Estimating the Future Supply of Shale Oil:
A Bakken Case Study

James L. Smith
Department of Finance
Southern Methodist University
jsmith@smu.edu

Abstract

We propose a new way to estimate the remaining volume of recoverable shale oil resources in the U.S. Our method applies the principle of "successive sampling without replacement" to derive from historical drilling data maximum likelihood estimates of the number and productivity of remaining drilling sites. Unlike existing techniques, this approach identifies the portion of "technically recoverable" resources that can be developed economically at alternative price levels. For the Bakken, we estimate that 50% of remaining technically recoverable resources—roughly 8 billion barrels—could be developed if the oil price remains near $50/barrel.

JEL Codes: L71, Q35, Q41

Keywords: shale oil supply, MLE, drilling productivity, sequential sampling

Acknowledgements

This research was supported by a grant from the MIT Center for Energy and Environmental Policy Research, Cambridge, MA, and presented at the CEEPR Fall 2016 Research Workshop. The author is grateful to participants in the workshop for comments on an earlier draft, and to Guanhuan Wang who provided research assistance. The author also thanks Timothy Fitzgerald, Brian Prest, and Petter Osmundsen for helpful comments on an earlier draft.

Introduction

The volume of future U.S. shale oil production depends directly on the number and productivity of potential drilling sites that remain to be tapped. Because shale oil is a depletable resource, past performance provides no assurance of continued abundance. Although estimates of future production must be forward looking, we show in this paper how historical drilling outcomes lead to maximum likelihood estimates of the number and productivity of remaining shale oil drilling sites. The method is illustrated with an empirical application to the Bakken shale oil play, located in the Williston Basin of North Dakota.

When compared to resource estimates developed independently by the U.S. Geological Survey (USGS) and the U.S. Energy Information Administration (EIA), our results appear plausible and potentially more useful. We show, for example, how our method facilitates analysis of the economic supply of remaining shale oil reserves—something that the USGS and EIA procedures do not permit. This point highlights the important distinction between two measures of resource abundance. The volume of "technically recoverable resources" refers to the portion of hydrocarbons in place that can be recovered using existing technology but without regard for economics. The volume of "economically recoverable resources" refers to the subset of technically recoverable resources that can be recovered economically at a given price level. We provide estimates of both measures of resource abundance.

The method we propose is based on two statistical hypotheses, both of which have been previously applied to the petroleum industry: sampling from the remaining population of potential drilling sites is assumed to occur without replacement and with probability of selection proportional to size. In the present case, the "size" of a drilling site refers to the productive potential of the associated well (barrels of oil produced per day). Sampling is without replacement because petroleum resources are finite in nature and subject to depletion. Resources that have been once produced cannot be produced again. Random sampling proportional to size captures the industry's general tendency to drill more productive prospects before less productive prospects, at least within the limits of imperfect knowledge regarding productivity. In other words, the historical sequence of drilling outcomes observed in a given play is assumed to reflect the results of selective sampling from an unknown parent population—one that is continuously diminished by successive efforts.[1] Our goal is to identify the original population of drilling sites that maximizes the likelihood of having observed the historical sequence of actual drilling outcomes. Having characterized the original population, we then estimate the remaining volumes of shale oil that can be produced economically at various price levels.

Although the application presented herein focuses on the Bakken region, the same analysis could be applied to any of the numerous shale oil plays that have emerged recently, and to shale gas plays as well. The Bakken happens to be one of the largest and most productive shale oil plays in the U.S., and is counted on to contribute substantially to future U.S. shale oil output.[2] In addition to illustrating a new methodological tool, our work therefore makes an important empirical contribution regarding future U.S. shale oil supply.

Related Studies

Our approach is an application of "successive sampling," the statistical properties of which have been studied previously by Horvitz and Thompson (1952), Rosén (1972), Gordon (1983), and others. Successive sampling models include two elements: a variable probability of selection and sampling without replacement. The variable probability of selection may be

[1] In industry parlance, a "play" is a group of fields or prospects in the same region that are controlled by the same set of geological circumstances.

[2] The Bakken play accounts for 25% of all the shale oil produced in the U.S. since 2000. The U.S. government forecasts that, by 2040, Bakken wells will be producing almost one-third of projected U.S. tight oil output (EIA 2016).

determined either purposefully as part of the sampling design (as in some survey research where certain segments of the population are deliberately over-weighted) or exogenously by some innate characteristic of the sampled member. Previous applications of successive sampling to the discovery of oil and gas fields are of the latter type: the probability of discovery of any particular field from among the remaining population is assumed to be proportional to its size.

Kaufman, Balcer, and Kruyt (1975), Barouch and Kaufman (1977), Smith (1980), O'Carroll and Smith (1980), Meisner and Demirmen (1981), Smith and Ward (1981), Lee and Wang (1983), and Bickel, Nair, and Wang (1992) each apply successive sampling models to estimate the size distribution and other characteristics of the distribution of remaining (undiscovered) conventional oil and gas fields. In that context, it is natural to imagine that the physical size of a field would influence the probability of discovery—even if wells were placed randomly across the surface.[3] Barouch and Kaufman (1977) impose the additional assumption that the distribution of field size is lognormal. Smith and Ward (1981) impose Weibull and exponential constraints in addition to the lognormal.

Andreatta and Kaufman (1992) develop an unbiased estimator and compare its performance with Smith and Ward's (1980) consistent maximum likelihood estimates of the North Sea oil field size distribution. The two sets of estimates are quite similar. One major difference, however, is that calculation of the unbiased estimator requires outside knowledge of at least one characteristic of the population (like the total number of fields extant), whereas the maximum likelihood approach (which is employed in this paper) does not, as Bickel, Nair, and Wang (1992) have demonstrated.

[3] For further discussion, see page 6 of Meisner and Demirmen (1981) and Menard and Sharman (1975).

The current paper relates closely to all of the works cited above, but differs in two material respects. First, our objective is to estimate the number and productivity of individual drilling sites (wells) as opposed to the size distribution of oil fields. The concept of "field" hardly applies in the case of shale oil and gas resources, where the hydrocarbons are laid down in continuous beds covering thousands of square miles. The economics of supply in such cases hinges on the productivity of individual wells and its variation throughout the play. Second, we adapt the traditional model of successive sampling without replacement to account for the impact of learning-by-doing. Operators are constantly improving the efficiency of fracking as experience with the technology accumulates. This increases output from individual wells and tends to mask the impact of resource depletion. Without taking into account improvements in fracking efficiency, consistent estimates of the shale oil resource base cannot be derived from the successive sampling framework.

The Model

We posit a stochastic relationship between drilling effort and drilling outcomes within a shale oil play. The drilling of each well constitutes a trial whose outcome reveals the productivity of a particular drilling site.[4] The relationship must account for substantial variation in the productivity of individual wells. USGS (2012) applied the lognormal distribution to model the variation in productivity of individual wells, and found that productivity of the best drilling sites within a given play typically exceeds that of the worst sites by two orders of magnitude. We adopt this approach and assume the productivity of a site chosen randomly from the population of all sites within a play follows a lognormal distribution with unknown mean and variance. To be clear, by "productivity" we refer to the initial production rate of the associated

[4] By "drilling site" we refer to the geographic area that is drained by an individual well.

well. If we let X denote productivity of the well and use $\Lambda(x|\mu,\sigma)$ to denote the lognormal probability of $X \leq x$, the density is then given by:[5]

$$d\Lambda(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\left\{-\frac{1}{2\sigma^2}(log x-\mu)^2\right\}} dx. \tag{1}$$

It will be convenient, as in earlier successive sampling studies, to group drilling outcomes into discrete ranges, or tiers, based on realized values of X. With K categories defined by successive cut points ($0 = C_1 < C_2 < \ldots < C_K$), the probability that a randomly chosen site is of the k^{th} type is:

$$\pi_k(\mu,\sigma) = \int_{C_k}^{C_{k+1}} d\Lambda(x|\mu,\sigma) \qquad \text{for } k = 1,2,\ldots,K-1, \text{ and}$$

$$\pi_K(\mu,\sigma) = 1 - \sum_1^{K-1} \pi_k(\mu,\sigma). \tag{2}$$

Suppose the total number of potential drilling sites included in the population is N, which like μ and σ is unknown. We may then describe the resource base by the number of drilling sites in each productivity tier:[6]

$$N_k(\mu,\sigma) = N \times \pi_k(\mu,\sigma). \qquad \text{for } k = 1,2,\ldots,K. \tag{3}$$

Our goal is to identify values $(\widehat{N},\hat{\mu},\hat{\sigma})$ that maximize the likelihood of having observed the particular historical sequence of drilling outcomes. These three parameters describe the initial endowment of technically recoverable resources that existed before drilling commenced.

[5] Aitchison and Brown (1966) provide a comprehensive description of the lognormal distribution.
[6] If the finite distribution is regarded as a random sample from a parent lognormal distribution, as in Smith and Ward (1981), then Eq. 3 defines the expected (not actual) number of sites within each tier. With a large sample (≈10,000), the variation around the mean is negligible (at most one-half of one percent) so we treat the N_k as actual counts.

Inferences regarding characteristics of the untapped portion of that population are obtained simply by deleting historical wells from the initial distribution.

To proceed, we formulate the likelihood function that describes the probability of observing any particular ordered sequence of drilling outcomes $(x_1, x_2, \dots, x_n)$. For convenience, we denote the productivity tier of the *i*th well by the symbol $I_i \in \{1,2,\dots,K\}$. An arbitrary sequence of *n* outcomes may then be represented by $\{I_1, I_2, \dots, I_n\}$. The probability that the first well will belong to the *k*th tier is determined by random sampling according to size, where S_k represents the average productivity of drilling sites included in the *k*th tier:

$$P(I_1 = k|N,\mu,\sigma) = \frac{N_k \times S_k}{\sum_{j=1}^{K} N_j \times S_j}, \qquad \text{for } k = 1,2,\dots,K. \tag{4}$$

As drilling proceeds, the probabilities of respective outcomes evolve according to the principle of sampling without replacement. In general, the number of remaining drilling sites in tier *k* is determined by the depletion that has gone before. The cumulative number of wells in the *k*th tier which have preceded the *i*th well is denoted m_{ik}. These well counts may be calculated directly from the observed drilling outcomes. The probability that the *i*th well will be in the *k*th tier, conditional on the outcome of preceding trials, may then be written:

$$P(I_i = k|I_1, I_2, \dots, I_{i-1}; N,\mu,\sigma) = \frac{(N_k - m_{ik}) \times S_k}{\sum_{j=1}^{K} (N_j - m_{ij}) \times S_j}, \qquad \text{for } k = 1,2,\dots,K. \tag{5}$$

Knowledge of the initial distribution of drilling sites (N,μ,σ) is sufficient to calculate the likelihood of any particular sequence of *n* drilling outcomes, which is given by the product of successive conditional probabilities:

$$L(I_1, I_2, \dots, I_n|N,\mu,\sigma) = \prod_{i=1}^{n} \frac{\left(N_{I_i} - m_{iI_i}\right) \times S_{I_i}}{\sum_{j=1}^{K} (N_j - m_{ij}) \times S_j}. \tag{6}$$

Alternatively, given a particular sequence $\{I_1, I_2, \dots, I_n\}$, the likelihood function can be evaluated via grid search to identify the parameter values $(\widehat{N}, \hat{\mu}, \hat{\sigma})$ that maximize the likelihood of the observed data. Estimates of the discrete number of sites within each productivity tier are given by $\widehat{N}_k = \widehat{N} \times \pi_k(\hat{\mu}, \hat{\sigma})$, for $k = 1,2, \dots, K$.

Before proceeding to the Bakken application, it is useful to consider how an observed sequence of wells is able to identify the parameters of the original distribution. A sequence that consists mostly of highly productive wells would, ceteris paribus, indicate a larger value of the mean productivity level (μ). The relative numbers of high versus low productivity wells reveals the skewness of the distribution (which for the lognormal depends only on σ). Finally, rapid decline in the relative frequency of high versus low productivity wells in the observed historical sequence indicates a relatively small number of potential drilling sites (N). The key to this last effect is the realization that the impact of depletion caused by selective sampling develops more quickly when fewer targets are available.

Illustration: The Bakken Play

The case study is based on drilling results observed in the North Dakota portion of the Bakken play during years 2006-2015. During that period, a total of 12,376 horizontal fractured shale oil wells were drilled in North Dakota. Data on each of these wells comes from the North Dakota Industrial Commission (NDIC).[7] NDIC reports the day on which each well was spudded (i.e., the date when drilling began), so we can place the wells in sequence. Unfortunately, 14% of the wells lack productivity data. Most of the missing values pertain to recent wells that NDIC had not released from "confidential" status when our data were collected. For all other wells, we know the reported Initial Production (*IP*) rate, which measures the average daily flow of oil from

[7] The data can be accessed through the NDIC website: https://www.dmr.nd.gov/oilgas/.

the well over the first 30 days of operation. An overview appears in Table 1.

[Table 1 to appear here]

Several general tendencies are apparent. First, the average productivity of wells has grown steadily since 2006. Selective sampling of the best sites should create the opposite pattern, so the raw data reflect something that has been extensively discussed and documented by petroleum engineers as well as economists: experience gained during these years has increased the effectiveness of fracking operations and allowed operators to extract a larger portion of the oil trapped within a given shale.[8] We will need to account for these efficiency gains before the model can be estimated. To do so requires an adaptation to the successive sampling model that separates the impact of learning from that of depletion.

A second trend visible in Table 1 is that the coefficient of variation in productivity across wells has declined steadily over time. This is what one would expect to result from selective sampling from a skewed population. As selective sampling disproportionately removes drilling locations from the upper-tail, the remaining distribution becomes less skewed over time. This trend should be (and is) visible despite the improvement in fracking efficiency that has occurred since 2006. Suppose, for example, that experience has a uniform impact, raising the productivity of all remaining drilling sites by the same factor. This would increase the mean and standard deviation of the distribution by the same factor, with no impact on the coefficient of variation.

The prevalence of missing productivity data for 2015 wells is also evident in the table. Because we have no way to check whether the few fully reported 2015 wells are representative

[8] Fitzgerald (2015), and Covert (2015) provide estimates of the extent to which learning how to frack has increased well productivity in the Bakken play. Kellogg (2011) and Redlinger, Lange, and Maniloff (2016) provide evidence that learning-by-doing may also reduce the time required to drill and complete an oil well.

of the whole, we exclude all 2015 wells from the estimation procedure. We also exclude pre-2009 wells, primarily because shale oil fracking was very much in its infancy then. Very few wells had been drilled and operators were presumably not so well able to distinguish high potential from low potential drilling sites, which means that our selective sampling postulate may not be applicable in the earliest years.[9]

The potential productivity and relative attractiveness of any given drilling site is determined by physical factors that remain constant through time: total carbon concentration in the underlying shale, plasticity of the rock, thickness of the seam, pressure of the formation, etc. The most desirable sites are endowed with favorable physical characteristics. However, the actual productivity of each well that gets drilled also depends on the choice of fracking techniques, which have evolved quite rapidly during the past decade.[10] Operators have learned through experimentation how production from a given shale can be increased by varying the length of the horizontal section, the number of fractures employed, the volume and pressure of water injected into the formation, and the type of proppant used in the procedure.[11] With the accumulation of experience, a site that produced 500 barrels per day when developed five years ago might produce 750 barrels if developed today. This does not mean the site has gained in relative attractiveness compared to other sites, only that operators' ability to unlock the oil has improved.

Observed *IP* rates signal the quality of the underlying shale but can be used to classify wells based on potential productivity only if these gains in efficiency are taken into account. We

[9] Our results are essentially unchanged if 2008 wells are included in the analysis. Those results are available by request.
[10] Chermak, Crafton, and Patrick (2012) investigate the systematic impact of design and engineering decisions on the productivity of shale gas wells.
[11] Proppant is a sand-like material carried by water and designed to prevent fractures from closing.

will assume a constant trend, with fracking efficiency increasing at the annual rate of $g\%$. Using January 1, 2008 as a benchmark, the quality or "standardized productivity" of each well in the data base is thus determined as follows:

$$x_i = IP_i \times (1+g)^{-w_i}, \tag{7}$$

where w_i measures the difference (in years) between the spud date of the i^{th} well and January 1, 2008. Although this standardized index of productivity is calibrated to 2008 capabilities, the potential productivity of any site can (and will) be restated by incorporating subsequent improvements in fracking efficiency to produce a forward-looking estimate of the remaining volume of technically recoverable shale oil. We assume, as a base case, that the rate of efficiency gains has been 15% per year, but later explore the sensitivity of results to that value.[12]

Our concept of "standardized productivity" is consistent with Covert's (2015) evidence that, given the underlying physical characteristics of each drilling site, the production function of Bakken shale oil wells has remained stable over time. On the other hand, operators' knowledge of that production function has evolved with the accumulation of experience. Covert documents a monotonic trend of increasing efficiency in operators' choice of fracking inputs that progresses at roughly 16% per year. Although there are conceptual differences between Covert's measure and our parameter g, his estimate may be the best indicator available of the rate at which operators' better informed choice of inputs has increased the observed productivity of wells.

Other estimates also indicate large annual efficiency gains due to learning. Looking at a sample of vertical wells in Texas during the pre-shale era (1991-2005), Kellogg (2011) estimates

[12] In principle, g could be estimated along with the other parameters of the model, but that raises a potential identification problem. Consider, for example, the hypothetical possibility that selective sampling causes *IP* to decline at the same rate that learning causes it to grow. In such cases, it could be difficult to estimate either effect. It seems preferable, where possible, to estimate g directly based on engineering studies or industry intelligence and then perform conditional MLE to obtain estimates of the resource parameters. We follow the latter approach.

that the drilling time to complete a well tends to decrease by 12.6% per year as an operator and drilling team accumulate experience together in sustained operations, and that operators accumulate field-specific experience that tends to decrease the time to drill by 5.4% per year independently of the drilling team they hire. Although these results do not pertain to the Bakken basin or to increasing the productivity of wells, they do provide strong evidence that learning-by-doing plays a substantial role in drilling operations.

In a study that focuses specifically on the productivity of hydraulically fractured horizontal wells drilled in the Bakken basin during 2011-2013, Fitzgerald (2015) also finds strong evidence of learning-by-doing. He estimates the elasticity of well productivity with respect to operator experience is 6.19%, and with respect to drilling contractor experience is 5.63%. These give an elasticity of 11.82% with respect to the combined experience of operator and contractor. Fitzgerald also reports that the observed annual increase in well productivity during 2007-2013 is close to 11%,[13] which may constitute a lower bound on the impact of learning-by-doing in the Bakken since depletion of the most productive sites would have offset some of the gains in productivity. Fitzgerald also observes that learning by doing is subject to diminishing returns, which means the annual rate of efficiency improvements may be declining through time.

Estimation Results

The complete sample of wells included in our analysis, categorized by year and productivity tier, is summarized in Table 2.[14] The relative frequencies of high productivity wells (tiers 5 and 6) decline steadily after 2009—due we argue to selective drilling and depletion of the

[13] Personal correspondence.
[14] As in Smith (1980), the tiers are defined by a geometric series, sized to ensure that at least 5% of the observed sample falls within each tier.

most favorable sites. The relative frequencies of lower productivity wells (tiers 1-4) increase steadily, but even at the end of the sample, the lowest productivity wells (tiers 1-2) are rare.

[Table 2 to appear here]

These trends are clear when the data are charted, as in Figure 1. Because the figure is based on standardized productivity, the impact of efficiency gains has been removed so the trends seen in Figure 1 are due solely to the effect of selective sampling without replacement.

[Figure 1 to appear here]

Using the complete sequence of wells, we calculate the likelihood of the sample conditional on selected values of the three population parameters (N, μ, σ). The surface is well behaved and a grid-search is used to identify the maximum likelihood estimates $(\widehat{N}, \hat{\mu}, \hat{\sigma})$, which are shown in Table 3 along with asymptotic standard errors and the estimated variance-covariance matrix.[15]

[Table 3 to appear here]

The number of potential drilling sites existing in the North Dakota portion of the Bakken shale as of January 1, 2009 is estimated to be about 57,000. The estimated moments of the distribution of $ln(X)$ are $\mu = 5.06$ and $\sigma = 1.06$. Converted into natural units (X), these values imply mean well productivity (IP) of 276 barrels per day, with standard deviation 397.[16] These results describe the distribution of well productivity stated in terms of the 2008 benchmark. Stated in terms of 2015 fracking efficiency, the mean and standard deviation rise by the factor (1.15^7), to 735 and 1,058 barrels per day, respectively. The estimated distribution is highly skewed with a long upper tail, as shown in Figure 2. The coefficient of variation is 1.44.

[15]The variance-covariance matrix is obtained from the observed information matrix, which is approximated by a discrete calculation evaluated at the ML estimate.

[16]For the lognormal: $E[X] = e^{\mu+\sigma^2/2}$ and $stdev[X] = E[X] \times \sqrt{e^{\sigma^2} - 1}$.

[Figure 2 to appear here]

Although we estimate that the North Dakota section of the Bakken is comprised of 57,000 drilling sites, that estimate is accompanied by great uncertainty with a standard error that approaches 900 sites. Indeed, the actual uncertainty is even greater due to uncertainty regarding the presumed rate of efficiency gains, which is ignored in our conditional MLE approach. However, great uncertainty also attends the latest USGS (2013a) assessment of the Bakken play, which puts the most likely number of remaining North Dakota shale oil drilling sites (as of January 2013) at roughly 43,000, but ranging anywhere from 30,000 to 70,000.[17] To compare to USGS, we must deduct from our 57,000 sites the number that had already been drilled by January 2013 (about 5,000), which leaves 52,000. This is 20% higher than USGS's modal estimate, but still near the center of their confidence interval.

In contrast to the USGS estimate, EIA (2015) indicates the remaining number of Bakken drilling sites may be as high as 155,000 (including Montana), or as low as 32,000.[18] The North Dakota Department of Mineral Resources estimated, as of mid-2013, that an additional 40,000-45,000 shale oil wells would eventually be drilled in the State.[19] Uncertainty regarding the number of drilling sites is manifest.

It would be good, of course, if our method could produce a precise estimate of the number of potential drilling sites, but it is hardly unique in failing to do so. Having said that, our estimates of μ and σ are quite robust with respect to even large variations in $\widehat{N}$. For example, the

[17] These numbers are calculated by the author from the USGS data (see USGS 2013b). The most likely estimate of productive acreage was divided respectively by the minimum, mode, and maximum estimates of drainage area per well.

[18] We obtain 155,000 by multiplying EIA's estimated "area with potential" by their reported "average spacing" (wells/square mile) as shown in Table 9.3 of EIA (2015). The lower figure comes from Table 9.5, a few pages later, which lists the "number of potential wells" in each North Dakota county.

[19] Information taken from: www.dmr.nd.gov/oilgas/presentations/NDOGCPC091013.pdf.

conditional estimate of $\hat{\mu}$ varies only between 5.10 and 5.04 as the assumed value of *N* is raised from 45,000 to 65,000, and $\hat{\sigma}$ is constant over that entire range. Our knowledge of average well productivity, the range of variation, and skewness across drilling sites is therefore much more precise than our estimate of the total number of sites.

To facilitate economic analysis, it is necessary to categorize by productivity the estimated number of drilling sites that remain. This is not possible using the methods employed by USGS and EIA. It is possible given our approach. Table 4 compares the estimated number of initial sites within each tier $\left(\hat{N}_k\right)$ to the number of wells that have already been drilled. The difference forms an estimate of the number of sites within each tier that remain to be exploited. Of the estimated 57,000 sites, more than 47,000 had not been tapped as of year-end 2014. The large majority of untapped sites belong to the lower tiers, but despite selective sampling and prior depletion of the most productive sites, 55% of the original tier 5 and 6 sites remain untapped.

[Table 4 to appear here]

Table 4 also shows our estimate of the total remaining volume of technically recoverable resources within each tier of the Bakken play, calibrated to the 2015 standard. These estimates are based on the relationship between "expected ultimate recovery" per well (*EUR*) and *IP*. EIA (2014) provides estimates of *EURs* and corresponding *IPs* in each Bakken county. Across all North Dakota counties, *EUR* averages 610 times the *IP* rate.[20] This provides the conversion factor for our calculation. By multiplying the average *IP* rate for each tier by the factor 610 and then multiplying that result either by the estimated number of initial or remaining sites within each tier, and again applying the factor (1.15^7) to account for efficiency gains, we estimate the

[20] Hypothetically, it would take 610 days of production at the initial rate to extract the entire recoverable reserve. In fact, it would take much longer because the rate of production declines as the well matures.

remaining volume of technically recoverable shale oil resources in North Dakota to be roughly 17.0 billion barrels based on 2015 fracking efficiency.[21]

Due to the extent and success of selective sampling, many of the remaining drilling sites belong to the lower tiers. Whereas 54% of the initial deposition fell into highly productive tiers 5 and 6, only 42% of the remaining resources are found there. It remains to be seen whether further improvements will raise future *IP* rates of tier 3 and tier 4 sites to match or exceed the historical productivity of the partially depleted tier 5 and 6 sites. If so, this would potentially offset the negative economic impact of resource depletion.

Our estimate of the aggregate volume of technically recoverable reserves remaining in the Bakken play (17.0 billion barrels) can be compared to assessments prepared using quite different methods by USGS and EIA. USGS (2013) estimates the remaining volume of technically recoverable reserves in the Bakken play to be 7.4 billion barrels. But that calculation is based on 2013 fracking efficiency and also includes the portion of the Bakken that is located in Montana. Applying two years of efficiency gains and eliminating 27% that originates in Montana moves the USGS figure to 7.1 billion barrels, which is still less than half our result.[22] USGS's 90% confidence interval, similarly adjusted, extends from 4.3 to 11.0 billion barrels. Even their upper bound is far below our estimate. In contrast, EIA (2016) estimates there are 23 billion barrels of technically recoverable reserves remaining in the Bakken formation, including both North Dakota and Montana. Assuming again that 27% of the total originates in Montana, the portion attributable to North Dakota would be 16.8 billion barrels, which is a close match to our result.

---

[21] Although EIA's conversion factor is three years old, ensuing technical progress has worked to increase both the *IP* rate and *EUR* of individual wells. We are assuming the ratio has remained constant.

[22] According to USGS's analysis, 27% of the potential drilling sites are located in Montana.

One additional check of external validity is possible. The lognormal coefficient of skewness depends solely on the coefficient of variation, which we estimate to be 1.44.[23] USGS (2012) published estimates of lognormal distributions of well productivity for twenty major shale oil basins in the U.S. The coefficient of variation in those distributions averaged 1.47, with standard deviation of 0.49 across basins. While it is true that those estimates refer to the lognormal distribution of *EUR* rather than *IP*, these two measures of well productivity are closely related so it is to be expected that they would be similarly skewed. In that sense, the skewness of our estimated distribution of well productivity falls in the middle of the range previously reported by the USGS.

To summarize, we provide a maximum likelihood estimate of the remaining volume of technically recoverable shale oil resources in the North Dakota portion of the Bakken basin that is informed by selective sampling without replacement and based solely on statistical analysis of historical drilling outcomes. That estimate appears plausible when compared to other estimates produced by altogether different methodologies. But, there is more. Whereas the USGS and EIA methodologies provide only an estimate of the aggregate volume of technically recoverable resources,[24] our method shows how the total volume is distributed between high versus low productivity wells. From an economic perspective, this information is indispensable if we are to estimate the potential volume of future supplies. Production from low productivity sites will not be economically viable if prices are not high enough to pay back the investment required to drill

[23] The coefficient of skewness is given by $\eta^3 + 3\eta$, where η is the coefficient of variation (see Eq. 2.12, Aitchison and Brown, 1966).
[24] USGS previously prepared estimates of the full distribution of well productivity, as seen in USGS (2012), but they have abandoned that practice. All USGS tight oil and gas resource assessments conducted since 2012 are based on the estimated productivity of an average well, with no analysis or estimate of the variance in productivity across wells. The explanation and rationalization of this change is discussed in Charpentier and Cook (2012).

those wells. In the next section, we extend our analysis by applying this principle to calculate the "supply curve" of economically recoverable resources that remain in the Bakken basin.

<u>The Remaining Volume of Economically Recoverable Resources</u>

The resources that remain within each tier of Table 4 will be economic to develop only if the market price of oil is high enough to generate a competitive rate of return. Smith and Lee (2016) have calculated the breakeven price required for Bakken wells as a function of productivity, and we apply those results here.[25] The last row in Table 4 reports the estimated breakeven price that would be required to achieve a competitive return on the least productive well included within each tier. If the market price exceeds the breakeven price, the entire volume of technically recoverable resources within that tier is economically recoverable.

When plotted, as in Figure 3, these data trace the "supply curve" of remaining shale oil resources in North Dakota. It appears that nearly half of remaining technically recoverable resources could be produced economically at prices of $50/barrel and above. A much higher price ($130/barrel) would be required to recover 75%. Beyond that point, marginal cost rises quickly and the curve becomes highly inelastic. A market price of $100/barrel would facilitate development of about 11 billion additional barrels of North Dakota shale oil. These conclusions presume that fracking efficiency remains at the 2015 level. Further advances in drilling and fracking procedures could increase the productivity of future wells, thereby lifting thousands of additional drilling sites above the breakeven price levels we show here. This is in addition to the

[25] Breakeven calculations are based on a conventional DCF model that includes all investment, operating, and fiscal costs (with the exception of land costs which are treated as resource rent) plus an 8.5% rate of return. We apply that model here with one revision: *EUR* is assumed to equal 610 times the *IP* rate for each well (which is based on the EIA analysis we cited and incorporated earlier in this paper) as opposed to 852 times the *IP* rate (which was based on another source). We have adopted Smith and Lee's "Bakken Core Dynamic Cost" scenario, which entails higher capital costs but lower operating costs than the "Bakken Non-Core" scenario, but this choice has no material impact on our results.

increase in production that would flow from those sites that are already above the bar. The method applied in this paper accounts for additional supplies that emanate from both margins and could be used to trace the implications of various hypotheses regarding the efficiency of future operations.

[Figure 3 to appear here]

Sensitivity Analysis

The results presented so far are based on the assumption that fracking efficiency has increased by 15% per year—roughly consistent with Covert's estimate of efficiency gains due to industry's progressively better informed choice of fracking inputs. In this section, we investigate the sensitivity of our results to this parameter. As shown in Table 1, average *IP* levels increased by 11.4% per year from 2008 to 2014. This represents a hard lower bound for the rate we seek because it compounds the positive impact of learning with the negative impact of resource depletion. We therefore focus on the upside: what are the implications for our results if annual gains in efficiency have actually been 20% rather than 15%? What would be the effect on the estimated number of remaining drilling sites and the volume of remaining resources? And what portion of those resources could be recovered economically at various price levels? The answers are indicated in Table 5.

[Table 5 to appear here]

The assumed rate of increase in fracking efficiency has a profound impact on the results. A higher rate causes the estimated number of potential drilling sites to fall from 57,000 to 37,500. Recall that the USGS confidence interval extends from 30,000 to 70,000, so neither result is implausible. The difference between the two can be explained by reference to the selective sampling hypothesis. When assuming that fracking efficiency has increased at a higher rate, we necessarily demote to lower tiers many of the recent wells in our database that would otherwise have appeared to be of higher standardized productivity. These wells are understood to have achieved high *IP* rates not due to superior physical characteristics, but because they benefitted from the cumulative effect of rapid learning. Thus, the number of upper tier wells in the historical drilling sequence drops off faster under a higher rate. As noted earlier, more rapid decline in the relative frequency of high productivity wells indicates fewer total drilling sites and faster progression towards ultimate depletion. Efficiency gains offset and tend to mask resource depletion, but given an estimate of those efficiency gains our estimation procedure adjusts the estimated number of sites accordingly.

The assumption of 20% growth has a less pronounced effect on the estimated shape of the distribution. Mean standardized *IP* (stated in terms of the 2008 benchmark) decreases from 276 to 224 barrels per day, which is to be expected because a greater portion of the observed productivity is attributed to technological progress rather than the inherent physical characteristics of the sites. The same factor tends to reduce the estimated coefficient of variation and skewness of the distribution.

As shown in Table 5, faster learning translates into a sharp reduction in the estimated number of original and remaining drilling sites. The initial volume of technically recoverable resources consequently drops from 25.0 to 18.3 billion barrels—a decline of 27%. More important going forward, the remaining volume of technically recoverable resources drops from 17.0 to 10.1 billion barrels—a decline of 40%, with most of that loss concentrated in the highest quality tiers.

Conclusion

The process of resource depletion creates an indirect record of the volume and quality of resources that remain to be developed. We have applied two familiar principles of petroleum exploration to interpret that record in the case of the Bakken shale oil basin. When resource development is undertaken with probability proportional to productivity and without replacement, the observed decline in the relative frequency of the most productive wells permits an estimate of both the number and productivity of remaining drilling sites.

Our estimates are consistent with estimates of the aggregate volume of technically recoverable reserves published by the USGS and the EIA, but provide a more detailed picture of the distribution of resources across high versus low productivity drilling sites. That information, when coupled with estimates of drilling and operating costs, determines the volume of resources that can be economically recovered at various oil price levels; i.e., the economic supply curve of shale oil reserves. The supply of remaining reserves from the Bakken shale appears large, even at the relatively low prices ($50/barrel) currently prevailing. The supply also appears to be quite inelastic—particularly at higher price levels.

Efficiency gains tend to mask the impact of resource depletion and therefore play a key role in our analysis. Operators have learned to extract larger volumes of shale oil from any given site by optimizing the combination of fracking inputs. To make a proper inference regarding the productivity of remaining resources, it is necessary to identify and net out the impact of learning on the observed sample of wells. We show that faster learning produces lower estimates of the number and productivity of remaining drilling sites. This may seem paradoxical unless one remembers that inherent productivity of the shale and enhanced operating procedures act as substitutes in determining the observed *IP* rate. When viewing any sequence of drilling outcomes, more of one implies less of the other.

Our forward-looking estimates of remaining resources are stated in terms of current (2015) efficiency levels. The choice of 2015 serves simply as a benchmark, not a prediction. It is likely that further experimentation and learning will continue to increase fracking efficiency and raise the volumes of shale oil that operators are able to extract from the remaining sites, which would shift the supply curve further to the right. Part of that shift comes from sites that would otherwise have been uneconomic to develop and part comes from incremental production from sites that would have been developed anyway. Because our method of analysis assesses the relative numbers of both types of sites, it accounts for both effects and could be used to explore the implications of faster or slower learning going forward.

The potential for further application and development of the method described here is large. It could be applied to estimate the potential resource supply from dozens of additional shale oil and shale gas plays within the U.S. The data requirements are not onerous since operators have to report on their drilling activities and outcomes to state regulatory agencies. Extensions of the method to embrace further hypotheses (e.g., Weibull or exponential distributions in place of the lognormal) regarding the variation in productivity across sites would be straightforward. A further extension to generalize the notion of "size" would also be of interest. For example, the probability of selection might be assumed to vary with the square root of productivity, etc. It is clear how each of these extensions could be handled within the current framework, but that work is left to the future.

The potential impact of oil prices on selection of drilling prospects might also be considered in future research. After oil prices crashed in late 2014, there was much talk in the industry about “high-grading” drilling programs, which signaled a more deliberate attempt to focus investment on the most profitable and least risky drilling sites. Because our sample ends in 2014 this could not have played an important role in the composition of our sample data. However, future research could investigate this hypothesis by testing to see whether the relative probability of selecting high productivity drill sites increased while prices remained depressed.

<u>References</u>

Aitchison, J. and A. C. Brown (1966). *The Lognormal Distribution*, Cambridge University Press.

Andreatta, G., and Gordon M. Kaufman (1986). "Estimation of Finite Population Properties when Sampling is Without Replacement and Proportional to Magnitude," *Journal of the American Statistical Association*, 81(395): 657-666.

Bickel, Peter J., Vijayan N. Nair, and Paul C. C. Wang (1992). "Nonparametric Inference Under Biased Sampling from a Finite Population," *The Annals of Statistics*, 20(2): 853-878.

Charpentier, R.R., and T. A. Cook (2012). "Improved USGS Methodology for Assessing Continuous Petroleum Resources," U.S. Geological Survey Data Series 547.

Chermak, Janie M., James W. Crafton, and Robert H. Patrick (2012). "Impacts of Hydraulic Fracturing and Completion Decisions on Shale Gas Well Productivity," working paper, October.

Covert, Thomas R. (2015). "Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale," working paper, February 22.

Fitzgerald, Timothy (2015). "Experiential Gains with a New Technology: An Empirical Investigation of Hydraulic Fracturing," *Agricultural and Resource Economics Review*, 44(2): 83-105.

Gordon, Louis (1983). "Successive Sampling in Large Finite Populations," *The Annals of Statistics*, 11(2):702-706.

Horvitz, D. G., and D. J. Thompson (1952). "A Generalization of Sampling without Replacement From a Finite Universe," *Journal of the American Statistical Association*, 47(260): 663-685.

Kaufman, Gordon M., Y. Balcer, and D. Kruyt (1975). "A Probabilistic Model of Oil and Gas Discovery," in (J.D. Haun, ed.), *Methods of Estimating the Volume of Undiscovered Oil and Gas Resources*, (Studies in Geology No. 1), American Association of Petroleum Geologists, Tulsa, OK, pp. 113-142.

Kellogg, Ryan (2011). "Learning by Drilling: Interfirm Learning and Relationship Persistence in the Texas Oilpatch," *Quarterly Journal of Economics*, 126: 1961-2004.

Lee, P. J., and P. C. C. Wang (1983). "Probabilistic Formulation of a Method for the Evaluation of Petroleum Resources," *Mathematical Geology*, 15(1):163-181.

Meisner, J. and F. Demirmen (1981). "The Creaming Method: A Bayesian Procedure to Forecast Future Oil and Gas Discoveries in Mature Exploration Provinces," *Journal of the Royal Statistical Society, Series A (General)*, 144(1):1-31.

Menard, H. W. and George Sharman (1975). “Scientific Uses of Random Drilling Models,” *Science,* 190(4212):337-343.

O’Carroll, Frank M. and James L. Smith (1980). "Probabilistic Methods for Estimating Undiscovered Petroleum Resources," *Advances in the Economics of Energy and Natural Resources*.

Redlinger, Michael, Ian Lange, and Peter Maniloff (2016). “Interfirm Learning Economies in Drilling and Environmental Safety,” Colorado School of Mines Working Paper 2016-10, October.

Rosén, Bengt (1972). “Asymptotic Theory for Successive Sampling with Varying Probabilities Without Replacement, I,” *The Annals of Mathematical Statistics*, 43(2):373-397.

Smith, James L. (1980). “A Probabilistic Model of Oil Discovery,” *The Review of Economics and Statistics*, 62(4): 587-594.

Smith, James L. and Thomas K. Lee (2016). “The Price Elasticity of U.S. Shale Oil Reserves,” working paper.

Smith, James L. and Geoffrey L. Ward (1981). “Maximum Likelihood Estimates of the Size Distribution of North Sea Oil Fields,” *Mathematical Geology,* 13(5); 399-413.

U.S. Energy Information Administration (2014). “Oil and Gas Supply Module of the National Energy Modeling System: Model Documentation 2014,” July.

U.S. Energy Information Administration (2015). “Assumptions to the Annual Energy Outlook 2015,” September.

U.S. Energy Information Administration (2016). “Future U.S. Tight Oil and Shale Gas Production Depends on Resources, Technology, Markets,” *Today in Energy,* August 22.

U.S. Geological Survey (2012). “Variability of Distributions of Well-Scale Estimated Ultimate Recovery for Continuous (Unconventional) Oil and Gas Resources in the United States,” Open-File Report 2012-1118, U.S. Department of the Interior.

U.S. Geological Survey (2013a). “Assessment of Undiscovered Oil Resources in the Bakken and Three Forks Formations, Williston Basin Province, Montana, North Dakota, and South Dakota, 2013,” Fact Sheet 2013-3013, April.

U.S. Geological Survey (2013b). “Input-Form Data for the U.S. Geological Survey Assessment of the Devonian and Mississippian Bakken and Devonian Three Forks Formations of the U.S. Williston Basin Province, 2013,” Open-File Report 2013-1094.

Table 1: Number and Productivity of ND Horizontal Oil Wells, by Year

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Total
Reporting Well Count	**209**	**259**	**541**	**475**	**1,085**	**1,434**	**1,961**	**2,142**	**2,241**	**281**	**10,628**
Average IP (b/d)	**222**	**377**	**617**	**856**	**966**	**970**	**1,116**	**1,292**	**1,180**	**1,320**	**1,063**
Annual Growth, Avg. IP	**na**	**70%**	**64%**	**39%**	**13%**	**0%**	**15%**	**16%**	**-9%**	**12%**	**22%**
Std. Dev. IP (b/d)	**227**	**412**	**648**	**655**	**798**	**776**	**868**	**883**	**784**	**798**	**na**
Coef. Var	**1.02**	**1.09**	**1.05**	**0.76**	**0.83**	**0.80**	**0.78**	**0.68**	**0.66**	**0.60**	**na**
Missing Well Count	**30**	**28**	**46**	**17**	**42**	**44**	**70**	**93**	**356**	**1,022**	**1,748**
Total well Count	**239**	**287**	**587**	**492**	**1,127**	**1,478**	**2,031**	**2,235**	**2,597**	**1,303**	**12,376**
% Missing	**13%**	**10%**	**8%**	**3%**	**4%**	**3%**	**3%**	**4%**	**14%**	**78%**	**14%**
Data from the North Dakota Industrial Commission, including all horizontal oil wells, by spud date, since January 1, 2006.											

Table 2: Productivity Distribution of ND Wells, by Year

Tier	IP Range (b/d)	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Total
1	<50	10%	7%	5%	3%	5%	6%	3%	5%	8%	6%	5%
2	50-100	9%	10%	7%	5%	5%	5%	7%	8%	8%	6%	7%
3	100-200	25%	17%	14%	9%	12%	15%	19%	19%	25%	30%	19%
4	200-400	29%	33%	30%	23%	26%	32%	32%	28%	34%	38%	31%
5	400-800	22%	19%	23%	36%	31%	27%	26%	32%	24%	21%	27%
6	>800	5%	14%	21%	25%	22%	15%	13%	8%	1%	0%	11%

Table 3: Maximum Likelihood Estimates

	μ	σ	N
point estimate:	5.06	1.06	57,000
standard error:	0.008	0.006	889
variance-covariance matrix:	6.45E-05		
	-2.56E-05	3.32E-05	
	-5.84E-06	-8.90E-06	7.90E+05

Table 4: Estimated ND Bakken Technically Recoverable Resource (2015 technology)

	Tier 1	Tier 2	Tier 3	Tier 4	Tier 5	Tier 6	Total
Standardized *IP* Range (barrels/day)	0-50	50-100	100-200	200-400	400-800	> 800	
Standardized *IP* Average (S_k)	27	76	153	299	574	1,171	
Estimated Sites (N_k)	7,946	11,088	14,536	12,613	7,244	3,573	57,000
Observed Sample	365	409	1,154	2,532	2,978	1,900	9,338
Remaining Sites	7,581	10,679	13,382	10,081	4,266	1,673	47,662
Initial Volume (million barrels)	343	1,367	3,603	6,109	6,749	6,787	24,958
Remaining Volume (million barrels)	327	1,317	3,317	4,883	3,975	3,178	16,996
Breakeven Price for Tier ($/barrel)	∞	$1,141	$374	$128	$47	$18	

Table 5: Impact of Assumed Efficiency Gains

Parameter Estimates	Efficiency Gains 15% p.a.	Efficiency Gains 20% p.a.
Lognormal Distribution		
N	57,000	37,500
μ	5.06	4.93
σ	1.06	0.98
2008 Fracking Efficiency		
E[*IP*], barrels/day	276	224
StDev[*IP*], barrels/day	398	284
2015 Fracking Efficiency		
E[*IP*], barrels/day	735	801
StDev[*IP*], barrels/day	1,059	1,018
Coefficient of Variation	1.44	1.27
Technically Recoverable Resources, 2015:		
Initial, billion barrels	25.0	18.3
Remaining, billion barrels	17.0	10.1

Figure 1: Frequency Distribution, by Productivity Tier and Year

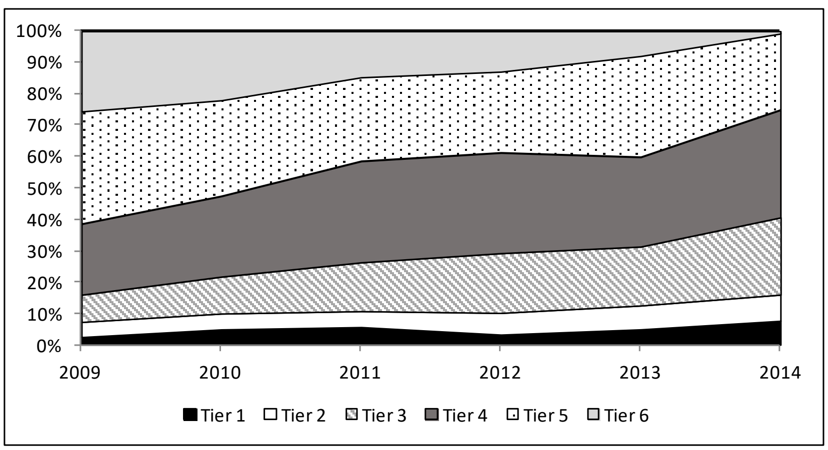

Figure 2: Estimated Productivity Distribution of Original Drilling Sites

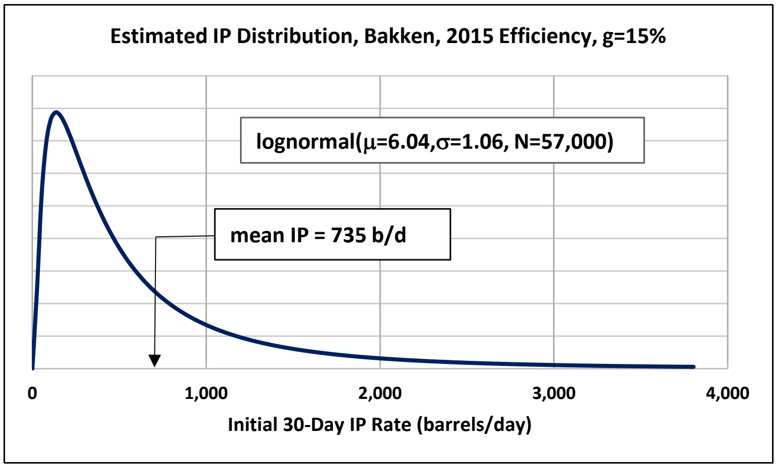

Figure 3: Volume of Economically Recoverable ND Shale Oil Resources