in corn. Based on traditional expected performance estimates, 27 out of the 34 soybean programs have positive expected performance, 25 programs are expected to outperform the benchmark by more than 1%, and 7 programs by more than 5%. Based on the Bayesian model, 33 programs have positive expected performance, 30 are expected to outperform the benchmark by more than 1%, and one program (#29) is expected to outperform the benchmark by more than 5%. In the model including skeptical beliefs, 33 programs have positive expected performance, 27 programs are expected to outperform the benchmark by more than 1% and none are expected to outperform the benchmark by more than 5%.

Panel C in figure 3 shows that there is more dispersion in traditional expected performance estimates across programs for wheat compared to the other crops. Based on traditional expected performance estimates, 14 out of the 28 wheat programs have positive expected performance, 11 programs are expected to outperform the benchmark by more than 1%, and three programs by more than 5%. Based on the Bayesian model, 21 programs have positive expected performance, 7 are expected to outperform the benchmark by more than 1%, and one program (#2) is expected to outperform the benchmark by more than 5%. In the model including skeptical beliefs, 21 programs have positive expected performance, 3 programs are expected to outperform the benchmark by more than 1% and none are expected to outperform the benchmark by more than 5%.

It is interesting to note that when the entire posterior distribution of expected performance in the Bayesian models is considered, not just the point estimates presented in figure 3, there is a 95% or larger probability that expected performance is greater than zero for those programs that are expected to outperform the benchmark by more than 5%. Also, for the majority of the programs that are expected to outperform the benchmark by more than 1% (with the exception of program #6 in soybeans and #7 and #27 in wheat) there is a 75% or more probability that expected performance is greater than zero.

Based on the results presented above, it is evident that the answer to the question of whether farmers should follow the advice of specific market advisory programs depends on the beliefs that the decision-maker is willing to include in the estimation model and the magnitude of expected pricing performance that he/she considers desirable. For instance, consider a decision maker who, based on market efficiency, has a skeptical prior and is willing to follow an advisory program only if it is expected to increase price received by more than 5%. This skeptical decision-maker would prefer to adopt a strategy that mimics the market benchmark rather than follow any advisory program in each of the three crops. Now consider a more "optimistic" decision-maker interested in following programs with expected performance greater than 1% based on the Bayesian model without skeptical beliefs. One of the corn programs, most of the soybeans programs, and a few wheat programs are better marketing alternatives than the market benchmark for this decision-maker.

Farmers and other market participants are also naturally interested in the pricing performance of the top advisory program. Consider the upper bound of the benefits from following advisory programs under the most conservative approach, which is the Bayesian model with skeptical beliefs. Expected gains from following the recommendations of the top performing program are 3¢/bu., for corn, 24¢/bu. for soybeans and 10¢/bu. for wheat when compared to the market benchmark. These values represent a gain of 1%, 4%, and 3% for the three crops, respectively. Expressing expected performance of the top-ranked programs on a per-acre basis provides additional perspective on the magnitude of the upper bound of benefits from following these programs. Consider farms with average yields of 150 bu./acre, 47 bu./acre, and 36 bu./acre for corn, soybeans and wheat, respectively. The expected annual dollar gains for a farmer with skeptical beliefs are around \$4.50/acre for corn, \$11/acre for soybeans and \$3.60/acre for wheat. These values are small, but non-trivial, relative to the net farm income per acre of crop producers (Irwin, Good, and Martines-Filho 2006).

In the decision of whether to follow an advisory program farmers should also compare expected pricing performance with associated subscription costs. These fees are charged annually on a per-farm basis and represent small values when expressed on a per-bushel basis for a medium-sized commercial farm. For the advisory programs in the sample, annual subscription ranges from \$100 to \$600 per farm. The subscription to the most expensive program represents a cost of only 0.8¢/bu. for a farmer growing 500 acres of corn with an average yield of 150 bu./acre. However, there are other costs related to following advisory programs, such as the cost of implementing, monitoring, and managing the marketing strategies recommended by advisory programs (Tomek and Peterson 2001). While these costs are difficult to measure, they may well be large enough to offset a considerable portion of expected benefits, if any, of following advisory programs.

The discussion of the value of advisory services so far is based on comparisons to the 24-month market benchmark. AgMAS performance evaluations consider multiple benchmarks. Expected performance for individual programs was also estimated using the alternative benchmarks. In corn, the 24-market benchmark had, on average, the highest price among the benchmarks. Therefore, advisory programs have somewhat more attractive performance when compared to the rest of the benchmarks. For instance, when compared against the other three benchmarks, more than half of the corn programs have positive expected performance greater than 1% of the benchmark price and one program had performance greater than 5% for the Bayesian model with skeptical beliefs (compare these values with the ones plotted in panel A of figure 3). In the case of soybeans, the 24- and 20-month market benchmarks have similar average prices, while the farmer benchmarks have slightly higher average prices. Therefore, the performance of advisory programs in soybeans is less attractive when compared to farmer benchmarks. Still, based on skeptical estimates, a few programs have expected performance greater than 1% of the average benchmark price for both farmer benchmarks in soybeans. In the wheat market, the average harvest price is the highest among the four benchmarks. In this case, wheat performance is less attractive when advisory prices are compared against harvest prices.

For instance, the pooled estimate for expected performance is negative for both types of wheat. Also, with the exception of one program, Bayesian point estimates for expected performance are negative for individual programs.

Finally, it can be argued that an optimistic prior about expected performance of market advisors should be considered for completeness. The main difficulty in this regard is that no theory is available to support a particular value for positive expected performance. In the case with a skeptical prior, the market efficient hypothesis supports a prior with expected performance equal to zero. It is still interesting to consider a particular case of an optimistic view on advisory program performance. Consider the belief that advisors with superior information tend to recommend more "active" marketing programs (i.e., a tendency to make large bets on price movements). This belief is consistent with the results presented in Cabrini, Irwin, and Good (2007) for corn and soybeans markets. Under this view, it is reasonable to group advisory programs into active and conservative categories and to estimate a Bayesian hierarchical model for each group. Not surprisingly, the expected gain from following top-performing programs is higher: 6% for corn, 8% for soybeans, and 9% for wheat.[11]

Summary and Conclusions

This article employs a Bayesian hierarchical approach to estimate expected performance of agricultural market advisory programs. This estimation procedure is a conservative approach compared to traditional estimation, since it reduces estimation error in the expected gains from following top-performing advisory programs. Two versions of the model are estimated. The first combines information across the sample, while the second includes skeptical beliefs based on the efficient market hypothesis. The data consist of past observations of performance for corn, soybean, and wheat advisory programs for the 1995 to 2004 crop years. The posterior distribution of individual expected performance is computed by simulation.

The Bayesian hierarchical model produces shrinkage estimators that are weighted averages of individual and pooled estimates. Adding a skeptical prior shrinks performance estimates further towards zero. The answer to the question of whether farmers should follow the advice of any market advisory program depends on the beliefs that the decision-maker is willing to include in the estimation model and the magnitude of expected pricing performance that he/she considers desirable. For a skeptical decision maker who is willing to follow an advisory program only if it is expected to increase price received by more than 1%, one of the corn programs, most of the soybean programs, and a few wheat programs may be valuable marketing alternatives. A skeptical decision maker who is willing to follow an advisory program only if it is expected to increase price received by more than 5% would prefer to adopt a strategy that mimics the market benchmark rather than follow any advisory program in each of the three crops. More specifically, the expectation under the skeptical model from following the single top-ranked program is a price increase of 1%, 4%, and 3% for corn, soybeans, and wheat, respectively. These values imply a combined expected annual gain around $4.50/acre for corn, $11/acre for soybeans, and $3.60/acre for wheat. Whether these gains offset the cost of implementing, monitoring, and managing the recommended marketing strategies is an open question.

A complete comparison of the estimation procedures employed in this study would include out-of-sample tests. However, more performance observations per program are needed for this type of analysis. Based on previous research (e.g. Jorion 1986; Grauer 1995) shrinkage estimators should provide more accurate estimates than traditional estimators when the number of programs is high and the number of observation per program is low, which is the case here. Therefore, marketing advice offered to farmers should be based on the Bayesian estimates reported in this article rather than on traditional ones.

An interesting extension of the current article would be the inclusion of risk in the performance evaluation. The hierarchical Bayesian approach can also be applied to evaluate the benefits from following different combinations of marketing advisory programs in a portfolio optimization context. In this case, the estimation of the covariance matrix for advisory pricing performance represents a challenge given the data availability restrictions. The Bayesian approach implemented in this article also provides an interesting framework for more general evaluations of grain marketing alternatives. For instance, a Bayesian hierarchical model with and without skeptical beliefs can be employed in the estimation of expected gains of different combinations of cash and derivatives transactions and marketing contracts offered by grain companies. These applications of Bayesian modeling in grain marketing represent interesting opportunities for further research.

References

Baks, K.P., A. Metrick, and J. Wachter. 2001. "Should Investors Avoid All Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation." *Journal of Finance* 56:45-85.

Batts, R.M., S.H. Irwin, and D.L. Good. 2009. "The Pricing Performance of Market Advisory Services in Wheat Over 1995-2004." AgMAS Project Research Report 2009-01, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign (forthcoming).

Brorsen, B.W., and K. B. Anderson. 1994. "Cash Wheat Marketing: Strategies for Real People." *Journal of Agribusiness* 12:85-94.

Busse, J.A., and P.J. Irvine. 2006. "Bayesian Alphas and Mutual Fund Persistence." *Journal of Finance* 61:2251-2288.

Cabrini, S.M., S.H. Irwin, and D.L. Good. 2007. "Style and Performance of Agricultural Market Advisory Services." *American Journal of Agricultural Economics* 89:607-623.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2003. "Hierarchical Models." Ch. 5 in *Bayesian Data Analysis, Second Edition*. New York, NY: Chapman and Hall/CRC.

Gelman, A. 2006. "Prior Distributions for Variance Parameters in Hierarchical Models." *Bayesian Analysis* 3:515-533.

Gehrt, D., and D.L. Good. 1993. "Evaluation of Market Advisory Services for Corn and Soybeans." *Journal of the American Society of Farm Managers and Rural Appraisers* 57:1-7.

Grauer, R.R. 1995. "Stein and CAPM Estimators of the Means in Asset Allocation." *International Review of Financial Analysis* 4:35-66.

Irwin, S.H., D.L. Good, J. Martines-Filho, and R. M. Batts. 2006. “The Pricing Performance of Market Advisory Services in Corn and Soybeans Over 1995-2004.” AgMAS Project Research Report 2006-02, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

Irwin, S.H., D.L. Good, and J. Martines-Filho. 2006. “The Performance of Market Advisory Services in Corn and Soybeans.” *American Journal of Agricultural Economics* 88:162-181.

Isengildina, O., J. M. E. Pennings, S. H. Irwin, and D. L. Good. 2006. “U.S. Crop Farmers’ Use of Market Advisory Services.” *Journal of International Food and Agribusiness Marketing* 18:67-86.

Jorion, P. 1986. “Bayes-Stein Estimation for Portfolio Analysis.” *Journal of Financial and Quantitative Analysis* 21:279-292.

Kastens, T.L., and T.C. Schroeder. 1996. "Efficiency Tests of July Kansas City Wheat Futures." *Journal of Agricultural and Resource Economics* 21:187-198.

Kosowski, R., N.Y. Naik, and M. Teo. 2007. “Do Hedge Funds Deliver Alpha? A Bayesian and Bootstrap Analysis” *Journal of Financial Economics* 84:229-264.

Michaud R.O. 1989. “The Markowitz Optimization Enigma: Is ‘Optimized’ Optimal?” *Financial Analysts Journal* 45:31-42.

Patrick, G.F., W.N. Musser, and D.T. Eckman. 1998. “Forward Marketing Practices and Attitudes of Large-Scale Midwestern Grain Farmers.” *Review of Agricultural Economics* 20:38-53.

Tomek, W.G., and H.H. Peterson. 2001. "Risk Management in Agricultural Markets: A Review." *Journal of Futures Markets* 21:853-985.

Tomek, W.G., and H.H. Peterson. 2005. "Implications of Commodity Price Behavior for Marketing Strategies." *American Journal of Agricultural Economics* 87:1258-1264.

Zulauf, C.R., and S.H. Irwin. 1998. "Market Efficiency and Marketing to Enhance Income of Crop Producers." *Review of Agricultural Economics* 20:308-331.

Appendix

Step by step computation of the Bayesian hierarchical models to estimate individual expected performance of market advisory programs in corn, soybeans and wheat.

For each of the three crops the following steps were followed:

1. Compute the performance for each advisory program in each crop year:

Subtract the benchmark price from the advisory price for each program in each crop year (Equation 1, advisory price is the price received by a farmer who markets the crop following advisory program recommendations)

2. Compute average performance for each program $(\bar{y}_j,$ equation 2$)$ and the standard errors $\hat{\sigma}^2_{\bar{y}_j}$ for these parameters.

3. Compute the posterior distribution for τ (Equation 11): Use the values of average performance for each program $(\bar{y}_j)$ and the corresponding standard error $(\hat{\sigma}_{\bar{y}_j})$ to compute the posterior probabilities for a range of values of τ according to equation (11). The computation of $p\left(\tau \middle| y\right)$ is applied to a grid of equally spaced values of τ within a range. The range is determined by computing $p\left(\tau \middle| y\right)$ for a wide array of values for τ and then selecting the range of τ with non-negligible positive probability. The distributions presented in figure 2 are obtained.

4. Compute the cumulative distribution of τ: A normalizing factor needs to be applied because the probability distribution presented in equation (11) is defined up to an unknown normalizing constant (note that the proportional sign , $\propto$, is employed in the equation). The normalizing factor (K) is calculated by adding up all values of $p(\tau|y)$ for the range of τ considered. Then, the values of $p(\tau|y)$ are divided by the normalizing factor and the cumulative density function is computed by summing the values of $p(\tau|y)/K$ for τ less than or equal to the given value.

5. Sample from the cumulative density function of τ using the uniform distribution. A thousand simulated values of τ are obtained.

6. Simulate μ by drawing from its conditional posterior normal distribution $p(\mu|\tau, y)$ (equation 10), given the simulated values of τ. A thousand values of μ are generated. Note that the mean and variance of the posterior distribution of μ can be computed based on the sample statistics $\overline{y}_j$ and $\hat{\sigma}_{\overline{y}_j}$ and a given value of τ.

7. Simulate θ_j by sampling from the relevant conditional posterior normal distribution (equation 9) given the simulated values for τ and μ. One thousand simulated values for expected performance for each program are generated.

8. For each program the 1000 values of θ_j are ordered by magnitude and the 0.05, 0.25, 0.50, 0.75 and 0.95 quantiles are computed. These quantiles describe the distribution of expected performance for each program. The Bayesian point estimates reported in the study are the median values of these 1000 simulated values for each program.

Bayesian model with skeptical beliefs

The estimation procedure for the model with skeptical beliefs is done following the steps 1 to 6 presented above and the additional steps presented next:

7'. Set values for the skeptical prior parameters: The subjective prior employed corresponds to a decision maker who believes that expected performance is most likely to be zero with a 1% probability that the expected difference between the advisory revenue and the benchmark is more than 5% of the benchmark. Therefore, this statement implies a mean of zero $(\mu_0 = 0)$ and a standard deviation of 5¢/bu., 13¢/bu., and 6¢/bu. for the three crops, respectively. To compute standard deviations, a value equal to 5% of the average benchmark price for corn, soybeans, and wheat (11¢/bu. for corn, 29¢/bu. for soybeans, and 15¢/bu. for wheat) is calculated, and then this number is divided by *z*:

(A.1) $$\tau_0 = (0.05 * \overline{BP}) / z \qquad \text{where} \quad P(Z \geq z) = 0.01 \quad \rightarrow \quad z = 2.33$$

where Z has standard normal distribution.

8'. Simulate θ_j^{skep} by sampling from the relevant conditional posterior normal distribution (equation 13), given the simulated values for τ and μ and the values for μ_0 and τ_0 presented above. One thousand simulated values for expected performance for each program, θ_j^{skep} are obtained.

9'. For each program the 1000 values of θ_j^{skep} are ordered by magnitude and the 0.05, 0.25 , 0.50, 0.75 and 0.95 quantiles are indentify. These quantiles describe the distribution of expected performance for each program based on the skeptical model. The Bayesian point estimates for the skeptical model are the median values of these 1,000 simulated values for each program.

Endnotes

[1] The term "advisory program" is used throughout the remainder of this article because several advisory services have more than one distinct marketing program

[2] A minimum of two observations is necessary to estimate the standard error of the separate estimates of expected performance.

[3] Two market benchmarks and two farmer benchmarks are used in AgMAS performance evaluations in corn and soybeans (Irwin et al. 2006). Three market benchmarks and one farmer benchmark are used in wheat performance evaluations (Batts, Irwin, and Good 2009). Market benchmarks measure the average price offered by the market over the marketing window of a representative farmer who follows advisory program recommendations. The average price is computed in order to reflect the returns to a naïve, "no-information" strategy of marketing equal amounts of grain each day during the marketing window. Farmer benchmarks measure the average price actually received by farmers for a crop.

[4] The estimation procedures employed in the study assumes that performance observations are independent for consecutive crop years. Since marketing windows overlap across crop years this may not be the case. However, Irwin et al. (2006) and Batts, Irwin, and Good (2009) report that rank correlations of net advisory prices in adjacent crop years average only 0.27 in corn, 0.25 in soybeans, and 0.12 in wheat. This small degree of positive correlation is unlikely to substantially impact estimation procedures.

[5] Confidence intervals for pooled estimates are much narrower compared to intervals for individual programs; note that markers in figure 1 for the pooled estimator cover the line representing the confidence intervals.

[6] The formula for the pooled estimate is presented in equation (10). It has a similar structure to the formula for $\hat{y}^{pool}$ (equation 3).

[7] Although this assumption is not true in actual applications, it is commonly used as a reasonable approximation (Gelman 2003). The assumption of known variances allows working with the normal distribution and focusing on the estimation of expected performance.

[8] The shrinkage coefficient, w, is equal to $\frac{1}{\tau^2} \Big/ \left(\frac{1}{\sigma^2_{\bar{y}_j}} + \frac{1}{\tau^2} \right)$.

[9] Note that $\hat{y}^{pool}$ in equation (3) is similar to $\hat{\mu}$, with the difference that the prior uncertainty, τ, is included in the computation of $\hat{\mu}$.

[10.] Baks, Metrick, and Wachter (2001) use a more complex functional form for the prior, an asymmetric distribution with a lower bound and a right tail of a normal distribution, which leads to a more complex posterior distribution of expected performance. The idea in this study is that fund managers are expected to loose at most the transaction costs, on average. The authors argue that losses greater than this limit imply consistently trading on misinformation or the existence of behavioral biases.

[11] A detailed description of the activeness measures considered for corn and soybeans is presented in Cabrini, Irwin, and Good (2007). Advisory programs are divided into two groups (conservative and active) employing a cluster method. Information on the activeness degree of advisory programs is not available for wheat, so the same two groups were considered for wheat performance estimation. This is reasonable since Cabrini, Irwin, and Good (2007) indicate that advisory programs have very similar degrees of activeness across corn and soybeans.

Table 1. List of Market Advisory Programs Tracked by the AgMAS Project over the 1995-2004 Crop Years

ID	Market Advisory Program	Crop Years
1	Ag Alert for Ontario	1996
2	Ag Financial Strategies	2001-2004
3	Ag Market Professional (cash only)	2004
4	Ag Market Professional (hedge)	2004
5	Ag Profit by Hjort	1995-1999
6	Ag Review	1995-2004
7	AgLine by Doane (cash only)	1995-2004
8	AgLine by Doane (hedge)	1996-2004
9	AgResource	1995-2004
10	Agri-Edge (cash only)	1995-1996
11	Agri-Edge (hedge)	1995-1996
12	Agri-Mark	1995-2000
13	AgriVisor (aggressive cash)	1995-2004
14	AgriVisor (aggressive hedge)	1995-2004
15	AgriVisor (basic cash)	1995-2004
16	AgriVisor (basic hedge)	1995-2004
17	Allendale (futures & options)	1996-2004
18	Allendale (futures only)	1995-2004
19	Brock (cash only)	1995-2004
20	Brock (hedge)	1995-2004
21	Cash Grain	1999
22	Co-Mark	2000-2002
23	Freese-Notis	1995-2004
24	Grain Field Marketing	2001-2004
25	Grain Field Report	1995
26	Grain Marketing Plus	2000-2001
27	Harris Weather/Elliott Advisory	1995-1996
28	North American Ag	1995
29	Northstar Commodity	2001-2004
30	Pro Farmer (cash only)	1995-2004
31	Pro Farmer (hedge)	1995-2004
32	Progressive Ag	1996-2004
33	Prosperous Farmer	1995
34	Risk Management Group (cash only)	1999-2004
35	Risk Management Group (futures & options)	1999-2004
36	Risk Management Group (options only)	1999-2004
37	Stewart-Peterson Advisory Reports	1995-2004
38	Stewart-Peterson Strictly Cash	1995-1999
39	Top Farmer Intelligence	1995-2004
40	Utterback Marketing Services	1997-2004
41	Zwicker Cycle Letter	1995-1998

Notes: The Allendale (futures & options) program is offered only for corn. Ag Alert for Ontario, Ag Market Professional (cash only), Co-Mark and Grain Marketing Plus were tracked only for corn and soybeans.

Table 2. Traditional Individual and Bayesian Estimates of Expected Performance versus Market Benchmark for Advisory Programs in Corn, 1995-2004 Crop Years

Program ID	Number of Observations	Traditional Individual Estimates	Bayesian Hierarchical Model Estimates	Bayesian Hierarchical Model Estimates with Skeptical Beliefs
			---$/bu.---	
9	10	0,298	0,015	0,010
11	3	0,242	0,015	0,011
40	8	0,174	0,020	0,016
13	10	0,127	0,041	0,032
6	10	0,091	0,018	0,015
32	9	0,077	0,011	0,008
14	10	0,057	0,015	0,012
7	10	0,051	0,020	0,017
12	6	0,047	0,007	0,005
8	9	0,043	0,018	0,017
15	10	0,027	0,011	0,009
16	10	0,023	0,010	0,008
39	10	0,022	0,007	0,004
34	6	0,013	0,006	0,007
35	6	0,013	0,006	0,005
24	4	0,012	0,008	0,009
41	4	0,008	0,007	0,005
36	6	0,003	0,005	0,004
19	10	-0,001	0,004	0,003
20	10	-0,003	0,005	0,003
29	4	-0,004	0,002	0,003
22	4	-0,009	-0,002	-0,001
10	3	-0,010	0,004	0,004
21	2	-0,013	-0,005	-0,003
23	10	-0,016	0,001	0,001
17	9	-0,018	0,000	0,001
38	6	-0,030	-0,006	-0,004
18	10	-0,046	0,001	0,004
27	2	-0,056	0,005	0,004
37	10	-0,082	-0,008	-0,006
30	10	-0,101	-0,013	-0,011
31	10	-0,104	-0,018	-0,015
26	3	-0,124	-0,003	-0,002
5	5	-0,131	-0,005	-0,003
2	4	-0,187	-0,012	-0,009

Note: Names of advisory programs corresponding to the ID numbers are listed in table 1.

Table 3. Traditional Individual and Bayesian Estimates of Expected Performance versus Market Benchmark for Advisory Programs in Soybeans, 1995-2004 Crop Years

Program ID	Number of Observations	Traditional Individual Estimates	Bayesian Hierarchical Model Estimates	Bayesian Hierarchical Model Estimates with Skeptical Beliefs
			---$/bu.---	
9	10	0,702	0,266	0,168
32	9	0,535	0,206	0,131
29	4	0,429	0,307	0,236
22	4	0,390	0,202	0,127
12	6	0,383	0,160	0,101
41	4	0,356	0,189	0,128
24	4	0,333	0,155	0,096
40	8	0,241	0,140	0,093
21	2	0,231	0,140	0,089
14	10	0,216	0,165	0,119
16	10	0,199	0,163	0,121
20	10	0,189	0,147	0,102
13	10	0,157	0,148	0,108
31	10	0,153	0,141	0,116
37	10	0,152	0,137	0,112
27	2	0,152	0,125	0,077
7	10	0,140	0,138	0,110
30	10	0,140	0,134	0,101
11	3	0,140	0,131	0,092
15	10	0,134	0,127	0,102
10	3	0,134	0,130	0,081
8	7	0,112	0,121	0,097
39	10	0,106	0,115	0,091
26	3	0,078	0,110	0,075
19	10	0,066	0,079	0,068
18	10	0,046	0,087	0,066
34	6	0,027	0,078	0,057
23	10	-0,006	0,020	0,018
38	6	-0,009	0,014	0,010
35	6	-0,009	0,072	0,049
36	6	-0,027	0,059	0,044
5	5	-0,065	0,079	0,054
2	4	-0,086	-0,041	-0,039
6	10	-0,294	0,065	0,044

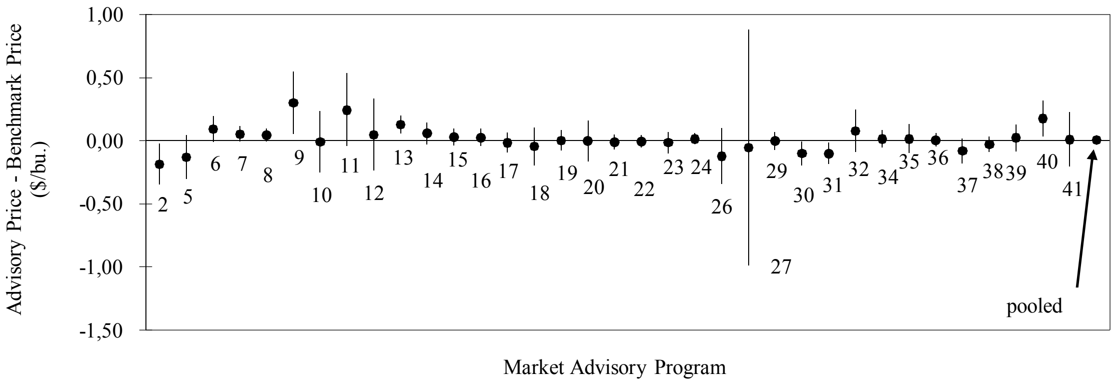
Note: Names of advisory programs corresponding to the ID numbers are listed in table 1.

Table 4. Traditional Individual and Bayesian Estimates of Expected Performance versus Market Benchmark for Advisory Programs in Wheat, 1995-2004 Crop Years

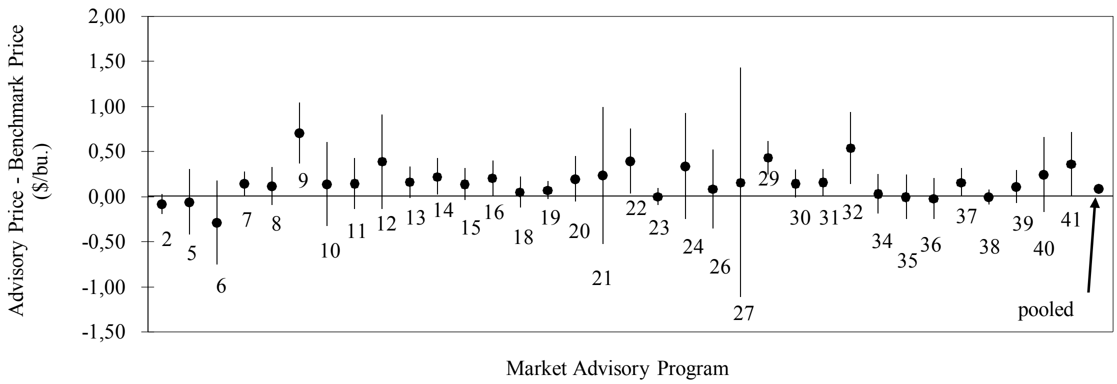
Program ID	Number of Observations	Traditional Individual Estimates	Bayesian Hierarchical Model Estimates	Bayesian Hierarchical Model Estimates with Skeptical Beliefs
			---$/bu.---	
27	2	0,546	0,033	0,009
2	4	0,343	0,192	0,096
25	2	0,180	0,030	0,014
18	10	0,141	0,080	0,043
23	10	0,132	0,060	0,028
29	4	0,132	0,026	0,015
9	10	0,090	0,027	0,008
6	10	0,088	0,029	0,012
7	10	0,087	0,042	0,018
24	3	0,073	0,050	0,033
20	10	0,059	0,042	0,024
31	10	0,008	0,017	0,009
38	6	0,008	0,013	0,011
5	5	0,005	0,016	0,008
14	10	-0,015	0,010	0,006
11	2	-0,020	0,020	0,011
19	10	-0,025	-0,008	-0,004
8	7	-0,033	0,005	0,001
30	10	-0,035	0,000	0,004
16	10	-0,050	0,004	0,001
37	10	-0,077	-0,032	-0,019
15	10	-0,099	-0,004	-0,001
13	10	-0,101	-0,002	0,000
40	8	-0,130	-0,019	-0,003
10	2	-0,136	0,018	0,006
32	3	-0,181	0,007	0,005
39	10	-0,254	-0,094	-0,041
41	4	-0,544	-0,034	-0,008

Note: Names of advisory programs corresponding to the ID numbers are listed in table 1.

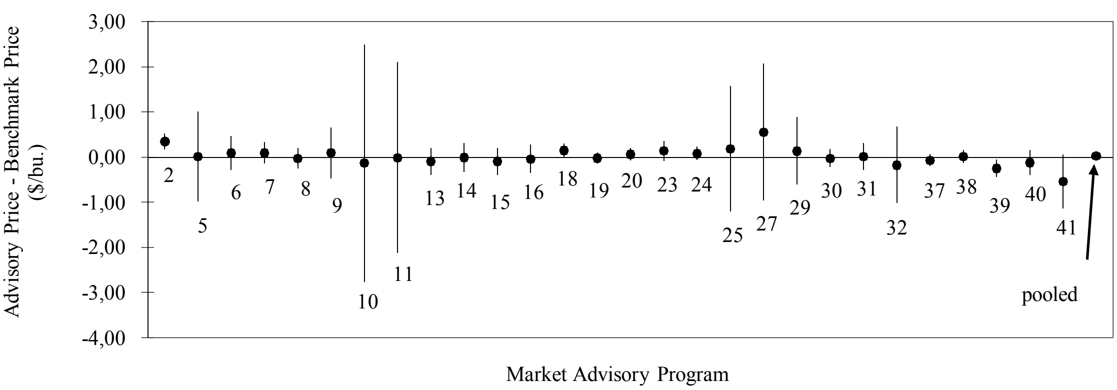
Panel A. Corn



Panel B. Soybeans



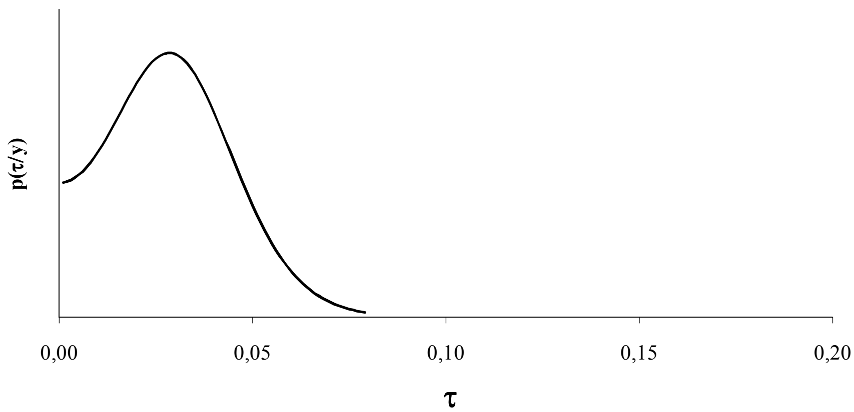
Panel C. Wheat



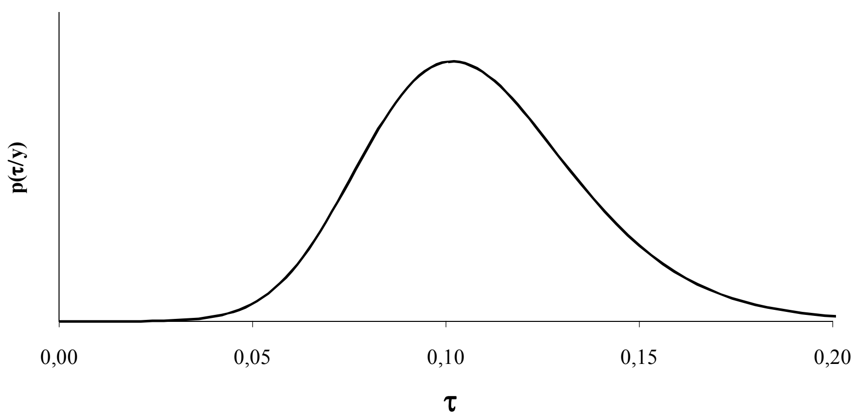
Note: Advisory program names are listed in table 1. The dots in the figures represent point estimates and lines the 90% confidence intervals for expected performance.

Figure 1. Expected performance of advisory programs versus market benchmark, traditional individual point estimates and 90% confidence intervals, 1995-2004 crop years

Panel A. Corn



Panel B. Soybeans



Panel C. Wheat

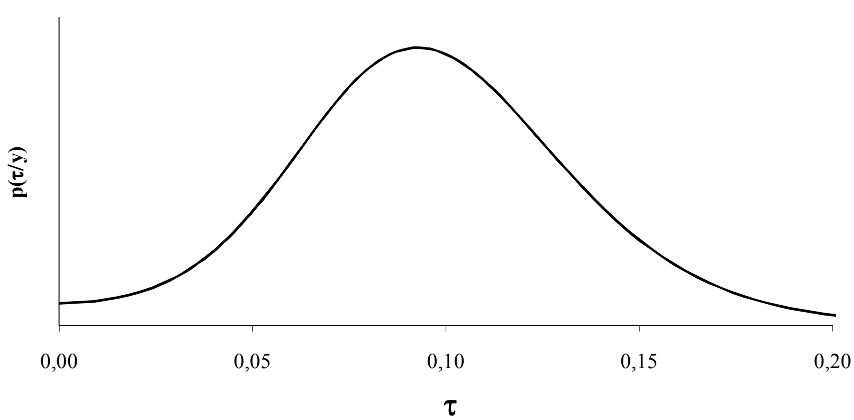
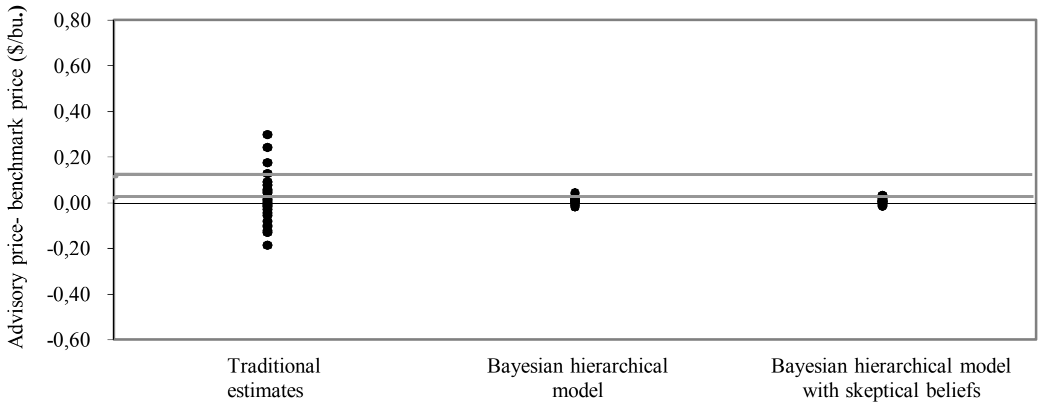
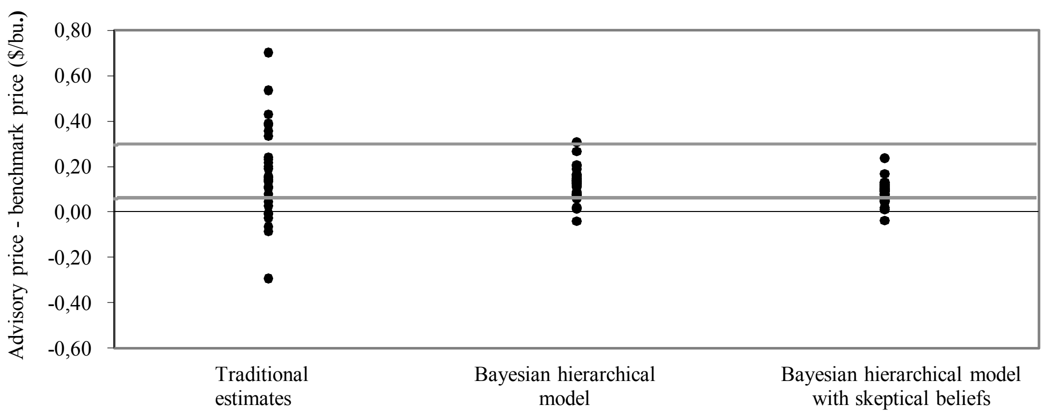


Figure 2. Marginal posterior density of τ ($/bu.) for advisory service performance in corn, soybeans, and wheat versus the market benchmark, 1995-2004 crop years

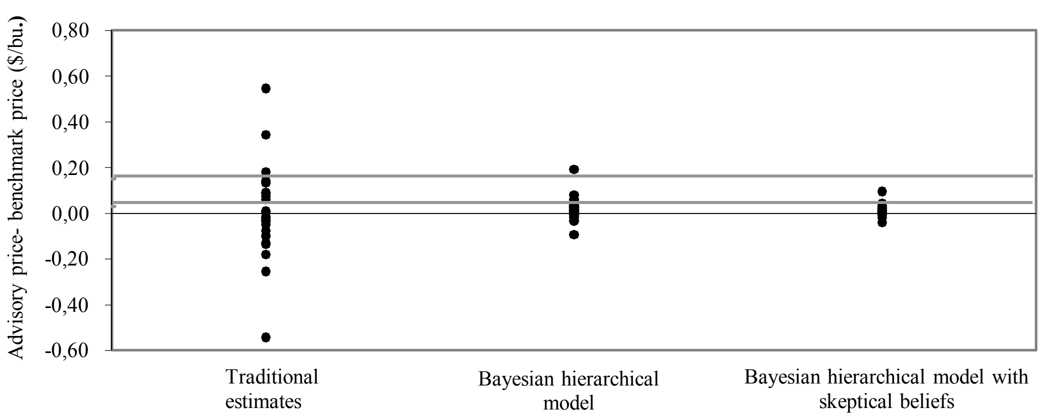
Panel A. Corn



Panel B. Soybeans



Panel C. Wheat



Note: The dots represent point estimates of expected pricing performance for each advisory program under the different estimation models. Dots above the lower gray line represent programs expected to outperform the benchmark by more than 1%. Dots above the higher gray line represent programs expected to outperform the benchmark by more than 5%.

Figure 3. Expected pricing performance for advisory programs versu the market benchmark, traditional individual and bayesian point estimates, 1995-2004 crop years