The Impact of Land Use on Water Quality: Evidence from California Wells

Scott Somerville, Pierre Mérel & Jeffrey Hadachek*

December 17, 2024

Abstract

Nitrate pollution threatens human health and ecosystems in many regions of the world. Although scientists agree that nitrogen compounds from human activity, notably agriculture, enter groundwater systems, empirical estimates of the impacts of land use on nitrate concentrations in well water are still lacking. We provide evidence of such impacts by combining nitrate concentration measurements from 6,016 groundwater wells with remotely sensed California land use data from 2007–2023. Results show that a 10 percentage point increase in the share of land used to grow high nitrogen hazard index crops within 500 meters of a well relative to undeveloped land is associated with a 12% increase in nitrate concentrations, while the same increase in low-intensity urban development is associated with a 10% increase. Local cattle populations also contribute to nitrate pollution. However, conditioning on initial nitrate measurements, we find limited evidence that human activity affects nitrate concentrations a decade later.

*Scott Somerville is a Ph.D. candidate and Pierre Mérel is a Professor, both in the Department of Agricultural and Resource Economics at the University of California, Davis. Jeffrey Hadachek is an Assistant Professor in the Agricultural and Applied Economics Department at the University of Wisconsin-Madison. Scott Somerville can be reached at ssomerville@ucdavis.edu, Pierre Mérel can be reached at merel@primal.ucdavis.edu, and Jeffrey Hadachek can be reached at hadachek@wisc.edu.

1 Introduction

Water pollution from human activity threatens drinking water supplies and ecosystems in many world regions (Rabotyagov et al., 2014; Van Meter, Van Cappellen, and Basu, 2018; Rahman, Mondal, and Tiwari, 2021; Abascal et al., 2022; Tozer, 2023; Jones et al., 2023).

In the United States (U.S.), water contamination has been a lingering cause of concern among the general public and, since the early 1970s, the object of much federal regulatory action following the adoption of the Clean Water Act and the Safe Drinking Water Act (Keiser and Shapiro, 2019b). By several measures, these laws have been successful at improving the quality of U.S. waterways (Keiser and Shapiro, 2019a). However, contaminants, such as nitrates, that enter water systems through non-point sources are excluded from the Clean Water Act because they are hard to directly regulate and increasingly contribute to water quality impairment (Olmstead, 2010; DeSimone, McMahon, and Rosen, 2014; Van Metre et al., 2016). For instance, Pennino, Compton, and Leibowitz (2017) estimate that in 2016, about 1.5 million people in the U.S. were supplied water from public systems in violation of the U.S. Environmental Protection Agency maximum contaminant level (MCL) for nitrates of 10 mg/L of nitrate-nitrogen. For households outside of public water system boundaries, a 2014 study of 3,621 randomly sampled private groundwater wells found that 4.1% exceeded the nitrate MCL (DeSimone, McMahon, and Rosen, 2014).

Nitrate contamination of U.S. groundwater resources poses large social costs, largely through damages to public health (Keeler et al., 2016). The regulatory MCL for nitrates in public water systems was established to mitigate the risk of infant methemoglobinemia, commonly known as "blue baby syndrome" (U.S. EPA, 1977). The limit was set based on public health studies conducted during the 1950s (Walton, 1951; Fewtrell, 2004) and did not comprehensively reflect other potential health hazards (Ward et al., 2018). Reviews of the epidemiological literature reveal that, beyond methemoglobinemia, nitrate-contaminated drinking water has been most strongly associated with gastric cancer (Picetti et al., 2022) along with colorectal cancer, thyroid disease, and neural tube defects (Ward et al., 2018).

Importantly, these health conditions exhibit positive correlations with nitrate concentrations even below established regulatory thresholds.

Nitrates in waterways originate from numerous sources, both anthropogenic such as urban runoff, wastewater treatment, or septic systems, and natural such as nitrogen fixation or atmospheric deposition (Wakida and Lerner, 2005; Lockhart, King, and Harter, 2013). However, studies point to nitrogen leaching from farmland as a leading cause of increasing ambient nitrate concentrations not only in surface water systems (Isbell et al., 2013; Hendricks et al., 2014; Paudel and Crago, 2021) but also in groundwater (Lockhart, King, and Harter, 2013; Rosenstock et al., 2014; Ransom et al., 2018), a primary source of drinking water for over 130 million U.S. residents (DeSimone, McMahon, and Rosen, 2014).

Groundwater nitrate pollution has received increased research attention in recent years. According to Abascal et al. (2022), the annual number of peer-reviewed studies on the topic increased from 20 in 1990 to 280 in 2021. Existing studies, most of them in the environmental science literature, tend to employ highly mechanistic (Kourakos et al., 2012) or non-parametric models (Ransom et al., 2017) of nitrate emissions, attenuation, and transport through soils and aquifers. Even economic studies of water contamination from agriculture tend to resort to mechanistic biophysical models to represent leaching from fields once farmer behavior has been accounted for (Mérel et al., 2014; Lark et al., 2022; Weng et al., 2024).

Our paper complements this literature by providing reduced-form evidence on the response of groundwater nitrate concentrations to local land use decisions based on observational data. To this end, we combine water-quality measurements from approximately 6,000 California wells over 17 years with cattle inventories and detailed remotely sensed land use data. Our main econometric model identifies groundwater quality impacts from spatial variation in land use shares and cattle populations that are plausibly exogenous to groundwater nitrate concentrations. Specifically, we regress mean nitrate concentrations measured from wells in recent years (2019–2023) on local cattle populations as well as mean land use shares measured twelve years prior (2007–2011) in the vicinity of each well, conditional on a suite of biophysical and regional controls. In an alternate set of regressions, we also control for initial nitrate concentrations.

Our approach provides novel insights into the contributions of various land use decisions to groundwater pollution. Notably, we find that a 10 percentage point increase in the share of land used to grow high-nitrogen crops—such as tree nuts or corn—within 500 meters of a well relative to undeveloped land is associated with an 11.6–19.7% increase in nitrate concentrations, while a 10 percentage point increase in the share of land used to grow low-nitrogen crops—such as rice and alfalfa— is associated with a 6.3–8.3% increase in nitrates. By comparison, a 10 percentage point increase in the share of land used for low-intensity (resp., high-intensity) urban development is linked to a 10.1–15% (resp., 10.5–16.2%) increase in nitrate concentrations. We also find that an additional 1,000 dairy cows, which falls short of the average dairy herd size in California,[1] within one kilometer of a well increases nitrate concentrations by 16–20%. Overall, these effects document the critical role of land development and human activities on groundwater quality.

Interestingly, these meaningful effects of land use shares or cattle populations on nitrate concentrations do not survive the inclusion of our control for baseline contamination. Specifically, regression coefficients decrease by an order of magnitude and become statistically insignificant once we add mean 2007–2011 nitrate concentrations as a covariate in the regression. We interpret this finding as arising from the fact that land use patterns, even if they determine nutrient leaching and ultimate groundwater contamination, tend to evolve very slowly over time. As a result, initial concentration levels may be highly correlated with both past and subsequent land use patterns, rendering identification of incremental contamination impacts challenging empirically. Indeed, our data show that over the period 2007–2023, land use patterns around wells have evolved very little relative to the existing cross-sectional variation across wells. To the extent that land use shares can be considered stationary, the regression that omits initial conditions may then be interpreted as reflecting

[1]As of 2019, California had about 1.7 million dairy cows, 1,330 dairy farms, and an average of 1,278 milking age cows per dairy (Marklein et al., 2021).

long-run contributions of land uses to groundwater contamination.

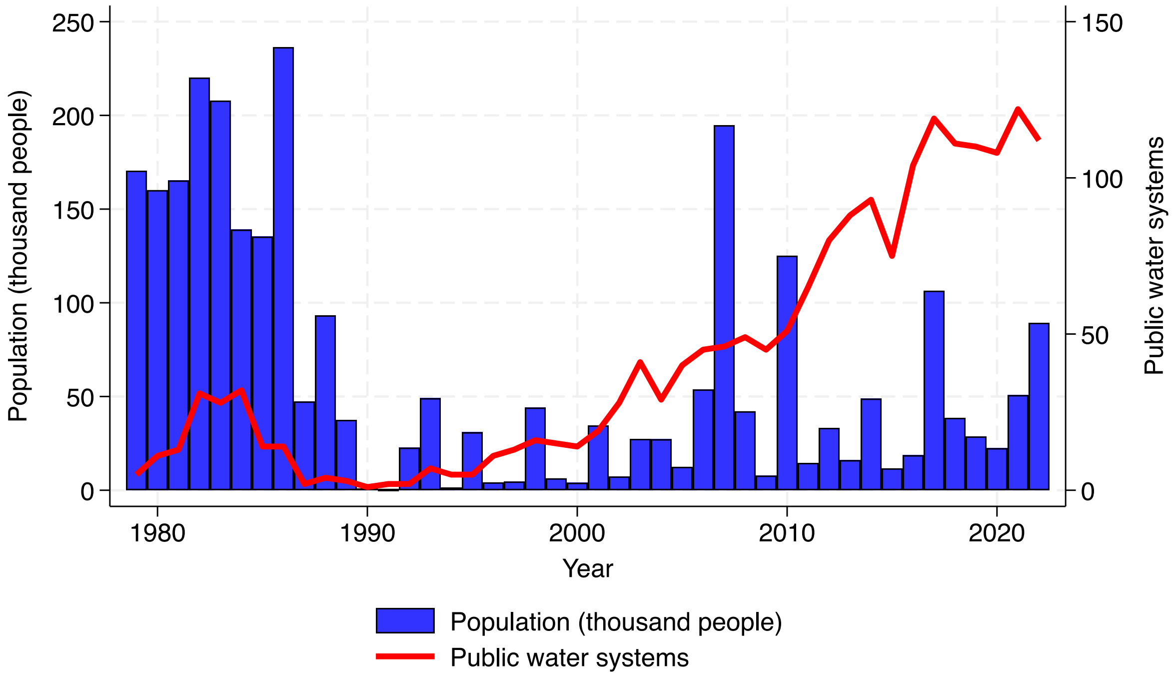


Figure 1: Count of California public water systems in violation of the nitrate maximum contaminant level, and affected population.

Source: U.S. Environmental Protection Agency (2023)

The intersection of heavy reliance on groundwater, urban development pressure, and proximity of wells to input-intensive and highly diversified agriculture makes California a compelling setting to measure pollution externalities from human activity. California has suffered from groundwater nitrate problems for decades (Harter et al., 2012) and studies indicate that the issue has gotten worse in recent times (Burow et al., 2013). Indeed, Figure 1 shows a rising trend in California public water systems violating the nitrate MCL since the 1990s.[2] The topic is central to many ongoing public policy deliberations regarding the competing interests of agricultural and residential groundwater users (Lubell, Blomquist, and Beutler, 2020) amidst efforts by California regulators to deter agricultural practices

[2]The California Health and Safety Code §116275 (2024) defines a public water system as delivering water for human consumption to at least 15 service connections used by yearlong residents or at least 25 people daily for 60 or more days per year.

believed to most significantly contribute to nitrate leaching. Notably, the Irrigated Lands Regulatory Program requires growers to implement nutrient management plans to prevent nitrogen fertilization in excess of crop uptake (California Water Boards, 2024b).[3]

Groundwater nitrate concentrations depend on the interplay of human, chemical, biological, and hydrological actions responsible for the nitrate emissions below the plant root zone and the subsequent transport of nitrates through subterranean systems (McMahon et al., 2008). The primary empirical challenge is controlling for the many features that determine the path—over both time and space—of nitrate molecules from the point at which they appear on the land surface. We address this challenge with the following strategies. First, we focus on land use within a 500-meter radius of the well, encompassing a region where nitrates accumulate within our effective lag length of about twelve years (California Department of Health Services, 2000; Harter, 2002; Boyle et al., 2012; DeSimone, McMahon, and Rosen, 2014). Second, we confine our analysis to variations within sub-basins—geographical areas of interconnected groundwater with shared characteristics influencing nitrate movement—and include additional controls for site-specific variables, such as soil composition, groundwater depth, and proximity to rivers, previously shown to be important factors in determining the fate of nitrates (Ransom et al., 2017). Third, we harness newly available comprehensive land use data sets, such as the Cropland Data Layer, which have been underutilized in existing literature. The detailed land use data allows for more precise land use groupings that reveal the relative importance of different land uses often missed by coarser aggregation procedures (Lichtenberg and Shapiro, 1997; Ransom et al., 2017). Lastly, given that the primary sources of variation in nitrate concentrations and land use allocations occur across space, our cross-sectional model offers the advantage of leveraging substantial variation for identification, in a context where variables evolve gradually over time. Because aquifer contamination occurs primarily through the leaching of contaminants deposited near the land surface, and we con-

[3] In addition, between 2000 and 2010, the Central Valley Regional Water Board implemented the Central Valley Dairy Order and the Central Valley Salinity Alternatives for Long-Term Sustainability program, which require land managers to engage in nitrate monitoring and reporting and adopt nutrient management plans (Harter, 2015).

trol for a variety of factors affecting their flow towards the water table, the scope for omitted variable bias remains limited in our setting, notably in comparison to hedonic analyses of climate change impacts on U.S. farmland values, which have been shown to be biased by the presence of public irrigation infrastructure (Schlenker, Hanemann, and Fisher, 2005) or non-farm influences (Ortiz-Bobea, 2020).

Our paper contributes to the understanding of critical questions surrounding the fate of nutrient emissions from land management (Galloway et al., 2008). Notably, we take a decadal view of the association between land use patterns and the concentration of nitrate in groundwater wells, where nitrates can potentially cause the most harm. Our empirical methodology relies on observational data, distinguishing it from approaches that integrate estimates derived from simulation models. Some recent observational studies have linked historical water quality data to crop and livestock land use decisions (Paudel and Crago, 2021; Raff and Meyer, 2022; Metaxoglou and Smith, 2024) and conservation efforts (Liu, Wang, and Zhang, 2023; Karwoski and Skidmore, 2024). These empirical studies, however, focus on surface water quality impacts, and analogous evidence for groundwater quality is warranted. Regarding the primary explanatory variables, we adopt an innovative approach for categorizing land uses according to their propensity to emit nitrogen below the root zone (Wu et al., 2005). Finally, quantifying the extent to which agricultural production contributes to nitrate pollution in groundwater can inform future conservation efforts targeting water quality.

The remainder of the paper is structured as follows. In the next section, we present our main econometric model. Next, we explain the construction of our data set, including data sources, and provide summary statistics. Subsequently, we present our main findings supported by several sensitivity checks. The last section concludes.

2 Empirical Methods

Our research design leverages cross-sectional variation in nitrate concentrations and land use surrounding well i in sub-basin b by regressing mean nitrate concentrations in recent years on mean land use shares measured in a prior period. Nitrate concentrations and land use shares are calculated as five-year averages and measured twelve years apart to account for the diffusion process. Our preferred regression equation takes the following form:

$$\ln N_{ibT_1} = \mathbf{L}_{ibT_0}{}'\boldsymbol{\beta} + \kappa C_{ib} + \mathbf{X}_{ib}{}'\boldsymbol{\gamma} + \lambda_b + \alpha \ln N_{ibT_0} + \varepsilon_{ibT_1} \tag{1}$$

where T_0 represents the period 2007–2011, T_1 represents the period 2019–2023, N_{ibT_0} and N_{ibT_1} represent the mean well-level nitrate concentrations in 2007-2011 and 2019-2023, respectively, and ln denotes the natural logarithm. We denote the land use shares by the column vector $\mathbf{L}_{ibT_0}$ and the local dairy cattle population by C_{ib}.[4] The column vector $\mathbf{X}_{ib}$ includes control variables for soil characteristics, the share of land with subsurface drains, distance to the nearest river, depth to groundwater, precipitation, and surface water deliveries. The symbol λ_b represents sub-basin fixed effects that control for unobserved characteristics of the aquifer and deep soils common to the wells within sub-basin b. In some regressions we include lagged nitrate concentrations to reflect the fact that nitrates accumulate in groundwater over time, so initial conditions matter.

We estimate Equation (1) using Ordinary Least Squares. The vector $\boldsymbol{\beta}$ captures land use impacts, with each element β^k representing an $(e^{\beta^k} - 1) \times 100$ percent change in groundwater nitrate concentrations associated with a one unit increase in the share of land dedicated to activity k. If we measure the land use share using an index from zero to one, then the effect is that of a change from a zero to one share. When discussing our results, we consider instead the effect of a 10 percentage point increase in the share, which is obtained by computing $(e^{\beta^k \times 0.1} - 1) \times 100$. Since land use shares sum up to one, we omit undeveloped land (i.e.,

[4] Although not ideal, we only observe one year of dairy cattle population data. We discuss this issue further in the data and results section.

natural lands such as forests, wetlands, and deserts), so the effect of an increase in land use k is to be understood as arising from replacing undeveloped land by activity k. The coefficient κ captures the relative impact of an additional one thousand dairy cattle located within a specific distance of well i.

In our main cross-sectional specifications, temporal averages in the recent and past periods are calculated over a five-year window, and we allow for a 12-year lag between the midpoints of recent and past periods, the longest lag afforded by our data set conditional on using five-year averages. This lag, as well as the averaging of land shares over multiple years, are meant to capture the fact that nutrient leaching into groundwater is generally thought to be a gradual, multi-year process. The average depth to groundwater across our sample of wells being 31 meters, it is unlikely that land uses could have a direct contemporaneous effect on groundwater concentrations (Boyle et al., 2012).[5]

The vector $\mathbf{X}_{ib}$ includes, among others, controls for depth to groundwater, precipitation, and surface water deliveries. Because these variables typically vary over time and their values plausibly affect the ultimate nitrate concentrations cumulatively over time, we define the depth to groundwater as the mean depth to groundwater over the length of our panel (2007 through 2023), precipitation as the cumulative precipitation from 2007 through 2023, and surface water deliveries as the sum of acre-feet of surface water delivered per acre of agricultural land in the water region in which well i is located from 2007 through 2021.[6]

Identification of the impacts of land use on nitrate concentrations requires adequate controls for biophysical and well characteristics, as these may also partially affect land uses. Our long list of controls includes variables highlighted in previous literature as meaningfully impacting nitrate concentrations in groundwater (Ransom et al., 2017). However, data limitations mean that we cannot control for every characteristic of the aquifers, wells, and

[5] For completeness, we conducted contemporaneous regressions of mean 2019–2023 nitrate concentrations on mean 2019–2023 land use shares. We present the results in Appendix Table A.4 and discuss them in Section 4.4.

[6] Water regions are Detailed Analysis Unit by County regions as defined by California's Department of Water Resources (California Department of Water Resources, 2019).

deep soils; thus, we adopt the common econometric approach of controlling for unobserved potential confounders using fixed effects, here at the sub-basin level.

The California Department of Water Resources defines 515 sub-basins that underlie approximately 42% of the state's land area, 82% of its population, and 97% of its agricultural land (California Department of Water Resources, 2019, 2021). On average, sub-basins underlay 480 square kilometers of land area. The sub-basin delineation of California groundwater systems divides the large San Joaquin and Sacramento Valley basins—both within the Central Valley, the heartland of California agriculture and the region where about half of the wells in our sample reside—into 35 regions based on geological features that slow, but might not entirely prevent, water from mixing over time.[7] The California Department of Water Resources subdivides 28 smaller basins into 86 sub-basins. The remainder of the state's basins are not divided into sub-basins and are treated as their own sub-basin in our analysis. Our sample includes 230 sub-basins because we do not observe wells in all groundwater basins.

Cross-sectional regressions are, by nature, susceptible to omitted variable bias (Deschênes and Greenstone, 2007), because cross-sectional variation in the regressor of interest might correlate with variation in the error term. This is particularly true with spatial data as geographical proximity may affect the outcome variable through a variety of channels that are difficult to control for. A common way to alleviate omitted variables concerns in such contexts is to include fixed effects that capture spatial proximity. Many studies use administrative delineations such as zip codes, county, agricultural district, or state boundaries (Schlenker, Hanemann, and Fisher, 2006; Ortiz-Bobea, 2020; Gammans, Mérel, and Ortiz-Bobea, 2024). In the present case where the outcome of interest, groundwater contamination, likely varies according to hydrology, we find it appropriate to use sub-basin fixed effects.

Sub-basin fixed effects capture local unobserved biophysical characteristics that might vary greatly throughout the large Central Valley basins, such as reduction/oxidation conditions—the chemical environment that determines the rate at which nitrates convert

[7] In comparison, basins map hydrologically isolated groundwater bodies.

to other nitrogen compounds—and groundwater age (Ransom et al., 2017). To the extent that these characteristics correlate with local land use patterns, using basin instead of sub-basin fixed effects would cause bias. The inclusion of sub-basin fixed effects is not costless, however. First, it removes a source of potentially useful cross-sectional variation, namely variation present within basins. Second, our data include 29 sub-basins with only one well, so using sub-basin fixed effects reduces the effective sample size. Fortunately, these 29 observations represent a small share of the 6,016 wells and 230 sub-basins present in the sample.

3 Data

We compile our data set using publicly available data from several government and academic institutions. In what follows, we describe our data sources, explain our approach to data cleaning and aggregation, and provide detailed descriptions and summary statistics of the variables used in our regressions.

3.1 Data Set Construction

3.1.1 Nitrate Concentrations

We obtain data on nitrate concentrations from the State Water Resources Control Board Groundwater Ambient Monitoring and Assessment (GAMA) Program, readily available from their online portal (California Water Boards, 2024a). GAMA staff compiled these data from multiple government, research, and local water organizations from private, public, irrigation, and monitoring wells throughout California. We exclude monitoring wells from our sample as many of these are located close to locations of unauthorized releases of pollutants into the environment, such as leaking underground storage tanks on industrial sites. In addition, monitoring wells have a different set of design, construction, and management standards compared to irrigation and drinking water wells, leading to water samples not representative of water extracted for human use.[8]

In the groundwater quality data, we find measurements of nitrate-nitrogen, nitrite-nitrogen, and nitrate-nitrogen plus nitrite-nitrogen. Nitrate-nitrogen refers to the weight of the nitrogen atom in the nitrate molecule.[9] For brevity, we refer to nitrate-nitrogen as nitrates. Nitrite, the other nitrogen molecule reported in the data, is an unstable molecule that readily oxidizes to nitrate and occurs in much smaller quantities in groundwater (Burkart and Stoner, 2008). In addition, for drinking water quality purposes, nitrate plus nitrite concentrations represent nitrate concentrations (California Water Boards, Central Coast Regional Water Quality Control Board, 2013). Therefore, we sum nitrate-nitrogen and nitrite-nitrogen concentrations and report this measure as nitrate.

The GAMA data includes geographic coordinates, sample collection date, and the laboratory minimum nitrogen concentration detection limit. Minimum detection limits range from 0.02 mg/L to 1 mg/L, depending on the laboratory. To deal with censored observations, we follow Keiser and Shapiro (2019a) and let concentrations below the minimum detection threshold equal the detection limit. As a sensitivity check, we let the censored concentration equal half the detection limit and report these regression estimates in Section 4.4.

Lastly, we find that the month and frequency of water sampling vary over wells. To deal with these issues, we calculate the annual average nitrate concentration within the calendar year. Our decision to calculate annual average nitrate concentration is supported by the fact that we find no evidence of seasonality in nitrate concentrations in the data. The lack of seasonal variation in nitrate concentrations is not surprising given that California groundwater experiences low annual recharge and because of groundwater mixing (Lockhart, King, and Harter, 2013).

[8] These well standards are described on the California Department of Water Resources (2024) website.

[9] Laboratories typically measure and report nitrate-nitrogen concentrations in water samples, and by convention, scholarly works focus on nitrate-nitrogen. We follow this convention in this paper.

3.1.2 Land Use

We use land use data from the Cropland Data Layer, a satellite data product of the U.S. Department of Agriculture National Agricultural Statistics Service readily available online. The Cropland Data Layer is an annual raster image of the U.S., with each cell classified into a land use or crop type based on remotely sensed data and verified by U.S. Department of Agriculture personnel by ground truth sampling to ensure accuracy. The first California Cropland Data Layer image was taken in 2007, with cells measuring 56 square meters. Subsequent images have cells measuring 30 square meters.

We focus on land use within a 500-meter radius of a well. Using a circular buffer zone is a standard approach used in the existing literature and is a good approximation to the land that contributes to groundwater recharge to a well (Johnson and Belitz, 2009). Furthermore, Kolpin (1997) used a range of buffers from 200 meters to 2 kilometers and showed that 500 meters provide the best correlation between land use and nitrate concentrations, and Koterba (1998) reviewed six articles focusing on nitrate and pesticide pollution and recommended a 500-meter buffer for empirical studies. More recently, Johnson and Belitz (2009) show that a 500-meter buffer is a robust proxy for the well supply area in the Central Valley of California.[10] We use the coordinates of the well locations from the groundwater quality data and extract land use data from the Cropland Data Layer using geospatial methods in R.

There are about 160 land use and crop types within 500 meters of our sample of wells. Therefore, to create a parsimonious regression model, we aggregate crops based on the Nitrogen Hazard Index (NHI), a protocol that uses expert opinion and data to assign a value of 1 through 4 to crops based on five factors. These factors are 1) rooting depth, 2) ratio of nitrogen in the crop tops to the recommended nitrogen application, 3) fraction of the crop nitrogen that is removed from the field with the marketed product, 4) magnitude of the peak nitrogen uptake rate, and 5) whether the crop is harvested at a time when nitrogen uptake rate is high (Wu et al., 2005). To date, researchers have used the NHI to aggregate

[10]In Section 4.4, we show results based on an alternative one-kilometer buffer.

crops in a handful of research and extension publications (e.g., Dzurella et al. (2015) and Beaudette and O'Geen (2009)). We provide a full list of the crops and non-agricultural land uses found within 500 meters of the sample wells and their assigned category in Appendix Table A.1. Crops with an NHI equal to 1 are assigned to the "Low-NHI crops" group, while crops with an NHI equal to 2, 3, or 4 are assigned to the "High-NHI crops" group. This categorization results in crop groups with land shares having a similar order of magnitude on average, as crops with an NHI of 3 or 4 tend to occupy a very small area in our sample.

3.1.3 Soils

Data on soil characteristics come from U.S. Department of Agriculture (2014) as a raster file, with each pixel representing an area of 90 square meters. Similar to the procedure for land use data, we extract soil data within 500 meters of the well and then calculate average sand, silt, clay, and organic matter percentages using weighted averages across pixels, where weights are based on the share of pixel area contained within the buffer.

The raw soil data from U.S. Department of Agriculture (2014) represents an area and depth weighted average of soil attributes within the raster cell measured from the land surface to a depth of about 1.5 meters.[11] Sand, silt, and clay percentages characterize the soil texture, and the percentages sum up to one. Organic matter, also called humus, refers to the percentage of decomposed plant and animal residues in the soil's dry weight. Organic matter, sand, silt, and clay constitute the major components of soil (Hillel, 2008).

3.1.4 Tile Drains

Subsurface tile drainage is often installed in poorly drained soil to divert water away from fields and into distant waterways. Therefore, the presence of tile drainage may meaningfully affect nitrate leaching into groundwater sources. We use a raster data set of subsurface tile drainage in the U.S. from Valayamkunnath et al. (2020) with cells measuring 30 square

[11] Wieczorek (2014) offers a useful overview of the soil data.

meters. This data set contains a variable indicating the presence of subsurface drains within a raster cell. Like land use and soil data, we extract drainage raster cells within 500 meters of the well and then calculate the share of land with drains using weighted averages.

3.1.5 Precipitation

Annual precipitation data come from Oregon State University PRISM Climate Group (2024). The popular PRISM data set provides a model-based estimate of precipitation for the U.S. at a resolution of 4 kilometers. For each well in our sample, we extract the precipitation data from 2007 to 2023 from the grid cell containing the well.

3.1.6 Rivers

We use the Major Rivers and Creeks maps from the United States Geological Survey National Hydrology data set available online from the California Department of Water Resources (2023). Using geospatial techniques in R, we measure the distance from the well to the nearest river.

3.1.7 Surface Water Deliveries

California agriculture heavily relies on irrigation via a network of surface water storage and distribution networks, since California's agricultural regions receive very little rainfall during the growing season. Surface water irrigation allotments, therefore, may affect both land use patterns and nitrate leaching, making it an important control variable. We use a data set of surface water deliveries compiled by Hagerty (2022) and available online. The data set contains reported surface water deliveries to agricultural users in California regions from 1993 to 2021 and acres of agricultural land in each region. The regions, called Detailed Analysis Unit by county, divide California's hydrological regions and planning areas into smaller geographic areas for agricultural land use and water balance analysis by California's Department of Water Resources.

3.1.8 Dairy Cattle Inventories

As part of the Waste Discharge Reports and Requirements Order, livestock enterprises are required to submit a Report of Waste Discharge that includes livestock inventories, and the California State Water Resources Control Board publishes an online database that includes the geographic coordinates of the facility and the livestock populations (California State Water Resources Control Board, 2022). The inventories are self-reported and periodically updated. Therefore, the available data represent the latest livestock inventories reported by farmers. Unfortunately, it is not possible to use this data set to recover historical inventories throughout the sample period. We focus on dairy cattle because milk is the largest agricultural commodity in California in terms of farm receipts, dairies are responsible for most of the nitrogen-rich manure produced by livestock species, and most dairies are located in the Central Valley.

3.2 Data Aggregation

In our preferred model specification, we group annual well-level observations of nitrate concentrations and land use shares into five-year windows at the endpoints of the sample period and calculate the mean of the respective variables. The following considerations inform our decision to construct averages over time. First, many wells have nitrate measurements in some years but not in others. Therefore, using a single year for determining nitrate concentrations would severely limit the number of observations that enter the regression, especially when we include a lagged dependent variable to capture initial conditions. Averaging over the available annual observations within a five-year window results in a larger sample size. In doing so, we implicitly assume that nitrate concentration data are missing at random. Second, we find that nitrate concentrations and land use shares evolve gradually over time, therefore we do not believe that we lose meaningful information by averaging over a five-year window. Third, nitrate concentrations at a point in time are the result of cumulative nutrient leaching from past years, excluding perhaps the most recent years as it takes time for nutrients to make their way to the groundwater table. Using a five-year average of land use over prior years partially captures the cumulative nature of the leaching process.

3.3 Data Summary and Statistics

Table 1 provides the variable descriptions. Figure 2 maps the relative size, shape, and location of 515 groundwater sub-basins. It further illustrates the spatial distribution of our effective sample of 6,016 wells across 230 sub-basins scattered across the state, with most observed wells located in the Central Valley and coastal and southern regions.[12] The figure shows that ten large Central Valley sub-basins contain 90–455 well observations and a further thirteen sub-basins contain more than 30 observed wells. Consequently, about 50% of observations lie within the 35 Central Valley sub-basins. In comparison, basins in the Northeast and along the North Coast regions are smaller and have fewer observations. On average, the observed sub-basins outside the Central Valley contain 17 wells. We do not observe nitrate concentrations in many of the basins in southeast California, a mostly desert and mountainous region with little agriculture and far from major population centers. White regions in Figure 2 represent non-basin regions. For instance, the large non-basin region east of the Central Valley follows the Sierra Nevada mountain range.

Table 2 reports summary statistics of the effective sample of 6,016 wells used in the regressions. The sample includes observations from 5,998 municipal wells, 3 domestic wells, 5 irrigation wells, and 10 other water supply wells. The summary statistics reveal that mean nitrate concentrations increased from 2.7 mg/L in 2007–2011 to 2.9 mg/L in 2019–2023, a meaningful increase of 7.4% that points to worsening water quality. This positive trend is consistent with earlier studies of nitrate concentrations in public well water supply (Burow, Shelton, and Dubrovsky, 2008; Pennino, Compton, and Leibowitz, 2017). Scientists have linked nitrate concentrations above 2 mg/L to contamination from anthropogenic sources such as agriculture and urban development (Mueller and Helsel, 1996; Harter, 2009; Lock-

[12] The Central Valley is a 700-kilometer-long region of flat land running north to south along the spine of California and is the top agricultural area in the state.

Table 1: Variable Descriptions

Variable	Description
Nitrate concentration	Milligrams of nitrate-nitrogen plus nitrite-nitrogen per liter of untreated water
Low-NHI crops	Share of land within buffer zone used to grow crops with a Nitrogen Hazard Index 1
High-NHI crops	Share of land within buffer zone used to grow crops with a Nitrogen Hazard Index 2, 3 and 4
Fallow	Share of cropland within buffer zone fallowed or idle
Pasture	Share of land within buffer zone used to grow pasture or grassland
Low-intensity development	Share of land within buffer zone with a mixture of vegetation, such as gardens and parks, and constructed materials where impervious surfaces account for less than 50% of the land cover, typically single-family housing
High-intensity development	Share of land within buffer zone where impervious surfaces account for more than 50% of land cover such as apartment, commercial, and industrial complexes
Undeveloped land	Share of land within buffer zone used for forest, wetland, and other natural land cover
Cattle within 1km	Inventory of lactating and dry cows and heifers on dairies located within 1 kilometer of the well (thousand)
Cattle within 1–5km	Inventory of lactating and dry cows and heifers on dairies located between 1 and 5 kilometers of the well (thousand)
Surface water deliveries	Cumulative acre-feet of surface water delivered to a DAUCO region per acre of agricultural land in 2007–2021 in one hundred feet increments
Precipitation	Cumulative precipitation in 2007–2023 in ten meter increments
Depth to groundwater	Distance from land surface to groundwater in ten meter increments
Distance to river	Distance from well to nearest river in ten kilometer increments
Sand, silt & clay	Mean share of soil textural fraction composed of sand, silt and clay, respectively, within buffer zone
Organic matter	Mean share of organic matter in soil within buffer zone
Drainage	Share of land within buffer zone with subsurface drains

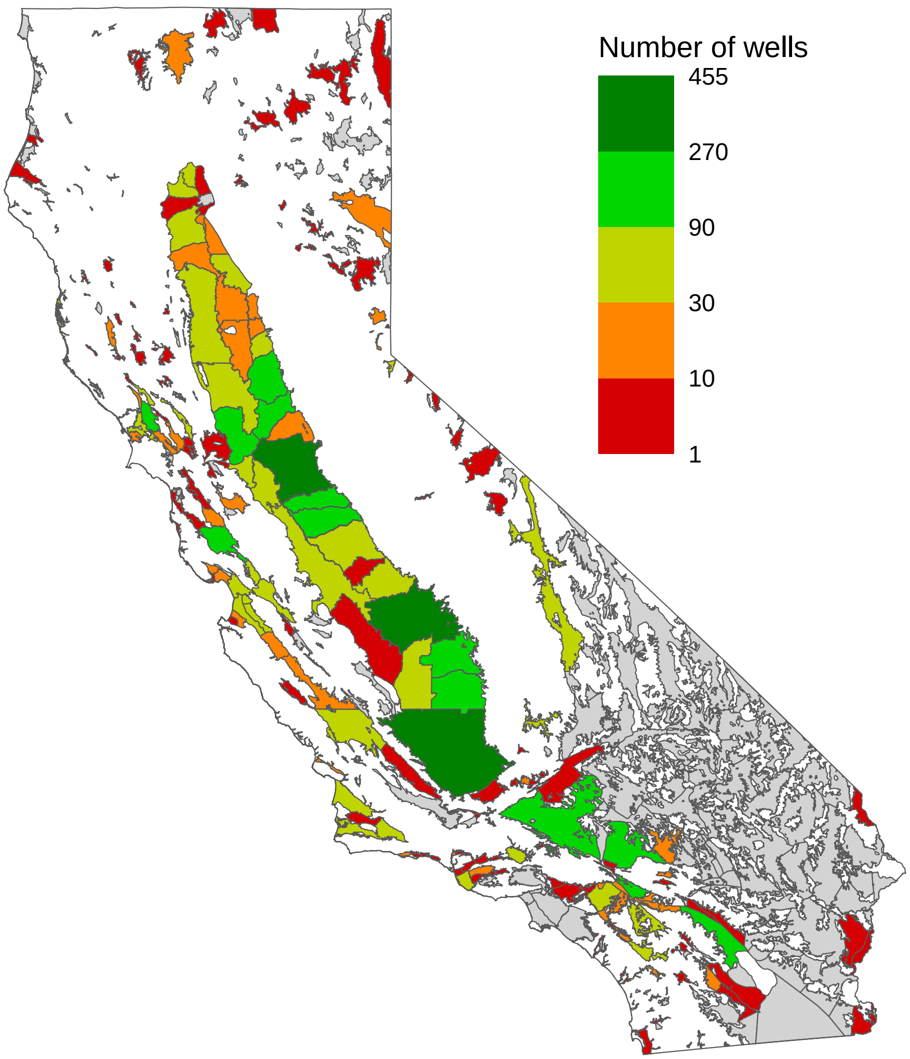


Figure 2: Number of wells observed in each sub-basin.

Note: The figure maps the location of 515 sub-basins and shows the locations of the 6,016 wells used in our regressions. Grey regions indicate sub-basins with no well data, and white regions represent non-basin areas.

hart, King, and Harter, 2013). Moreover, epidemiological studies reveal an increased risk of thyroid (Ward et al., 2010) and ovarian (Inoue-Choi et al., 2015) cancer from drinking water containing 2–3 mg of nitrates per liter relative to water with less than 0.5 mg/L. Out of our 6,016 sample wells, 45.2% have 2007–2011 nitrate concentrations above 2 mg/L, and 2.7% have baseline concentrations above the federal MCL of 10 mg/L. These shares rise to 45.4% and 3.6%, respectively, for the period 2019–2023.

Table 2: Summary Statistics

	Mean	Std. Dev.	Min.	Max.
Nitrate concentration in 2019–2023	2.90	3.61	0.004	67.28
Nitrate concentration in 2007–2011	2.70	3.06	0.02	56.44
Low-NHI crops	0.05	0.11	0	0.90
High-NHI crops	0.09	0.15	0	0.94
Fallow	0.03	0.06	0	0.69
Pasture	0.17	0.20	0	0.96
Low-intensity development	0.30	0.19	0	0.95
High-intensity development	0.23	0.25	0	0.99
Undeveloped land	0.13	0.21	0	1
Cattle within 1km	0.01	0.17	0	8.69
Cattle within 1–5km	0.38	1.67	0	23.03
Surface water deliveries	0.06	0.85	0	20.21
Precipitation	0.65	0.35	0.10	2.50
Depth to groundwater	3.10	2.87	0.05	20.97
Distance to river	0.15	0.19	0	1.92
Sand	0.55	0.20	0.04	0.98
Silt	0.26	0.11	0.01	0.70
Clay	0.19	0.11	0.01	0.65
Organic matter	0.01	0.02	0.00005	0.37
Drainage	0.01	0.05	0	0.82

Note: Effective sample size is 6,016 wells. Low NHI crops, high NHI crops, pasture, fallow, developed, and undeveloped represent mean land use shares from 2007 through 2011.

The maximum nitrate concentration observed in our data, 67 mg/L, is high but within the expected range reported by other researchers. For instance, Pennino, Compton, and Leibowitz (2017) studied California public water systems that exceeded the 10 mg/L MCL threshold for safe drinking water between 1994 and 2016 and found that well water contained an average nitrate concentration of 64 mg/L.

Table 2 reveals that low- and high-NHI crops account for 5% and 9% of total land use, respectively. Fallow agricultural land has the smallest share (3%) among the land use categories defined in our analysis. Pasture, which includes grazing grasslands and pasture harvested for hay, makes up the largest share of agricultural land surrounding wells (17%).

Urban developments represent the largest share of land use, with low- and high-intensity developments occupying 30% and 23% of land within 500 meters of wells, respectively. Low-intensity developments are areas with less than 50% of the land covered in impervious surfaces like concrete and asphalt. Examples include urban open spaces like parks and golf courses and areas with a mixture of constructed materials and vegetation, like housing units with gardens. In high-intensity developments, impervious surfaces account for more than 50% of the land cover, and examples include apartment, commercial, and industrial complexes. The dominance of urban land use in our sample is not surprising, considering that many of the wells are used to extract water for urban uses, particularly municipal wells.

Undeveloped land comprises, on average, 13% of total land use, and encompasses forests, herbaceous scrubland, and barren regions with little vegetation, such as rocky mountains and deserts. Undeveloped land serves as the base land use category in our regressions.

Table 2 reveals that the sample has, on average, 10 dairy cattle within 1 km of wells and 380 cattle within 1–5 km. We find that 73% of wells with a dairy within 1 km lie inside the Central Valley, with most in Stanislaus, Tulare, and Kings counties. Similarly, Central Valley wells account for 81% of wells with a dairy within 1–5 km, with wells in Stanislaus and Tulare accounting for the largest share.[13] Outside of the Central Valley, we observe wells with a dairy located less than five kilometers away in the North Coast region, particularly Sonoma and Marin counties, and a handful in the Southern California and Central Coast regions, both minor dairy-producing areas. The predominance of Central Valley wells in the set of wells with nearby dairies is unsurprising considering that approximately 90% of milk cows reside in Central Valley dairies, with Tulare county containing 28% of the state herd,

[13]Figure B.1 in Appendix B provides a map of California counties.

followed by Merced and Stanislaus (U.S. Department of Agriculture, 2022). Our sample includes a Tulare-based well located within 1 km of a large 8,690-cow dairy (the maximum cattle population within 1 km of wells), and a Stanislaus-based well with 22 dairies managing 23,030 cattle within 1–5 km (the maximum cattle population within 1–5 km).

Focusing on the control variables in Table 2, the mean cumulative 2007–2023 precipitation equals 6.5 meters. The mean depth to groundwater of 31 meters is consistent with the depth to groundwater found by others for Central Valley wells (Nolan et al., 2014). The means of the textural components of soil surrounding wells, namely sand, silt, and clay equal 55%, 26%, and 19%, respectively, and the mean organic matter equals 1%. Subsurface field drainage is concentrated in the Imperial Valley and Northern San Joaquin Valley regions and, on average, underlies 1% of land within 500 meters of sample wells.

4 Results and Discussion

Table 3 shows estimates of mean 2019–2023 nitrate concentrations regressed on mean 2007–2011 land use shares within 500 meters of the well and various controls, based on Equation (1). The table contains four columns of results. Columns 1 and 2 display the results of regressions without initial nitrate concentrations, while the regressions in columns 3 and 4 control for mean 2007–2011 nitrate concentrations. Therefore, the land use coefficients in columns 3 and 4 represent impacts on nitrates approximately twelve years later. Columns 2 and 4 show results of regressions that include sub-basin fixed effects. We estimate coefficients for land use relative to undeveloped land, and sand and silt relative to clay. To account for the potential correlation of the error term across wells within sub-basins, we estimate cluster-robust standard errors clustered by sub-basin.

Model fit comparisons indicate that sub-basin fixed effects serve as important controls. In models excluding initial nitrate concentrations, the R-squared rises from 0.14 to 0.32 upon the inclusion of sub-basin fixed effects. The model fit improves markedly when including initial nitrate concentrations, with the R-squared increasing to 0.77 in column 3 and further to 0.79 when sub-basin fixed effects are added in column 4.

4.1 Land Use and Cattle Effects

We begin the discussion by focusing on estimates of regressions without initial nitrate concentrations reported in columns 1 and 2 of Table 3. Estimates reveal positive and significant relationships between nitrate concentrations and the share of land used for agriculture and urban development relative to undeveloped land. The inclusion of sub-basin fixed effects leads to a marked decrease in the magnitude of land use and cattle population coefficients and a loss of significance at the 5% level on cattle population within 1 km compared to the regression without fixed effects in column 1. The coefficients on fallow and cattle population within 1–5 km are exceptions, although they do not exhibit statistical significance in columns 1 or 2.

The coefficients on high-NHI crops are the largest point estimates among the land use shares in specifications without initial nitrates, equalling 1.8 without sub-basin fixed effects (column 1) and 1.1 when controlling for sub-basin fixed effects (column 2). The coefficients from columns 1 and 2 imply that a 10 percentage point increase in the share of land used for high-NHI crops is associated with a 19.7% and 11.6% increase in nitrate concentrations, respectively, relative to undeveloped land. For example, using the sample means reported in Table 2 and the coefficient from high-NHI crops from column 2, increasing the mean share of land dedicated to high-NHI crops from 9% to 19% increases nitrate concentrations from 2.9 mg/L to 3.2 mg/L, all else constant.

The coefficients on low-NHI crops from columns 1 and 2 imply that a 10 percentage point increase in the share of land used for low-NHI crops is associated with nitrate concentrations 8.3% and 6.3% higher, respectively, than undeveloped land. The finding that the coefficients on low-NHI crops are smaller than the coefficients on high-NHI crops is consistent with the crop categories' propensity for leaching nitrates below the root zone (Wu et al., 2005).

Table 3: Impacts of Mean 2007–2011 Land Use Shares Within 500 meters of Wells on Mean 2019–2023 Nitrate Concentrations

	Dependent variable:			
	Log nitrate concentration 2019–2023			
	(1)	(2)	(3)	(4)
Low-NHI crops	0.80***	0.61**	-0.23**	-0.01
	(0.26)	(0.31)	(0.10)	(0.13)
High-NHI crops	1.8***	1.1***	0.21	0.14
	(0.34)	(0.31)	(0.14)	(0.14)
Fallow	-0.38	0.88	0.02	0.48*
	(0.60)	(0.74)	(0.16)	(0.26)
Pasture	1.3***	0.73***	0.11	0.07
	(0.17)	(0.25)	(0.07)	(0.10)
Low-intensity development	1.4***	0.96***	0.11	0.10
	(0.19)	(0.16)	(0.07)	(0.07)
High-intensity development	1.5***	1.0***	0.08	0.13
	(0.23)	(0.22)	(0.07)	(0.09)
Cattle within 1km	0.18**	0.15*	0.02	0.009
	(0.07)	(0.08)	(0.07)	(0.07)
Cattle within 1–5km	0.01	-0.01	0.004	0.003
	(0.02)	(0.02)	(0.009)	(0.01)
Surface water deliveries	0.05*	0.05***	-0.001	-0.0003
	(0.03)	(0.003)	(0.003)	(0.001)
Precipitation	-0.57***	-0.11	-0.19***	0.03
	(0.14)	(0.38)	(0.05)	(0.12)
Depth to groundwater	0.006	0.002	0.002	0.007
	(0.01)	(0.02)	(0.004)	(0.005)
Drainage	-1.6***	-0.85	-0.33*	-0.30**
	(0.35)	(0.58)	(0.18)	(0.13)
Sand	0.35	-0.27	-0.16	-0.38**
	(0.33)	(0.31)	(0.16)	(0.17)
Silt	-0.17	-0.68	-0.45*	-0.74**
	(0.55)	(0.56)	(0.27)	(0.30)
Organic matter	-5.1***	-4.2**	-0.43	-0.42
	(1.7)	(1.9)	(0.62)	(0.66)
Distance to river	0.24	0.51	0.16**	0.16*
	(0.28)	(0.34)	(0.07)	(0.09)
Initial nitrate concentration			1.0***	0.98***
			(0.02)	(0.02)
Sub-basin FE	No	Yes	No	Yes
Observations	6,016	6,016	6,016	6,016
R^2	0.13855	0.32333	0.77203	0.79430

Note: Undeveloped land is the default land use and clay is the default soil textural fraction. Initial nitrate concentration equals log nitrate concentration 2007–2011. Standard errors are clustered by sub-basin * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Estimates of nitrate emissions reaching groundwater from the hydrology literature further support the difference between low- and high-NHI crops. For instance, Ransom et al. (2018) estimate that emissions of nitrate-nitrogen into California Central Valley groundwater from rice and alfalfa fields (low-NHI crops) total about 4 kg of nitrogen per hectare per year (kg N/Ha/yr), while emissions from high-NHI crops like citrus equal 65 kg N/ha/yr, vegetables and berries equal 49 kg N/ha/yr, and tree nuts equal 25 kg N/ha/yr.

The coefficient on pasture in column 1 equals 1.3, meaning that a 10 percentage point increase in the share of pasture is linked with a 13.9% rise in nitrate concentrations relative to undeveloped land. From column 2, a 10% increase in the share of pasture is linked with a 7.6% rise in nitrates. The large coefficients on pasture may be due to the inclusion of grass harvested for hay and silage in the Cropland Data Layer's pasture and grasslands category. Growers typically fertilize harvested grasslands with synthetic fertilizer and manure if the field is located close to a dairy (Ransom et al., 2018). Consequently, excess nitrogen compounds can leach below the shallow root zone and into groundwater if they surpass the grass crop's nutrient requirements. Indeed, Ransom et al. (2018) find that nitrate emissions from manure fertilized forage crops in the Central Valley equal 46 kg N/Ha/yr, roughly equivalent to emissions from some high-NHI crops like vegetables and berries. Nonetheless, researchers postulate that nitrogen deposits onto pastures from grazing livestock and synthetic fertilizer do not typically exceed the nitrogen requirements of the grass crop (Rosenstock et al., 2014). Furthermore, Ransom et al. (2018) observed low nitrogen emissions below the pasture root zone, which were comparable to emissions from natural vegetation.

Increasing the cattle population within 1 km of a well by 1,000, which falls short of the average dairy herd size in California, is associated with a 16–20% increase in nitrate concentrations. As dairy operations became increasingly concentrated in California, especially through the 1990s, manure production in Central Valley counties has exceeded the acreage available for application (Kellogg et al., 2000). Thus, dairy managers face a choice between applying excess nutrients to suitable crops or costly processing, such as drying manure, to transport it to other areas.[14] We find no meaningful effect of dairies located further (1–5 km) from wells. The evidence from prior literature on the extent to which cattle populations contribute to nitrate contamination of groundwater is inconclusive. On the one hand, Ransom et al. (2018) find that nitrate emissions from manure-fertilized forage are twice as large as emissions from equivalent fields treated with synthetic fertilizer, and Harter et al. (2002) show that cattle housing and manure lagoons act as point sources of nitrates. On the other hand, using a sample of Central Valley wells, Lockhart, King, and Harter (2013) find no clear correlation between the presence of a dairy farm within 2.4 km of wells and nitrate concentrations. However, Lockhart, King, and Harter (2013) focus on correlations and do not control for biophysical features like depth to water or soils. Using regression models that include biophysical and land uses at the county level in Maryland, Lichtenberg and Shapiro (1997) find no statistically significant association between dairy cattle and nitrate concentrations. However, typical dairy production in Maryland in the early 1990s involved small pasture-based herds rather than concentrated feeding operations that dominate in California (Somerville et al., 2020).

The positive coefficient on fallow in column 2 might be explained by the fact that without a crop to use residual nutrients, nitrates leach deeper into the soil and below the root zone. Results of field experiments by Bauder, Sinclair, and Lund (1993), Campbell et al. (2006), and John et al. (2017) reveal positive associations between fallow cropland and nitrate emissions into groundwater. However, the coefficient on fallow is not statistically significant.

Urban developments are positively and significantly associated with nitrate concentrations in the first two specifications, although the differences in the estimated effect between the two types of development considered are modest. The coefficients from column 2 indicate that a 10 percentage point increase in low-intensity (resp., high-intensity) development is associated with a 10.1% (resp., 10.5%) increase in nitrate concentrations. Some key ni-

[14]Since 2013, the Irrigated Lands Regulatory Program (ILRP) has required dairies to submit an annual nutrient management plan to the Central Valley Water Control Board. The ILRP technical standards for nutrient management stipulate that dairies should not apply nitrogen in excess of 1.4 times the nitrogen removed from the field in the harvested portion of the crop.

trate sources in urban landscapes include septic and sewage systems, gardens, and parklands (Lichtenberg and Shapiro, 1997; Pennino, Compton, and Leibowitz, 2017; Ransom et al., 2018).

Nitrate contamination of groundwater through leaching is a cumulative process, characterized by potentially long time lags between the absorption of nutrients in the topsoil and their reaching of the water table as well as the persistence of contaminants in groundwater over time. To understand the extent to which land use may affect nitrate concentrations over a twelve-year period, we include initial concentrations in our analysis and present the regression estimates in columns 3–4 of Table 3. In general, these specifications show that the land use coefficients are an order of magnitude smaller compared to columns 1–2 and insignificant even at the 10% level, indicating that land use has little relationship to changes in nitrate concentrations twelve years later. Point estimates for low-NHI crops and fallow are two exceptions that are weakly significant. In column 3, the coefficient on low-NHI crops is -0.23, meaning that a 10% increase in the share of land used for low-NHI crops leads to a 2.3% drop in nitrate concentrations twelve years later relative to undeveloped land. However, the estimate on low-NHI crops decreases in magnitude and loses significance when controlling for sub-basin fixed effects. Column 4 reveals a relatively large coefficient on fallow (0.48), but this estimate is only significant at the 10% level.

4.2 Controls

Our regressions include biophysical controls that encompass some of the most important factors affecting nitrate concentrations in groundwater (Kolpin, 1997; Castaldo et al., 2021). However, perhaps due to low variation within sub-basins for some variables and correlation with other regressors, some control coefficients are insignificant. Additionally, some exhibit unexpected signs, which we discuss in the following with reference to the underlying biophysical mechanisms.

Water from irrigation and precipitation may transport nitrates through the soil and into groundwater. However, important differences exist between irrigation and precipitation in terms of timing, quantity, and location of water supplied to the landscape, impacting the estimated coefficients. Irrigation water supply in excess of evapotranspiration is the primary source of groundwater recharge in the semi-arid Central Valley (Burow, Shelton, and Dubrovsky, 2008) and contributes to leaching of nitrates to groundwater (Botros et al., 2012). The coefficients on surface water deliveries are positive, as expected, in columns 1 and 2. Controlling for initial nitrates, the coefficients become trivially small and close to zero. In contrast, precipitation is liable to surface runoff, diverting nitrates from groundwater toward surface waters, recharging groundwater below undeveloped land, and diluting nitrates in groundwater (Ransom et al., 2017; Abascal et al., 2022). In models without sub-basin fixed effects, precipitation exhibits a significant negative association with nitrate concentrations, whether initial nitrates are included as a covariate or not. However, once sub-basin fixed effects are included, the association becomes insignificant, likely due to the low variation in precipitation within sub-basins.

Across all specifications, we find positive coefficients on depth to groundwater that are not statistically different from zero. Deeper groundwater is expected to have lower nitrates because deep aquifers tend to contain a larger share of older water that entered the water table before historic increases in nitrate emissions from intensive agricultural production (Castaldo et al., 2021). Our positive estimates are inconsistent with earlier works that reveal a negative association between depth to groundwater and nitrate concentrations (Lichtenberg and Shapiro, 1997; Castaldo et al., 2021). However, the large standard errors suggest that depth to groundwater has little importance in this setting after controlling for other biophysical factors.

Subsurface drains intercept and divert nitrates away from groundwater (Kladivko et al., 1991). The negative and significant coefficient on the share of land with subsurface drainage corroborates the findings of Pennino, Compton, and Leibowitz (2017) showing a lower probability of wells exceeding the maximum threshold for nitrate concentrations as the share of land with drainage increased.

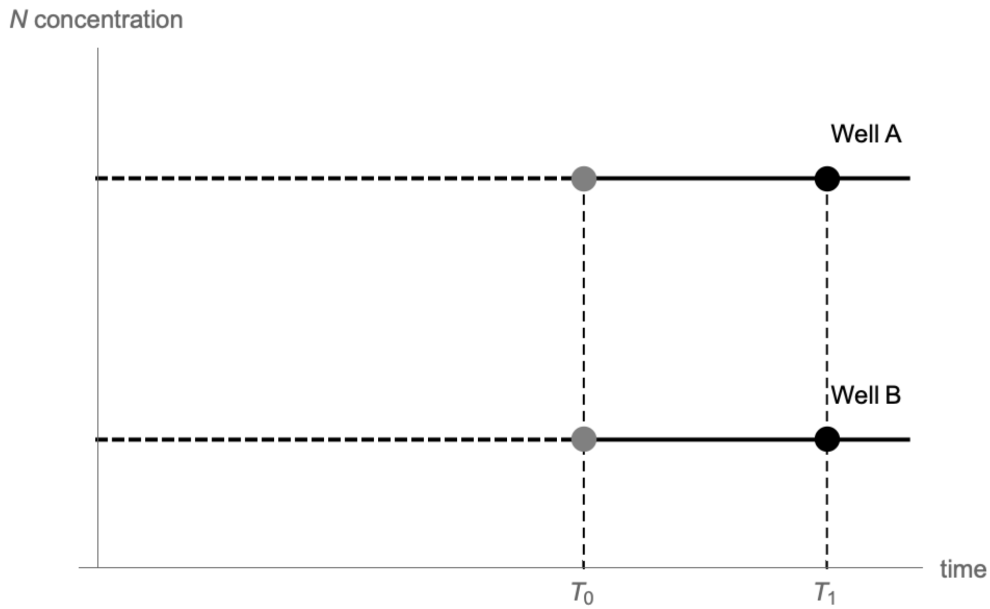
The relative proportions of sand, silt, and clay describe the soil textural fraction. Clay is relatively impermeable compared to sand and silt and can chemically bind nitrates. Therefore, *a priori*, we expect the coefficients on sand and silt to be positive in columns 1 and 2 of Table 3. We find mostly negative coefficients (except for sand in column 1) that are not significantly different from zero. Given that soil texture does not change over time and affects the initial nitrate concentrations, one may not expect sand and silt to have any explanatory power in columns 3 and 4. Nonetheless, we find negative coefficients on sand and silt that are significant at the 5% level in column 4.

Our results reveal negative coefficients on soil organic matter, consistent with previous works that find lower nitrate concentrations in regions with high organic matter (Spalding and Exner, 1993; Pennino, Compton, and Leibowitz, 2017). Organic matter percentage is indicative of high carbon content and oxidizing potential that encourages processes to convert nitrates into nitrogen gas through denitrification (Spalding and Exner, 1993). Finally, the positive coefficients on the distance to rivers are consistent with Castaldo et al. (2021), who find that infiltration of river water into the aquifer has a meaningful dilution effect on nitrate concentrations.

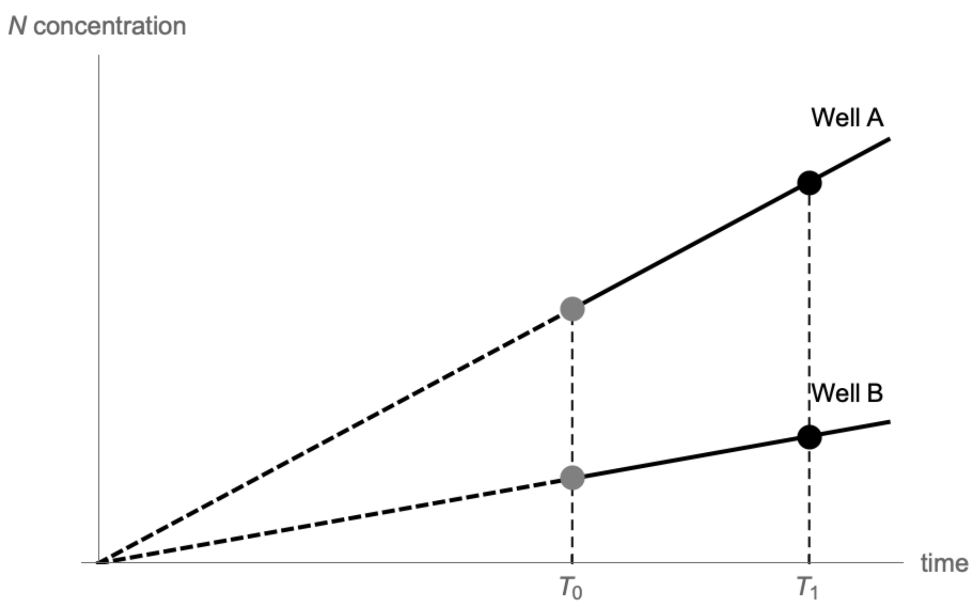
4.3 Initial Nitrates

The coefficient on initial nitrate concentrations in column 3 is estimated to be one, indicating that a 1% increase in mean 2007–2011 nitrate concentrations is associated with a 1% increase in 2019–2023 concentrations, conditional on land use and biophysical factors. Introducing sub-basin fixed effects in column 4, the coefficient slightly decreases to 0.98, with a 95% confidence interval that includes one. Another striking feature of the regressions reported in columns 3 and 4 of Table 3 is their relatively large R-squared, 0.77 and 0.79, respectively.

Given that our model involves the logarithm of nitrate concentration, as opposed to the concentration level, a coefficient on initial nitrates equal to one, controlling for features



(a)



(b)

Figure 3: Possible Evolution of Nitrate Concentrations Over Time

that determine the flow of nitrate pollutants, could be consistent with constant nitrate concentrations within the time frame considered, as depicted in panel (a) of Figure 3 for two wells, A and B, whose concentrations are observed at two time periods, T_0 and T_1. Constant nitrate concentrations over time could be obtained if the twelve-year period considered was too short for nitrates leached during the initial period to reach the water table (which would explain why the coefficients on land use are not statistically significant) and if there was negligible attenuation of the stock of nitrates initially present in groundwater over a decade or so. The absence of attenuation would align with research by Landon et al. (2011) showing that denitrification—the process by which natural processes convert nitrates to nitrogen gas—has minimal impact on nitrate concentrations in the San Joaquin Valley. Similarly, other attenuation mechanisms, such as the assimilation of nitrates into microbial biomass, appear to have a limited effect in reducing nitrate levels in this context (Rivett et al., 2008).

However, the summary statistics in Table 2 indicate that, on average, nitrate pollution has worsened during the period. Panel (b) of Figure 3 depicts nitrate accumulation pathways for two wells that are driven by stationary nitrate emissions, resulting in linear trajectories, where the slope of the trajectory is a direct function of land use and biophysical characteristics at each well. Because the relative difference in concentrations between wells is constant over time and our cross-sectional approach essentially compares wells subject to different conditions, the process depicted in panel (b) of Figure 3 is consistent with our findings that (i) nitrate concentrations have increased during the period of investigation, (ii) the coefficient on initial concentrations is close to one, (iii) land use and biophysical factors have no explanatory power once initial concentrations are controlled for, and (iv) the explanatory power of the regression with initial concentrations is large.

Of course, linear accumulation of nitrates in groundwater over time would require the leaching process to be stationary, but our data suggest that this assumption may actually be plausible. Table 4 reports the R-squared of a series of panel regressions of land use shares on well fixed effects. Observations represent 5-year averages of land uses during the recent and initial periods, therefore there are two time periods per panel. The R-squared measures the proportion of variation in land use shares that is due to cross-sectional differences, the rest being attributable to variations over time. The large R-squared values reported in the table indicate that most of the variation in land use shares is cross-sectional, suggesting that the assumption of stationarity may be justified in our context.[15] In that case, initial nitrate concentrations would act as a sufficient statistic for land use patterns, resulting in small and insignificant effects of land use variables.

Table 4: Land Use Shares Regressed on Well Fixed Effects

	Dependent variable:						
	Low NHI Crops	High NHI Crops	Fallow	Pasture	Low Int. Devel.	High Int. Devel.	Undevel.
Obs.	12,032	12,032	12,032	12,032	12,032	12,032	12,032
R^2	0.92196	0.94098	0.82213	0.88985	0.94807	0.98056	0.93987

Note: Table shows the R^2 of a two-period panel of land use shares regressed on well fixed effects using mean land use shares in 2007–2011 and 2019–2023.

These considerations point to a critical trade-off inherent in observational studies seeking to identify the contributions of land use to groundwater contamination. On the one hand, clean identification of land use impacts requires controlling for initial concentrations. On the other hand, because land uses change slowly over time, land use determinants of incremental contamination may be highly correlated with those that explain initial concentrations, resulting in *de facto* collinearity and the impossibility of identifying land use contributions conditional on initial conditions. Under that interpretation, regressions that omit initial concentrations, such as those reported in columns 1–2 of Table 3, indicate long-run land use impacts, provided that the land use shares used as explanatory variables proxy for long-run patterns.

4.4 Sensitivity Analysis

To assess the sensitivity of the results to our assumptions regarding the nitrate travel times and nitrate concentrations below the detection limit, we employ three alternative data strategies. First, we double the buffer zone radius from 500 meters to one kilometer. Second, we halve the imputed lower detection limit for well observations with nitrate concentrations be-

[15] Since we only have one observation of cattle populations per well, we cannot test for stationarity of cattle populations.

low the lower detection limit. Third, we regress nitrate concentrations on contemporaneous land use shares.

An immediate concern is that our preferred 500-meter buffer is too narrow. We thus construct a 1 km buffer zone and calculate the mean land use shares, soil characteristics, and drainage share using similar spatial methods as those described in the data section. Summary statistics are provided in Appendix Table A.2 and reveal slightly higher mean 2007–2011 land use shares within 1 km of wells used for low- and high-NHI crops, fallow and undeveloped land, and a lower share of land used for urban development.

We report the results using the 1 km buffer zone in Table 5, which has the same structure as Table 3. Estimates are similar in magnitude and significance to those in Table 3, with some exceptions. Column 3 reveals significant coefficients on high-NHI crops, pasture, and low-intensity urban development at the 5% level, providing some evidence that these land use shares are associated with increased nitrate concentrations twelve years later. By including sub-basin fixed effects, as in column 4, we find a positive coefficient on low-intensity development that is significant at the 10% level.

In the second sensitivity test, we follow Keiser and Shapiro (2019a) and halve nitrate concentrations previously set at the detection limit for concentrations below the laboratory detection limit. Summary statistics of these data are provided in Appendix Table A.3 and show small decreases in mean nitrate concentrations compared to Table 2. Results of the regressions using these revised data are provided in Table 6. The results reveal no notable changes in the magnitude or significance of the regression coefficients compared to the estimates in Table 3.

In a final sensitivity test, we regress nitrate concentrations on contemporaneous (2019–2023) land use shares and present the results in Appendix Table A.4. To the extent that land use shares in regressions without initial nitrate concentrations proxy for long-run patterns, contemporaneous land use likely provides an inferior approximation of land use around the time of emissions of nitrates compared to land use measured approximately 12 years prior.

Table 5: Impacts of Mean 2007–2011 Land Use Shares Within 1 km of Wells on Mean 2019–2023 Nitrate Concentrations

	Dependent variable: Log nitrate concentration 2019–2023			
	(1)	(2)	(3)	(4)
Low-NHI crops	0.61**	0.52*	-0.19*	0.05
	(0.26)	(0.32)	(0.10)	(0.14)
High-NHI crops	2.0***	1.5***	0.31**	0.24
	(0.31)	(0.29)	(0.13)	(0.15)
Fallow	-0.88	0.84	-0.13	0.38
	(0.64)	(0.83)	(0.18)	(0.34)
Pasture	1.3***	0.80***	0.14**	0.11
	(0.18)	(0.26)	(0.07)	(0.10)
Low-intensity development	1.5***	1.1***	0.18**	0.15*
	(0.24)	(0.18)	(0.09)	(0.09)
High-intensity development	1.4***	1.0***	0.07	0.15
	(0.26)	(0.27)	(0.07)	(0.11)
Cattle within 1km	0.17**	0.15**	0.02	0.010
	(0.07)	(0.08)	(0.07)	(0.07)
Cattle within 1–5km	0.001	-0.01	0.002	0.001
	(0.02)	(0.02)	(0.009)	(0.01)
Surface water deliveries	0.05*	0.05***	-0.0006	-9.7×10^{-5}
	(0.03)	(0.004)	(0.003)	(0.001)
Precipitation	-0.53***	-0.05	-0.19***	0.04
	(0.14)	(0.38)	(0.05)	(0.13)
Depth to groundwater	0.001	-0.004	0.0008	0.006
	(0.01)	(0.02)	(0.004)	(0.005)
Drainage	-2.0***	-1.5*	-0.40*	-0.50***
	(0.48)	(0.86)	(0.21)	(0.17)
Sand	0.40	-0.19	-0.12	-0.38*
	(0.35)	(0.40)	(0.18)	(0.23)
Silt	-0.25	-0.70	-0.46	-0.82**
	(0.64)	(0.77)	(0.32)	(0.41)
Organic matter	-6.7***	-6.7***	-0.74	-0.64
	(2.1)	(2.5)	(0.74)	(0.66)
Distance to river	0.25	0.51	0.17**	0.16*
	(0.27)	(0.34)	(0.07)	(0.09)
Initial nitrate concentration			1.0***	0.98***
			(0.02)	(0.02)
Sub-basin FE	No	Yes	No	Yes
Observations	6,016	6,016	6,016	6,016
R^2	0.14928	0.32807	0.77251	0.79431

Note: Undeveloped land is the default land use and clay is the default soil textural fraction. Initial nitrate concentration equals log nitrate concentration 2007–2011. Standard errors are clustered by sub-basin * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Impacts of Mean 2007–2011 Land Use Shares Within 500 meters of Wells on Mean 2019–2023 Nitrate Concentrations Using Nitrate Concentrations Equal to Half the Detection Limit for Concentrations Below the Detection Limit

	Dependent variable: Log nitrate concentration 2019–2023			
	(1)	(2)	(3)	(4)
Low-NHI crops	0.93***	0.72**	-0.31***	-0.03
	(0.30)	(0.34)	(0.11)	(0.14)
High-NHI crops	2.0***	1.2***	0.20	0.16
	(0.40)	(0.37)	(0.15)	(0.16)
Fallow	-0.58	0.92	-0.01	0.53*
	(0.79)	(0.91)	(0.21)	(0.32)
Pasture	1.6***	0.92***	0.09	0.09
	(0.20)	(0.29)	(0.08)	(0.11)
Low-intensity development	1.7***	1.2***	0.14	0.14*
	(0.22)	(0.18)	(0.08)	(0.08)
High-intensity development	1.7***	1.2***	0.06	0.16*
	(0.26)	(0.26)	(0.08)	(0.09)
Cattle within 1km	0.21**	0.17*	0.01	0.007
	(0.09)	(0.10)	(0.07)	(0.07)
Cattle within 1–5km	0.006	-0.01	0.001	-0.0002
	(0.02)	(0.02)	(0.01)	(0.01)
Surface water deliveries	0.05*	0.06***	0.003	0.004
	(0.03)	(0.007)	(0.004)	(0.004)
Precipitation	-0.71***	-0.03	-0.25***	0.03
	(0.16)	(0.46)	(0.06)	(0.15)
Depth to groundwater	0.010	0.005	0.002	0.007
	(0.02)	(0.02)	(0.005)	(0.006)
Drainage	-2.1***	-1.1	-0.40	-0.37*
	(0.42)	(0.69)	(0.25)	(0.21)
Sand	0.53	-0.23	-0.13	-0.39**
	(0.40)	(0.37)	(0.18)	(0.19)
Silt	-0.005	-0.68	-0.43	-0.78**
	(0.66)	(0.68)	(0.30)	(0.33)
Organic matter	-6.9***	-5.4**	-0.06	-0.13
	(1.9)	(2.3)	(0.55)	(0.64)
Distance to river	0.27	0.57	0.20***	0.20**
	(0.31)	(0.38)	(0.07)	(0.09)
Initial nitrate concentration			0.97***	0.96***
			(0.02)	(0.02)
Sub-basin FE	No	Yes	No	Yes
Observations	6,016	6,016	6,016	6,016
R^2	0.14863	0.33352	0.78150	0.80274

Note: Undeveloped land is the default land use and clay is the default soil textural fraction. Initial nitrate concentration equals log nitrate concentration 2007–2011. Standard errors are clustered by sub-basin * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Consistent with this expectation, we find that the land use coefficients in columns 1–2 of Table A.4 are smaller in magnitude compared to those in Table 3, with a corresponding loss in statistical significance on the low-NHI crop coefficients. Controlling for initial nitrates, as in columns 3–4, we also find a decrease in the magnitude of most coefficients. The coefficients on low-NHI crops in columns 3 and 4 and fallow in column 3 are exceptions. Compared to Table 3, the coefficients on the control variables in Table A.4 do not change much, and the R-squared values are very similar.

5 Concluding Remarks

In this study, we examine the relationship between land use activity and groundwater nitrate concentrations using reduced-form regressions applied to a dataset of 6,016 groundwater wells in California. Our findings indicate significant associations between nitrate concentrations and agricultural and urban land use shares. The most pronounced effects are observed in areas with high-NHI (Nitrogen Hazard Index) crops, where a 10 percentage point increase in land use share is associated with an 11.6% rise in nitrate concentrations relative to undeveloped land. Similarly, urban land shares contribute to a 10.1% and 10.5% increase in nitrate concentrations for low- and high-intensity urban developments, respectively, for every 10 percentage point increase in land use share. These significant associations underscore the importance of directing groundwater quality programs toward emissions from both agricultural and urban land uses in regions where wells exhibit poor or deteriorating water quality. Focusing solely on agricultural emissions could overlook cost-effective opportunities to improve well water quality through better management of urban developments.

A key challenge uncovered by our analysis is the difficulty of identifying the impact of land use on nitrate concentrations a decade or so later using observational data. Our results demonstrate that initial nitrate concentrations almost entirely explain the observed nitrate concentrations, likely due to limited temporal variation in land use surrounding wells over time, at least within the temporal range of our data. That is, land use patterns that explain nitrate concentrations at the beginning of our period are likely close to those we use to explain subsequent contamination, so that conditioning on initial concentrations renders our land use variables redundant, explaining the overall decrease in size and statistical significance for their estimated effects, as well as the near-unit coefficient estimate on initial concentrations. Addressing this empirical challenge may require a multi-decade study relating long-term variation in land use on nitrate concentrations, or a focus on shallow wells, although current data constraints prevent such analyses.

Importantly, the near-unit coefficient estimate on initial concentrations should not be construed as implying that the environment exhibits no capacity to lose nitrates over a 12-year period through, for example, conversion to other nitrogen compounds. Indeed, the fact that nitrates tend to increase in proportion to their initial levels during the study period could simply mean that nutrient accrual through leaching and nutrient disappearance both occur in proportion to initial concentrations, since we only observe net effects. Nutrient accrual would occur in proportion to initial concentrations if the nitrate emissions process were stationary and contamination started at the same time across sample wells. Nutrient disappearance would occur in proportion to initial concentrations if a fixed share of nitrates present at a well were lost due to horizontal diffusion, for instance.

Interpreting our land use variables as capturing a stationary source of nitrates, that is, not conditioning on initial concentrations, we are able to document meaningfully smaller coefficients on low-NHI crops compared to high-NHI crops, consistent with the definition of the Nitrogen Hazard Index. To our knowledge, this is the first study to apply the NHI within a regression analysis, offering a novel approach to aggregating crops based on characteristics that influence nitrate emissions below the root zone. The NHI provides a parsimonious method for researchers to relate land use to nitrogen pollution, though further research is needed to compare the performance of NHI-based estimates against other crop aggregation methods commonly used in economics, such as those based on nitrogen fertilizer application rates or traditional crop classifications (e.g., tree nuts, vegetables, and row crops). Our regression analysis also confirms the role of dairy cattle as a meaningful determinant of nitrate concentrations.

This analysis also underscores two important observations for economic researchers. First, while a growing economics literature quantifies the reduced-form effects of land use or policy intervention on *surface water* bodies (Keiser and Shapiro, 2019a; Paudel and Crago, 2021; Metaxoglou and Smith, 2024), parallel analyses have not been widely applied to U.S. groundwater. Here, we provide evidence of the groundwater quality consequences of diverse land uses. We also show, however, that there are inherent empirical challenges in measuring the fate of nitrate pollution across the landscape due to the gradual nature of nitrate leaching. Second, regulatory attempts to rectify the anthropogenic contribution to groundwater nitrate may not be identifiable for decades. To that end, the valuation of groundwater quality degradation (or improvements) must account for the discounted payoff of land use externalities many years into the future.

References

Abascal, E., L. Gómez-Coma, I. Ortiz, and A. Ortiz. 2022. "Global diagnosis of nitrate pollution in groundwater and review of removal technologies." *Science of the Total Environment* 810, 152233, https://doi.org/10.1016/j.scitotenv.2021.152233.

Bauder, J., K. Sinclair, and R. Lund. 1993. "Physiographic and Land Use Characteristics Associated with Nitrate-Nitrogen in Montana Groundwater." *Journal of Environmental Quality* 22:255–262, https://doi.org/10.2134/jeq1993.00472425002200020004x.

Beaudette, D., and A. O'Geen. 2009. "Soil-Web: an online soil survey for California, Arizona, and Nevada." *Computers & Geosciences* 35(10):2119–2128, https://doi.org/10.1016/j.cageo.2008.10.016.

Botros, F.E., Y.S. Onsoy, T.R. Ginn, and T. Harter. 2012. "Richards Equation–Based Modeling to Estimate Flow and Nitrate Transport in a Deep Alluvial Vadose Zone." *Vadose Zone Journal* 11(4), vzj2011.0145, https://doi.org/10.2136/vzj2011.0145.

Boyle, D., A. King, G. Kourakos, K. Lockhart, M. Mayzelle, G.E. Fogg, and T. Harter. 2012. *Groundwater Nitrate Occurrence, Technical Report 4: Addressing Nitrate in California's Drinking Water with a Focus on Tulare Lake Basin and Salinas Valley Groundwater. Report for the State Water Resources Control Board Report to the Legislature*. Center for Watershed Sciences, University of California, Davis. 78 p, https://groundwaternitrate.ucdavis.edu/files/139106.pdf.

Burkart, M.R., and J.D. Stoner. 2008. "Nitrogen in Groundwater Associated with Agricultural Systems." In J. Hartfield and R. Follett, eds. *Nitrogen in the Environment: Sources, Problems, and Management*. Academic Press, 2nd ed., pp. 177–202, https://doi.org/10.1016/B978-0-12-374347-3.00007-X.

Burow, K.R., B.C. Jurgens, K. Belitz, and N.M. Dubrovsky. 2013. "Assessment of regional change in nitrate concentrations in groundwater in the Central Valley, California, USA, 1950s–2000s." *Environmental Earth Sciences* 69:2609–2621, https://doi.org/10.1007/s12665-012-2082-4.

Burow, K.R., J.L. Shelton, and N.M. Dubrovsky. 2008. "Regional Nitrate and Pesticide Trends in Ground Water in the Eastern San Joaquin Valley, California." *Journal of Environmental Quality* 37:S–249–S–263, https://doi.org/10.2134/jeq2007.0061.

California Department of Health Services. 2000. *Drinking Water Source Assessment and Protection (DWSAP) Program*. Santa Rosa, January. https:

//www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/documents/dwsapguidance/DWSAP_document.pdf (accessed September 27th, 2024).

California Department of Water Resources. 2019. *California Water Plan Update 2018: Managing Water Resources for Sustainability*. Sacramento, June. https://cawaterlibrary.net/document/california-water-plan-update-2018/.

—. 2021. *California's Groundwater Update 2020*. Bulletin 118, Sacramento, November. https://water.ca.gov/programs/groundwater-management/bulletin-118.

—. 2024. *Monitoring Well Construction*. https://water.ca.gov/Programs/Groundwater-Management/Wells/Well-Standards/Combined-Well-Standards/Monitoring-General (accessed September 27th, 2024).

—. 2023. *NHD Major Rivers and Creeks Geodatabase*. [Data file and code book]. Retrieved from https://data.cnra.ca.gov/dataset/511528b2-f7d3-4d86-8902-cc9befeeeed5/resource/fea1bd34-49b6-4ade-91d2-bb97d7767928/download/nhd_major_rivers_and_creeks.gdb.zip (accessed April 7th, 2023).

California State Water Resources Control Board. 2022. "Confined Animal Facilities Regulated by the Water Boards." [Data file and code book]. Retrieved from https://data.ca.gov/dataset/surface-water-water-quality-regulated-facility-information (accessed January 3rd, 2023).

California Water Boards. 2024a. *GAMA Groundwater Information System*. [Data file and code book]. Retrieved from https://gamagroundwater.waterboards.ca.gov/gama/datadownload (accessed March 15th, 2024).

—. 2024b. *Irrigation and Nitrogen Management Plan (INMP) & INMP Summary Report*. https://www.waterboards.ca.gov/centralvalley/water_issues/irrigated_lands/annrpt_brochure.pdf (accessed September 30th, 2024).

California Water Boards, Central Coast Regional Water Quality Control Board. 2013. *Fact Sheet: Nitrate/Nitrite in Drinking Water*. San Luis Obispo. https://www.waterboards.ca.gov/centralcoast/water_issues/programs/gap/docs/salinas_pajro_valley_proj_2012-2013/gap_nitrate_fact_sheet_121613.pdf (accessed September 27th, 2024).

Campbell, C., F. Selles, R. Zentner, R. De Jong, R. Lemke, and C. Hamel. 2006. "Nitrate leaching in the semiarid prairie: Effect of cropping frequency, crop type, and fertilizer after

37 years." *Canadian Journal of Soil Science* 86:701–710, https://doi.org/10.4141/S05-008.

Castaldo, G., A. Visser, G.E. Fogg, and T. Harter. 2021. "Effect of Groundwater Age and Recharge Source on Nitrate Concentrations in Domestic Wells in the San Joaquin Valley." *Environmental Science & Technology* 55:2265–2275, https://doi.org/10.1021/acs.est.0c03071.

Deschênes, O., and M. Greenstone. 2007. "The Economic Impacts of Climate Change: Evidence From Agricultural Output and Random Fluctuations in Weather." *American Economic Review* 97(1):354–385, https://www.aeaweb.org/articles?id=10.1257/aer.97.1.354.

DeSimone, L.A., P.B. McMahon, and M.R. Rosen. 2014. *The Quality of Our Nation's Waters–Water Quality in Principal Aquifers of the United States, 1991–2010*. U.S. Geological Survey, Circular 1360, https://dx.doi.org/10.3133/cir1360.

Dzurella, K., G. Pettygrove, A. Fryjoff-Hung, A. Hollander, and T. Harter. 2015. "Potential to assess nitrate leaching vulnerability of irrigated cropland." *Journal of Soil and Water Conservation* 70:63–72, https://doi.org/10.2489/jswc.70.1.63.

Fewtrell, L. 2004. "Drinking-Water Nitrate, Methemoglobinemia, and Global Burden of Disease: A Discussion." *Environmental Health Perspectives* 112:1371–1374, https://doi.org/10.1289/ehp.7216.

Galloway, J.N., A.R. Townsend, J.W. Erisman, M. Bekunda, Z. Cai, J.R. Freney, L.A. Martinelli, S.P. Seitzinger, and M.A. Sutton. 2008. "Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions." *Science* 320:889–892, https://doi.org/10.1126/science.1136674.

Gammans, M., P. Mérel, and A. Ortiz-Bobea. 2024. "Double cropping as an adaptation to climate change in the United States." *American Journal of Agricultural Economics*, pp. 1–26, https://doi.org/10.1111/ajae.12491.

Hagerty, N. 2022. "Adaptation to Surface Water Scarcity in Irrigated Agriculture." Working paper, Department of Agricultural Economics & Economics, Montana State University, https://hagertynw.github.io/webfiles/Surface_Water_Adaptation.pdf.

Harter, T. 2009. "Agricultural Impacts on Groundwater Nitrate." *Southwest Hydrology* 8(4):22–23.

—. 2015. "California's agricultural regions gear up to actively manage groundwater use and protection." *California Agriculture* 69:193–201, https://doi.org/10.3733/ca.E.v069n03p193.

—. 2002. *Delineating Groundwater Sources and Protection Zones*. Sacramento, April, https://ucanr.edu/sites/groundwater/files/136261.pdf.

Harter, T., H. Davis, M.C. Mathews, and R.D. Meyer. 2002. "Shallow groundwater quality on dairy farms with irrigated forage." *Journal of Contaminant Hydrology* 55:287–315, https://doi.org/10.1016/S0169-7722(01)00189-9.

Harter, T., J.R. Lund, J. Darby, G.E. Fogg, R. Howitt, K.K. Jessoe, G.S. Pettygrove, J.F. Quinn, J.H. Viers, D.B. Boyle, H.E. Canada, N. DeLaMora, K.N. Dzurella, A. Fryjoff-Hung, A.D. Hollander, K.L. Honeycutt, M.W. Jenkins, V.B. Jensen, A.M. King, G. Kourakos, D. Liptzin, E.M. Lopez, M.M. Mayzelle, A. McNally, J. MedellinAzuara, and T.S. Rosenstock. 2012. *Addressing Nitrate in California's Drinking Water with a Focus on Tulare Lake Basin and Salinas Valley Groundwater. Report for the State Water Resources Control Board Report to the Legislature*. Center for Watershed Sciences, University of California, Davis. 78 p.

Hendricks, N.P., S. Sinnathamby, K. Douglas-Mankin, A. Smith, D.A. Sumner, and D.H. Earnhart. 2014. "The environmental effects of crop price increases: Nitrogen losses in the U.S. Corn Belt." *Journal of Environmental Economics and Management* 68:507–526, https://doi.org/10.1016/j.jeem.2014.09.002.

Hillel, D. 2008. *Soil in the Environment*. Burlington, MA: Academic Press.

Inoue-Choi, M., R.R. Jones, K.E. Anderson, K.P. Cantor, J.R. Cerhan, S. Krasner, K. Robien, P.J. Weyer, and M.H. Ward. 2015. "Nitrate and nitrite ingestion and risk of ovarian cancer among postmenopausal women in Iowa." *International Journal of Cancer* 137:173–182, https://doi.org/10.1002/ijc.29365.

Isbell, F., D. Tilman, S. Polasky, S. Binder, and P. Hawthorne. 2013. "Low biodiversity state persists two decades after cessation of nutrient enrichment." *Ecology Letters* 16:454–460, https://doi.org/10.1111/ele.12066.

John, A.A., C.A. Jones, S.A. Ewing, W.A. Sigler, A. Bekkerman, and P.R. Miller. 2017. "Fallow replacement and alternative nitrogen management for reducing nitrate leaching in a semiarid region." *Nutrient Cycling in Agroecosystems* 108:279–296, https://doi.org/10.1007/s10705-017-9855-9.

Johnson, T.D., and K. Belitz. 2009. "Assigning land use to supply wells for the statistical characterization of regional groundwater quality: Correlating urban land use and VOC occurrence." *Journal of Hydrology* 370:100–108, https://doi.org/10.1016/j.jhydrol.2009.02.056.

Jones, E.R., M.F.P. Bierkens, P.J.T.M. van Puijenbroek, L.P.H. van Beek, N. Wanders, E.H. Sutanudjaja, and M.T.H. van Vliet. 2023. "Sub-Saharan Africa will increasingly become the dominant hotspot of surface water pollution." *Nature Water* 1:602–613, https://doi.org/10.1038/s44221-023-00105-5.

Karwoski, N., and M. Skidmore. 2024. *Nature's Kidneys: the Role of the Wetland Reserve Easements in Restoring Water Quality*. Working paper. Available online at https://drive.google.com/file/d/1HzJ7EYEXVwWj-K2GtmzKcV44lR2oDvj0/view (accessed October 7th, 2024).

Keeler, B.L., J.D. Gourevitch, S. Polasky, F. Isbell, C.W. Tessum, J.D. Hill, and J.D. Marshall. 2016. "The social costs of nitrogen." *Science Advances* 2(10), e1600219, https://www.science.org/doi/abs/10.1126/sciadv.1600219.

Keiser, D.A., and J.S. Shapiro. 2019a. "Consequences of the Clean Water Act and the Demand for Water Quality." *The Quarterly Journal of Economics* 134:349–396, https://doi.org/10.1093/qje/qjy019.

—. 2019b. "US Water Pollution Regulation Over the Past Half Century: Burning Waters to Crystal Springs?" *Journal of Economic Perspectives* 33(4):51–75, https://www.aeaweb.org/articles?id=10.1257/jep.33.4.51.

Kellogg, R.L., C.H. Lander, D.C. Moffitt, and N. Gollehon. 2000. *Manure Nutrients Relative to the Capacity of Cropland and Pastureland to Assimilate Nutrients: Spatial and Temporal Trends for the United States*. Washington DC: U.S. Department of Agriculture, December.

Kladivko, E., G. Van Scoyoc, E. Monke, K. Oates, and W. Pask. 1991. "Pesticide and Nutrient Movement into Subsurface Tile Drains on a Silt Loam Soil in Indiana." *Journal of Environmental Quality* 20:264–270, https://doi.org/10.2134/jeq1991.00472425002000010043x.

Kolpin, D.W. 1997. "Agricultural Chemicals in Groundwater of the Midwestern United States: Relations to Land Use." *Journal of Environmental Quality* 26:1025–1037, https://doi.org/10.2134/jeq1997.00472425002600040014x.

Koterba, M.T. 1998. *Ground-Water Data-Collection Protocols and Procedures for the National Water-Quality Assessment Program: Collection, Documentation, and Compilation of Required Site, Well, Subsurface, and Landscape Data for Wells*. Water–Resources Investigations Report 98–4107, Baltimore, MA: U.S. Geological Survey.

Kourakos, G., F. Klein, A. Cortis, and T. Harter. 2012. "A groundwater nonpoint source pollution modeling framework to evaluate long-term dynamics of pollutant exceedance probabilities in wells and other discharge locations." *Water Resources Research* 48, W00L13, https://doi.org/10.1029/2011WR010813.

Landon, M.K., C.T. Green, K. Belitz, M.J. Singleton, and B.K. Esser. 2011. "Relations of hydrogeologic factors, groundwater reduction-oxidation conditions, and temporal and spatial distributions of nitrate, Central-Eastside San Joaquin Valley, California, USA." *Hydrogeology Journal* 19:1203–1224, https://doi.org/10.1007/s10040-011-0750-1.

Lark, T.J., N.P. Hendricks, A. Smith, N. Pates, S.A. Spawn-Lee, M. Bougie, E.G. Booth, C.J. Kucharik, and H.K. Gibbs. 2022. "Environmental Outcomes of the US Renewable Fuel Standard." *Proceedings of the National Academy of Sciences* 119:e2101084119, https://doi.org/10.1073/pnas.2101084119.

Lichtenberg, E., and L.K. Shapiro. 1997. "Agriculture and Nitrate Concentrations in Maryland Community Water System Wells." *Journal of Environmental Quality* 26:145–153, https://doi.org/10.2134/jeq1997.00472425002600010022x.

Liu, P., Y. Wang, and W. Zhang. 2023. "The influence of the Environmental Quality Incentives Program on local water quality." *American Journal of Agricultural Economics* 105:27–51, https://onlinelibrary.wiley.com/doi/abs/10.1111/ajae.12316.

Lockhart, K., A. King, and T. Harter. 2013. "Identifying sources of groundwater nitrate contamination in a large alluvial groundwater basin with highly diversified intensive agricultural production." *Journal of Contaminant Hydrology* 151:140–154, https://doi.org/10.1016/j.jconhyd.2013.05.008.

Lubell, M., W. Blomquist, and L. Beutler. 2020. "Sustainable Groundwater Management in California: A Grand Experiment in Environmental Governance." *Society & Natural Resources* 33:1447–1467, https://doi.org/10.1080/08941920.2020.1833617.

Marklein, A.R., D. Meyer, M.L. Fischer, S. Jeong, T. Rafiq, M. Carr, and F.M. Hopkins. 2021. "Facility-scale inventory of dairy methane emissions in California: implications for mitigation." *Earth System Science Data* 13:1151–1166, https://doi.org/10.5194/essd-13-1151-2021.

McMahon, P.B., J. Böhlke, L. Kauffman, K. Kipp, M. Landon, C. Crandall, K. Burow, and C. Brown. 2008. "Source and transport controls on the movement of nitrate to public supply wells in selected principal aquifers of the United States." *Water Resources Research* 44, W04401, https://doi.org/10.1029/2007WR006252.

Mérel, P., F. Yi, J. Lee, and J. Six. 2014. "A regional bio-economic model of nitrogen use in cropping." *American Journal of Agricultural Economics* 96:67–91, https://doi.org/10.1093/ajae/aat053.

Metaxoglou, K., and A. Smith. 2024. *Agriculture's Nitrogen Legacy*. Working paper. Available online at https://files.asmith.ucdavis.edu/water_draft_legacy.pdf (accessed September 29th, 2024).

Mueller, D.K., and D.R. Helsel. 1996. *Nutrients in the Nation's Waters: Too Much of a Good Thing?*. U.S. Geological Survey, Circular 1136, https://doi.org/10.3133/cir1136.

Nolan, B.T., J.M. Gronberg, C.C. Faunt, S.M. Eberts, and K. Belitz. 2014. "Modeling Nitrate at Domestic and Public-Supply Well Depths in the Central Valley, California." *Environmental Science & Technology* 48:5643–5651, https://doi.org/10.1021/es405452q.

Olmstead, S.M. 2010. "The Economics of Water Quality." *Review of Environmental Economics and Policy* 4:44–62, https://doi.org/10.1093/reep/rep016.

Ortiz-Bobea, A. 2020. "The Role of Nonfarm Influences in Ricardian Estimates of Climate Change Impacts on US Agriculture." *American Journal of Agricultural Economics* 102:934–959, https://doi.org/10.1093/ajae/aaz047.

Paudel, J., and C.L. Crago. 2021. "Environmental Externalities from Agriculture: Evidence from Water Quality in the United States." *American Journal of Agricultural Economics* 103:185–210.

Pennino, M.J., J.E. Compton, and S.G. Leibowitz. 2017. "Trends in Drinking Water Nitrate Violations Across the United States." *Environmental Science & Technology* 51:13450–13460, https://doi.org/10.1021/acs.est.7b04269.

Picetti, R., M. Deeney, S. Pastorino, M.R. Miller, A. Shah, D.A. Leon, A.D. Dangour, and R. Green. 2022. "Nitrate and nitrite contamination in drinking water and cancer risk:

A systematic review with meta-analysis." *Environmental Research* 210, 112988, https://doi.org/10.1016/j.envres.2022.112988.

PRISM Climate Group. 2024. *PRISM Gridded Climate Data*. Oregon State University, [Data file and code book] https://prism.oregonstate.edu (accessed January 5th, 2024).

Rabotyagov, S.S., C.L. Kling, P.W. Gassman, N.N. Rabalais, and R.E. Turner. 2014. "The Economics of Dead Zones: Causes, Impacts, Policy Challenges, and a Model of the Gulf of Mexico Hypoxic Zone." *Review of Environmental Economics and Policy* 8:58–79, https://doi.org/10.1093/reep/ret024.

Raff, Z., and A. Meyer. 2022. "CAFOs and Surface Water Quality: Evidence from Wisconsin." *American Journal of Agricultural Economics* 104:161–189, https://onlinelibrary.wiley.com/doi/abs/10.1111/ajae.12222.

Rahman, A., N.C. Mondal, and K.K. Tiwari. 2021. "Anthropogenic nitrate in groundwater and its health risks in the view of background concentration in a semi arid area of Rajasthan, India." *Scientific Reports* 11, 9279, https://doi.org/10.1038/s41598-021-88600-1.

Ransom, K.M., A.M. Bell, Q.E. Barber, G. Kourakos, and T. Harter. 2018. "A Bayesian approach to infer nitrogen loading rates from crop and land-use types surrounding private wells in the Central Valley, California." *Hydrology and Earth System Sciences* 22:2739–2758, https://doi.org/10.5194/hess-22-2739-2018.

Ransom, K.M., B.T. Nolan, J.A. Traum, C.C. Faunt, A.M. Bell, J.A.M. Gronberg, D.C. Wheeler, C.Z. Rosecrans, B. Jurgens, G.E. Schwarz, et al. 2017. "A hybrid machine learning model to predict and visualize nitrate concentration throughout the Central Valley aquifer, California, USA." *Science of the Total Environment* 601:1160–1172, https://doi.org/10.1016/j.scitotenv.2017.05.192.

Rivett, M.O., S.R. Buss, P. Morgan, J.W. Smith, and C.D. Bemment. 2008. "Nitrate attenuation in groundwater: A review of biogeochemical controlling processes." *Water Research* 42:4215–4232, https://doi.org/10.1016/j.watres.2008.07.020.

Rosenstock, T.S., D. Liptzin, K. Dzurella, A. Fryjoff-Hung, A. Hollander, V. Jensen, A. King, G. Kourakos, A. McNally, G.S. Pettygrove, J. Quinn, J.H. Viers, T.P. Tomich, and T. Harter. 2014. "Agriculture's Contribution to Nitrate Contamination of Californian Groundwater (1945–2005)." *Journal of Environmental Quality* 43:895–907, https://doi.org/10.2134/jeq2013.10.0411.

Schlenker, W., W.M. Hanemann, and A.C. Fisher. 2006. "The Impact of Global Warming on U.S. Agriculture: An Econometric Analysis of Optimal Growing Conditions." *Review of Economics and Statistics* 88:113–125, https://doi.org/10.1162/rest.2006.88.1.113.

—. 2005. "Will U.S. Agriculture Really Benefit from Global Warming? Accounting for Irrigation in the Hedonic Approach." *American Economic Review* 95:395–406.

Somerville, S., D. Sumner, J. Fadel, Z. Fu, J. Hart, and J. Heguy. 2020. "By-Product Use in California Dairy Feed Has Vital Sustainability Implications." *ARE Update* 24:5–8, University of California Giannini Foundation of Agricultural Economics.

Spalding, R.F., and M.E. Exner. 1993. "Occurrence of Nitrate in Groundwater–A Review." *Journal of Environmental Quality* 22:392–402, https://doi.org/10.2134/jeq1993.00472425002200030002x.

Tozer, L. 2023. "Water pollution 'timebomb' threatens global health." *Nature*, July, https://doi.org/10.1038/d41586-023-02337-7 (accessed August 20th, 2023).

U.S. Department of Agriculture. 2022. *California Cattle County Estimates*. Sacramento CA, May, https://www.nass.usda.gov/Statistics_by_State/California/Publications/County_Estimates/2022/CATCNTYE2022.pdf.

—. 2014. "Web Soil Survey." [Data file and code book]. Retrieved from https://websoilsurvey.nrcs.usda.gov/ (accessed November 1st, 2022).

U.S. Environmental Protection Agency. 1977. *National Interim Primary Drinking Water Regulations*. EPA–570/9–76–003, Washington DC.

—. 2023. "Safe Drinking Water Information System." [Data file and code book]. Retrieved from https://ordspub.epa.gov/ords/sfdw_rest/r/sfdw/sdwis_fed_reports_public/9?p9_report=VIO (accessed October 1st, 2023).

Valayamkunnath, P., M. Barlage, F. Chen, D.J. Gochis, and K.J. Franz. 2020. "Mapping of 30-meter resolution tile-drained croplands using a geospatial modeling approach." *Scientific Data* 7, 257, https://doi.org/10.1038/s41597-020-00596-x.

Van Meter, K.J., P. Van Cappellen, and N.B. Basu. 2018. "Legacy nitrogen may prevent achievement of water quality goals in the Gulf of Mexico." *Science* 360:427–430, https://doi.org/10.1126/science.aar4462.

Van Metre, P.C., J.W. Frey, M. Musgrove, N. Nakagaki, S. Qi, B.J. Mahler, M.E. Wieczorek, and D.T. Button. 2016. "High nitrate concentrations in some Midwest United States streams in 2013 after the 2012 drought." *Journal of Environmental Quality* 45:1696–1704, https://doi.org/10.2134/jeq2015.12.0591.

Wakida, F.T., and D.N. Lerner. 2005. "Non-agricultural sources of groundwater nitrate: a review and case study." *Water Research* 39:3–16, https://doi.org/10.1016/j.watres.2004.07.026.

Walton, G. 1951. "Survey of Literature Relating to Infant Methemoglobinemia Due to Nitrate-Contaminated Water." *American Journal of Public Health and the Nations Health* 41:986–996, https://doi.org/10.2105%2Fajph.41.8_pt_1.986.

Ward, M.H., R.R. Jones, J.D. Brender, T.M. De Kok, P.J. Weyer, B.T. Nolan, C.M. Villanueva, and S.G. Van Breda. 2018. "Drinking Water Nitrate and Human Health: An Updated Review." *International Journal of Environmental Research and Public Health* 15, 1557, https://doi.org/10.3390/ijerph15071557.

Ward, M.H., B.A. Kilfoy, P.J. Weyer, K.E. Anderson, A.R. Folsom, and J.R. Cerhan. 2010. "Nitrate intake and the risk of thyroid cancer and thyroid disease." *Epidemiology* 21:389–395, https://doi.org/10.1097/ede.0b013e3181d6201d.

Weng, W., K.M. Cobourn, A.R. Kemanian, K.J. Boyle, Y. Shi, J. Stachelek, and C. White. 2024. "Quantifying co-benefits of water quality policies: An integrated assessment model of land and nitrogen management." *American Journal of Agricultural Economics* 106:547–572, https://doi.org/10.1111/ajae.12423.

Wieczorek, M.E. 2014. "Area- and Depth-Weighted Averages of Selected SSURGO Variables for the Conterminous United States and District of Columbia." Washington DC: U.S. Geological Survey, https://doi.org/10.5066/P92JJ6UJ. (accessed November 1st, 2022).

Wu, L., J. Letey, C. French, Y. Wood, and D. Birkle. 2005. "Nitrate leaching hazard index developed for irrigated agriculture." *Journal of Soil and Water Conservation* 60:90A–95A.

A Data Appendix: Summary Statistics and Variable Descriptions

Table A.1: Land Use Categories

Categories	Land uses
Low NHI crops; Nitrogen Hazard Index 1	Rice, alfalfa, peas, dry beans, vetch, apricots, grapes, olives, soybeans, lentils, chickpeas, and Christmas trees.
High NHI crops; Nitrogen Hazard Index 2	Barley, canola, corn, sugar beets, oats, rye, cotton, safflower, sorghum, sunflower, winter wheat, spring wheat, durum wheat, buckwheat, almonds, apples, avocados, cherries, citrus, nectarines, oranges, peaches, pears, pecans, pistachios, plums, pomegranates, prunes, walnuts, carrots, sweet potatoes, triticale, millet, other small grains, other tree crops, camelina, sod grass seed, other hay non-alfalfa, other crops, sugarcane. In addition, the Cropland Data Layer defines the following double-cropped combinations: Winter wheat and corn, triticale and corn, winter wheat and sorghum, oats and corn, barley and corn, winter wheat and cotton, lettuce and cotton, barley and sorghum, and durum wheat and sorghum.
High NHI crops; Nitrogen Hazard Index 3	Asparagus, cantaloupes, sweetcorn, cucumbers, eggplant, garlic, honeydew melons, mint, potatoes, pumpkins, radishes, squash, tomatoes, turnips, watermelons, popcorn, and ornamental corn.
High NHI crops; Nitrogen Hazard Index 4	Broccoli, cabbage, cauliflower, celery, lettuce, onions, mustard, herbs, peppers, greens, strawberries, cranberries, blueberries, cane berries, miscellaneous fruits and vegetables, and the following double-cropped crops: lettuce and cantaloupe, lettuce and barley, and lettuce and durum wheat.
Low intensity development	Developed open space and developed low intensity.
High intensity development	Developed high intensity, developed medium intensity, and aquaculture.
Undeveloped land	Clover wildflowers, barren, shrubland, evergreen forest, woody wetlands, herbaceous wetlands, deciduous forest, mixed forest, forest, perennial ice and snow, and wetlands.
Fallow	Fallow and idle cropland.
Pasture	Pasture and grassland.

Note: the table includes land uses that appear within 500 meters of the sample of California wells.

Table A.2: Summary Statistics of Land Use Shares, Soil Characteristics, and Drainage Within One Kilometer of Wells

	Mean	Std. Dev.	Min.	Max.
Low NHI crops	0.06	0.11	0	0.90
High NHI crops	0.11	0.17	0	0.84
Fallow	0.03	0.05	0	0.59
Pasture	0.18	0.20	0	0.97
Low intensity development	0.27	0.17	0	0.91
High intensity development	0.20	0.24	0	0.95
Undeveloped land	0.15	0.22	0	1.00
Sand	0.55	0.19	0.06	0.97
Silt	0.26	0.10	0.01	0.69
Clay	0.19	0.10	0.01	0.60
Organic matter	0.01	0.01	0	0.31
Drainage	0.01	0.04	0	0.74

Note: effective sample size is 6,016 wells. Mean land use shares in 2007 through 2011.

Table A.3: Summary Statistics Using Nitrate Concentrations Equal to Half the Detection Limit for Concentrations Below the Detection Limit

	Mean	Std. Dev.	Min.	Max.
Nitrate concentration in 2019–2023	2.80	3.64	0.002	67.28
Nitrate concentration in 2007–2011	2.60	3.10	0.01	56.44

Note: effective sample size is 6,016 wells.

Table A.4: Impacts of Mean 2019–2023 Land Use Shares Within 500 meters of Wells on Mean 2019–2023 Nitrate Concentrations

	Dependent variable:			
	Log nitrate concentration 2019–2023			
	(1)	(2)	(3)	(4)
Low-NHI crops	0.32	0.37	-0.32***	-0.16
	(0.24)	(0.27)	(0.08)	(0.10)
High-NHI crops	1.4***	1.0***	0.18**	0.13
	(0.24)	(0.21)	(0.09)	(0.10)
Fallow	-0.98	0.04	-0.11	0.27
	(0.70)	(0.78)	(0.24)	(0.33)
Pasture	1.1***	0.63***	0.07	-0.04
	(0.19)	(0.21)	(0.08)	(0.09)
Low-intensity development	0.98***	0.81***	0.06	0.05
	(0.19)	(0.14)	(0.07)	(0.06)
High-intensity development	1.2***	0.83***	0.06	0.07
	(0.20)	(0.16)	(0.06)	(0.07)
Cattle within 1km	0.19***	0.16**	0.02	0.01
	(0.07)	(0.07)	(0.07)	(0.07)
Cattle within 1–5km	0.006	-0.01	0.002	0.001
	(0.02)	(0.02)	(0.01)	(0.01)
Surface water deliveries	0.04	0.05***	-0.002	-0.0003
	(0.03)	(0.004)	(0.003)	(0.001)
Precipitation	-0.48***	-0.18	-0.17***	0.010
	(0.15)	(0.38)	(0.05)	(0.12)
Depth to groundwater	0.010	0.003	0.002	0.008
	(0.02)	(0.02)	(0.004)	(0.005)
Drainage	-1.3***	-0.71	-0.18	-0.25*
	(0.38)	(0.55)	(0.16)	(0.13)
Sand	0.27	-0.29	-0.17	-0.40**
	(0.36)	(0.31)	(0.16)	(0.17)
Silt	-0.25	-0.72	-0.47*	-0.76***
	(0.55)	(0.57)	(0.26)	(0.29)
Organic matter	-4.9***	-4.3**	-0.31	-0.51
	(1.6)	(1.9)	(0.64)	(0.69)
Distance to river	0.19	0.47	0.16**	0.15*
	(0.28)	(0.35)	(0.06)	(0.09)
Initial nitrate concentration			1.0***	0.98***
			(0.02)	(0.02)
Sub-basin FE		Yes		Yes
Observations	6,016	6,016	6,016	6,016
R^2	0.12841	0.32444	0.77276	0.79448

Note: Undeveloped land is the default land use and clay is the default soil textural fraction. Initial nitrate concentration equals log nitrate concentration 2007–2011. Standard errors are clustered by sub-basin $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

B Map of California Counties

Figure B.1: California counties.