Benefits of Reducing Domestic Well Nitrate Contamination from Concentrated Animal Feeding Operations: A National Model of Groundwater Contamination

Jeffrey K. Lazo,¹ Donald M. Waldman,² Thomas D. Ottem,¹
and William J. Wheeler³
¹Stratus Consulting Inc., Boulder, CO
²University of Colorado, Boulder, CO
³U.S. EPA, Washington, DC


Copyright 2003 by Jeffrey K. Lazo, Donald M. Waldman, Thomas D. Ottem, and William J. Wheeler. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
Abstract
This paper presents an analysis of benefits to private drinking water well users from regulatory changes for concentrated animal feeding operations (CAFOs). Combining a statistical model of groundwater quality with benefit estimates based on values available from the literature, we develop aggregate national benefit estimates for reduced well water contamination from changes in CAFO regulations. The statistical model is developed to explore truncation and selection issues. We conduct a sensitivity analysis of aggregate benefit estimates to model estimation and benefits transfer values.

Keywords
benefits, concentrated animal feeding operations, nitrates, sensitivity analysis, selection truncation model

Acknowledgements
Thanks are given to Robert S. Raucher, Marca Hagenstad, Megan Harrod, and Amy Gage of Stratus Consulting; Maurice Hall of CH2M Hill of Redding, California; and Linda Chappell of U.S. EPA for their support, input, and guidance on this work. The authors are solely responsible for any remaining errors.

Disclaimer
The views expressed in this paper are those of the authors and do not necessarily represent those of the U.S. Environmental Protection Agency. In addition, although the research described in this paper may have been funded entirely or in part by the U.S. Environmental Protection Agency, it has not been subjected to the Agency’s required peer and policy review. No official Agency endorsement should be inferred.
**JEL/EconLit Categories**

Q000 — Agricultural and Natural Resource Economics: General

Q250 — Renewable Resources and Conservation; Environmental Management: Water; Air; Climate; Noise

C24 - Econometric Methods: - Single Equation Models: Truncated and Censored Models
**Introduction**

Animal feeding operations (AFOs) are agricultural businesses where animals are kept and raised in confined situations thus focusing animal wastes (i.e., manure, urine, and dead animals) within a small area of land. Concentrated animal feeding operations (CAFOs) represent a small portion of AFOs that are regulated by the U.S. Environmental Protection Agency (EPA) as point sources under the Clean Water Act (CWA).\(^1\) On December 15, 2002, the EPA revised and updated the National Pollutant Discharge Elimination System (NPDES) provisions that define and establish permit requirements for CAFOs, and the effluent limitations guidelines (ELGs) for feedlots, which establish the technology-based effluent discharge standard that applies to regulated CAFOs.\(^2\)

In developing the revised CAFO rule, EPA undertook an economic assessment of a number of scenarios and options for potential revisions to the NPDES and ELGs for CAFOs. One concern of CAFO effluents is the effect of nitrogen leaching into groundwater as nitrates and causing human health risks from nitrate contamination. This paper discusses the portion of EPA’s benefit-cost analysis of the regulatory revisions that developed benefit estimates for reductions in nitrate concentrations in private drinking water wells from groundwater contamination.

In this paper we focus on two aspects of the groundwater benefits analysis: (1) statistical models of the nitrogen-nitrate relationship — models developed specifically for this analysis dealing with data issues of selection and truncation (explained below); and (2) national aggregation and sensitivity analysis of the benefits estimates.

For this benefit analysis we estimated the environmental impact on groundwater resources from pollutants associated with the operation of CAFOs, primarily nitrogen. The physical relationship
between nitrogen loadings from CAFOs and private well nitrate concentrations was statistically modeled, controlling for other sources of nitrogen and well characteristics. The statistical model was developed so that we could generate benefit estimates for a range of potential regulatory options. We determined that our model had to be a national model and predict actual nitrate levels because the valuation literature is linked to concentrations of 10 mg/L. Existing models of nitrate contamination were not sufficient to link changes in the CAFO regulatory structure with nitrate contamination in wells for policy analysis at the national level. Studies that estimate nitrate concentrations in groundwater wells as a function of local characteristics are generally detailed scientific studies that focus on a very specific geographic region. These studies were generally regional or local in scope, and obtained their data by sampling the wells directly — an approach not feasible for the work reported here. The two studies identified (Sparco; Lichtenberg and Shapiro) that did develop statistical models to predict nitrate concentrations in groundwater were not adaptable for estimating the effects of national CAFO regulations because they incorporated either spatial or temporal data that are not available for a national level assessment. Other models that are national in scope do not predict actual nitrate levels, but provide screening or other indicators of the risk of nitrate contamination (e.g., Letson, Gollehon, and Mose). Therefore, the statistical model was developed to model nitrate concentrations as a function of several factors, including nitrogen from animal waste from CAFOs.

Two issues that arose with the statistical analysis and modeling were (1) the model of well nitrate concentrations is truncated at zero (negative nitrate concentrations are not possible) and (2) the data set was limited in terms of the number of counties across the country that were represented
in the data set. Concerns that the included counties were different than the excluded counties led us to develop a selection model to account for potential selection bias.

Using estimates of the change in nitrogen loadings from CAFOs under different regulatory options, the statistical model was used to calculate changes in well nitrate concentrations in households with private wells. Using benefit-transfer methods, the annualized national benefits of reducing CAFO impacts on groundwater were estimated for each of the regulatory options. Because these benefit estimates are based on a statistical model of the relationship between surface nitrogen deposition and groundwater nitrate concentrations, there is variance in the model parameter estimates and thus uncertainty in the resulting benefit estimates. There is also uncertainty in the value estimates used in the benefits transfer. We thus discuss additional analysis undertaken to explore how sensitive the benefit estimates are to model variance and to variance in values from the benefits transfer.

**Background**

*Well Nitrates and Health Effects*

Nitrogen from CAFOs can leach through soils and ultimately reach groundwater as nitrates. The federal health-based National Primary Drinking Water Standard for nitrate is 10 mg/L. This Maximum Contaminant Level (MCL) applies to all community water supply systems, but not to households that rely on private wells. As a result, households served by private wells are at risk of exposure to nitrate concentrations above 10 mg/L, which EPA considers unsafe for sensitive subpopulations (e.g., infants). Nitrate above concentrations of 10 mg/L can cause methemoglobinemia (“blue baby syndrome”) in bottle-fed infants (National Research Council, 1997), which causes a blue-gray skin color, irritableness or lethargy, and potentially long-term
developmental or neurological effects. Generally, once nitrate intake levels are reduced, symptoms abate. If the condition is not treated, however, methemoglobinemia can be fatal. No other health impacts are consistently attributed to elevated nitrate concentrations in drinking water; however, other health effects are suspected (Weyer). Based on USGS data, between 9 and 10% of the 13.5 million U.S. households (about 1.3 million households) on wells located in counties with CAFOs exceed the EPA-mandated 10 mg/L MCL for nitrates.

Relation of Nitrogen Deposition to Groundwater (Well) Nitrate Concentrations

Nitrate is found in groundwater because of surface applications of two forms of the nutrient nitrogen: nitrate and amine groups (of which nitrogen is a component). Generally nitrogen from fertilizer is already in the nitrate form, which leaches more readily into the soil. Nitrogen from manure and septic systems generally occurs as large organic molecules called amine groups. Once in the soil, these large molecules convert to nitrate and ammonia as microbes break down the organic matter. The ammonia then volatizes as a gas into the atmosphere, and the nitrate leaches through the soil and potentially into groundwater. This process takes a few hours to a few weeks, depending on the soil conditions (M. Hall, CH2M Hill, pers. comm., Sept. 15, 2000). It is estimated that AFOs produced 291 billion pounds of manure in 1997 (U.S. EPA, 2001, p. 1-3). Before regulatory changes, it is estimated that CAFOs generated over 555 million tons of nitrogen, which was disposed of largely through land application of manure. Under the potential regulatory revisions reviewed by EPA, this could be reduced by as much as 19.5%, to 447 million pounds of nitrogen.

Factors Affecting Well Nitrate Concentrations

We reviewed studies that observed statistical relationships between groundwater nitrate concentrations and various other hydrogeological and land use factors to determine what data
and variables were needed to model the nitrogen-nitrate relationship. Although the groundwater monitoring and modeling studies reviewed covered different geographic areas and focused on varying nitrogen sources (septic systems, agricultural fertilizers, animal feedlots), certain variables were significant across many of the studies. Physical characteristics such as well depth, soil type, and near-well land use all influenced well nitrate concentrations. Well depth was frequently inversely related to nitrate concentrations in wells, regardless of nitrate source (Detroy, Hunt, and Holub; Ritter and Chirnside; Kross, Hallberg, and Bruner; Spalding and Exner; Swistock, Sharpe, and Robillard; Sparco; Lichtenberg and Shapiro; Ham et al.; North Carolina Division of Water Quality). A number of studies identified geological characteristics such as soil types as a significant factor affecting nitrate concentrations. Two studies found unconfined aquifers to be associated with elevated nitrate in groundwater (Lichtenberg and Shapiro; Lindsey). Other studies found higher nitrate levels associated with more permeable, well-drained soils (Ritter and Chirnside; Spalding and Exner; Sparco; Burrow; Chen; Ham et al.; Nolan et al.; Kerr-Upal, Van Seters, and Stone). Different types of land use near wells are also associated with higher groundwater nitrate. Several studies found agricultural land use in general to be associated with higher groundwater nitrate than other land uses (Rausch; Spalding and Exner; Swistock, Sharpe, and Robillard; Mueller et al.; Sparco; Carleton; Clawges and Vowinkel; Richards, Baker, and Wallrabenstein; Nolan et al.). Several studies identified relationships between various sources of nitrogen and well nitrate concentrations. Nitrogen application rates, whether from agricultural fertilizers, animal wastes, or private septic systems, were the most consistent and significant factor affecting nitrate levels in wells (Rausch; Spalding and Exner; Clawges and Vowinkel; Richards, Baker, and
Studies that investigated the effects of animal manure production on groundwater nitrate concentrations found manure to be positively correlated with groundwater nitrate. Animal waste lagoons were associated with elevated groundwater nitrate concentrations, particularly as the distance to the water table decreased (Miller, Robinson, and Gallagher; Ritter and Chirnside; North Carolina Division of Water Quality). Farms that applied manure as fertilizer tended to have higher nitrate concentrations in groundwater as well (Rausch; Swistock, Sharpe, and Robillard; Clawges and Vowinkel; Richards, Baker, and Wallrabenstein; Lindsey; Letson, Gollehon, and Mose; Kerr-Upal, Van Seters, and Stone).

The proximity of septic systems to wells was a small, but significant, contributing factor to elevated nitrate concentrations in groundwater in several studies (Carleton; Richards, Baker, and Wallrabenstein; CDC; Nolan et al.).

Analysis

Model Specification and Data

Based on the review of the factors influencing well nitrate concentrations, we developed a statistical model of the relationship between nitrogen loadings from CAFOs and other nitrogen sources and well nitrate concentrations. Analysis of the relationship between nitrogen loadings and well nitrate concentrations is based on the following linear model:

\[
\text{Nitrate (mg/L)} = \beta_0 + \beta_1 \text{Agricultural Dummy} + \beta_2 \text{Soil Group} + \beta_3 \text{Well Depth} + \beta_4 \text{Septic Ratio} + \beta_5 \text{Atmospheric Nitrogen} + \beta_6 \text{Loadings Ratio} + \beta_i \text{Regional Dummies}.
\]

The dependent variable in this model, nitrate, is the observed nitrate concentration in a groundwater well. These data comes from the USGS Retrospective Database (USGS), which
contains water quality and land use data from 10,426 wells sampled between 1969 and 1992 in 725 counties in 38 states. These data indicate that 9.45% of domestic wells in the sample have nitrate levels above the MCL of 10 mg/L.

The independent variables used to explain nitrate concentrations in well water comprise two data types relevant to the modeling: county-specific data and well-specific data. The well-specific characteristics are all from the USGS Retrospective Database: agricultural dummy is equal to 1 for wells in highly agricultural areas and 0 otherwise; soil group is a continuous variable measuring the well’s susceptibility to nitrates (lower values indicate higher susceptibility); well depth is the total depth of the well.

County-specific data come from a variety of sources. Although nitrogen loadings from CAFOs are significant contributors to nitrate concentrations in wells, they are not the only important factor. Therefore an analysis that does not incorporate these other factors, and assumes that the relationship between nitrate concentrations and nitrogen loadings is directly proportional, will overestimate the potential changes in nitrate concentrations due to decreased loadings. We included two other sources of nitrogen deposition in the analysis to control for these additional impacts. Atmospheric nitrogen is a measure of atmospheric deposition of nitrogen near the well that is included in the USGS Database. The septic ratio is the number of homes with septic systems per acre in the county where the well is located, which is calculated from U.S. 1990 Census data. The loadings ratio (pounds of leached nitrogen per acre which is county specific data) is the policy-relevant variable and is derived from a model that estimates per acre loadings of leached nitrogen from manure and fertilizer by county (National Pollutant Loading Analysis, or NPLA, TetraTech). The NPLA provided estimates of leached nitrate from animal feedlot operations.
under different regulatory options. The NPLA developed a national estimate of pollutant load reductions expected once the full benefits of the revised animal feeding operation effluent guidelines are realized. The estimate is based on loadings for the current effluent guidelines (preregulation baseline) and after the implementation of revised effluent guidelines (postregulation modeling scenarios).

Figure 1 shows the baseline nitrogen loadings in pounds per acre for the counties in the NPLA. The counties in the upper Midwest, central Midwest, and throughout the South up into the lower mid-Atlantic states have the highest concentrations of nitrogen loadings from CAFOs in the country.

Regional dummies were included as defined by EPA in developing the new CAFO standards. Every well is identified by one of the five regional dummies: Central Region, 6% of the wells in the sample; Mid-Atlantic Region 39%; Pacific Region 12%; South Region 7%; Midwest Region 36%. In the regression analysis, the Midwest Region is the excluded dummy variable and thus parameters on other regional dummies are estimates relative the Midwest Region. Table 1 presents summary statistics for the dependent and independent variables for the 2,985 observations used in the regression analysis.

Selection Issue

The NPLA dataset contains information on nitrogen loadings for 2,637 counties with AFOs amongst the roughly 3,083 counties in the United States. We combined data from 374 of these 2,637 counties with data from the USGS Retrospective Database to estimate the statistical models. An issue is whether the counties used for the modeling are different in some manner from those (2,263) not used for estimating the nitrogen-nitrate relationship. It was determined that mean values for the average loadings and various sociodemographic data from these two
groups of counties had some significant differences. In general, the average county nitrogen loadings in the model counties are higher than the excluded counties. In addition, the included counties are somewhat smaller (39% smaller) and have larger populations, higher median income, and greater numbers of housing units. The included counties also have a larger portion of their land area in farms.

Figure 2 shows the 2,637 counties in the United States that are included in the NPLA. The darker shaded counties are the 374 counties we included in the statistical modeling because they are in the USGS database and have estimated loadings in the NPLA. Counties that do not have estimated loadings are assumed by EPA not to have any CAFOs. As can be seen, the “included” counties seem to be clustered in a limited number of regions across the country and do not (visually) appear to be randomly distributed. This raised a concern that counties that are included in the statistical analysis are not representative of the range of counties containing CAFOs. We address this issue in the selection/truncation modeling described below.

*Modeling Well Nitrate Concentrations*

We explored different modeling approaches to test the importance of statistical issues such as skewness, truncation, and selection with leached nitrogen loadings and well nitrate concentration data. These models included OLS, Tobit, and gamma models. The Tobit and gamma models are modeling approaches for dealing with the non-negativity of the nitrate well concentrations. Finally we developed a model combining the Tobit model with a sample selection model (the selection/truncation model).

The selection/truncation model was developed specifically for this project to account for (1) dependent variable truncation and (2) sample selection issues. Truncation issues arise because well nitrate concentrations are nonnegative. As discussed above, selection issues arise because
counties for which well nitrate information is available represent a small portion of the total number of counties with CAFOs. We suspect that there may be selection of wells for sampling based on a desire to identify well nitrate levels exceeding safe levels.

The selection/truncation model was developed as follows. Let $y_{ij}^*$ be the theoretical nitrate concentration of the $i$-th well in county $j$, for $j = 1, \ldots, J (374)$ and $i = 1, \ldots, n_j (\sum n_j = 2985)$. Well nitrate concentration is modeled as a function of well characteristics and county characteristics (see Table 1), both observed and unobserved:

$$y_{ij}^* = \mathbf{\beta}' x_{ij} + z_j \gamma + u_j + \epsilon_{ij},$$

where the $z_j$ are the observed covariates common to all wells in county $j$, the $x_{ij}$ are the observed characteristics of the particular well, and the county and well-specific random unobserved factors are $u_j \sim N(0, \sigma_u^2)$ and $\epsilon_{ij} \sim N(0, \sigma^2)$, assumed mutually independent and independent of one another. That is,

$$E(u_j \epsilon_{ij}) = 0 \quad \forall i, j \quad \text{and} \quad E(u_j u_{j'}) = 0 \quad \forall j \neq j'. $$

Suppose nitrate concentration data for a well exists only if the county-specific component, $z_j \gamma + u_j$, is sufficiently high. Furthermore, we can measure concentration only down to 0.05 per mg/L. The selection model then is

$$y_{ij} = \begin{cases} y_{ij}^* & \text{if } y_{ij}^* > 0.05 \text{ and } u_j > -z_j \gamma \text{ (regime I)} \\ 0.05 & \text{if } y_{ij}^* \leq 0.05 \text{ and } u_j > -z_j \gamma \text{ (regime II)} \end{cases}$$

and there is information on $y_{ij}$ only if $u_j > -z_j \gamma$. This is a selection/truncation regression model (truncated by the detection limit; selected by the county rule). Let $w_{ij} = u_j + \epsilon_{ij}$. The likelihood is
\[ L(y_j) = \prod_i f(y_j, u_j | y_j > 0.05, u_j > -z_j' \gamma) \times P(y_j > 0.05, u_j > -z_j' \gamma) \]
\[ \times \prod_u P[y_u \leq 0.05, u_j > -z_j' \gamma] , \]

which can be written

\[ L(y_j) = \prod_i \int_{-z_j' \gamma}^{\infty} f(y_j, u) \, du \prod_u P[w_j < 0.05 - (\beta x_j + z_j' \gamma), u_j > -z_j' \gamma] . \]

where \( f(\cdot, \cdot) \) is the bivariate normal density, and the bivariate probability is from the normal distribution:

\[ \begin{pmatrix} w_j \\ u_j \end{pmatrix} \sim N \begin{pmatrix} 0 \\ 0 \end{pmatrix} , \begin{pmatrix} \sigma_w^2 + \sigma_e^2 & \sigma_w^2 \\ \sigma_w^2 & \sigma_e^2 \end{pmatrix} . \]

The contribution to the likelihood in the first data regime can be written

\[ \frac{1}{\sigma_w} \phi \left( \frac{y_j - \beta' x_j - z_j' \gamma}{\sigma_w} \right) \Phi \left( \frac{\gamma' + \rho \frac{\sigma_e}{\sigma_w} (y_j - \beta' x_j - z_j' \gamma)}{\sigma_w \sqrt{1 - \rho^2}} \right) , \]

where \( \sigma_w^2 = \sigma_u^2 + \sigma_e^2 \). For the second data regime \((y_j < 0.05 \text{ and } u_j > -z_j' \gamma)\), we have

\[ P[w_j < 0.05 - (\beta' x_j + z_j' \gamma), u_j > -z_j' \gamma] = \]
\[ P[w_j < 0.05 - (\beta' x_j + z_j' \gamma), u_j < z_j' \gamma] = \]
\[ P \left[ \frac{w_j}{\sqrt{\sigma_u^2 + \sigma_e^2}} < \frac{0.05 - (\beta' x_j + z_j' \gamma)}{\sqrt{\sigma_u^2 + \sigma_e^2}}, \frac{-u_j}{\sigma_u} < z_j' \gamma \right] = \]
\[ \Phi \left[ \frac{0.05 - (\beta' x_j + z_j' \gamma)}{\sqrt{\sigma_u^2 + \sigma_e^2}}, \frac{z_j' \gamma}{\sigma_u}; -\rho \right] , \]

where \( \Phi \) is the cdf of the cumulative distribution function for the standardized bivariate normal random vector with correlation

\[ \rho = \frac{\sigma_u}{\sqrt{\sigma_u^2 + \sigma_e^2}} . \]
The results from the estimation of the selection/truncation model are presented in Table 2. The selection/truncation results are very similar to those from the Tobit model. All the significant parameters have intuitive signs. The only variable in the model with a counterintuitive sign is atmospheric nitrogen deposition, which has a negative but not statistically significant sign. All of the other parameters are significant at the 1% level of significance except the septic ratio. Two of the regional dummies are significant (Pacific and South), indicating that these two areas have differences (e.g., climate) that make them significantly different from the excluded region (the Midwest).

In general we did not develop a statistical indication of which model performed the best. We prefer not to use the OLS model because it allows for negative well nitrate concentrations; we present it here only for comparative purposes. We discuss in more detail below how the benefit estimates varied depending on which model is used, but, again, this did not provide a statistical reason for preferring any one model. Given the theoretical reasons for developing the selection/truncation model (i.e., the possibility of a bias being introduced through the selection of counties included in the dataset), we focus the discussion on results from the selection/truncation model.

*Estimating Changes in Well Nitrate Concentrations Under Baseline and Scenarios/Options*

After estimating the model using the baseline loading information, expected values for well nitrate concentrations were calculated using baseline loadings from 2,985 observations and loadings from the regulatory scenarios. The regulatory scenarios are based on different manure application rates, manure management practices, and monitoring requirements.

EPA considered a number of different scenarios and options for defining and regulating CAFOs, and we discuss two of them here for illustrative purposes. The regulatory choices evaluated in this analysis are based on different combinations of two factors: limits for land application of
manure (options), and variations on how many facilities will be subject to the regulation (scenarios). The land application options are based on either total nitrogen applied (Option 1), total phosphate applied (Option 2), or total phosphate applied plus covered lagoons (Option 5). In this paper we focus on just Option 2 — total phosphates.\textsuperscript{7,8}

The scenarios determine how many small, medium, and large AFOs will be defined as CAFOs under the NPDES regulation, and thus become subject to the nitrogen-based or phosphate-based limits. The size categories are based on the number of animals at the facility and vary by animal type (e.g., a large beef farm is defined as more than 1,000 head of cattle, whereas a swine farm is large if it has more than 2,500 pigs, and a turkey farm is large if it has more than 55,000 turkeys).

We focus on two scenarios that provided the largest range of nitrogen loadings changes. Scenario 6 would regulate 100\% of large and medium AFOs, and Scenario 8 would regulate all large and some medium where new NPDES conditions would be used for identifying medium-sized CAFOs, plus a number of qualifying dry poultry and immature swine and heifer operations.\textsuperscript{9}

Table 3 shows total national nitrogen loadings at baseline and for Scenarios 6 and 8. Figure 3 shows the percentage change in per acre nitrogen loadings from baseline for Scenario 6, Option 2. As can be seen the largest percent reductions would be experienced in the western half of the United States and in the Northeast. Referring to Figure 1, which shows the baseline loadings, it is interesting to note that these are not the areas with the highest per acre loadings. A similar spatial pattern of nitrogen loadings reductions is seen in Scenario 8, but in general the reductions are smaller because this is a somewhat less restrictive regulatory option.

Loadings for the scenarios were input into the model to estimate post-regulatory well nitrate concentrations under these scenarios. In the impact analysis, the loadings ratio is the only variable that changes across scenarios.
Expected well nitrate concentrations under the different scenarios were compared with the expected well nitrate concentrations using the baseline loadings. We then used the changes projected from the model to calculate percentage differences in expected well nitrate concentrations under the different regulatory options and scenarios. These were calculated by dividing the difference from baseline for the expected values from the different scenarios by the expected values from the baseline loadings. These percentage differences were then applied to the actual nitrate concentrations, the observed well nitrate concentrations from the USGS Retrospective Database, to estimate well nitrate concentrations under the various scenarios.

Census data show that approximately 13.5 million households in the United States use domestic wells and are located in counties with animal feedlot operations. Based on the USGS Retrospective data, 9.45% of these wells currently exceed 10 mg/L. This is roughly 1.3 million domestic wells. Applying the percentage reductions calculated using the statistical model and EPA regulatory scenarios, it is estimated that between 107,000 and 149,000 households that are above the MCL at baseline are expected to be brought under 10 mg/L. As shown in Table 3, there are 121,000 and 149,000 for scenarios 8 and 6 respectively.

Many households on wells with nitrate concentrations below the MCL at baseline may also gain benefits from incremental changes in nitrate concentrations below the 10 mg/L level and above the natural level, which is assumed to be 1 mg/L. We assumed that these incremental benefits are gained only for wells beginning with concentrations between 1 and 10 mg/L. We thus did not calculate values for incremental changes where well concentrations begin and remain above the MCL because reliable value estimates do not exist for changes in incremental nitrate concentrations above the MCL.
For households that start above the MCL preregulation and move below the MCL post-regulation, we also did not calculate values for incremental changes below the MCL. The available valuation literature gives no reliable estimates for valuing incremental changes below the MCL in addition to valuing changes reductions to the MCL; thus counting both values could double count some portion of the benefits for these households. Table 3 also shows the average reduction in nitrate concentrations for wells between 1 and 10 mg/L at baseline, for each of the scenarios. Approximately 5.77 to 5.81 million households will benefit from these incremental reductions.

**Estimation of National Values**

*Benefits Transfer*

Given limited time and budget constraints, collecting primary data for a nationwide benefits study was not feasible. We thus applied a benefits transfer approach to existing studies of household values for reduced well nitrate contamination. “Benefits transfer” refers to the “application of existing valuation point estimates or valuation function estimates and data that were developed in one context to value a similar resource and/or service affected by the discharge of concern” [59 FR 1183] (see also Desvousges, Naughton, and Parsons; Loomis; Walsh, Johnson, and McKean). In other words, benefits transfer entails applying empirical results obtained from a primary research effort conducted at one site and set of circumstances to another (similar) site and set of circumstances. In this manner, existing research findings from a “study site” can be used as an expeditious means of drawing inferences regarding the magnitude of benefits or damages associated with a change in resource conditions at a “policy site.”

Three general steps were used to identify and apply values for benefits transfer. First, a literature search identified potentially applicable primary studies. We identified 11 studies that focused on
values for reductions in or prevention of increases in nitrate contamination for drinking water wells. Second, we evaluated the validity and reliability of the studies identified. Primary evaluation criteria included the applicability and quality of the original study, each evaluated on multiple criteria such as sample size, response rates, significance of findings in statistical analysis, etc. And, third, we selected and adjusted values for application to CAFO impacts. We selected three studies to provide the primary values used for the benefit transfer:

- Poe and Bishop conducted a contingent valuation study in rural Portage County, Wisconsin, to estimate conditional incremental benefits of reducing nitrate levels in household wells. We used results from this study for per household values for changes in well nitrate concentrations from above the MCL to below the MCL.

- Crutchfield, Cooper, and Hellerstein evaluated the potential benefits of reducing or eliminating nitrates in drinking water by estimating average WTP for safer drinking water using a survey responses of 819 people in rural and nonrural areas in four regions of the United States (Indiana, Nebraska, Pennsylvania, Washington). We used results from this study for valuing incremental changes in nitrate concentrations below the MCL.

- De Zoysa used a contingent valuation study to estimate the benefits from three environmental services in the Maumee River basin in northwestern Ohio, including stabilization and reduction of nitrate levels. Rural and urban areas in the river basin were sampled and one out-of-basin urban area was sampled, with 427 returned questionnaires. We also used results from this study for valuing incremental changes in nitrate concentrations below the MCL.
The Consumer Price Index (CPI) was used to convert the annual mean household willingness-to-pay values obtained from these studies to 2001 dollars. Table 4 shows the point value estimates used for benefits transfer.

Aggregation of National Values

Table 5 shows the benefit estimates derived using the two different scenarios for three of the models (gamma, Tobit, and selection/truncation). Aggregate annual benefit estimates using these models range from $72 to $94 million. As seen in Table 5, there is no specific pattern of the magnitude of benefit estimates, depending on which model is used. For the “all large and medium” facilities scenario, the Tobit model generates the largest benefit estimates. For the “all large and select medium” facilities scenario, the selection/truncation model generates the largest benefit estimates.

Given the nature of this relationship and due to data limitations, it was important to explore truncation and selection issues in this analysis. Analysis of potential selection bias did not reveal a significant impact on the benefit estimates, thus increasing our confidence in the use of these benefit estimates for regulatory decisions.

Sensitivity Analysis

EPA and other federal agencies are under increasing pressure to expand their use of quantitative uncertainty analysis when evaluating regulations. The Office of Management and Budget (OMB) in their Draft Guidelines for the Conduct of Regulatory Analysis encourages agencies to attach probability distributions to costs and benefits “whenever possible.” The guidelines discuss several possible approaches of increasing complexity: qualitative discussion, sensitivity analysis, and formal probabilistic analysis. However, rules that cost more than $1 billion per year will be required to undertake formal treatment of uncertainty. The National Research Council of the
National Academy of Sciences has also recommended that EPA expand its treatment of uncertainty when evaluating the benefits of air pollution regulations (National Research Council, 2002).

The sensitivity analysis discussed here incorporates the random sampling error from the studies that generate the concentration-response function in the primary analysis. NRC recommended that additional sources of uncertainty be included in the primary analysis by specifying probability distributions for uncertain components of its benefits analysis. We thus also derive an estimate of the variance of one of the value estimates used in the benefits transfer and examine the sensitivity of aggregate benefit estimates with respect to this uncertainty as well. In doing so we felt it was important to explore the robustness of the benefit estimates to variance in (1) the model estimation and (2) in the per household benefit numbers used in the benefits transfer.

Variance in model estimation is inherent in the parameter estimates shown in Table 2. The variance-covariance matrix that is generated during model estimation provides information on the uncertainty in the parameter estimates and is used for exploring the sensitivity of benefit estimates to model uncertainty.

Variance in the household benefit numbers from Poe and Bishop is because their benefits estimates are also derived from a statistical model that has variances. Assessing the sensitivity to uncertainty in the per household benefit numbers used in the benefits transfer was undertaken by deriving the variance of the benefit estimates in Poe. Poe provides a 95% confidence interval for the household WTP estimates used in the benefit transfers — an estimate of $484 per household per year in 1991 dollars with a 95% confidence interval of $347-$655 (Table VII.2.4.2 in Poe). Adjusting to the mean of the two Poe and Bishop estimates used in this analysis and to 2001 dollars, this represents a mean WTP of $583 (95% confidence interval of $518-$789 in 2001
dollars). Assuming a standard normal distribution of the WTP estimates, the 95% confidence interval implies a standard deviation of $95.

It should be noted that we did not explore benefit estimation uncertainty with respect to variance in benefits for changes below the MCL because the benefits from this portion of the benefit analysis represent only a small percentage of the total benefits. It should also be noted that we assumed that there is no correlation between the variance in Poe and Bishop’s benefit estimates and the variance of the selection truncation model because these are derived from entirely different sources.

Confidence intervals for the aggregate benefit measures were generated using the Krinsky-Robb (Krinsky and Robb) method with 5,000 draws. As shown in Table 6, we did this for (1) the variance of just the model parameter estimates, (2) just for the value estimates from Poe, and (3) for both combined — using the same 5,000 draws from the normal distribution to maintain consistency for comparing the variation.

Overall, the 90% confidence intervals from the selection truncation model are within 22% of the mean estimate from the model. The 90% confidence interval is plus or minus about 27% when accounting for just the uncertainty of the Poe and Bishop benefit estimates. Combining these two sources of uncertainty generates a 90% confidence interval that is plus or minus about 32% of the mean benefit estimate. As shown in Table 6, the standard error of the aggregate benefit estimates from the selection truncation model is about 11% of the mean aggregate benefit estimate. The standard error of the aggregate benefit estimates from variation in the Poe and Bishop value estimates is about 16% of the mean aggregate benefit estimate.
Conclusions

In the analysis we discuss an approach to generating national benefit estimates for different regulatory options that EPA explored in revising CAFO regulations. The reduced impact of CAFOs on nitrates in drinking water wells was potentially a significant portion of the benefits derived from tighter controls on CAFO effluents, and thus it was useful to develop quantitative benefit estimates and to understand the uncertainty and variability inherent in these benefit estimates.

To conduct this assessment, we developed a statistical model of the surface nitrogen-well nitrate relationship using a variety of extant data sources. To account for potential difficulties in the statistical analysis, we developed the selection/truncation model to account for the nonnegativity of well nitrate concentrations and for the limited number of counties represented in the data set. Sensitivity analysis suggests that the aggregate benefit estimates are more sensitive to the variation in the benefit estimates used for the benefits transfer than to uncertainty inherent in the nitrogen-nitrate model. In either case the degree of uncertainty, as indicated by 90% confidence intervals, does not seem excessive and could provide additional insight when comparing potential benefits and costs from different regulatory options.

While the benefit estimates discussed here represent the best estimates available given the resources and data sources available, the analysis did indicate areas where improved information may significantly reduce uncertainty in this approach.

First, there is limited information on the scope and severity of the well nitrate problem. There have been a significant number of studies in specific regions, but to our knowledge there has been no concerted effort to test for and analyze the full range of potential contaminants in
drinking water wells on a national scale. With 20% of more of the population using wells as their primary water source, this would appear to be a significant area of concern.

Second, although the studies used in the benefit transfers are high quality efforts, they are limited geographically and temporally. They are limited geographically to the extent that most studies of benefits from well-nitrate reductions focus on very specific regions, often at the county or subcounty level. They are limited temporally in that as future benefit analysis is undertaken, the existing literature will be more and more outdated and subject to question in benefits transfer exercises. There would thus appear to be a potential significant benefit to continued research on both of these fronts, i.e., national analysis of the groundwater contamination problems and widespread studies of the benefits of reducing groundwater contamination.
Footnotes

1. “The term ‘point source’ means any discernible, confined and discrete conveyance, including but not limited to any pipe, ditch, channel, tunnel, conduit, well, discrete fissure, container, rolling stock, concentrated animal feeding operation, or vessel or other floating craft, from which pollutants are or may be discharged. This term does not include agricultural stormwater discharges and return flows from irrigated agriculture” [CWA Section 502(14)].

2. Under the old regulations, 12,410 of the approximately 450,000 AFOs were classified as CAFOs. With the revised NPDES scenarios considered by EPA, this could have risen to as many as 45,140.

3. We also considered a number of nonlinear specifications. These added little to the explanatory power and thus we stay with the linear model.

4. County level source of water data did not appear to be available from the 2000 U.S. Census.

5. Regions are defined as:
   - Midwest: ND, SD, MN, WI, IA, IL, MI, IN, MO, NE, KS
   - Mid-Atlantic: ME, NH, VT, NY, MA, RI, CT, OH, NJ, PA, DE, MD, VA, WV, KY, TN, NC
   - Pacific: CA, OR, WA, AK, HA
   - Central: ID, MT, WY, NV, UT, CO, AZ, NM, TX, OK
   - South: AR, LA, MS, AL, GA, SC, FL

6. The term “gamma model” is used because the regression is based on a gamma distribution, rather than the normal distribution (as is used in ordinary least squares regression), or another type of distribution. The gamma distribution allows for fitting of nonnegative, right skewed distributions (no observations are assumed to be censored in the gamma models). The gamma distribution has the density function:
We used the gamma distribution instead of the more commonly used exponential distribution since it is more general that the exponential model (includes the exponential specification as a special case).

7. Full details on all of the options and scenarios considered for the rule are in U.S. EPA (2003).

8. Although the regulation is based on total phosphates in applied manure, nitrogen loadings are closely linked to phosphate concentrations and thus vary as well under Option 2.

References


Table 1. Summary Statistics (n = 2985)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable (Well-Specific Data)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well Nitrate Concentration</td>
<td>3.57</td>
<td>6.51</td>
<td>0.05</td>
<td>84.30</td>
</tr>
<tr>
<td>Well-Specific Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ag Dummy</td>
<td>0.78</td>
<td>0.42</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Soil Group</td>
<td>2.42</td>
<td>0.66</td>
<td>1.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Well Depth</td>
<td>170.07</td>
<td>136.11</td>
<td>1.00</td>
<td>1,996.00</td>
</tr>
<tr>
<td>County-Specific Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loadings Ratio</td>
<td>2.11</td>
<td>4.32</td>
<td>0.00</td>
<td>19.64</td>
</tr>
<tr>
<td>Atmospheric Nitrogen</td>
<td>5.07</td>
<td>1.87</td>
<td>0.54</td>
<td>8.92</td>
</tr>
<tr>
<td>Septic Ratio</td>
<td>0.03</td>
<td>0.03</td>
<td>0.00</td>
<td>0.15</td>
</tr>
<tr>
<td>Variable</td>
<td>OLS(^a) Estimate (t-ratio)</td>
<td>Tobit Estimate (t-ratio)</td>
<td>Gamma Estimate (t-ratio)</td>
<td>Truncation Estimate (t-ratio)</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Constant</td>
<td>4.90 (6.38)</td>
<td>5.20 (5.53)</td>
<td>2.20 (11.74)</td>
<td>7.53 (7.13)</td>
</tr>
<tr>
<td>Nitrogen Loading</td>
<td>0.17 (1.50)</td>
<td>0.22 (5.98)</td>
<td>0.05 (7.34)</td>
<td>0.21 (5.33)</td>
</tr>
<tr>
<td>Well Depth</td>
<td>-0.63 (-7.29)</td>
<td>-0.73 (-7.19)</td>
<td>-0.17 (-13.73)</td>
<td>-0.72 (-6.84)</td>
</tr>
<tr>
<td>Soil Group</td>
<td>-1.24 (-6.70)</td>
<td>-1.62 (-6.95)</td>
<td>-0.38 (-9.84)</td>
<td>-2.33 (-10.52)</td>
</tr>
<tr>
<td>Agricultural</td>
<td>1.71 (5.93)</td>
<td>2.23 (6.01)</td>
<td>0.69 (10.42)</td>
<td>2.23 (4.50)</td>
</tr>
<tr>
<td>Dummy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmospheric</td>
<td>0.19 (6.04)</td>
<td>0.22 (1.59)</td>
<td>0.03 (1.16)</td>
<td>-0.12 (-0.85)</td>
</tr>
<tr>
<td>Nitrogen</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Septic Ratio</td>
<td>-0.14 (-0.02)</td>
<td>4.19 (0.47)</td>
<td>1.62 (0.94)</td>
<td>5.68 (0.53)</td>
</tr>
<tr>
<td>Central</td>
<td>0.06 (0.08)</td>
<td>0.38 (0.48)</td>
<td>-0.08 (-0.49)</td>
<td>0.07 (0.06)</td>
</tr>
<tr>
<td>Region</td>
<td>Mean Log-Likelihood</td>
<td>Confidence Interval</td>
<td>Mean Log-Likelihood</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Mid-Atlantic</td>
<td>-0.36</td>
<td>(-0.77)</td>
<td>-0.17</td>
<td>(-2.49)</td>
</tr>
<tr>
<td>Pacific</td>
<td>3.29 (6.62)</td>
<td>3.45 (6.13)</td>
<td>0.81 (6.95)</td>
<td>3.59 (5.88)</td>
</tr>
<tr>
<td>South</td>
<td>-1.89 (-3.56)</td>
<td>-4.15 (-5.65)</td>
<td>-0.91 (-7.47)</td>
<td>-4.51 (-5.02)</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.50 (50.64)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma</td>
<td>3.65 (140.83)</td>
<td>1.91 (0.014)</td>
<td>3.863 (0.012)</td>
<td></td>
</tr>
<tr>
<td>Mean Log-Likelihood</td>
<td>-2.32</td>
<td>-1.91</td>
<td>-1.86</td>
<td>-2.90</td>
</tr>
</tbody>
</table>

*a. Estimated using maximum likelihood.*
Table 3. Total Nitrogen Loadings and Changes in Households Affected by Well Nitrates

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Total N (lb)</th>
<th>Number of Households with Well Nitrate Concentrations above 10 mg/L&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Expected Reductions in [N] (mg/L)</th>
<th>Households Benefiting from Incremental Nitrate Reductions</th>
<th>Total Expected National Nitrate Reduction (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>555,025,318</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Option 2/3 —</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 6</td>
<td>447,034,179</td>
<td>148,705</td>
<td>0.15</td>
<td>5,813,446</td>
<td>927,730</td>
</tr>
<tr>
<td>Option 2/3 —</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 8</td>
<td>470,918,383</td>
<td>120,823</td>
<td>0.13</td>
<td>5,771,623</td>
<td>788,305</td>
</tr>
</tbody>
</table>

<sup>a</sup> Based on results from the Gamma model.
<table>
<thead>
<tr>
<th>Study</th>
<th>Value</th>
<th>Value 2001$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poe and Bishop</td>
<td>Annual WTP per household for reducing nitrates from above the MCL to the MCL</td>
<td>$583.00</td>
</tr>
<tr>
<td>Average of Crutchfield,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooper, and Hellerstein and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Zoysa</td>
<td>Annual WTP per mg/L between 10 mg/L and 1 mg/L</td>
<td>$2.09</td>
</tr>
</tbody>
</table>
Table 5. Benefits Estimates

<table>
<thead>
<tr>
<th>Facilities Regulated</th>
<th>Model</th>
<th>Result of Rule</th>
<th>Total mg/L</th>
<th>Number of Wells Crossing the MCL as Result of Rule</th>
<th>Reduction in Wells below MCL in Baseline</th>
<th>Groundwater Benefits$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option 2 — Scenario 6 All</td>
<td>Gamma</td>
<td>148,705</td>
<td></td>
<td>927,730</td>
<td></td>
<td>$88.6</td>
</tr>
<tr>
<td>Large and Medium</td>
<td>Tobit</td>
<td>157,999</td>
<td></td>
<td>935,138</td>
<td></td>
<td>$94.1</td>
</tr>
<tr>
<td></td>
<td>Select/Truncate</td>
<td>148,705</td>
<td></td>
<td>1,190,218</td>
<td></td>
<td>$89.2</td>
</tr>
<tr>
<td>Option 2 — Scenario 8 All</td>
<td>Gamma</td>
<td>120,823</td>
<td></td>
<td>788,305</td>
<td></td>
<td>$72.1</td>
</tr>
<tr>
<td>Large, Select Medium</td>
<td>Tobit</td>
<td>120,823</td>
<td></td>
<td>775,856</td>
<td></td>
<td>$72.1</td>
</tr>
<tr>
<td></td>
<td>Select/Truncate</td>
<td>130,117</td>
<td></td>
<td>992,566</td>
<td></td>
<td>$77.9</td>
</tr>
</tbody>
</table>

$^a$ In million dollars, not present value or annual.
Table 6. Sensitivity of Benefit Estimates (Selection/Truncation Model). a Option 2 —

Scenario 6 All Large and Medium

<table>
<thead>
<tr>
<th></th>
<th>Std. Deviation of Aggregate Benefit Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Variation from Selection/Truncation</td>
</tr>
<tr>
<td></td>
<td>Model</td>
</tr>
<tr>
<td></td>
<td>Variation from Poe and Bishop Benefit</td>
</tr>
<tr>
<td></td>
<td>Estimates</td>
</tr>
<tr>
<td></td>
<td>Variation from Model and Poe and Bishop</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Calculated from sample of 5,000 sets of model parameters, generated using Krinsky-Robb method (Krinsky and Robb).
Figure 1. Baseline Nitrogen Loadings (lbs per acre).
Figure 2. Counties Included in NPLA and Included in Statistical Analysis.
Figure 3. Percent Changes in Loadings Per Acre — Scenario 6, Option 2.