# On the linkages in U.S. public R&D spending, knowledge capital and agricultural productivity growth: A Bayesian approach

Uris Lantz C. Baldos[1*], Frederi G. Viens[2], Thomas W. Hertel[1] and Keith O. Fuglie[3]

[1]Department of Agricultural Economics
Purdue University
West Lafayette, IN
[*]E-mail: ubaldos@purdue.edu

[2]Department of Statistics
Purdue University
West Lafayette, IN

[3]Economic Research Service
U.S. Department of Agriculture
Washington, DC

*Selected Paper prepared for presentation at the*
*2015 Agricultural & Applied Economics Association and*
*Western Agricultural Economics Association Annual Meeting,*
*San Francisco, CA, July 26-28.*

PRELIMINARY DRAFT AND RESULTS – DO NOT CITE

## I. Overview

Aggressive investments in agricultural research have led to the historic rise in productivity growth and subsequent expansion in U.S. farm production. Indeed over the period 1910 to 2007, growth in total output of U.S. agriculture outpaced that of total input use – at 1.6% versus 0.2% per annum, respectively – owing primarily to productivity growth (Pardey & Alston, 2012). Understanding the linkages between R&D spending and agricultural productivity is therefore crucial in informing future investment decisions particularly at a time wherein public research budgets are heavily constrained. However, quantifying the returns to R&D spending is confounded by the temporal nature of R&D knowledge stock formation. Clearly the process of technological innovation – from development to commercial adoption – takes time. For example, starting from experiments in field stations it took almost four decades for U.S. hybrid corn to be adopted commercially by farmers (Pardey & Beddow, 2013). Equally important is the uncertainty involved in the accumulation of knowledge resulting from R&D spending. Past scientific breakthroughs can be easily outdated by new discoveries while some remain relevant up to this day. Furthermore, the pool of knowledge capital stock is quite sensitive to the types of scientific research being conducted as well as the prevailing institutional and economic systems (Pardey & Beintema, 2001).

Innovations and technologies that we use today – which directly boost farm productivity – are inspired by accumulated knowledge capital which among other factors hinges on investments in agricultural R&D. In empirical studies which examined the relationship between U.S. R&D expenditures and agricultural productivity growth, the transformation of R&D spending to knowledge capital stocks is captured by lag weights (Alston, Andersen, James, & Pardey, 2011; Alston, James, Andersen, & Pardey, 2010b; Heisey, Wang, & Fuglie, 2011;

Huffman, 2009). These lags has been structured to embody the stages of technological innovation starting from initial research efforts towards commercial development and adoption of new technologies in the production system (Alston et al., 2010a). Although this strategy particularly useful in modeling the temporal transformation of R&D spending to knowledge stocks, current methods clearly ignore the uncertainty in knowledge stock accumulation. Indeed, it is widely recognized that knowledge stocks are quite sensitive to the chosen lag weight structure and its corresponding model parameters (Alston et al., 2010b). At the moment, experts broadly address this issue by considering different lag structures and distributional parameters (Alston et al., 2010a). However, such *ad hoc* approach fails to formally incorporate the assessment of uncertainty in the conversion of R&D expenditures to knowledge capital stocks and the subsequent gains in agricultural productivity from increased R&D investments.

In this paper, we employ Bayesian hierarchical modeling to better capture and communicate the uncertainties surrounding the transformation of U.S. public agricultural R&D expenditures to knowledge capital stock as well as its contribution to U.S. agricultural total factor productivity growth (TFP). Compared to traditional approaches founded on classical statistics, analytical methods grounded on Bayesian inference explicitly incorporate existing information and permits continuous revision of our knowledge regarding the distribution of the unknown model parameters. Since we are explicitly estimating posterior distributions, we can make stronger and informative probabilistic statements regarding our findings.

This paper is structured as follows. We start with a primer which summarizes the advantages of Bayesian inference over classical statistics. We then continue on with the development of the hierarchical model relating U.S. public R&D spending to knowledge capital stocks and finally to U.S. agricultural productivity growth. We carefully define the prior distributions of unknown model parameters and derive the required posterior distributions. We then outline our strategy in implementing the Bayesian hierarchical model in the context of the historical U.S. experience. Finally, we conclude the paper by discussing results and outlining some areas for future work.

## II. Methodology

***A primer on Bayesian inference:*** At the core of Bayesian analysis is the transformation of uncertainty to tractable probabilities through meticulous modelling of posterior distributions of unknown model parameters. Bayes' theorem links conditional and marginal probabilities of stochastic events (or parameters).

$$p\left(\mathrm{A}|\mathrm{B}\right)=\frac{p\left(\mathrm{B}|\mathrm{A}\right)p\left(\mathrm{A}\right)}{p\left(\mathrm{B}\right)} \qquad p\left(\mathrm{A}|\mathrm{B}\right)\propto p\left(\mathrm{B}|\mathrm{A}\right)p\left(\mathrm{A}\right)$$

Posterior distribution of unknown parameters ( $p\left(\mathrm{A}|\mathrm{B}\right)$ ) can be expressed as a proportion of the likelihood of data ( $p\left(\mathrm{B}|\mathrm{A}\right)$ ) and the prior distribution ( $p\left(\mathrm{A}\right)$ ). Prior distribution summarizes our initial beliefs and information regarding the unknown parameters while the likelihood of data contains all information relevant for inference. We then update our knowledge about the posterior distributions using observable data and iteratively evaluating the fitness of the model (Gelman, Carlin, Stern, & Rubin, 2003). Since inferences are grounded on posterior distributions of parameters and are exclusive on the data being evaluated, the probability statements based on Bayesian analysis are more informative. Bayesian methods also naturally favor parsimonious compared to over-parameterized models (Berger, 2006).

In contrast, classical statistics is geared towards making general inferences about the unknown population parameter given a random sample distribution; thus, any probability statement is strictly constrained to what we can say about the random distribution. To illustrate the weakness of the classical approach, take the case of a simple linear regression relating R&D stocks to agricultural TFP. Suppose that we are interested in formulating confidence intervals regarding the estimator of the partial effect of R&D stocks to agricultural TFP. Since the estimator is a random variable, we can then repeatedly calculate confidence intervals. Assuming that these intervals are centered on the unknown population parameter, we can then infer that 95% of these intervals will contain the said parameter. Note that we cannot make direct probability statements since the intervals will either contain the population parameter or not. Furthermore, probability statements under the classical approach cannot be applied to specific a dataset due to the underlying assumption that any data used is a sample from a random distribution (Bolstad, 2007; Robert, 2007).

**Bayesian hierarchical framework:** We now motivate our hierarchical model and examine key equations relating public R&D spending to R&D knowledge stocks and to agricultural TFP growth for both the linear and log specification:

$$T_t = \alpha_0 + \alpha_1 RD_t + \alpha_2 CI_t + \alpha_3 t + \varepsilon_{T_t}\sigma_T \qquad \text{Eq.1.a}$$

$$RD_t = \sum_{i=0}^{49} \beta_{RD,i} XD_{t-i} + \varepsilon_{RD_t}\sigma_{RD} \qquad \text{Eq.2.a}$$

$$\log T_t = \alpha_0 + \alpha_1 \log RD_t + \alpha_2 \log CI_t + \alpha_3 \log t + \varepsilon_{T_t}\sigma_T \qquad \text{Eq.1.b}$$

$$\log RD_t = \log\left(\sum_{i=0}^{49} \beta_{RD,i} XD_{t-i} + \varepsilon_{RD_t}\sigma_{RD}\right) \qquad \text{Eq.2.b}$$

Following the literature, we defined agricultural total factor productivity ( $T_t$ ) as a function of R&D knowledge capital stocks ( $RD_t$ ), corn moisture stress index ( $CI_t$ ), time trend ( $t$ ) and error terms ( $\varepsilon_{T_t}\sigma_T$ ) (Eq.1). Total factor productivity captures the changes in overall output which cannot be explained by total input use. Corn moisture index serve as a proxy for drought and incorporates information on the intensity of drought in agricultural areas cropped with corn (NOAA, 2015) while the time trend incorporates other drivers of agricultural TFP which changes overtime. The accumulation of R&D knowledge stocks is defined in Equation 2. Present stocks of R&D knowledge capital is derived from the stream of annual R&D spending ( $XD_t$ ) plus error terms ( $\varepsilon_{RD_t}\sigma_{RD}$ ). Note that R&D lag parameters govern the transmission of R&D spending to knowledge capital stocks and summarizes the stages of technological innovation and adoption. In calibrating the R&D lag parameters, we follow Alston et al (2010a) and adapt a gamma distribution structure with a 50-year lag span. Under this distribution, R&D lags can be parameterized by two factors $\delta$ and $\lambda$ :

$$\beta_{RD,i} = \frac{(i+1)^{\delta/1-\delta}\left(\lambda\right)^{i}}{\sum_{i=0}^{L}(i+1)^{\delta/1-\delta}\lambda^{i}} \text{ and } \sum_{i=0}^{49}\beta_{RD,i} = 1 \qquad \text{Eq.3}$$

Implementation of Bayesian hierarchical modeling requires information on the prior distribution of key model parameters. We define our prior distributions for parameters $\boldsymbol{\alpha} := (\alpha_0, \alpha_1, \alpha_2, \alpha_3), \sigma_T, \sigma_{RD,} \lambda, \sigma$ .

| Linear model | Log model |
|---|---|
| $\boldsymbol{\alpha} \sim N(0,1)$ | $\boldsymbol{\alpha} \sim N(0,1)$ |
| $\sigma_T^2 \sim IG(2,0.1); \sigma_{RD}^2 \sim IG(2,0.1)$ | $\sigma_T^2 \sim IG(2,0.1); \sigma_{RD}^2 \sim IG(2,0.1)$ |
| $\lambda \sim TN(0.75, 0.85, 0.65)$ | $\lambda \sim TN(0.70, 0.80, 0.65)$ |
| $\delta \sim TN(0.80, 0.80, 0.65)$ | $\delta \sim TN(0.90, 0.90, 0.85)$ |

In specifying the priors for the gamma distribution for the R&D lag parameters, we rely on the subset of parameter values identified by Alston et al (2010a) which best fit U.S. state-level analysis. With the priors and model at hand, we then carefully derive the posterior distributions of each parameter and then calculate the updated R&D capital stocks.

**Model implementation and data requirements:** Advances in computational techniques in recent years have made practical Bayesian inference possible. Following, Barboza et al (2014), we employ using the Markov-Chain Monte Carlo (MCMC) methods particularly the standard Gibbs Sampler to come up with the approximated posterior distributions. In addition to this, we use the Metropolis-Hastings algorithm to sample the parameters governing the gamma distribution of R&D lags parameters. In total, around 10,000 samples are drawn but we only use half of the samples to allow convergence in the results.

External data used in the study are outlined in Table 1. Our analysis covers a 63-period starting from 1949 up to 2011. We take advantage of long-run national-level data collected by USDA-ERS (2012) on U.S. agricultural total factor productivity growth. For annual R&D public expenditures starting from 1900 to 2011, we calibrate and extend the time series constructed by Huffman and Evenson (2008) and USDA-ERS (2012). We also use the corn moisture index produced by NOAA (2015) to incorporate drought information in our analysis. In addition to the Bayesian estimates of the linear and log models, we also implement both models using the ordinary least squares given the hypothesized values of R&D stocks. The results from the OLS will greatly help us in checking the robustness of our Bayesian estimates.

## III. Results and Discussion

**Posterior distributions of model parameters and TFP-RD stock elasticities:** Figure 1 illustrates the posterior distributions of the parameters in the TFP growth equation (Eq.1) under the linear (left panel) and log (right panel) models, respectively. Mean values of the posterior distributions are highlighted by yellow dashed lines while the grey dashed lines represent the least squares estimates. Both log and linear models are in agreement regarding the positive impact of U.S. public R&D stocks and time trend on U.S. agricultural total factor productivity. For corn moisture stress index, the posterior mean is centered at zero under the log model while it is clearly negative under the linear model. When we contrast between Bayesian and OLS mean estimates, we observe that the least squares estimates are typically exaggerated. For example, the mean estimates of U.S. public R&D stocks on agricultural TFP index under the OLS case are well above the mean of the posterior distribution for both log and linear models. Here, we even find that the OLS estimates are outside the 80% and the 95% Bayesian credible intervals (dashed and solid red lines, respectively). This suggests that – given the underlying posterior distribution – the postulated mean impact of U.S. public R&D stocks on agricultural TFP under the least squares method is clearly exaggerated and is unlikely. Likewise, the OLS estimates on the impact of time trend on agricultural productivity are well below the posterior means and are beyond the Bayesian credible intervals. However, we do see some agreement between the Bayesian and OLS mean estimates regarding the impact of drought on U.S. agricultural productivity.

It is important to note that interpretation of Bayesian credible intervals are strictly different from the confidence intervals in classical statistics since we are directly making inferences with respect to the posterior distributions of model parameters. For example, with the 95% credible intervals, we know that there is a 95% probability that the true value of a parameter is between the credible intervals based on our probability model and all the data. In contrast, inferences in classical statistics are grounded on random samples. Therefore, a 95% confidence interval requires iterative construction of such ranges and from these we know that 95% of these intervals will contain the true population parameter. Since we cannot make direct probability statements with the classical methods, the probability statements which we can make under with the Bayesian inference is superior and more informative.

We now evaluate the uncertainties in the elasticities of agricultural TFP with respect to U.S. public R&D knowledge capital stocks (Figure 2). Given the posterior distribution and data, we know there is a 95% probability that the true value of the elasticity of U.S. agricultural TFP given U.S. public R&D stocks is between (0.21, 0.50) and (0.08, 0.37) under the log and linear models, respectively. Of course, by taking the 80% credible intervals we can further narrow down these ranges to (0.25, 0.45) and (0.13, 0.32) for the log and linear models. In general, the elasticities estimated using Bayesian inference are much more conservative than those under the ordinary least square which are unlikely based on the credible intervals in Figure 2. The responsiveness of U.S. agricultural TFP to public R&D stocks are also higher under the log specification with posterior mean at around 0.35 compared to 0.22 in the linear model. This suggests that the expected gains from U.S. public R&D investments are likely to be smaller under the linear model, a point which we will revisit later. Of course, the superiority of Bayesian inference lies on the measurement of uncertainty which is currently being ignored in the literature. Indeed, given our posterior distribution and data we can confirm that published estimates are within the likely range of the true elasticity of U.S. agricultural TFP with respect to public R&D stock.

**Evaluating uncertainties in U.S. public R&D stocks and agricultural TFP:** Although there is consensus about the uncertainties in R&D knowledge capital stocks, little has been done to address this issue. Here, we introduce uncertainty in the estimates of U.S. R&D knowledge capital stocks by calculating the posteriors distributions of the gamma and U.S. public R&D lag parameters. Figure 3 illustrates the distribution of the underlying parameters which shape the gamma distribution. We clearly see that the likely range of parameter values governing the gamma distribution have been narrowed down in the linear model while the spread remains unchanged under the log model. Although prior values of these parameters are still close to the 95% credible intervals, we see that our posterior means are quite different from these priors. Specifically, posterior means for the delta parameter are greater than the priors while for the lambda parameter the posterior means are much smaller. To better illustrate the implications of uncertainties in these parameters, we plot the distribution of the U.S. public R&D lags in Figure 4. The black dots represent the lag parameters based on the hypothesized gamma distribution using priors from Alston et al (2010a) while the red dots correspond to our posterior means. The yellow and grey bands represent the 95% and 80% credible intervals, respectively. Across models, we observe that the posterior means of U.S. public R&D lag parameters are flatter than the hypothesized values. The figure also illustrate significant uncertainties in the R&D lag parameters during the early and late years which highlights in the ambiguous near-term and long-term impacts of public R&D spending on knowledge stock accumulation. The uncertainty varies across model specifications with narrower credible intervals under the linear than in the log model. In fact, we observe several points which are outside the 80% credible intervals under the linear model. The expected maximum impact of U.S. public R&D spending on stocks –namely the lag peak value – is also varies across model specification with the peak years at around year 23 and 18 for the log and linear case, respectively.

Using the distribution of the R&D lag parameters we can now introduce uncertainty in estimates of U.S. public R&D knowledge capital stocks (Figure 5). With the log model (top panel, Figure 5), the mean of the reconstructed R&D stocks are generally close to the hypothesized values and are within the 95% and 80% credible intervals which suggest that hypothesized values belong to the likely range of true R&D stocks. However, we observe a different picture with the linear model (bottom panel, Figure 5) with the hypothesized U.S. public R&D stocks clearly above the posterior mean values. Some of the points are even outside the 80% credible intervals which suggest that the hypothesized R&D stocks are far from the likely range of the true value of R&D stocks at least for the linear model.

Finally, we reconstruct the historical U.S. agricultural TFP index and visually examine how well our hierarchical linear and log model captures the uncertainties in the actual data (Figure 6 and 7, respectively). We also present our historical reconstruction using the OLS estimates. The OLS confidence intervals show poor coverage probability since all of the historical movements in U.S. agricultural TFP are outside the confidence intervals while we can clearly see that the Bayesian credible intervals have better coverage of probabilities. Comparing across model specifications, the Bayesian reconstruction under the linear model visually has a better hold of the uncertainty in U.S. agricultural TFP because under the logistic model none of the historical TFP data is outside the credible intervals. Of course, proper evaluation of model fitness requires a more formal analysis and we are planning to incorporate this in our future work.

**A counterfactual analysis of historical U.S. public R&D investments:** Having computed the key posterior distributions of our hierarchical model, we now demonstrate the advantages of Bayesian inference in evaluating the returns from increased investments in public R&D. We design a counterfactual scenario given a 30% increase in annual U.S. public spending on R&D starting from 1950 up to 1959. Following Alston et al (2011), we use 3% as our discount rate to convert the streams of costs and benefits into their present value at year 2011. The present value of total investment cost is around 23 B USD $2005. To compute for the present value of total benefits , we take the percentage increase in U.S. agricultural TFP under the counterfactual scenario relative to the historic baseline and then scale up historical U.S. agricultural output using data from USDA-ERS (2012) and converted these to present value. Figure 8 summarizes the increase in U.S. agricultural output given the counterfactual increase in U.S. public R&D spending. In the figure, we see that the gains from the increase in investments follow the expected trend with most of the expansion in output occurring between year 10 and year 30. We also see that the expansion in output is larger under the logistic model than in the linear model. Of course this is expected since the elasticity of U.S. agricultural TFP with respect to public R&D stock is much larger in the logistic than in the linear model.

The sensitivity of the investment returns to the model specification is more evident in the net present value and benefit-cost ratios (Figure 9). With the logistic specification, the posterior mean net present value of the counterfactual investment is around 432 B USD $2005. We know that there is a 95% probability that the true net present value of the project is around 235 to 680 B USD $2005 or around 295 to 583 B USD $2005 if we use the 80% credible intervals. The mean value of the benefit-cost ratio is around 20 and it is likely that the true benefit cost ratio is around 11 to 31 (with a probability of 95%). With the linear model, we find lower returns given the same increase in public R&D spending. Indeed, the expected net present value is just around 224 B USD $2005 and we know that the true net present value is between 67 to 404 B USD $2005 or between 117 to 336 B USD $2005 if we use the 95% and 80% credible intervals, respectively. The mean value of the benefit-cost ratio is around 11 and it is likely that the true benefit cost ratio is around 4 to 19 (with a probability of 95%). Our results highlight the importance of identifying the correct model specification since the returns to public R&D investments are more conservative in the linear case. However, we can clearly see disagreement between the Bayesian and OLS estimates since the net present values and benefit cost ratios calculated from the OLS method are clearly overstated and unlikely when compared to our Bayesian mean estimates and credible intervals. This highlights the potential issue of overstating the returns to R&D spending under the OLS case.

## References


Alston, J. M., Andersen, M. A., James, J. S., & Pardey, P. G. (2011). The Economic Returns to U.S. Public Agricultural Research. *American Journal of Agricultural Economics*, *93*(5), 1257–1277. http://doi.org/10.1093/ajae/aar044

Alston, J. M., James, J. S., Andersen, M. A., & Pardey, P. G. (2010a). 8. Research Lags and Spillovers. In *Persistence Pays* (pp. 239–269). Springer New York. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4419-0658-8_8

Alston, J. M., James, J. S., Andersen, M. A., & Pardey, P. G. (2010b). 9. Models of Research and Productivity. In *Persistence Pays* (pp. 271–311). Springer New York. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4419-0658-8_8

Barboza, L., Li, B., Tingley, M. P., & Viens, F. G. (2014). Reconstructing past temperatures from natural proxies and estimated climate forcings using short- and long-memory models. *Ann. Appl. Stat.*, (4), 1966–2001. http://doi.org/10.1214/14-AOAS785

Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Anal.*, (3), 385–402. http://doi.org/10.1214/06-BA115

Bolstad, W. M. (2007). *Introduction to Bayesian statistics* (Second). USA, Canada: John Wiley & Sons.

Burney, J. A., Davis, S. J., & Lobell, D. B. (2010). Greenhouse gas mitigation by agricultural intensification. *Proceedings of the National Academy of Sciences*, *107*(26), 12052–12057. http://doi.org/10.1073/pnas.0914216107

Evenson, R. E. (2003). Production impacts of crop genetic improvement programmes. In R. E. Evenson & D. Gollin (Eds.), *Crop Variety Improvement and its Effect on Productivity: The Impact of International Agricultural Research* (pp. 447–472). Oxon, UK; Cambrigde, USA: CAB International.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2003). *Bayesian Data Analysis, Second Edition*. Boca Raton London New York Washington, D. C.: Chapman and Hall/CRC.

Heisey, P., Wang, S. L., & Fuglie, K. (2011). *Public Agricultural Research Spending and Future U.S. Agricultural Productivity Growth: Scenarios for 2010-2050* (Economic Brief No. EB-17). Washington D.C., USA: Economic Research Service, U.S. Department of Agriculture.

Huffman, W. E. (2009). *Measuring Public Agricultural Research Capital and Its Contribution to State Agricultural Productivity* (No. 09022). Ames, IA, USA: Department of Economics, Iowa State University.

Huffman, W. E., & Evenson, R. E. (2008). *Science for agriculture: A long-term perspective*. John Wiley & Sons.

NOAA. (2015, April). Crop Moisture Stress Index. Retrieved from http://www.ncdc.noaa.gov/societal-impacts/cmsi/

Pardey, P. G., & Alston, J. M. (2012). Global and US trends in agricultural R&D in a global food security setting. In *Improving Agricultural Knowledge and Innovation Systems: OECD Conference Proceedings*. Paris, France: OECD Publishing. Retrieved from http://dx.doi.org/10.1787/9789264167445-4-en

Pardey, P. G., & Beddow, J. M. (2013). *Agricultural Innovation: The United States in a Changing Global Reality*. Chicago, IL, USA: The Chicago Council on Global Affairs.

Pardey, P. G., & Beintema, N. M. (2001). *Slow Magic: Agricultural R&D A Century After Mendel* (No. 36). International Food Policy Research Institute (IFPRI).

Pardey, P. G., Beintema, N. M., Dehmer, S., & Wood, S. (2006). *Agricultural research: a growing global divide?*. International Food Policy Research Institute (IFPRI). Retrieved from http://ideas.repec.org/p/fpr/fprepo/17.html

Robert, C. P. (2007). *The Bayesian Choice : From Decision-Theoretic Foundations to Computational Implementation*. New York, NY: Springer.

USDA ERS. (2012). Agricultural Research Funding in the Public and Private Sectors. Retrieved from http://www.ers.usda.gov/data-products/agricultural-productivity-in-the-us.aspx

Table 1. Description of Data

| Variables | Description | Mean Value | Sources and coverage |
|---|---|---|---|
| U.S. Agricultural Total Factor Productivity Index: 1949-2011 | Difference in annual growth rates of total agricultural output and inputs. Growth rates are then used to create TFP indices (1949 = 100). Rescaled index to 1.0 | 1.68 | USDA-ERS (2015) |
| U.S. Public Agricultural R&D Spending: 1900-2011 | In Millions 2005 PPP$. Rescaled to Billion 2005 PPP$ | 2.19 | Huffman & Evenson (2008) and USDA-ERS (2012) |
| U.S. Corn Moisture Stress Index 1949-2011 | Combined information on the intensity of drought as measured by the Palmer drought index and agricultural areas planted with corn crops. Rescaled index to 1.0 | 0.18 | NOAA (2015) |

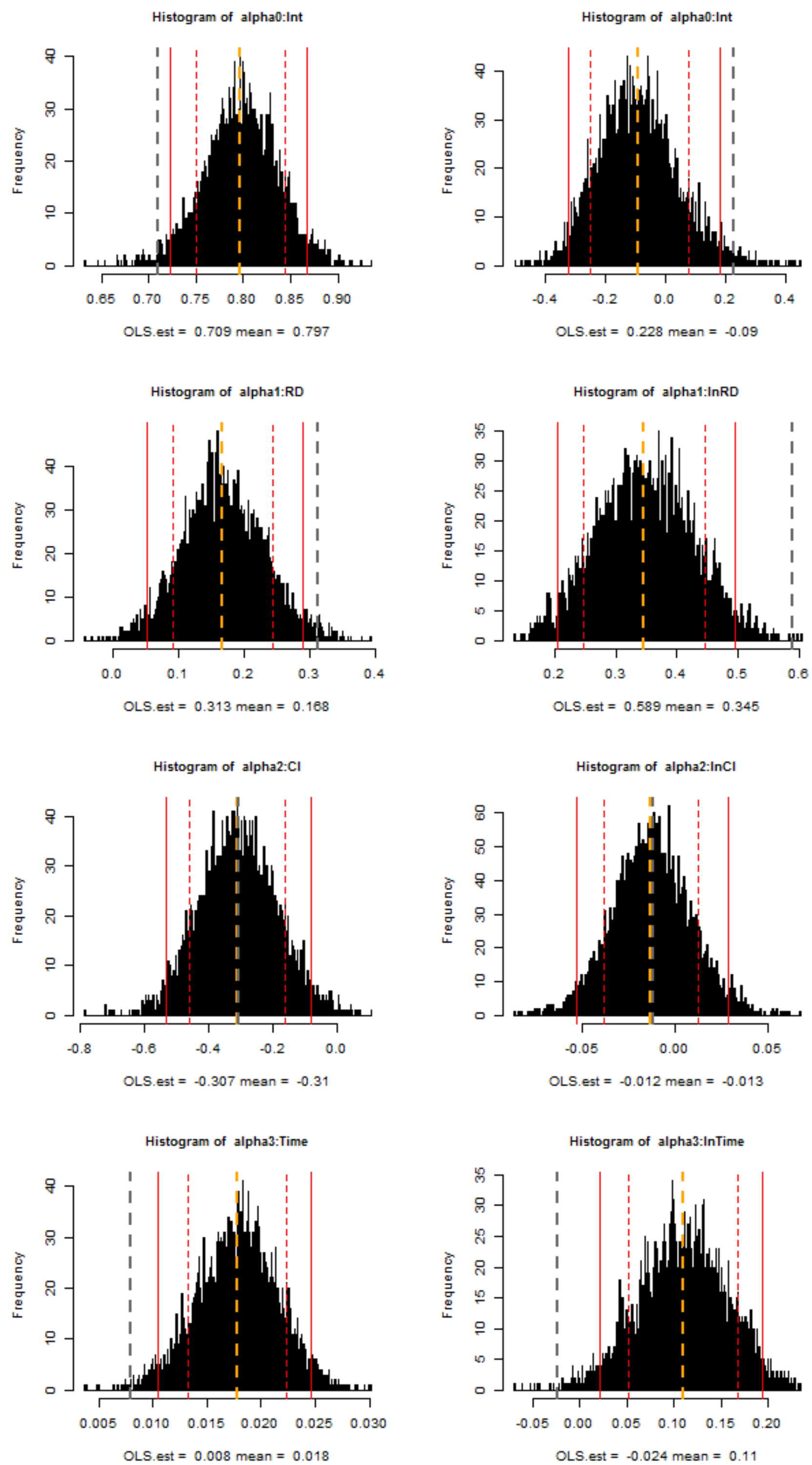



Figure 1. Posterior distributions of parameters explaining U.S. agricultural TFP indices: Linear (left) vs. Log (right) models. Mean values are highlighted by yellow dashed lines while the grey dashed lines represent the least squares estimates. The dashed and solid red lines show the 80% and the 95% Bayesian credible intervals, respectively.

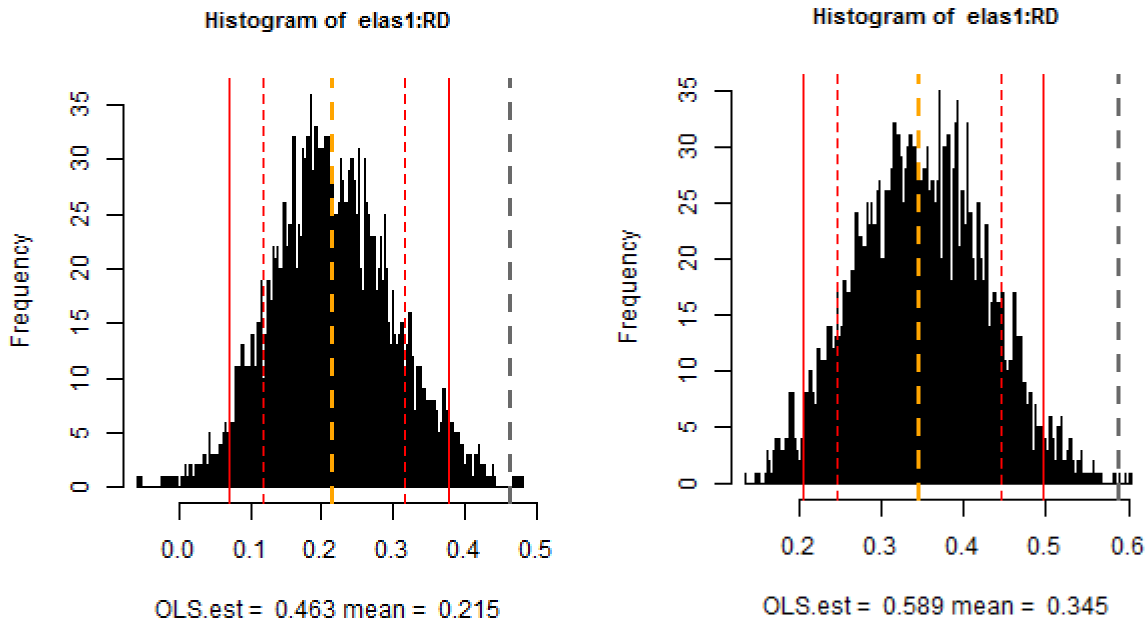



Figure 2. Posterior distributions of the elasticity of U.S. agricultural TFP index with respect to U.S. public R&D stocks: Linear (left) vs. Log (right) models. Mean values are highlighted by yellow dashed lines while the grey dashed lines represent the least squares estimates. The dashed and solid red lines show the 80% and the 95% Bayesian credible intervals, respectively.

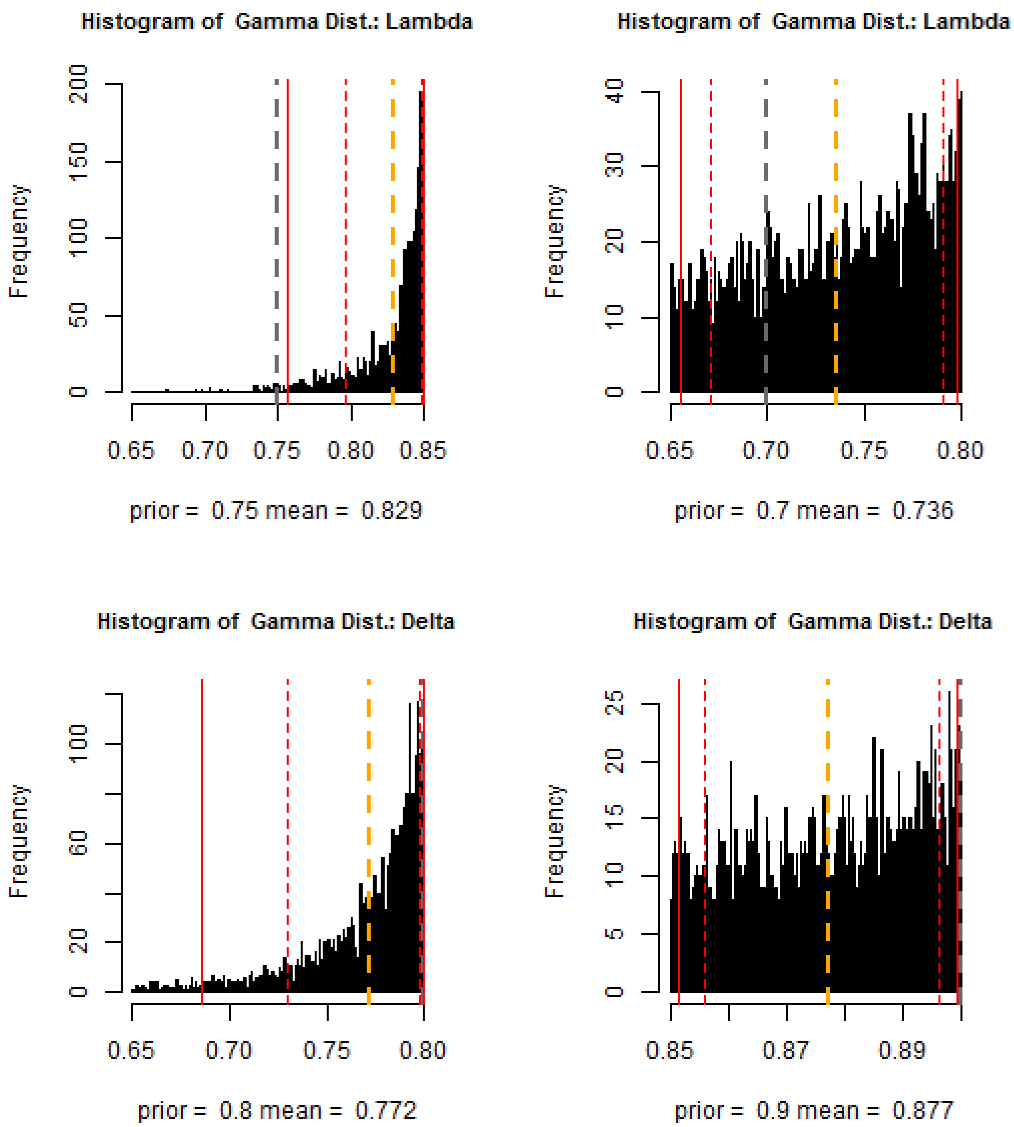



Figure 3. Posterior distributions of lambda and delta parameters in the gamma lag distribution: Linear (left) vs. Log (right) models. Mean values are highlighted by yellow dashed lines while the grey dashed lines represent prior values. The dashed and solid red lines show the 80% and the 95% Bayesian credible intervals, respectively.

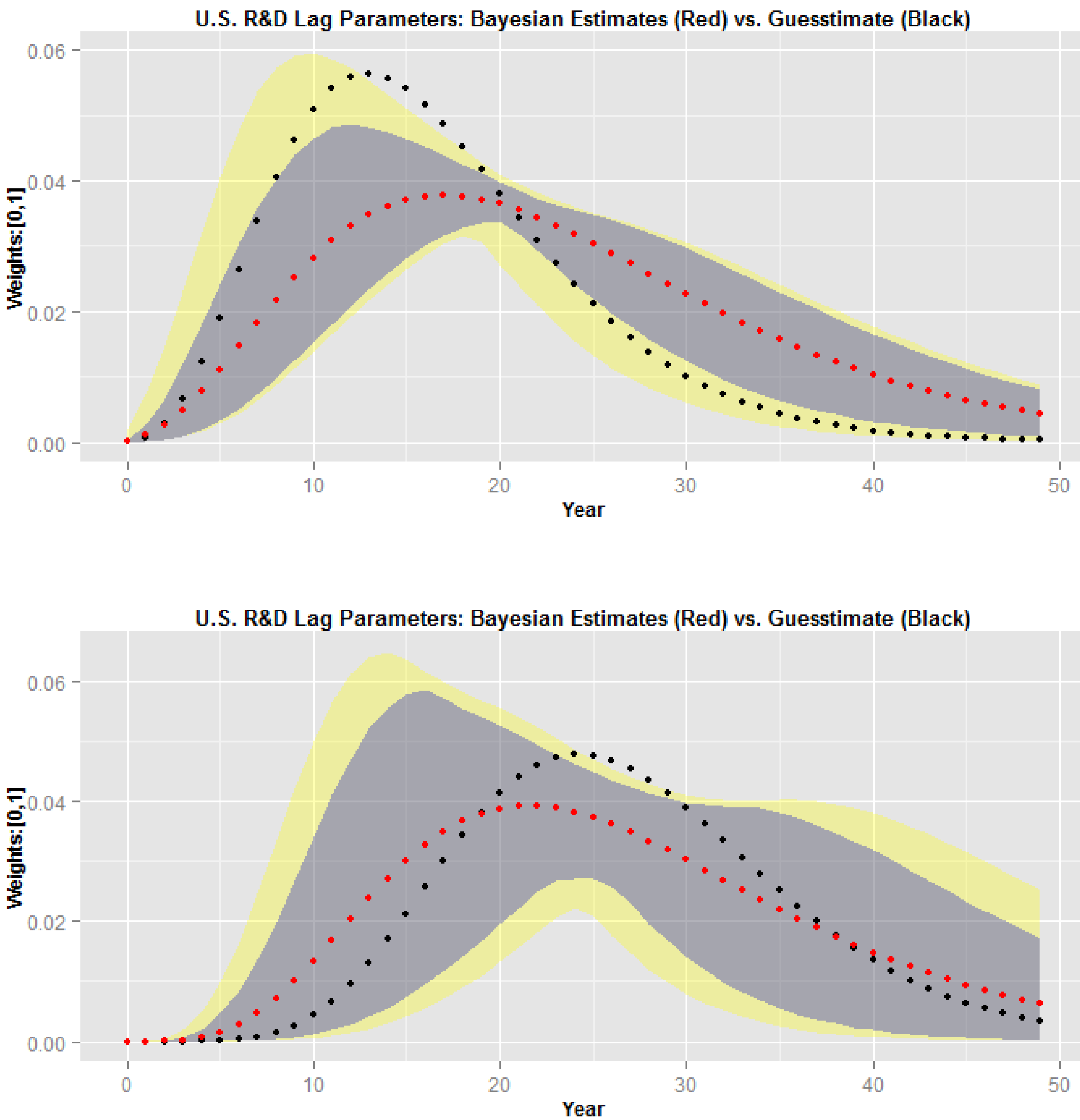
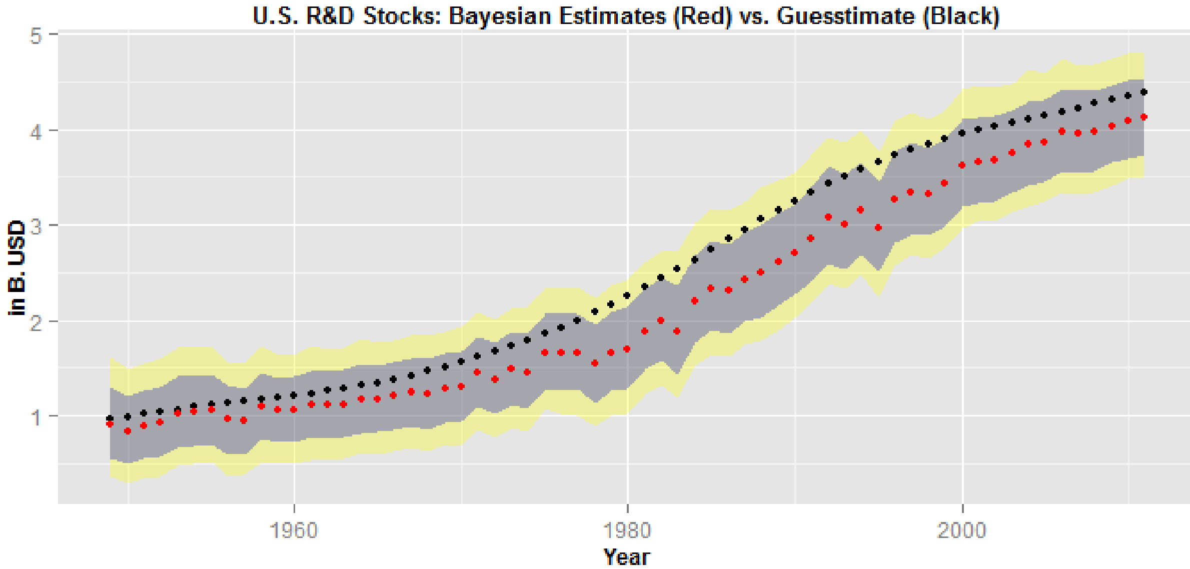



Figure 4. Posterior distributions of U.S. public R&D lag parameters: Linear (top) vs. Log (bottom) models. Mean values are highlighted by the red dots while the black dots represent the prior values based on Alston et al (2010). The grey and yellow areas show the 80% and the 95% Bayesian credible intervals, respectively.



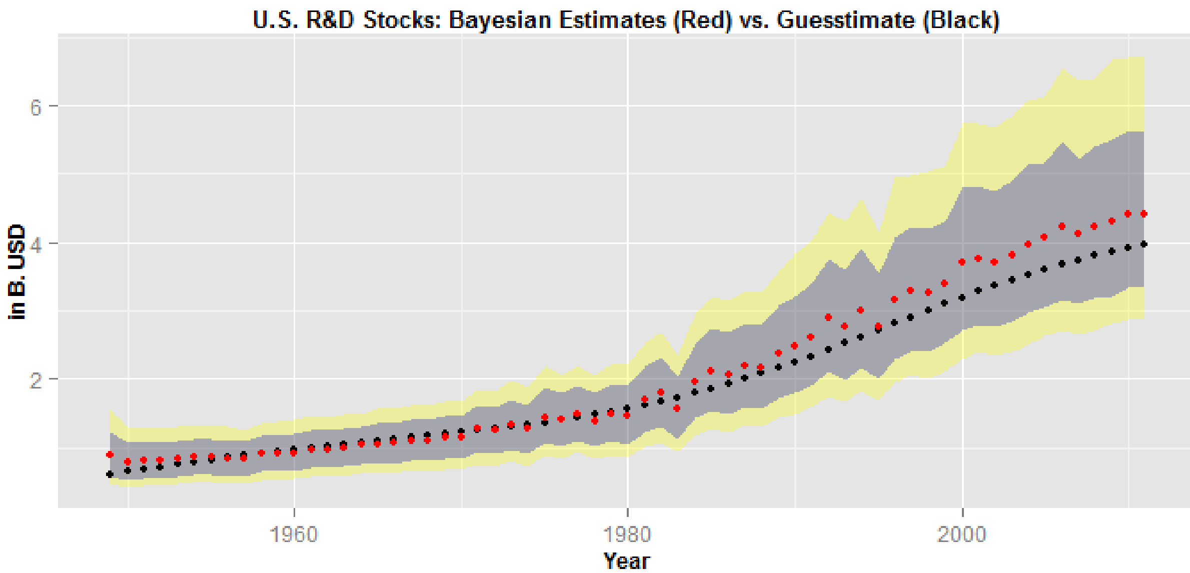
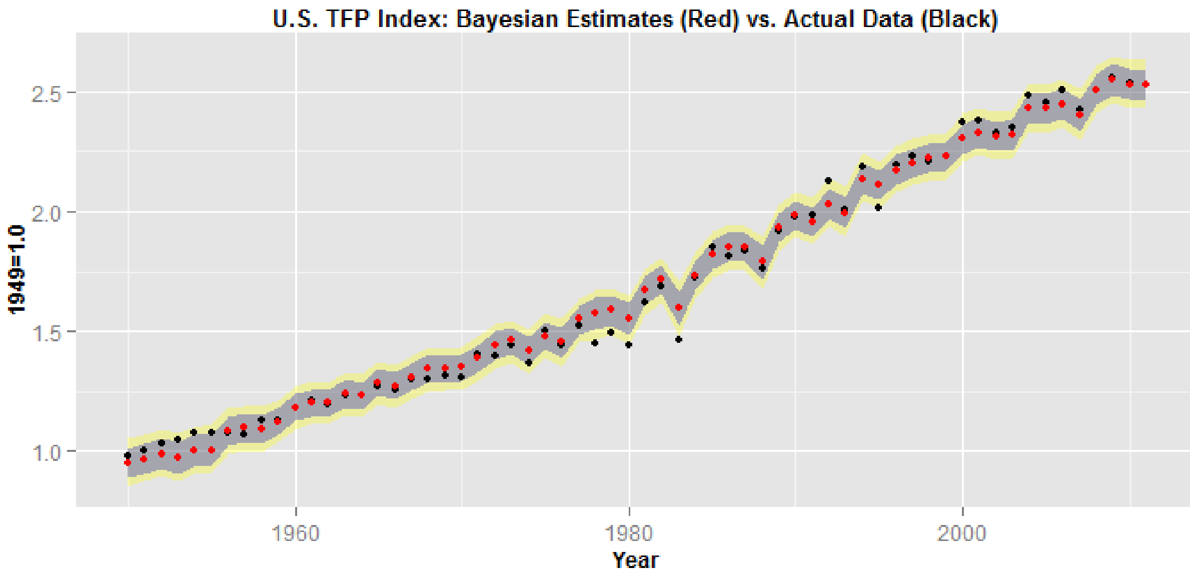




Figure 5. Posterior distributions of U.S. public R&D stocks: Linear (top) vs. Log (bottom) models. Mean values are highlighted by the red dots while the black dots represent the prior values. The grey and yellow areas show the 80% and the 95% Bayesian credible intervals, respectively.



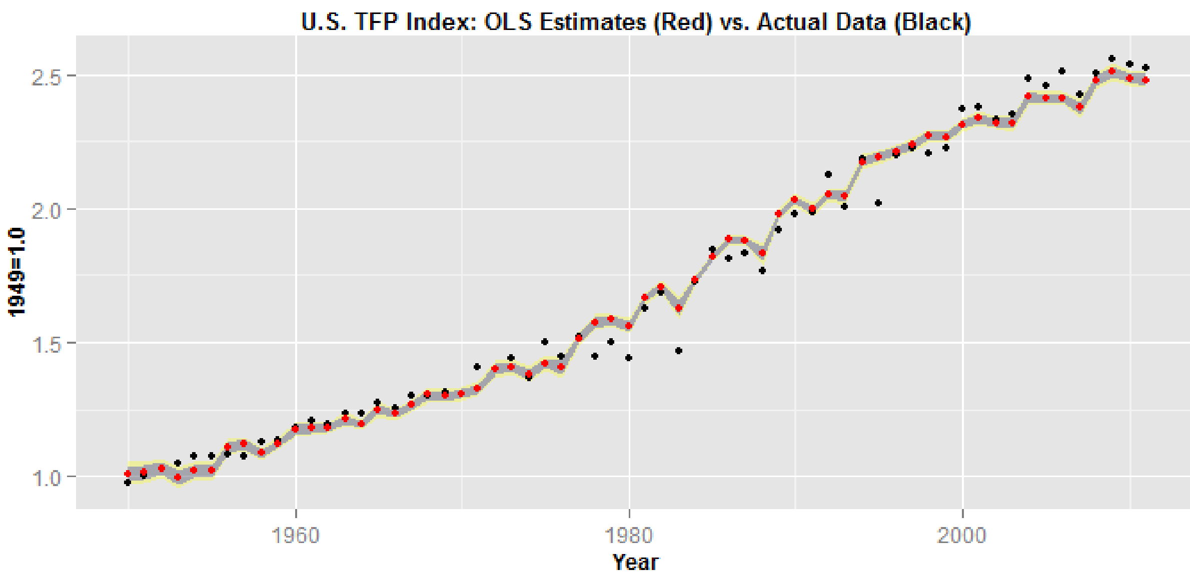




Figure 6. Posterior distributions of the reconstructed U.S. agricultural TFP indices under the linear model: Bayesian (top) vs. OLS (bottom) estimates. Mean values are highlighted by the red dots while the black dots represent the prior values. The grey and yellow areas show the 80% and the 95% Bayesian credible intervals, respectively.

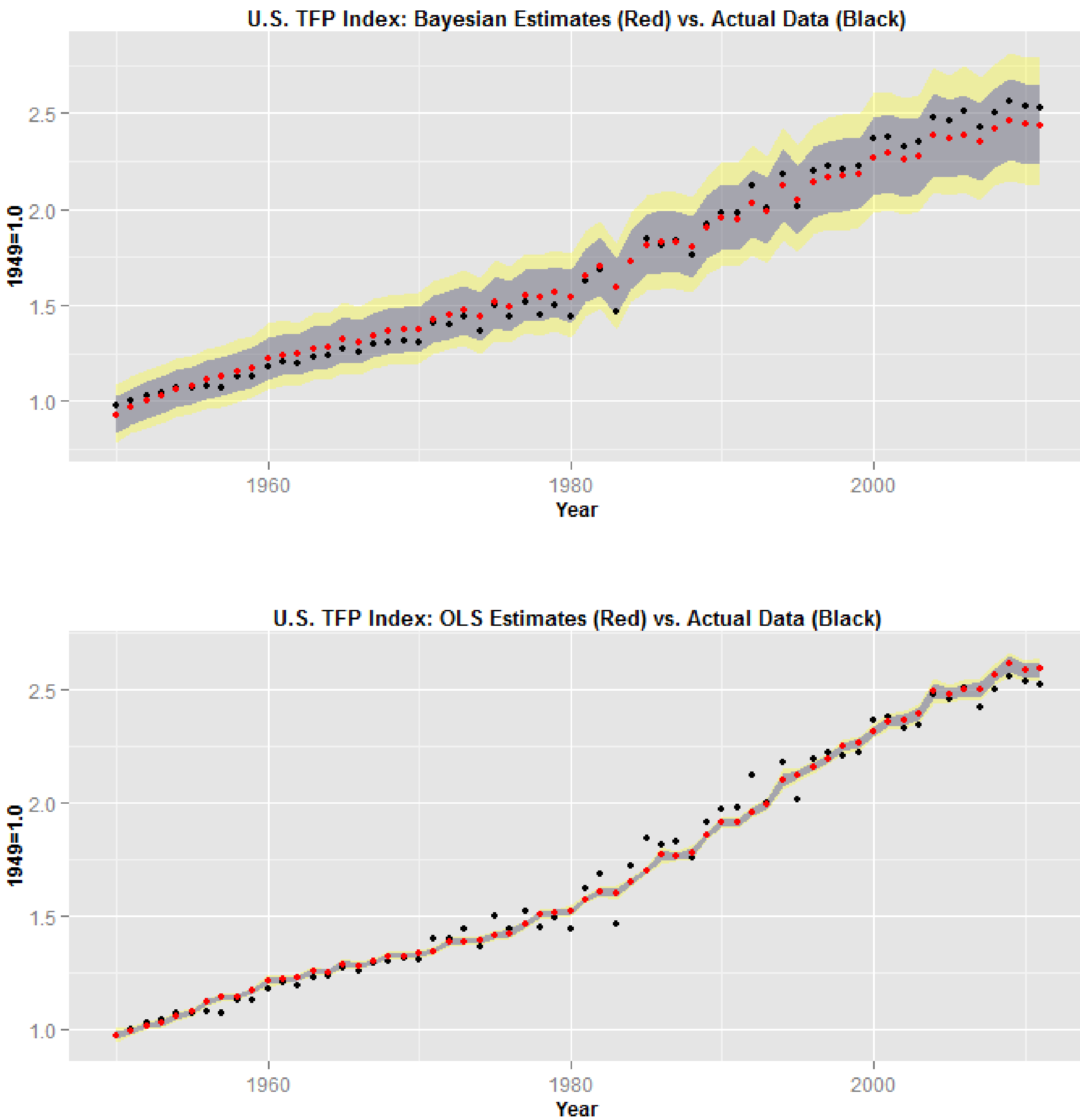



Figure 7. Posterior distributions of the reconstructed U.S. agricultural TFP indices under the log model: Bayesian (top) vs. OLS (bottom) estimates. Mean values are highlighted by the red dots while the black dots represent the prior values. The grey and yellow areas show the 80% and the 95% Bayesian credible intervals, respectively.

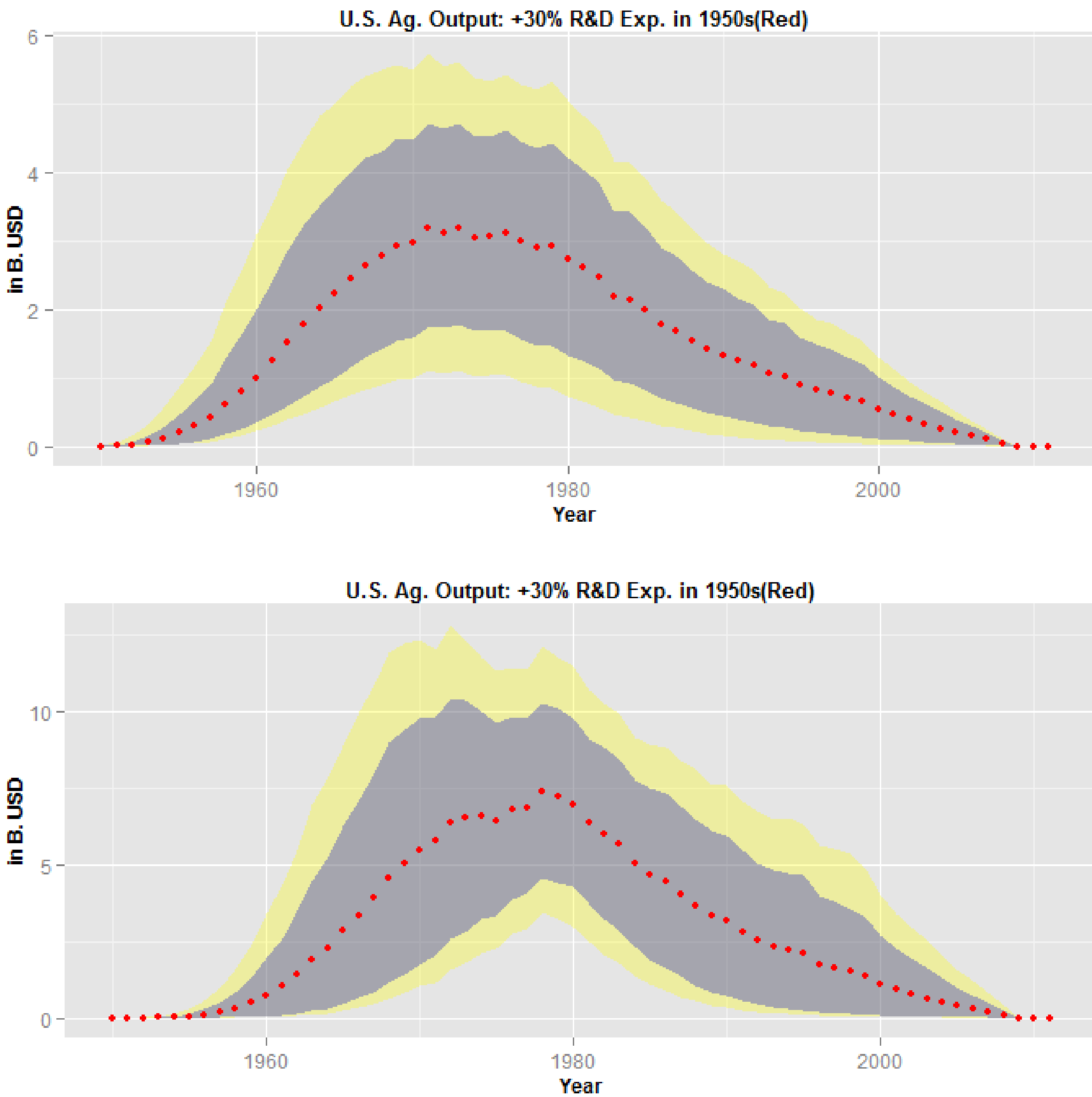



Figure 8. Posterior distributions of the increase in U.S. agricultural output given a 30% increase in U.S. public R&D spending starting from 1950 up to 1959: Linear (top) vs. Log (bottom) models. Mean values are highlighted by the red dots. The grey and yellow areas show the 80% and the 95% Bayesian credible intervals, respectively.

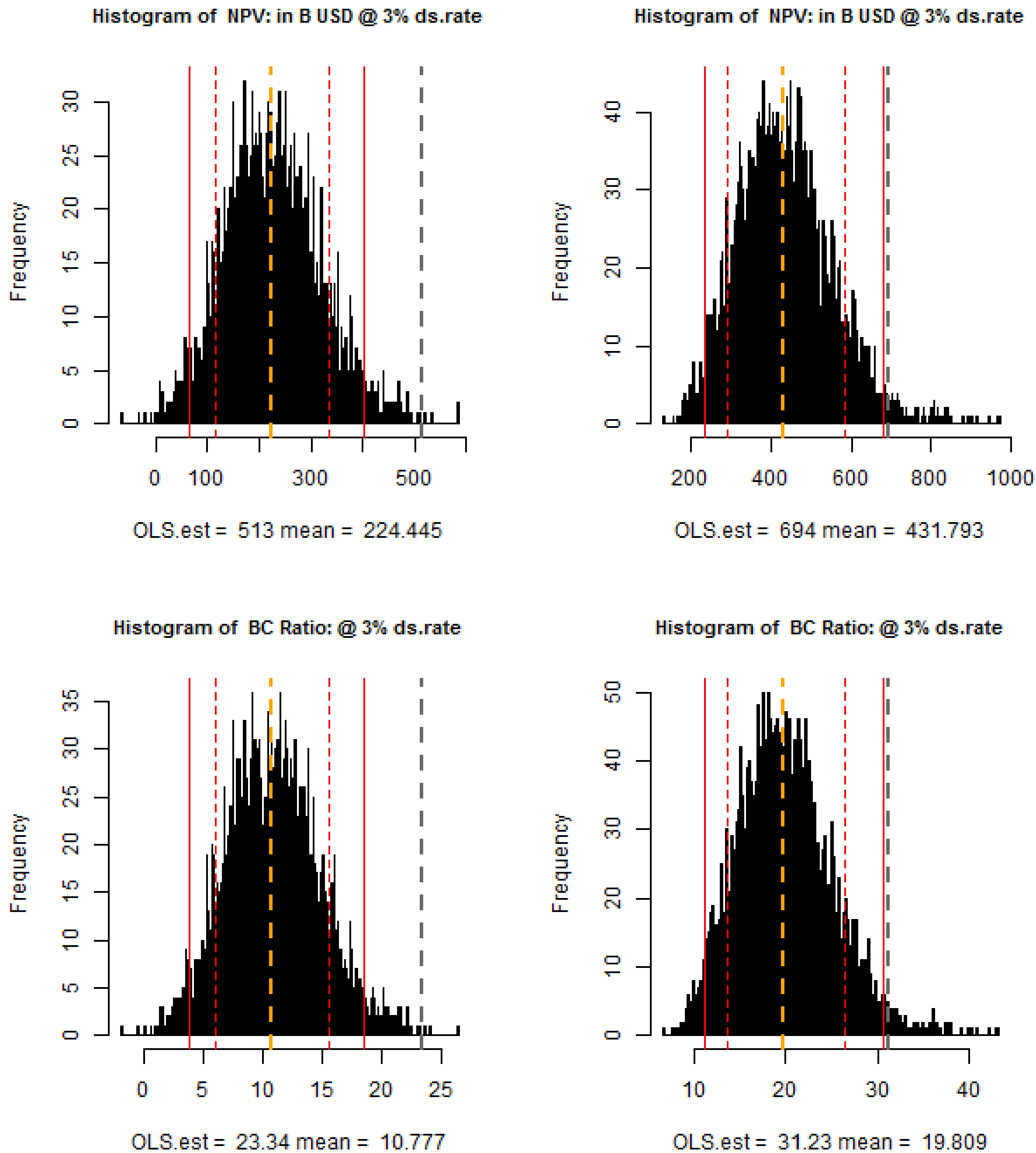



Figure 9. Posterior distributions of net present value and benefit-cost ratio given a 30% increase in U.S. public R&D spending starting from 1950 up to 1959: Linear (left) vs. Log (right) models. Mean values are highlighted by yellow dashed lines while the grey dashed lines represent OLS values. The dashed and solid red lines show the 80% and the 95% Bayesian credible intervals, respectively.