Geography of the Orphan Well Burden and the Fiscal Alignment of Impact Fees

Elyse Adamic[1]

[1]University of British Columbia

November 24, 2025

1 Introduction

Orphan oil and gas wells represent a growing environmental and fiscal liability across the United States. These wells, which have been abandoned by their operators without proper plugging or reclamation, pose serious risks to air, water, and soil quality and impose significant cleanup costs on state governments. Pennsylvania alone is home to more than 300,000 documented wells with tens of thousands are currently classified as orphaned or abandoned and many more are idle or inactive with uncertain futures. The scale of the problem has prompted unprecedented public investment: under the 2021 Bipartisan Infrastructure Law, Pennsylvania received over $400 million for well-plugging programs. Yet despite these efforts, policymakers still lack credible forecasts of how many additional wells are likely to become orphaned in the coming decades and where those liabilities will arise.

This paper addresses that gap by developing a predictive framework to identify the subset of Pennsylvania's existing wells most at risk of becoming orphaned. Using detailed well-level production and ownership data from the Pennsylvania Department of Environmental Protection (PADEP), I construct a naive, production-based classification of the current well stock. I can aggregate these categories to the operator level to characterize the structure of firm ownership and operational capacity which highlights the concentration of high-risk wells among small, low-production firms. This provides a transparent depiction of the state's aging well portfolio.

Second, I estimate the probability that a given well becomes orphaned within a fixed time horizon, using logistic models that incorporate well age, cumulative production, recent production decline, operator size, drilling vintage, and local geological and regulatory controls. The results indicate that age, operator scale, and production history are strong predictors of orphaning.

Finally, I can map the spatial distribution of these high-risk wells to provide a statewide view of emerging environmental liabilities. These at-risk wells are disproportionately located in older conventional oil fields in northwestern and southwestern Pennsylvania, rather than in the newer shale-gas regions of the state. I evaluate whether the Act 13 impact fee aligns with these future risks. Since the impact fee is assessed only on active unconventional wells, despite generating substantial revenues, only counties with ongoing shale development benefit. In contrast, the regions most exposed to future orphaning are older conventional areas.

By integrating production-based forecasts of future orphan wells with fiscal data on state and local revenues, this paper quantifies both the environmental and budgetary dimensions of Pennsylvania's orphan-well problem. The findings underscore the need for policy mechanisms that more effectively internalize legacy environmental risks—such as differentiated bonding requirements or risk-adjusted impact fees—so that current extraction more fully accounts for future cleanup obligations. More broadly, the framework developed here offers a scalable approach for other producing states to anticipate and manage long-term liabilities from the fossil fuel industry.

The remainder of the paper is organized as follows. section 2 describes the institutional and policy background of orphan well regulations and the structure of Pennsylvania's Act 13 impact fee. section 3 outlines the data sources and presents descriptive evidence on well life-cycles. section 4 introduces the empirical framework for the naive classification and predicting orphanhood risk. section 5 provides the results and maps the spatial distribution of at-risk wells. section 6 discusses the fiscal alignment and compares the results to the current known landscape of orphan wells and section 8 concludes.

2 Background

2.1 Legal Definitions and Well Status Classification in Pennsylvania

A well is considered active when it produces positive quantities of oil or gas, or zero quantity with expected future production. If a well ceases production for twelve consecutive months but remains under an operator's control, it is classified as inactive. Wells that are non-producing and lack an approved inactive status become abandoned. Abandonment may occur for a variety of reasons including economic, technical, or corporate, but it signifies that the well is no longer maintained and may pose environmental risk. Wells have to be properly decommissioned at the end of their useful life. This is the point at which a well is uneconomical, (which is different than just producing no revenue based on future price forecasts, it is much more complicated) This is state specific, but often requires plugging to prevent fluid mitigation, and land restoration. The large upfront cost of decommissioning gives well operators incentive to postpone. Often regulators allow operators indefinite postponement of commissioning allowing operators this indefinite "inactive" status. This means operators try to keep inactive status although the well will never produce again. Within this broader set, Pennsylvania defines "orphan wells" as a specific subset of abandoned wells for which no responsible operator exists and the well was abandoned prior to April 18, 1985. My point here is that the focus is always on orphan wells, but these abandoned wells are the exact same burden, if not more. These wells often change ownership multiple times, ultimately ending up with small or insolvent operators. When maintenance costs or transfer requirements eventually exceed expected revenues, these wells are abandoned.

This legal structure creates a continuum of well statuses rather than discrete categories. At one end are active and inactive wells that remain in private hands; at the other are orphan wells, which represent realized public liabilities. Between them lies a large and poorly monitored stock of abandoned wells that are legally operator-owned but effectively unmaintained—and, in many cases, environmentally hazardous. In this sense, the policy discourse's narrow focus on "orphan" wells understates the broader scope of the problem: abandoned wells without current production are the same environmental and fiscal burden, differing only in the state's ability to assign legal responsibility. This paper focuses on the set of active and inactive wells as the universe of potential future liabilities. By classifying these wells according to their probability of transitioning into abandonment or orphanhood, the analysis identifies where and for whom Pennsylvania's next generation of environmental liabilities is likely to emerge. Abandoned and orphan wells, by contrast, represent realized liabilities—the visible endpoint of a longer lifecycle of regulatory and economic decline.

My point here is that I have production quantity and production days. I turn to Weber for the minimum production threshold. And Muelenbachs who tells me that it is highly unlikely wells begin production again. There is also this aspect of wasted land - unlikely that there is investment surrounding these wells.

2.2 Production Profiles and the Changing Composition of Pennsylvania's Well Stock

The production lifecycle of an oil or gas well follows a characteristic decline pattern that shapes both its economic viability and its eventual environmental legacy. Conventional and unconventional wells differ sharply in the magnitude and timing of that decline, with important implications for the trajectory of orphan-well liabilities in Pennsylvania.

Conventional wells dominated Pennsylvania's drilling landscape for over a century, typically exhibit low initial production rates followed by a gradual exponential decline. Many continue to produce marginal quantities of oil or gas for decades to maintain "active" regulatory status.This long tail of low-output wells is a defining feature of Pennsylvania's legacy conventional sector and underpins much of the states current official orphan-well inventory.

Unconventional wells, by contrast, exhibit a dramatically different production profile. Horizontal drilling and hydraulic fracturing techniques generate very high initial output followed by a steep hyperbolic or harmonic decline. Gas production in the Marcellus and Utica shales typically falls by 60–80 percent within the first two years of operation and stabilizes at a small fraction of peak output after a decade. These wells are capital-intensive but have relatively short productive lifespans; once decline sets in, continued operation is viable only if fixed surface infrastructure and gathering systems are in place. The industry's rapid technological adoption and scale mean that nearly all new wells since 2010 consists of such unconventional gas wells.

This structural shift from long-lived conventional oil wells to short-lived unconventional gas wells has two competing implications for future orphan burdens. On one hand, the shorter economic lifespan of unconventional wells could accelerate the pace of decommissioning, leading to a faster turnover of wells reaching the end of productive life. If operators meet their plugging obligations, this could actually limit the long-run growth of the orphan stock. And the stock of orphans could remain conventional. On the other hand, the sheer number of high-cost, high-decline unconventional wells now in operation—combined with the likelihood of corporate defaulting creates the potential for a future wave of orphaning once these wells mature. The accelerated decline compresses the window between peak revenue and end-of-life, meaning that economic and regulatory decisions about plugging will arise much sooner than for the conventional wells of the past.

From a policy perspective, Pennsylvania's historical legacy ensures that the current orphan problem is concentrated in older conventional fields, while the emerging risk lies with newer unconventional gas wells that may reach the end of their productive lives within the next twenty years. The temporal and spatial overlap of these two regimes—legacy oil in the northwest and modern gas in the northeast and southwest—means that the state faces a dual challenge: managing a backlog of century-old liabilities while anticipating a shorter-cycle wave of future ones. Understanding the contrasting production dynamics of these well types is therefore essential for forecasting both the timing and geography of Pennsylvania's long-run environmental burden.

2.3 Fiscal Policy

Act 13 of 2012 fundamentally reshaped the fiscal treatment of unconventional gas development in Pennsylvania. Rather than adopting a conventional severance tax on the value or volume of extracted gas, the legislature created an annual "impact fee" levied on each unconventional well that is drilled or producing in a given year. The fee is set in a schedule that depends on well age and the annual average natural gas price, with higher fees in early years and declining obligations as wells age. Operators remit the fee to the Pennsylvania Public Utility Commission (PUC), which is responsible for assessing, collecting, and distributing revenues. Impact fee revenues are distributed according to a formula specified in Act 13. Roughly 60 percent of collections flow to counties and municipalities through the Unconventional Gas Well Fund, with allocations based on the number of unconventional wells located within each jurisdiction, adjusted for population and other factors. The remaining share is deposited into several state-level funds, including the Marcellus Legacy Fund and various environmental, infrastructure, and emergency-response programs. Distributions began in 2012 and have fluctuated with drilling activity and gas prices, but cumulatively have exceeded \$1.4 billion in their first seven years; in recent years, annual collections have been on the order of \$150–\$250 million. The fee is tied to the existence of an unconventional well, not to current production volumes per se, revenues are closely linked to the geography and timing of the shale gas boom.

Pennsylvania's impact fee thus differs in important ways from the severance taxes commonly used in other producing states. Severance taxes are typically levied as an ad valorem percentage of the value of oil and gas at the wellhead or as a specific tax per unit of production (e.g., per Mcf of gas). In contrast, the Act 13 impact fee is essentially a per-well charge that declines with well age and is only applied to unconventional wells. As a result, its effective tax rate on gas value is highest in early years and falls as production declines, and it applies neither to the vast legacy stock of conventional wells nor to their current or future output. Moreover, Act 13 explicitly links the fee to unconventional development: if the state were to adopt a broad-based severance tax, the statute provides that the impact fee would be terminated.

A separate but related instrument is well bonding, which is intended to ensure that operators internalize at least a portion of future plugging and reclamation costs. Under Pennsylvania law, operators must post bonds that vary by well type and operator scale. For conventional wells, firms may post a relatively modest per-well bond (e.g., $2,500 per well) or a blanket bond covering all wells up to a capped amount, while for unconventional wells the required bond depends on total wellbore length and the number of wells operated, with statutory caps that keep total bonding obligations well below estimated plugging costs.

Recent analyses have concluded that these bond amounts fall far short of the expected costs of plugging and site restoration, particularly for older conventional wells, implying substantial residual risk for taxpayers if operators default or dissolve.

Taken together, this fiscal architecture has a clear spatial and temporal incidence. Because the impact fee is assessed only on unconventional wells, it concentrates revenue flows in counties with recent shale gas development, largely in northeastern and southwestern Pennsylvania. Legacy conventional regions, where the current stock of orphan and abandoned wells is highest, receive relatively little Act 13 revenue because they host few unconventional wells. Bonding requirements, meanwhile, are too low and too loosely enforced to guarantee full cost recovery in the event of widespread operator default. The result is a policy design that decouples the geography of revenue collection from the geography of environmental liability: current unconventional drilling funds local governments in shale counties, while cleanup obligations for both historic conventional wells and future orphaned unconventional wells are likely to be borne by a broader set of counties and by the state as a whole.

This paper exploits that structure in two ways. First, by forecasting which active and inactive wells are most likely to become orphaned, it provides a forward-looking measure of expected environmental liabilities at the county level. Second, by comparing these expected liabilities to historical and contemporaneous Act 13 allocations, it assesses the degree of spatial misalignment between where the state collects drilling-related revenues and where long-run cleanup costs will fall. This misalignment is not merely a distributional concern; it also shapes incentives for local governments and operators, potentially weakening the link between present extraction decisions and future environmental responsibilities.

From a policy-design standpoint, this institutional architecture produces a temporal and spatial mismatch. Act 13 channels revenues to places currently experiencing unconventional development, whereas the environmental liabilities from both historic conventional wells and the future retirement of unconventional wells will arise decades later and often in different regions. The fee thus performs its intended function—addressing short-run local disruptions—but leaves the long-term cleanup margin unpriced. In economic terms, the fiscal regime internalizes contemporaneous externalities but not intertemporal ones: the costs of decommissioning and environmental restoration are largely deferred to future taxpayers.

This paper builds on that insight. By forecasting which active and inactive wells are most likely to become orphaned and comparing those projected liabilities with the geographic distribution of Act 13 revenues, I quantify the size of this missing fiscal margin—the gap between the short-term design of extraction-related revenues and the long-term costs of environmental remediation.

3 Data

3.1 Oil and Gas Production Reports

The oil and gas production data comes from Pennsylvania's Department of Environmental Protection (PADEP). I build on the compiled dataset constructed in Chapter 1, which aggregates and merges annual, semi-annual, and monthly production reports over 1980 - 2024. I balance the panel with the assumption that any year after the spud year in which a well is missing a production observation is sitting idle unless a plugging year is stated. I restrict the sample to unplugged oil, gas, or combined oil and gas wells. As natural gas is reported in thousand cubic feet (MCF) and oil in barrels, I natural gas quantity to barrels of oil equivalent (BOE) which allows direct comparison of production volumes across wells and operators. Unless otherwise stated, all quantites are in units of BOE. Each well in a given year can be described using well-specific characteristics (e.g. type, configuration, age), production variables summarizing current and historical activity, and operator characteristics constructed by aggregating and tracking the production of all wells managed by that operator. These variables are all constructed from the annual production recorded for each well and are described in detail in Table 1.

The majority of wells are non-producing at some point over the period of interest. Production decline is defined as the percentage change in production relative to a previous period. This measure becomes undefined when the lagged value is zero, however this could represent a transition from 0 to 0 or 0 to any amount of positive production. While these two cases are very different - either persistent inactivity or reactivation - here the distinction does not matter since periods of lagged zero production do not provide meaningful information on a well's depletion trajectory. Only when a well is producing does a percentage decline meaningfully signal diminishing reservoir pressure or resource exhaustion. I recode these values to zero so that the observations are retained in the regression.

Overall, production variables capture the key dimensions of a well's lifecycle and provide the foundation for the predictive analysis that follows. Figure 1 shows total production over the time period.

Operator are of particular importance as the behavior and operational capacity of firms ultimately determine whether wells are maintained, transferred, plugged, or abandoned. Tracking

Table 1: Variables for Classification and Prediction

Variable	Description
Well Characteristics	
Unconventional	Unconventional indicator
Age Variables	
Age	Well age in years
Number of Productive Years	Number of years since first production
Production Variables	
Annual Production	Annual oil or gas production in BOE
Daily Production	Average oil or gas production per day in BOE
Production Days	Total oil or gas production days in the year
Production Decline	% change in production from previous year
Cumulative Production	Sum of production up to date
Total Production (5 yr.)	Total production over previous 5 years
Non-Zero Production (5 yr.)	Number of non-zero production years over previous 5 years
Operator Characteristics	
Number of Wells	Annual number of wells operated
Total Quantity	Total annual quantity (BOE) over all wells operated
Total Quantity (5 yr.)	Total production by operator over previous 5 years
Percentile	Annual rank of an operator's production among producing operators
Number of High-Producing Wells	Total annual active well count
Number of Low-Producing Wells	Total annual marginal well count
Number of Idle/Zero Production Wells	Total annual idle well count
Number of Orphan Wells	Total annual orphan well count

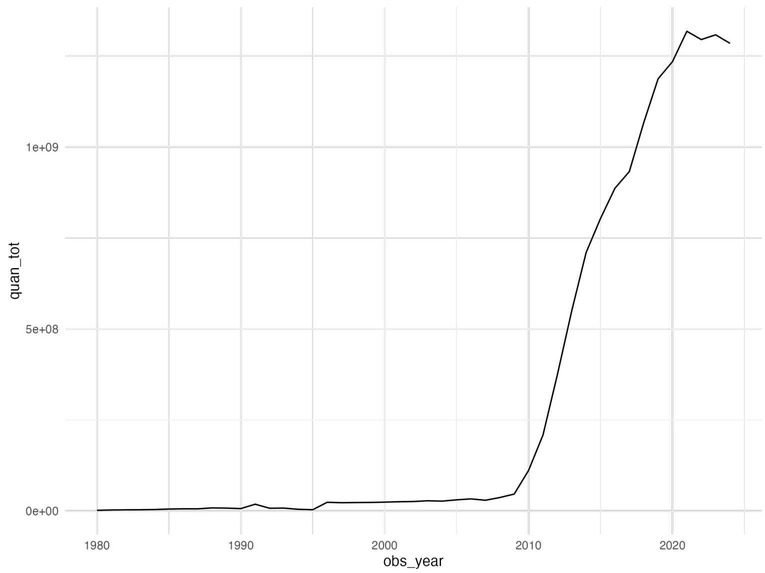

Figure 1: Total Production Over Time

operator characteristics over time (e.g. how many wells they own, how many of each status, production intensity) provides insight into both their economic viability. Operators managing large, actively producing portfolios are more likely to have the resources to comply with plugging obligations, whereas small or declining operators face greater bankruptcy risk and may leave wells orphaned. There are over 3700 unique operators over time, yet in any given year a larger number of them do not own producing wells as shown in Figure 2. This is a defining feature of Pennsylvania's oil and gas industry.

?? Unconventional wells exhibit a markedly different production profile than older, conventional wells: they generate extremely high initial output followed by a steep, predictable decline, whereas conventional wells tend to produce at lower but more stable rates over time. This contrast reflects the broader evolution of Pennsylvania's oil and gas industry. For more than a century, all drilling in the state was conventional and vertical, characterized by small operators and long-lived, low-output wells. Beginning around 2008, however, horizontal drilling and hydraulic fracturing in the Marcellus and Utica formations transformed the industry, with unconventional wells now dominating new development and driving the state's production levels. As a result, today's landscape reflects an aging stock of legacy conventional wells and a newer, highly productive generation of horizontal shale wells as demonstrated in Figure 3. Given the compressed production cycle of unconventional wells, the question is whether they will progress to non-viable status sooner, potentially accelerating the rate at which they are abandoned, or whether recent plugging efforts have successfully slowed the accumulation of new orphan wells, leaving older, legacy wells as the primary contributors to the remaining burden.

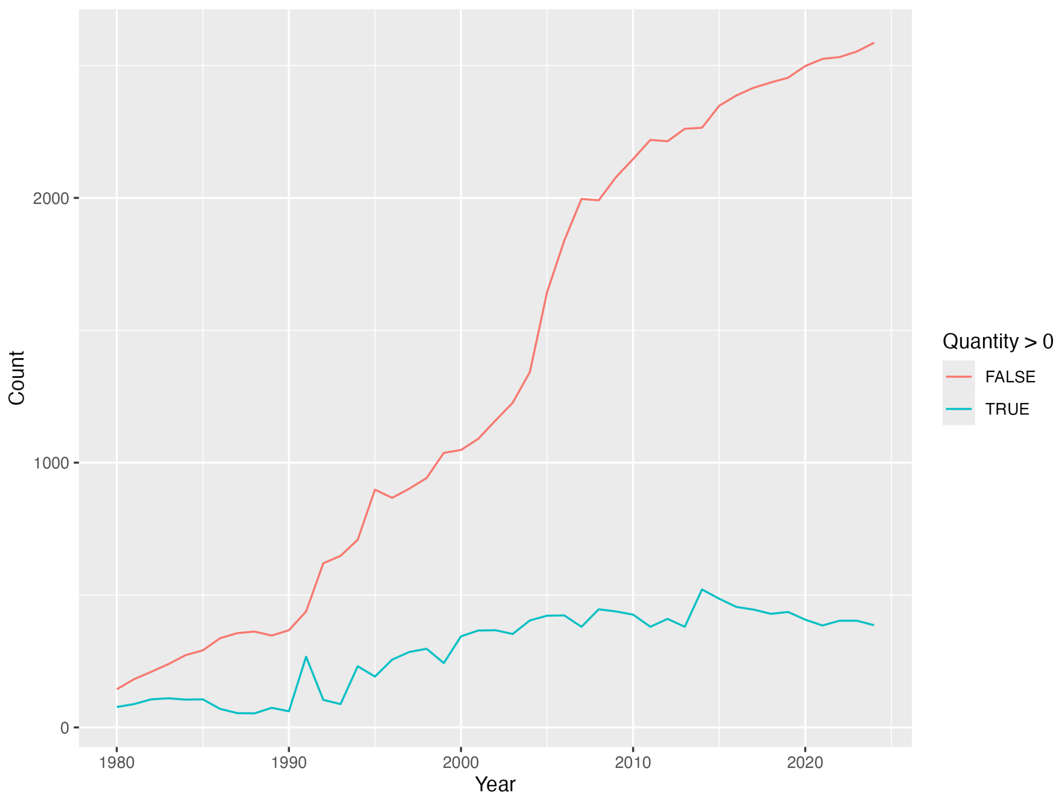

Figure 2: Operator Count Over Time

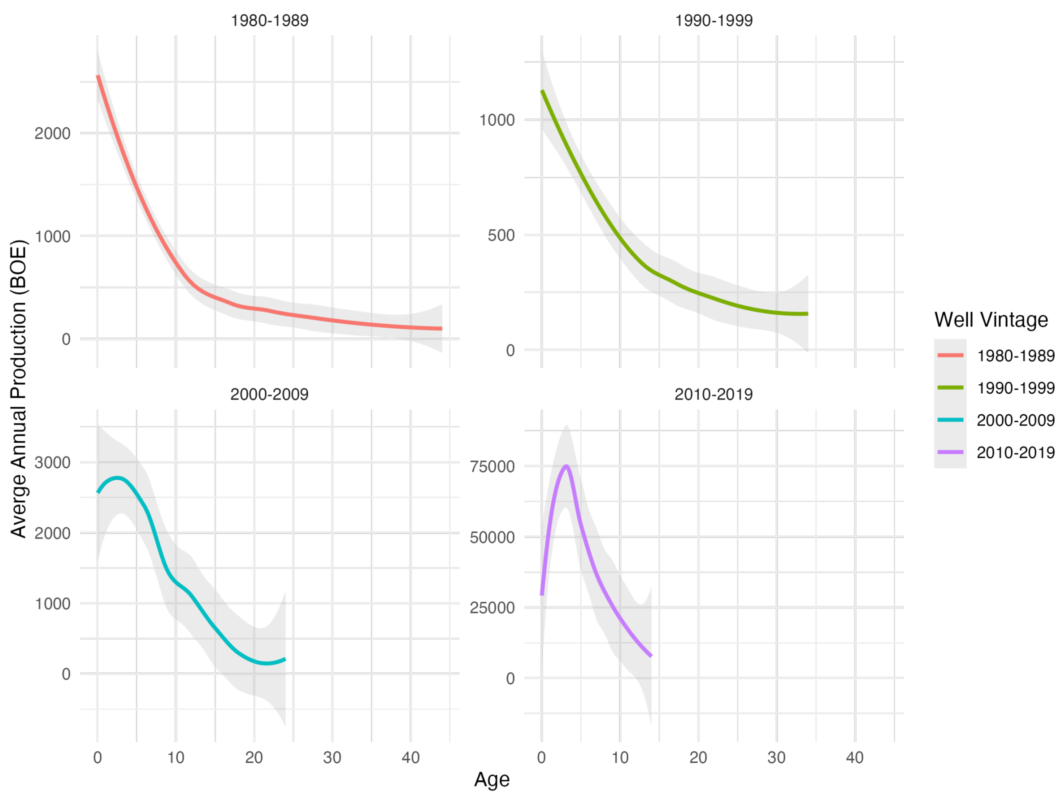

Figure 3: Average Annual Production by Well Vintage

3.2 Act 13 Impact-Fee Data

To examine the spatial distribution of revenues, I compile annual impact-fee collection and disbursement data from the Pennsylvania PUC. The dataset provides total fee payments by operator and detailed county and municipality level allocations by year since 2012. Each jurisdiction's allocation is linked to the number of unconventional wells located within its boundaries. I aggregate these data to the county-year level, measuring: total impact-fee revenue received, per-well and per-dollar-of-liability revenue ratios, and shares of allocations directed to environmental, infrastructure, and legacy programs.

4 Empirical Strategy

4.1 Naive Classification

For a naive classification of the current stock of unplugged wells, I develop a rough screen for wells similar to Boomhower (2018). I assign a well to one of the following categories: likely already orphaned, high risk of becoming orphan, marginal, idle, and active. First of all, I take all wells with zero production over the last five years. If the operator also has no production over the last five years, the well is assumed to be orphan. If the operator has positive production, but the average production rate over all wells is less than 5 BOE/day and primarily operates marginal or idle wells, the well is classified as a high risk of becoming orphan. The remaining wells are classified using current production. Marginal wells are those between 1 and 5 BOE per day, and idle wells are less than 1 BOE per day with many completely inactive producing 0. High-producing wells with over 5 BOE per day are considered active. Daily output is the preferred measure of production intensity, since wells that produce at high rates for only part of the year can appear low-producing in the annual sum. However, because many wells have incomplete or missing data on production days, I also construct an annual production classification using a threshold of 1825 BOE per year (equivalent to 5 BOE per day) to ensure comparability across wells with less reliable operating-day records. I also explore alternative cutoffs as a robustness check.

4.2 Logit Model

The naive classification labels wells based on recent production patterns and is inherently descriptive. While it provides an informative picture of the current landscape, it offers little insight into which active or inactive wells are likely to become future liabilities. To forecast Pennsylvania's future orphan-well burden, I estimate a predictive model that uses the full production history data to relate current well characteristics to the probability of abandonment within the next five years.

I use historical patterns in production profiles, decline trajectories, and operator attributes of likely orphan wells to predict future patterns. While the DEP maintains an inventory of verified legacy orphan and abandoned wells, these records lack production histories and therefore provide little basis for predicting future orphan-well risk. However these wells contribute to the overall plugging burden and fall under state responsibility so when explicitly stated I will incorporate counts into the overall estimates of orphaned wells alongside my model-based predictions.

First, I create a binary indicator Y_{it} equal to one if the well becomes abandoned in the next five years of the observation year. This model is conditional on being observed for the past 5 years for certain explanatory variables and for the future 5 years for the dependent variable so the first and last four observations per well are omitted. This essentially restricts the sample to 1984 - 2019 [1].

Then I estimate a logistic regression that maps well-level and operator-level covariates to the probability of near-term abandonment. The covariates as outlined above include production attributes that are known to correlate with well retirement decisions: well age, recent annual production decline, cumulative lifetime production, daily output, five-year production history, and operator characteristics such as size and well status portfolio. Note year or county fixed effects are not included as that would absorb the cross-sectional variation essential for forecasting. Since the objective is to identify structural drivers of orphan risk rather than short-run market dynamics, I also exclude resource prices from the baseline specification, yet will explore in an alternative specification. Specifically, I estimate the following:

$$\begin{aligned}\Pr(Y_{it+5}) = \beta_0 + \beta_1\,\text{Age}_{it} + \beta_2\,\text{Decline}_{it} + \beta_3\,\log(\text{CumProd}_{it})\\ + \beta_4\,\log(\text{Prod5yr}_{it}) + \beta_5\,\text{DailyProd}_{it} + \beta_6\,\text{SmallOp}_{it} + \epsilon_{it}.\end{aligned}$$

I apply the estimated coefficients to all active and inactive wells in 2024, generating well-level predicted probabilities of orphan status in the next five years. These predicted risks form the foundation of the subsequent mapping and aggregation analysis. These probabilities range from near zero for recent unconventional wells operated by large firms to substantially higher values for older, low-producing conventional wells managed by small operators.

This approach fully reflects uncertainty, incorporates the joint influence of age, decline patterns, operator characteristics, and production history, and avoids arbitrary threshold choices present in the naive classification. It also makes the estimates additive and comparable across regions enabling a direct link between expected future liabilities and current policy instruments.

4.3 County Level Aggregation

The predicted probabilities generated by the logistic model provide a well-level measure of near-term abandonment risk. To translate these individual predictions into a statewide assessment of Pennsylvania's future orphan-well burden, I aggregate the well-level probabilities to the county level and map the resulting spatial distribution. This section presents the geographic distribution of predicted risk.

$$ExpectedOrphans_c = \sum_{i \in c} p_{it+5}$$

[1] If a well is spud in a later year, the first four years of production still need to be removed.

Table 2: Naive Classification of 2024

Status	N	% Horiz.	Annual	Daily	Total 5-Year	Age	Prod. Yrs.
Likely orphan	14,689	2.41	0.00	0.00	0.00	33.21	27.43
Idle	65,057	0.81	99.40	0.29	227.38	28.23	24.47
Marginal	9,927	3.18	532.54	1.64	641.64	25.74	23.67
Active	11,831	97.98	106945.25	362.49	86951.90	9.65	9.38

Notes: Columns (2) and (3) use raw well counts. Columns (4) - (6) refer to average production in BOE among wells of the given status. Column (7) refers to average years since spud year, while column (8) is average years since first production.

5 Results

5.1 Naive Classification

I present the results of this naive classification for 2024 in Table 2. The four categories demonstrate different operational profiles. Current active wells are overwhelmingly horizontal (98 percent), relatively young with an average age of 9.65 years, and produce substantial volumes, averaging over 100,000 BOE and 360 BOE per day. By construction marginal wells average fewer than 2 BOE per day, while idle wells produce virtually nothing by construction. These wells are considerably older, with mean ages of over 25 years. Wells that are likely orphan are the oldest in the inventory with an average of 32 years and have the longest operational histories.

Although the naive classification is primarily descriptive, it offers simple forward-looking insights. Idle and marginal wells are substantially older and produce far less than active wells, and together represent more than 75,000 wells statewide. If even 5-10 percent of these wells follow the same trajectory as current orphan wells. Pennsylvania could face 3,000-7,500 additional orphan wells in the coming decade. Age profiles also suggest substantial future risk as idle and marginal wells are only a few years younger than today's orphan wells. Finally, large numbers of wells already produce less than 1 BOE/day, a threshold at which wells historically exhibit high abandonment rates. These naive projections underscore the scale of potential future liabilities and motivate the more formal predictive modeling approach that will use this information.

I present operator summary statistics in Table 3 which use the naive classification and reveal an interesting industry structure in Pennsylvania. The group of operators with above-median production is small at 369 firms, but on average they manage more than 230 wells each and account for virtually all statewide production. Nearly 90 percent have at least one idle well, over half manage marginal wells, yet only a quarter hold high-producing wells. Importantly, 51 percent of them also have at least one well that has not produced over the past five years suggesting a likely orphan well if the producer could not be held liable. The below-median production operators average fewer than five wells and overwhelmingly consisting of single-well entities where 264 firms own exactly one well. Their production volumes are negligible, and nearly all of them operate exclusively idle wells. The zero-production operators represent small legacy operators that col-

Table 3: Operator Summary Stats for 2024

Stat	large	small	zero
Number of Operators	363.00	363.00	2246.00
Mean Number of Wells	243.13	4.54	5.16
Number of Wells = 1	23.00	263.00	1593.00
Total (1000 BOE)	3517.94	0.03	0.00
% Have Orphan Well	51.52	6.89	92.30
% Have Idle Well	90.08	100.00	9.17
% Have Marginal Well	53.17	0.00	0.00
% Have Active Well	24.24	0.00	0.00
% Only Idle	22.87	93.11	7.70

Notes: Above (below) median refer to operators with annual quantity in the top (bottom) half, and zero production refer to operators with zero annual quantity.

lectively manage wells with no measurable output. These firms own an average of 5–6 wells and 90 percent of these operators have a likely orphan well. This underscores the strong association between very small operators, aging conventional wells, and abandoned infrastructure.

Taken together, these patterns show that Pennsylvania's future orphan well burden is highly concentrated among the thousands of very small, very low-production operators who collectively manage the bulk of aging, nonproducing wells, while nearly all economically meaningful production is concentrated among a few hundred large operators.

Again, the results above classify wells in 2004. I can also use this classification for all wells over time. Figure 4 shows how wells transition between states and highlight the strong persistence of well conditions over time with movement reflecting declining production. However the transition matrix underscores that historical status alone has limited predictive content for identifying future orphan-well risk, reinforcing the need for a modeling framework that incorporates richer well-level and operator-level characteristics.

5.2 Prediction for 2024

Table 4 reports the logistic regression estimates linking well and operator characteristics to the probability of becoming "likely orphaned" within five years. The signs of the coefficients are consistent with expected patterns in well life cycles. Wells with greater age, fewer recent production observations, and steeper production declines exhibit sharply higher abandonment risk. Higher cumulative and recent production are strongly protective, reflecting that productive wells are rarely left idle long enough to orphan. Operator characteristics also matter: wells owned by smaller operators face systematically higher near-term abandonment risk. Unconventional wells have markedly lower predicted orphan risk, consistent with their younger age and higher annual output.

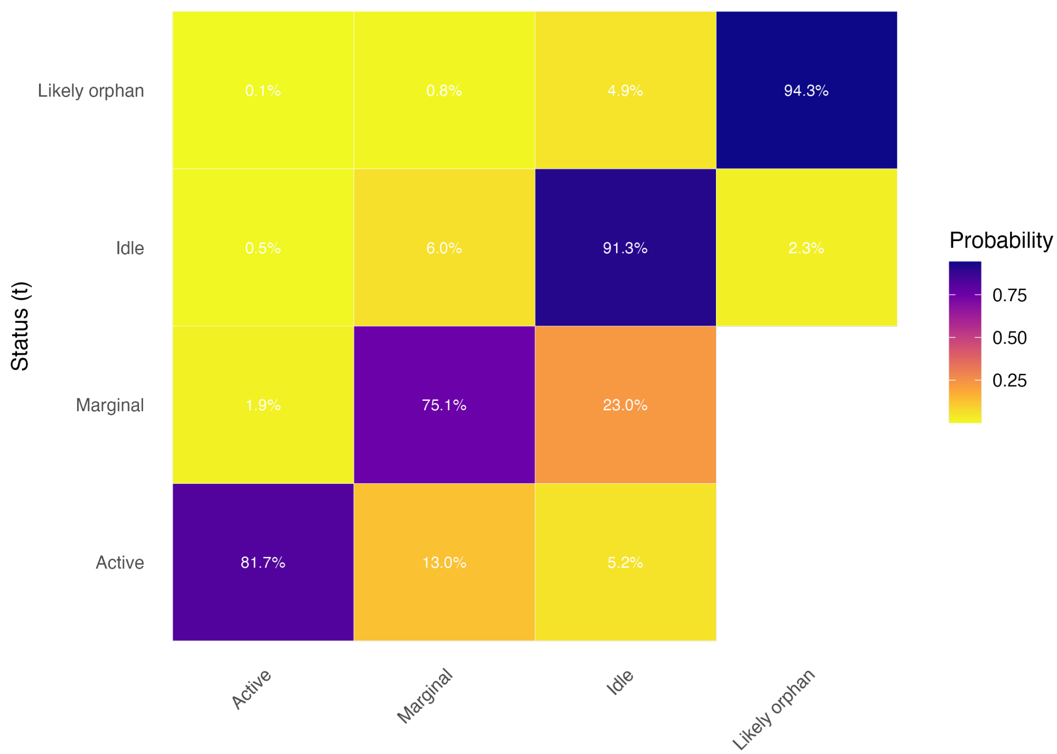

Figure 4: Transition Matrix

Table 4: Results from Logistic Regression

	(Orphan in 5 Years = 1)
	(1)
Well age	0.003***
	(0.000)
Production years	-0.017***
	(0.001)
One-year production decline (%)	-0.003***
	(0.000)
Log cumulative production	-0.012***
	(0.004)
Log 5-year production	-0.404***
	(0.004)
Mean Daily production	-0.009***
	(0.002)
Operator active-well count	-0.001***
	(0.000)
Operator production percentile	-2.023***
	(0.012)
Observations	1,248,444

Notes: Standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

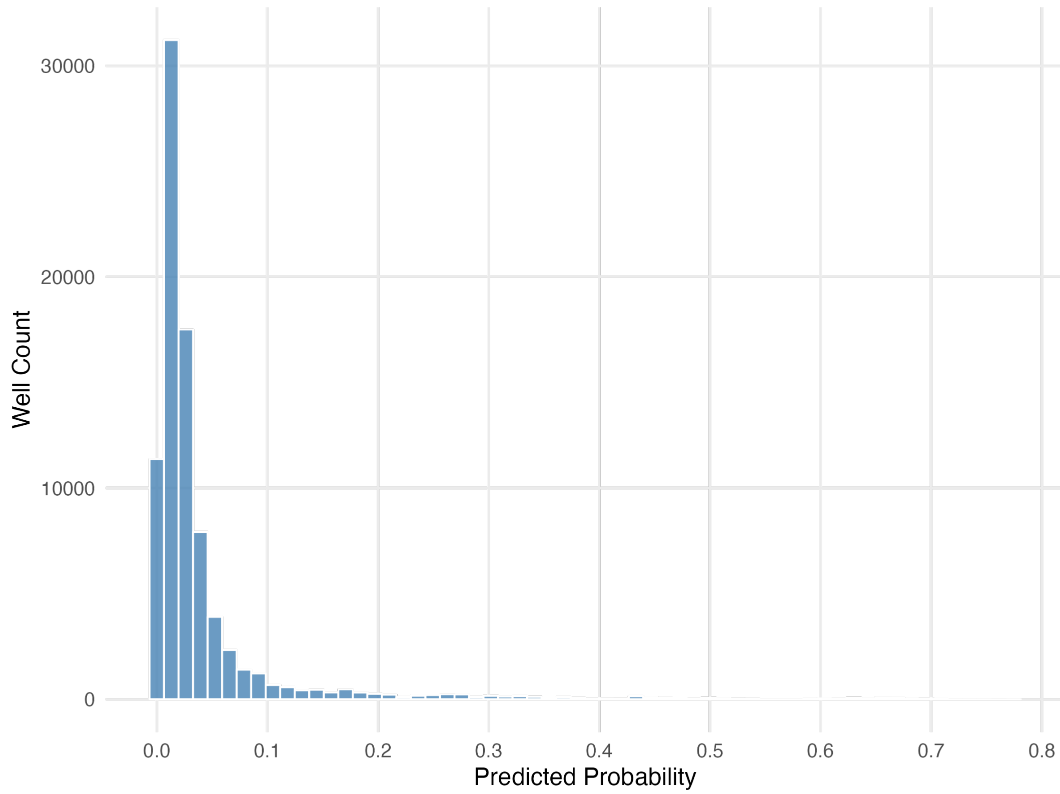

Figure 5: Predicted Probabilities for all Wells

The predictive model is then used to forecast orphaning risk for the current inventory of wells in 2024. Using the characteristics of each active or inactive well and the coefficients from the logit equation, the model yields a well-specific probability of becoming orphaned within the next five years. Logistic regression estimates effects in terms of log-odds, which are not directly interpretable as probabilities. After estimating the model, I apply the logistic inverse-link function to transform the linear predictor into a probability between zero and one, yielding a well-level measure of the chance that the well becomes orphaned within five years. These predicted risks summarize the combined information contained in production trends, operational intensity, operator features, and geological controls, providing a forward-looking measure of each well's vulnerability to abandonment. I present the distribution of predicted probabilities in **??**. While the majority of wells have a low chance of abandonment, the probability is non-zero.

In Table 5 I summarize the logit results by my previous classification. Current high-producing wells have the lowest probability of abandonment, highlighting how it they are managed by the largest operators with little risk of bankruptcy. Idle wells have the highest probability as expected due to the persistance of low output and smaller operators. Overall, within the next five years the expected number of additional orphans added to the current legacy burden is over 3400.

Table 5: Prediction based on Logit Results

Status 2024	Mean	Max	Std. Dev.	Future Expected Orphans
Idle	0.051	0.774	0.084	3238.508
Marginal	0.018	0.298	0.038	170.645
Active	0.001	0.081	0.003	7.056

Notes: Predicted probabilities between 0 and 1 calculated from the logit model for each well based on model characteristics.

5.3 Spatial Aggregation

Figure 6 maps the spatial distribution of current "likely" orphans and expected orphans based on these predicted probabilities. Current orphan wells can be thought of as having a probability equal to one of remaining an orphan well, and all other unplugged wells have a range of probabilities.

As expected, the projected orphan-well burden is highest in counties with extensive drilling activity, simply reflecting the larger underlying stock of wells. This baseline pattern is not surprising: regions with dense historical development necessarily contain more wells at risk. However, the spatial concentration of the most likely future orphans are not uniformly distributed but rather clustered in specific parts of the conventional drilling region.

6 Discussion

6.1 Fiscal Alignment under Act 13

Although the impact fee was originally designed to address short-run externalities associated with unconventional drilling, comparing these revenues to the long-term orphan-well liabilities remains informative. The contrast highlights potential mismatches between where short-term revenues are generated and where long-term remediation burdens ultimately fall. The distribution of fee revenues is tied to the location and timing of active unconventional wells, with counties such as Washington, Susquehanna, Bradford, and Greene receiving the largest annual allocations. However, the predictive results show that the expected future orphan burden is concentrated in a completely different part of the state. Northwestern counties—McKean, Venango, Warren, Forest, and Clarion—harbor the oldest and least productive conventional wells, and therefore accumulate the highest predicted probabilities of abandonment in the coming years. Yet these counties receive little, if any, impact fee revenue because they host very few unconventional wells. The result is a policy mechanism that steers substantial funding toward counties with robust shale activity, where wells are newer, better capitalized, and at low risk of orphaning while the counties facing the highest remediation costs are disproportionately left without dedicated fiscal resources. Together, the predictive orphan-risk analysis and the geographic distribution of impact fee revenues suggest that Pennsylvania's existing funding framework is ill-suited to address the long-term costs well abandonment. A dedicated policy mechanism for long-term environmen-

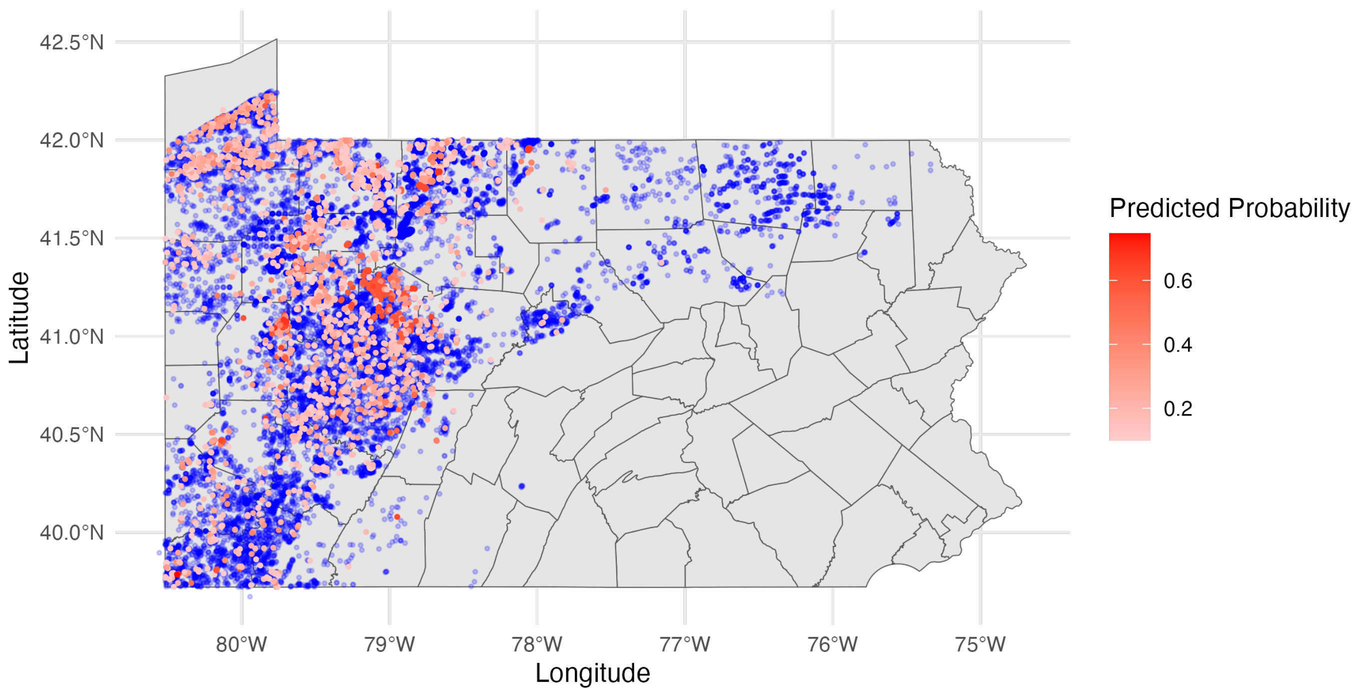

Figure 6: Future Orphan Well Burden across Pennsylvania

tal liabilities is needed to ensure adequate and sustainable funding for orphan-well remediation. Effective policy must align funding streams with the spatial and temporal distribution of those liabilities.

6.2 Comparison to DEP Inventory

Here I compare my results to the official DEP Inventory which contains data on all current and legacy wells known to the DEP. A comparison of the official reported status and the production-based naive classification reveals large inconsistencies in wells recorded as "Active" and how wells actually behave. A significant number of these wells exhibit production patterns far more consistent with idle or even orphan wells. Among wells recorded as Active, fewer than 12,000 also appear as Active under the naive production-based classification. In contrast, more than 66,000 DEP Active wells are classified as Idle, and an additional 10,700 are classified as Likely Orphan wells producing zero output over the past five years. Taken together, over four-fifths of wells labeled Active by DEP show little or no recent production, suggesting that the operational meaning of "active" in regulatory records diverges sharply from the economic reality of well behavior.

This mismatch most likely reflects structural features of Pennsylvania's regulatory regime where operators face limited pressure to either produce from or properly decommission non-producing wells, and many wells remain in a nominally "active" status long after production has ceased.

If DEP records define tens of thousands of wells as Active despite negligible output, the formal inventory dramatically overstates the operational viability of the state's well stock and understates both the scale and urgency of long-term liabilities. Figure 8 shows what the likely landscape of abandoned wells actually looks like compared to the DEP categories. The DEP abandoned and orphan lists remain small relative to the much larger pool of wells that meet behavioral criteria for abandonment but are not administratively recognized as such.

Using the predictive probabilities among these wells in the DEP "Active" category, the expected number of orphan wells that could ultimately be added to the burden within five years is between 10,000 and 15,000. The lower bound comes from summing the predictive probabilities across all wells, and the upper bound gives all wells classified as "likely orphan"s" a probability of one. Together, these comparisons underscore that Pennsylvania's regulatory inventory substantially understates the scale of its non-producing well problem.

7 Robustness

7.1 Threshold of Activity

This explores whether 5 BOE per day is too stringent of a requirement for high-producing.

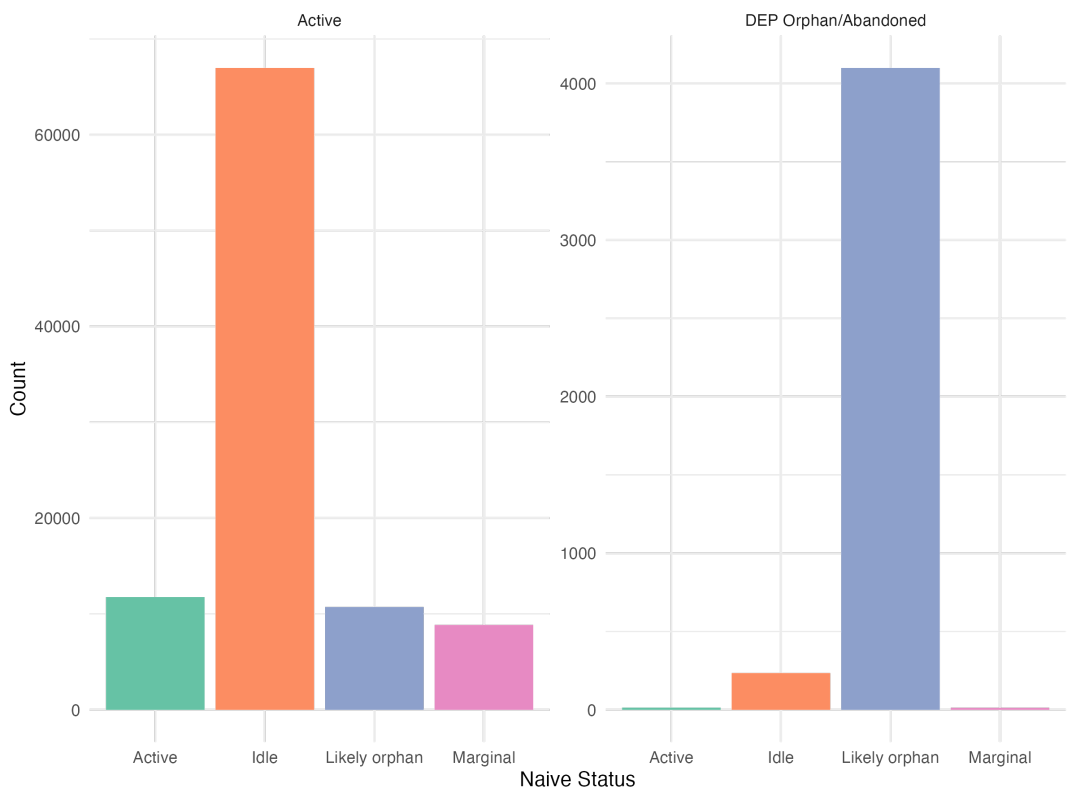

Figure 7: Naive status distribution of DEP Inventory Categories

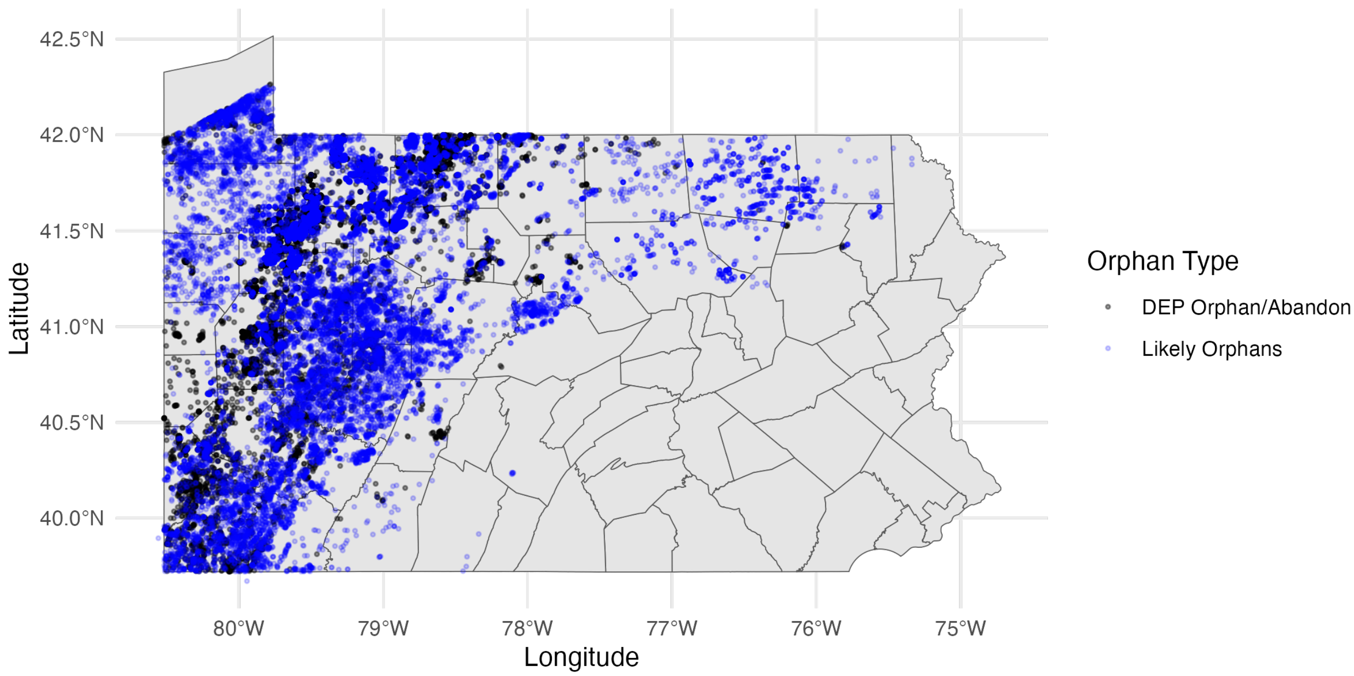

Figure 8: DEP Orphan and Abandoned Wells vs. “Likely Orphan” Classification

7.2 Short and Long Run Forecasts of Well Abandonment

In addition to the baseline five year prediction window, I examine how abandonment risk evolves under alternative time horizons, namely one and ten years.

8 Conclusion

This paper develops a forward-looking assessment of Pennsylvania's orphan well liabilities by linking detailed production histories, operator characteristics, and regulatory classifications to the probability that currently active or inactive wells will become orphaned in the near future. Although the predictive model sharpens the identification of wells at highest risk of abandonment, the broader implications of the analysis are ultimately structural rather than statistical. The core patterns uncovered by the model align closely with intuitive features of well life cycles: long-idle, low-production, and aging conventional wells exhibit the highest abandonment risk, while recent unconventional wells operated by large firms rarely approach orphanhood in the short run. In this sense, sophisticated prediction validates that Pennsylvania's abandoned well problem is driven not by failures among active, high-producing wells, but by the vast, persistent stock of legacy wells that have effectively ceased to be economically viable which are not recognized by the DEP. By quantifying the expected number of orphan wells embedded in the current well inventory, this paper highlights the extent to which Pennsylvania's long-term environmental liabilities stem from wells that are, for all practical purposes, already orphaned but not administratively recognized as such.

The fiscal analysis further reveals a structural misalignment between the geography of future liabilities and the distribution of revenues under Act 13. Impact-fee collections are concentrated in counties with ongoing unconventional development, while the burden of future orphaning lies predominantly in older conventional regions that receive little fee revenue. This divergence between short-run revenue flows and long-run cleanup costs implies that Pennsylvania's existing fiscal architecture internalizes immediate drilling externalities but leaves intertemporal environmental liabilities largely unfunded. Addressing this gap will require complementary policy mechanisms.

Taken together, the findings demonstrate that credible forecasts of orphan-well risk are feasible using naive classifications and predictive models, but the deeper value lies in demonstrating the scale and geography of a longstanding legacy problem. Predictive frameworks of this kind can help policymakers anticipate emerging liabilities, target resources effectively, and design fiscal institutions that better align revenue collection with long-term environmental responsibility.