The Salient Distribution of Unconventional Oil and Gas Well Productivity

An MIT Energy Initiative Working Paper

December 2015

Francis M. O'Sullivan[1]*
frankie@mit.edu

Justin B. Montgomery[1,2]
jbm@mit.edu

*Corresponding author

[1]MIT Energy Initiative, Massachusetts Institute of Technology, USA
[2]Department of Civil & Environmental Engineering, Massachusetts Institute of Technology, USA

MIT Energy Initiative, 77 Massachusetts Ave., Cambridge, MA 02139, USA

MITEI-WP-2015-05

Index

Abstract

The economic sustainability of production from shale gas and tight oil resources is coming into sharper focus as oil and gas prices remain low, relative to the prices that spurred extensive drilling programs in recent years. Large variability in the initial production rates of wells—even within similar areas of drilling—is a point of particular concern. In addition to the economic risk this creates for oil and gas companies, the high degree of well-to-well variability makes it a challenge to reliably assess the scale of recoverable resources, the environmental impacts of development, and ultimately the future role of shale gas and tight oil. Here we show that there is a consistent scale-invariant pattern to initial well production rates in major U.S. shale plays, which can accurately be described by a lognormal distribution. This characterization is valid for spatially contiguous well ensembles in both core and non-core areas of fields, for ensembles of different vintages (year of initial production), and for well portfolios of individual operating companies. Recognition of this distribution is an essential step for accurately characterizing the short-term economics, long-term recoverable resource, and environmental impacts of these resources' development.

Introduction

Over the last decade, the combination of horizontal drilling and hydraulic fracturing has greatly expanded North American shale gas and tight oil production, leading to increased estimates of domestic recoverable resources and expectations of a more hydrocarbon-abundant future. [1-5] The term "unconventional," which has long been used to describe these resources, no longer seems appropriate. In spite of this, concerns abound regarding the environmental impact and long-term economic viability of these resources' development. [6-12] With the continuing slump in the oil price, there is growing concern about some of the economic characteristics of the resources, specifically the large well-to-well variability of productivity and rapid well production decline rates. [11-19]

It is becoming increasingly apparent that unconventional resources present the oil and gas sector with a different type of economic risk than it is used to dealing with. [12,15,16] Economic risk in conventionals—reservoirs in which oil and gas freely flows into wellbores without the need for techniques like hydraulic

fracturing—is driven primarily by uncertainty about the size and presence of hydrocarbon accumulations. [20,21] This geological risk is most prevalent early in a field's exploration and appraisal, but as initial wells are drilled, geological data is collected, subsurface models are improved, and uncertainty about production is substantially reduced. [20-24] Economic risk in shale and tight resource plays is more statistical than geological—resources are known to be present continuously throughout the region referred to as a "play" but reservoir rock quality is highly heterogeneous and rates of production from individual wells are unpredictable, even in mature fields with many existing wells. [25-28] The ability to characterize the distribution of well productivity in shale gas and tight oil fields is critical to better understanding the production variability, and thus the economic risk associated with these burgeoning resources. [16]

Historically, the lognormal distribution—which, when its natural logarithm is taken, is Gaussian—has played a prominent role in the extractive industries in describing resource variability, for example the variability of mining deposits or of conventional oil and gas field sizes. [29-31] In fact, lognormality is commonly found in many natural systems due to the salient role of multiplicative effects (as opposed to additive effects which generate more Gaussian distributions), and the lower bounding at zero of values that cannot be negative. [32,33] An example of such multiplicative effects is the amount of petroleum in a reservoir, which is the product of bulk rock volume, porosity, and pore petroleum saturation. A normal distribution for each of these parameters leads to a lognormal distribution for the product. [33] Here, we use probability plotting of historical well production datasets to demonstrate the validity of a lognormal distribution for describing variability in well productivity across several levels of categorization for shale gas and tight oil wells. The utility of this observation is primarily in guiding forecasting efforts and providing a distributional assumption for statistical methods of resource evaluation.

Methods

We consider the distribution of individual well productivity in shale gas and tight oil fields, using each well's peak monthly production rate as our response variable. Although there is additional uncertainty in the production decline rate for each well, this is a widely used metric for early-life well productivity and is

correlated with total lifetime recovery from wells. [13] The first twelve months of production follow the same distribution, and we include an extension of our analysis to this alternate metric in S4-S5 Figs. Shale gas and tight oil resource economics are highly dependent on early-life well production rates: production falls substantially in wells early in their life, by some estimates as much as 50% of the peak production rate after only one year. [34] In addition to absolute productivity of wells, we analyze the specific productivity—a length-normalized index defined as the absolute productivity divided by the length of the productive lateral section.

In order to compare the empirical productivity distribution of well ensembles with the ideal lognormal distribution, we use theoretical quantile-quantile plots, also called probability plots. Specifically, we use a normal quantile-quantile plot to graph the log-transformed standard scores of well peak production rates. This test for distribution shape is not affected by the normalization we performed because it is invariant to location and spread. The normalization allows multiple ensembles to be compared simultaneously on a common scale. We provide a more detailed description of our dataset and methods in the Supporting Information section.

Results and Discussion

The Barnett shale play, the most extensively developed (and thus sampled) shale resource play, exemplifies the lognormality of well productivity (Fig. 1A, 1B). The same pattern is also apparent for the Marcellus, Bakken, Eagle Ford, and Haynesville plays (Fig. 1C, 1D; see also S1-S4 Figs.). For all of these, the correspondence of the empirical distribution to an ideal lognormal distribution is exceptional between the fifth and ninety-fifth percentile. The deviations at the extremes of the data are unsurprising given the tapering of a lognormal probability distribution toward zero in the tails, which tends to exaggerate differences in these low probability regions. [35] However, there may also be physical and economic factors more prevalent at the ends of the distribution. For instance, high density development in productive "sweet-spot" areas is known to lead to interference in which neighboring wells share a pressure drawdown area, reducing individual well output for some of the highest performing wells. [36] This distributional assumption may prove useful for identifying such interactions.
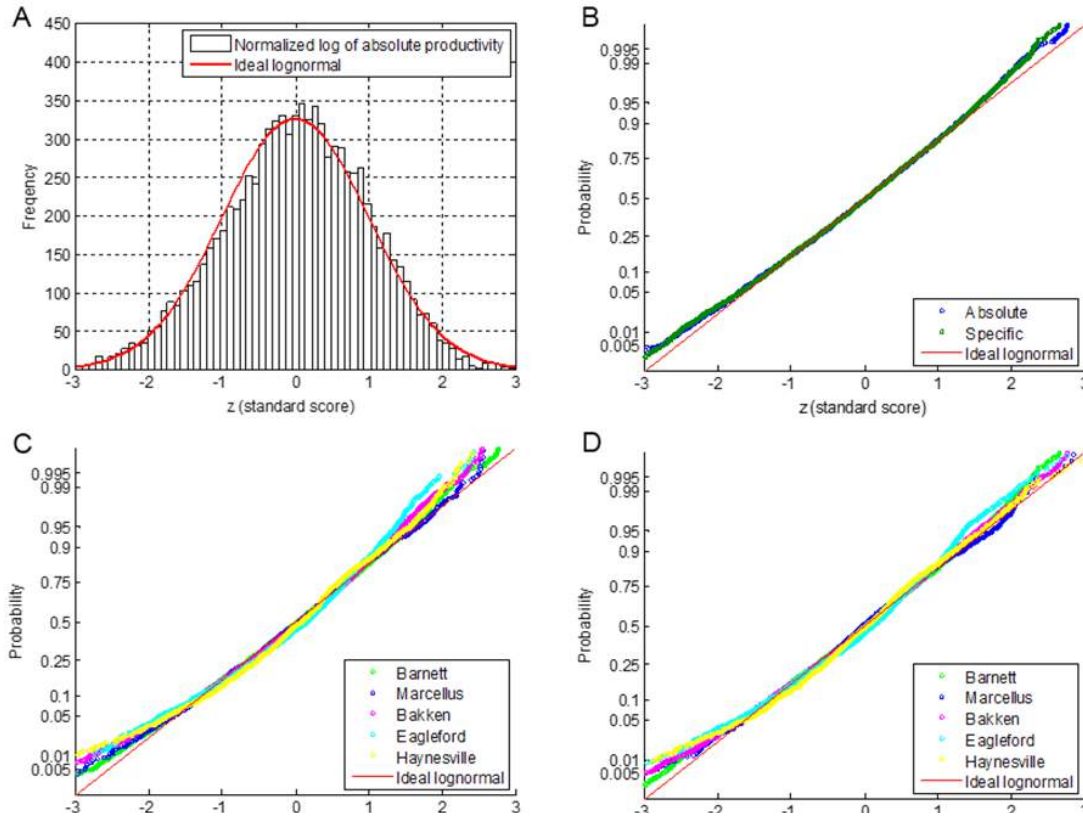
Fig. 1. Lognormal distribution of absolute and specific productivity in Barnett shale. **(A)** Lognormal histogram of well productivity in the Barnett. **(B)** Probability plot comparing normalized absolute and specific productivity in the Barnett to an ideal lognormal distribution. **(C)** Probability plot comparing normalized absolute productivity of five major shale plays to an ideal lognormal distribution. **(D)** Probability plot comparing normalized specific productivity of five major shale plays to an ideal lognormal distribution.

The characterization of production variability as lognormal also applies at various levels within plays. It is a reliable description of variability within individual counties of a play, like the Barnett, including core areas such as Johnson County and lower performing non-core areas like Parker County (Fig. 2A). Although some counties have higher median well productivity, the shape and spread (measured as the P90-P10 ratio) of the distribution is remarkably consistent. Categorizing wells in a play by the length of the perforated section (the length of the productive horizontal section of well) or by operating company yields the same productivity distribution (Fig. 2B, 2C). The inability of operators

to overcome the stochasticity of these resources means the most important strategies left to them are to acquire the best acreage—in order to have a higher average well productivity—and to minimize individual well costs. Additionally, given the probabilistic sampling nature of development, operators will seek to hedge themselves by drilling an abundance of wells, event though many will disappoint, knowing that a larger sample size increases the expected outcome across the total portfolio. [25]
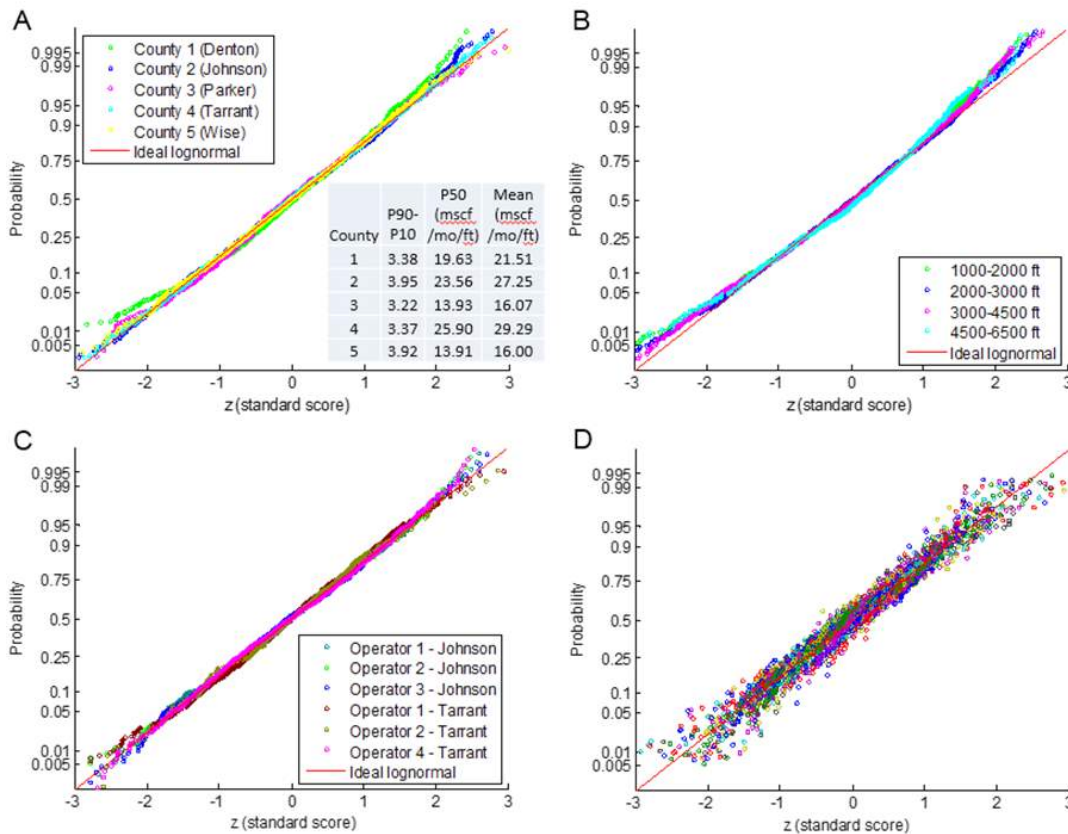

Fig. 2. Lognormal distribution of productivity for different well ensembles within the Barnett shale. (A) Probability plot comparing normalized specific productivity in different counties of the Barnett to an ideal lognormal distribution. Legend includes P90-P10 ratio, P50 (i.e. median), and mean production rates for each county. Units of mscf/mo/ft are 1000 standard cubic feet per month per foot. **(B)** Probability plot comparing normalized absolute productivity for Barnett wells with different perforated well-lengths to an ideal lognormal distribution. **(C)** Probability plot comparing normalized specific productivity for different operators in the Barnett to an ideal lognormal distribution. See S7 Fig. for an extension to other plays. **(D)** Probability plot comparing normalized specific productivity of Barnett wells in ten-square-mile blocks having 8 or more wells to an ideal lognormal distribution.

This probabilistic nature of shale and tight resource development has been recognized by industry in recent years [25] (although it has not made it into the public dialogue about the resource). An unnoticed but important aspect of this resource variability is the scale-invariance of the lognormal pattern—even at the scale of 10 square mile sections (Fig. 2D). The immense heterogeneity present in these rocks precludes the ability to accurately predict the output of a well deterministically, even when many similar wells have been drilled nearby. [25] Based on the universality of the distribution we have found, however, geostatistical techniques like kriging may be useful for assessing and identifying optimal development locations, as has been done in mining to address local variability in ore quality of a similar pattern. [16,37] This will reduce the number of wells required to access a given quantity of resource, improving the economics and reducing the local disturbance and environmental impacts of development.

Many of the most productive areas in shale and tight resource plays have already been developed. [12,15] Thus, much of the debate about the longevity and economic viability of these resources hinges on the amount of improvement that can be made to extraction techniques. Some assert that progress has already been made and point to highly productive recent wells. [38] These gains in efficiency, however, appear to be driven primarily by longer lateral sections and increased stimulation volumes. Specific productivity has generally been decreasing over time and there are diminishing productivity gains (on a per foot basis) in all plays, while wells of all development vintages in a play exhibit the same lognormal distribution of productivity (Fig. 3). It has not been possible for operators to alter the shape of the distribution toward more high-performing wells through either improved hydraulic fracturing design or more adept selection of drilling location. This lack of learning is especially interesting in the Barnett, a mature play that has shown little deviation from this pattern over nearly a decade. There are economic and physical limits on the productivity improvements that can be achieved through the scaling up of development wells and these constraints need to be understood to anticipate future well productivity. It is already apparent in the Barnett that the decline in resource quality as top acreage is expended is outstripping the pace at which the effectiveness of extraction improves.
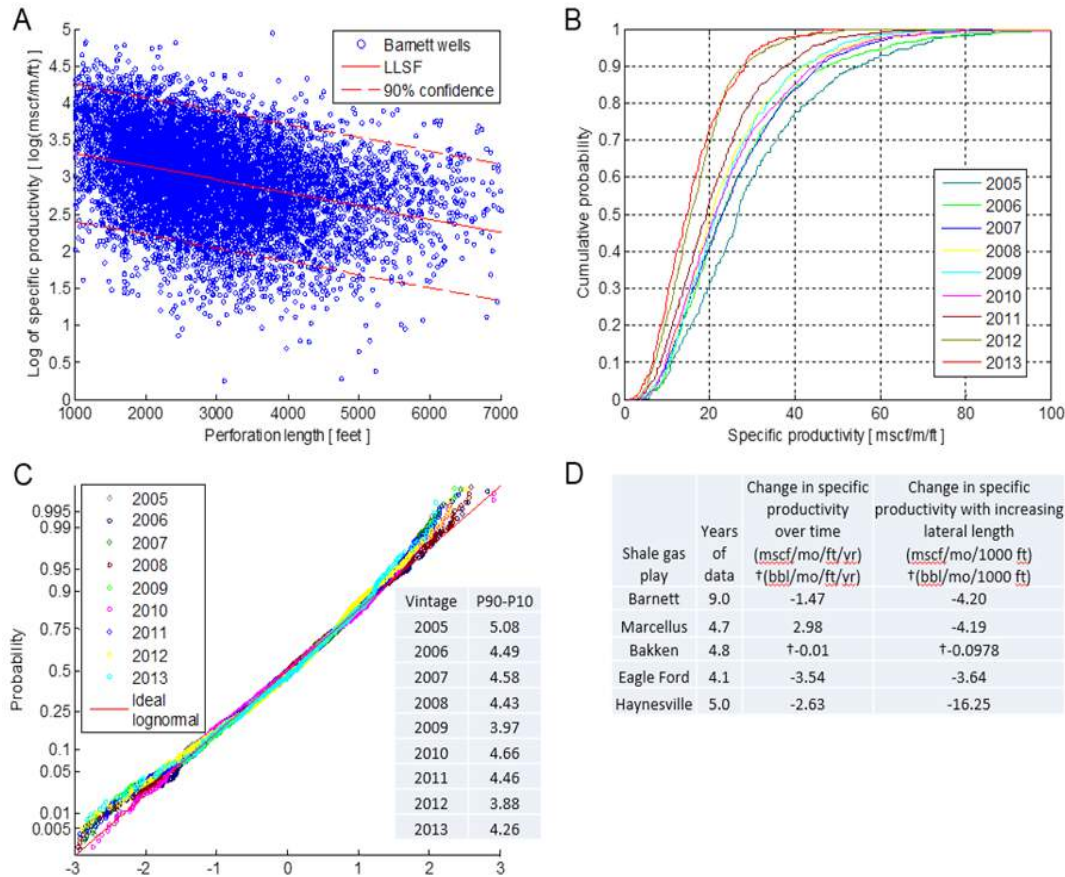
Fig. 3. Longer perforated sections in wells have diminishing marginal output; over time specific productivity is generally declining without a change in distribution. **(A)** Scatter plot with Linear Least Squares Fit (LLSF) and 90% confidence bands for the log of specific productivity and the perforation length for Barnett well ensemble. **(B)** Cumulative distribution plot of specific productivity by vintage in Barnett shows a general shifting of curves toward lower values. **(C)** Probability plot comparing normalized specific productivity of different vintages in the Barnett to an ideal lognormal distribution. Legend includes P90-P10 ratio for each vintage. See S8 Fig. for an extension to other plays. **(D)** Specific productivity has been declining in each play over time (except in Marcellus). Specific productivity declines with increasing perforation length in each play.

Conclusions

Unconventional oil and gas development must be understood in terms of the production variability. Claims that these resources can be "manufactured" understate the economic risk associated with development. [39] A more apt analogy may be slot machines, in which it is critical to quickly and accurately understand the "table odds." Under low oil and gas prices, operating companies will likely focus on fully exploiting fields that have already been "de-risked" through drilling and that have the most favorable distribution parameters, while avoiding risky investments in new areas.

This situation carries with it stark implications from the international to local level regarding resource size, economics, and environmental impact. It is critical to take into account the productivity distribution for different areas when assessing the resource size in unconventionals, and this will not be well understood until a number of wells have been drilled. This casts some doubt on the reliability of shale gas and tight oil production forecasts and has implications for state-level and regional actors planning long-term energy infrastructure. The right-skewed nature of the distribution of well productivity may lead to a tendency to overestimate production, especially when only a few wells are to be drilled. This is particularly important for individual landowners determining the value of signing a lease with drillers on their property. The expected production per well for a few wells on their property is much lower than the expected production per well that the operating company can expect from drilling many wells across the area. Finally, the variability in productivity shown here underlines the importance of considering more than one measure of a field, such as the mean, when assessing the environmental footprint of development. Each field will have some higher-performing wells and a larger number of low-performing wells and the expected environmental impact per well depends on how many wells will ultimately be drilled. The pervasive uncertainty about the economics of fully developing a field may also hinder companies' investment in infrastructure like pipelines—leading to more flaring of stranded gas.

Our analysis demonstrates that large and consistent production variability is a salient feature in shale gas and tight oil. It is time for a renewed consideration of the proper strategies for managing these resources.

Acknowledgements

References

1. *Annual Energy Outlook 2014 with projections to 2040*. 2014, US Energy Information Administration.
2. *Potential supply of natural gas in the United States.*, in *Potential Gas Comm. Bienn. Rep.* 2013, Colo. Sch. Mines: Golden, CO.
3. Deutch, J., *The good news about gas-The natural gas revolution and its consequences.* Foreign Affairs, 2011. 90: p. 82.
4. Joskow, P.L., *Natural gas: from shortages to abundance in the United States.* The American Economic Review, 2013. 103(3): p. 338-343.
5. Yergin, D., *The Quest: energy, security, and the remaking of the modern world*. 2011, New York: Penguin.
6. Jackson, R., et al., *The environmental costs and benefits of fracking.* Annual review of Environment and Resources, 2014.
7. Howarth, R.W., A. Ingraffea, and T. Engelder, *Natural gas: Should frack ing stop?* Nature, 2011. 477(7364): p. 271-275.
8. Brandt, A., et al., *Methane leaks from North American natural gas sys tems*. Science, 2014. 343: p. 733-35.
9. Zoback, M., *Managing the seismic risk posed by wastewater disposal.* Earth, 2012. 57: p. 38-43.
10. Hughes, J.D., *Drill, baby, drill: can unconventional fuels usher in a new era of energy abundance?* 2013.
11. *Documents: Leaked Industry E-mails and Reports*, in *New York Times*. Accessed 18 Aug. 2014, http://www.nytimes.com/interactive/us/ natural-gas-drilling-down-documents-4.html.
12. Inman, M., *The fracking fallacy.* Nature, 2014. 516(7529): p. 28-30.
13. Lee, J. and R. Sidle, *Gas-reserve estimation in resource plays.* SPE Econ. Manag., 2010. 2: p. 86-91.
14. Adams, C., *Oil price fall threatens $1tn of projects*. 15 Dec. 2014, Financial Times.
15. Crooks, E., *US Shale: What lies beneath,* in Financial Times. 26 Aug. 2014.
16. McGlade, C., J. Speirs, and S. Sorrell, *Methods of estimating shale gas resources – Comparison, evaluation and implications*. Energy, 2013. 59(0): p. 116-125.

17.   Adams, C., *Oil price slide puts producers under pressure*, in F. Times. 24 May 2015.

18.   Hilaire, J., N. Bauer, and R.J. Brecha, *Boom or bust? Mapping out the known unknowns of global shale gas production potential*. Energy Economics, 2015. 49(0): p. 581-587.

19.   Ikonnikova, S., et al., *Profitability of shale gas drilling: A case study of the Fayetteville shale play*. Energy, 2015. 81(0): p. 382-393.

20.   Rose, P.R., *Risk Analysis and Management of Petroleum Exploration Ventures*. 2001: American Association Of Petroleum Geologists.

21.   Suslick, S.B. and D.J. Schiozer, *Risk analysis applied to petroleum exploration and production: an overview*. Journal of Petroleum Science and Engineering, 2004. 44(1–2): p. 1-9.

22.   Dake, L.P., *Fundamentals of reservoir engineering*. 1983, New York: Elsevier.

23.   Prensky, S.E., *A survey of recent developments and emerging technology in well logging and rock characterization*. The Log Analyst, 1994. 35(2): p. 15-45.

24.   Mustafiz, S. and M.R. Islam, *State-of-the-art petroleum reservoir simulation*. Petrol. Sci. & Tech., 2008. 26(10-11): p. 1303-1329.

25.   Hall, R., et al., *Guidelines for the practical evaluation of undeveloped reserves in resource plays*. 2010, Society of Petroleum Evaluation Engineers.

26.   Dong, Z., et al., *Probabilistic Assessment of World Recoverable Shale-Gas Resources*. SPE Econ. Manag., 2015. 7(2): p. 72-82.

27.   Charpentier, R.R. and T.A. Cook, *Variability of oil and gas well productivities for continuous (unconventional) petroleum accumulations*. 2013, U.S. Geological Survey.

28.   Cook, T. and D. Van Wagener, *Improving well productivity based modeling with the incorporation of geologic dependencies*. 2014, U.S. EIA working paper.

29.   Kaufman, G.M., Y. Balcer, and D. Kruyt, *A probabilistic model of oil and gas discovery*, in Methods of Estimating the Volume of Undiscovered Oil and Gas Resources. 1975, Amer. Assoc. Pet. Geol. p. 113-142.

30.   Schuenemeyer, J.H. and L.J. Drew, *A procedure to estimate the parent population of the size of oil and gas fields as revealed by a study of economic truncation*. Mathem. Geol., 1983. 15(1): p. 145-161.

31.   Krige, D.G., A *statistical approach to some basic mine valuation problems on the Witwatersrand*. Journ. of the Chemical, Metallurgical, and Mining Society of South Africa, 1951. 52(6): p. 119-139.

32.   Limpert, E., W.A. Stahel, and M. Abbt, *Log-normal distributions across the sciences: keys and clues*. BioScience, 2001. 51(5): p. 341-352.

33.   Blackwood, L.G., *The lognormal distribution, environmental data, and radiological monitoring*. Environmental Monitoring and Assessment, 1992. 21: p. 193-210.

34.   Strickland, R., D. Purvis, and T. Blasingame, *Practical aspects of reserves determinations for shale gas*. Society of Petroleum Engineers North American Unconventional Gas Conference and Exhibition, 2011.

35.   Chambers, J.M., et al., *Graphical Methods For Data Analysis*. 1983, Boston: Wadsorth International Group and Duxbury Press.

36.   Ikonnikova, S., et al., *Well recovery, drainage area, and future drill-well inventory: empirical study of the Barnett shale gas play*. Reserv. Eval. & Eng., 2014. 17(4): p. 484-496.

37.   Cressie, N., *The origins of kriging*. Mathematical Geology, 1990. 22(3): p. 239-252.

38.   Gold, R., *Fracking gives U.S. energy boom plenty of room to run*, in Wall Street Journal. 14 Sep. 2014.

39.   Forbes, B., J. Ehlert, and H. Wilczynski, *The Flexible Factory: The*

Supporting Information

Dataset

The production data in this study was accessed on July 3, 2014 from the drilling info HPDI online database of U.S. oil and gas production data. This is a service that aggregates production data for oil and gas wells in 33 U.S. states, drawing on publicly available repositories managed by the respective states. The information in these databases has been reported by oil and gas operating companies according to state requirements. There are challenges associated with utilizing large well production databases like HPDI because the requirements for the type and quality of data reported vary from state to state. This led us to focus our study on monthly production rates and perforation locations (to determine length of the productive well section), which were reporting requirements in the states we analyzed.

The raw HPDI dataset also contains some erroneous wells which had to be excluded from the analysis because of obvious misreporting (i.e. unreasonable values for some or all data fields). In order to address this issue, relevant data filters were used to obtain a cleaned dataset of active wells producing from the plays of interest. The criteria we used for this is summarized in S1 Table.

Plays in this study

The plays we selected for inclusion in this study—Barnett, Marcellus, Bakken, Eagle Ford, and Haynesville—are key U.S. shale plays and have had high levels of drilling activity in recent years, providing abundant production. They also have been developed similarly enough to warrant comparison, yet have known geological differences. They include a range of reservoir fluid types, from the dry gas of Haynesville to the liquids-heavy gas condensate of the Eagle Ford (we analyzed the gas production rates from Eagle Ford to avoid complicating our analysis with consideration of varying gas-oil-ratios) and the black oil of the Bakken. Some descriptive characteristics of these plays are summarized in S2 Table with additional references.

Normalization of data

The probability density function of a lognormal distribution is:

$$f_X(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad x > 0 \tag{1}$$

In our case, x is the well's peak production rate, μ is the arithmetic mean of the log-transformed peak production rates in the well ensemble, and σ is the standard deviation of the log-transformed peak production rates in the well ensemble.

Although the shape of the productivity distribution for wells is consistent across all well ensembles, the central tendency and spread may vary. In order to compare the productivity distribution shape for different well ensembles, we normalized the production rates within well ensembles. We took the log-transformation of the peak production rate for each well and then calculated the "standard score" for each well relative to the other wells in the ensemble under consideration. The standard score, z, is calculated using the equation,

$$z = \frac{\ln x - \mu}{\sigma} \tag{2}$$

In addition to the absolute peak production rate, we considered the specific peak production rate. In order to calculate the specific peak production rate we used the equation,

$$Q_{spec.,peak} = \frac{Q_{peak}}{D_{perf,lower} - D_{perf,upper}} \tag{3}$$

where Qspec.,peak is the monthly specific production rate in the peak month, Qpeak is the (absolute) monthly production rate, Dperf.,lower is the measured depth of the lowest perforation in the well, and Dperf,upper is the measured depth of the highest perforation in the well. The denominator in this expression represents the productive length of the well that has been perforated.
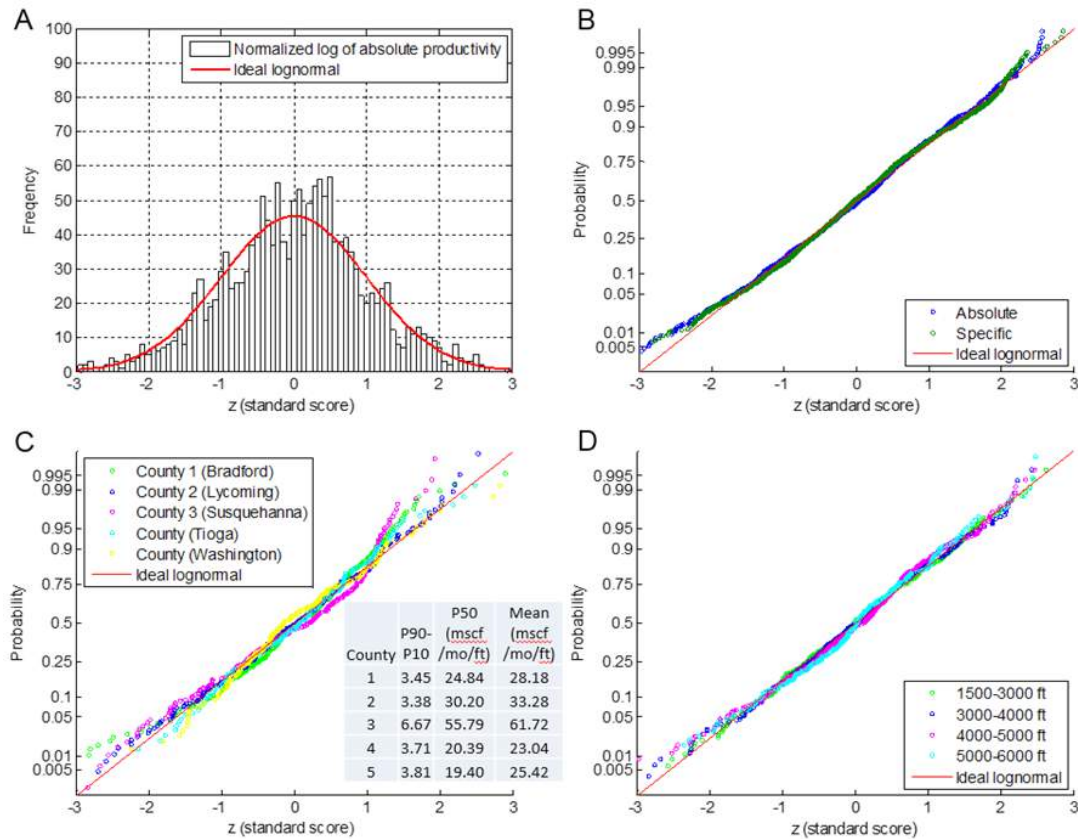
Probability plots

Constructing a probability plot involves rank ordering the samples and plotting each sample's actual value against the theoretical, "ideal" distribution value for the observation. We use the y-axis for the theoretical distribution and the x-axis for the data values. For ease of comparison, the plots have been set to a standard horizontal scale of $z = -3$ to $z = 3$. This pushed some left-hand extreme values off the graph but allows for easier visual inspection over a suitably wide range of the data. All of the data plotting was carried out in MATLAB, using the "probplot" function with the default midpoint probability plotting positions.

The interpretation of probability plots is explored in literature, and we provide only a brief discussion here.[35] A good fit of data with a distribution is indicated by straightness of the points in the plot. In addition to checking for goodness of distributional fit, systematic departures from the line may reveal important information. Often, outliers exist at either end of the data. Additionally, there tends to be greater variation in distribution tails with density that gradually tapers to zero at extreme values (such as normal or lognormal distributions). Defined and systematic curvature at the ends may indicate longer or shorter tails in the data than the ideal distribution. With our axis selection, curvature upward at the left tail or downward at the right tail indicate longer tails at those ends of the distribution. The opposite orientations indicate shorter tails at either end. Asymmetry can also be identified. Convexity of the plotted data indicates that the empirical distribution is more left-skewed than the ideal distribution (and contrarily, concavity indicates right-skewness).[35]

Analysis of blocks of wells

In order to group wells into ten square mile sections, such as those used to generate Fig. 2D, we created evenly-spaced divisions within Johnson County in the Barnett. Uniformly-dimensioned "blocks" of ten square miles were established based on lines of latitude and longitude and the location coordinates for each well in that county were used to associate wells with one of these blocks. Only blocks that had 8 or more wells were included in the analysis. The wells in each block were then treated as a separate well ensemble and normalized prior to comparison in a probability plot.
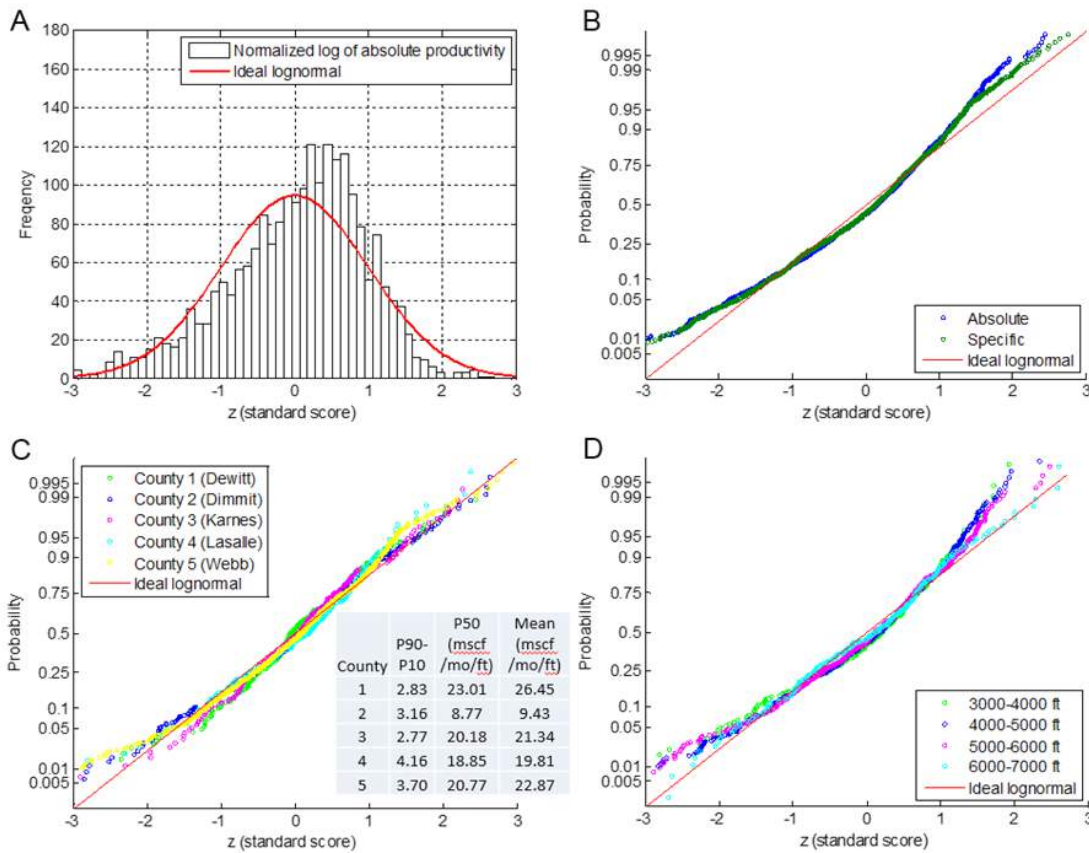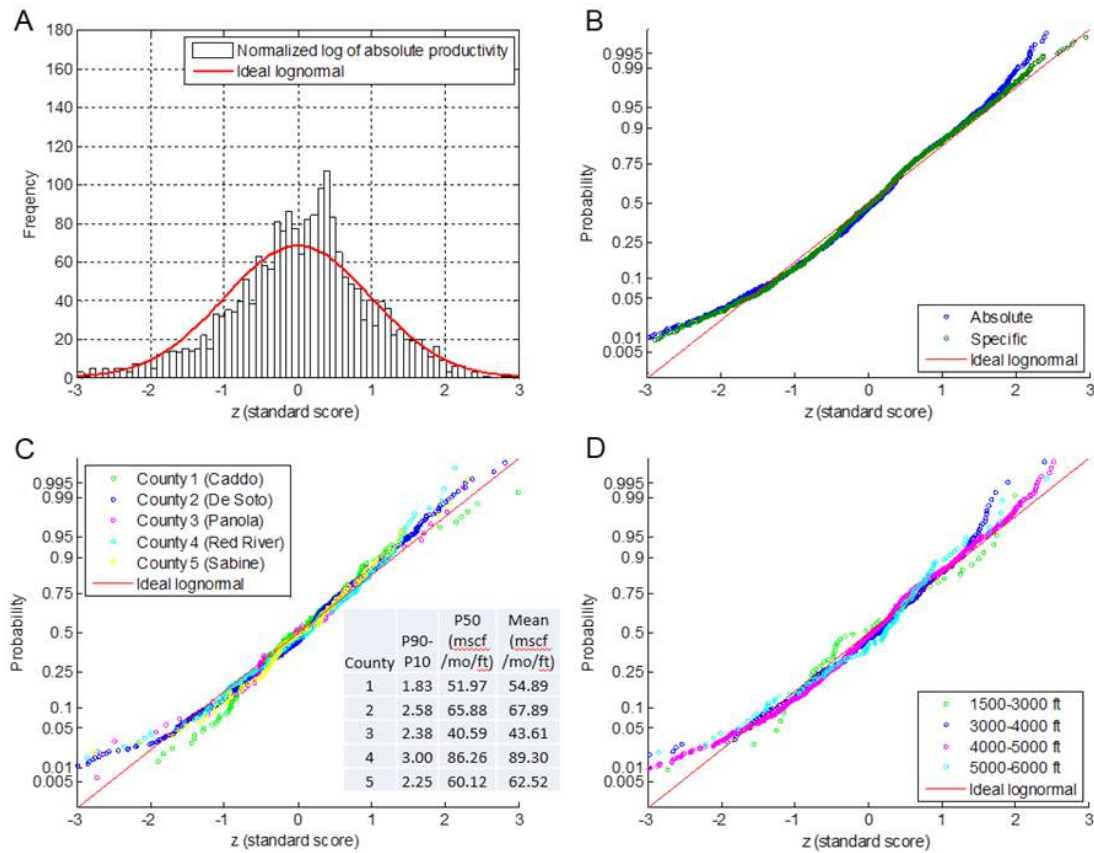
Figures

S1 Fig. Lognormal distribution of absolute and specific productivity in the Marcellus shale. (A) Lognormal histogram of well productivity in the Marcellus. **(B)** Probability plot comparing normalized absolute and specific productivity in the Marcellus to an ideal lognormal distribution. **(C)** Probability plot comparing normalized specific productivity in different counties of the Marcellus to an ideal lognormal distribution. Legend includes P90-P10 ratio, P50 (i.e. median), and mean production rates for each county. Units of mscf/mo/ft are 1000 standard cubic feet per month per foot. **(D)** Probability plot comparing normalized absolute productivity for Marcellus wells with different perforation lengths to an ideal lognormal distribution.
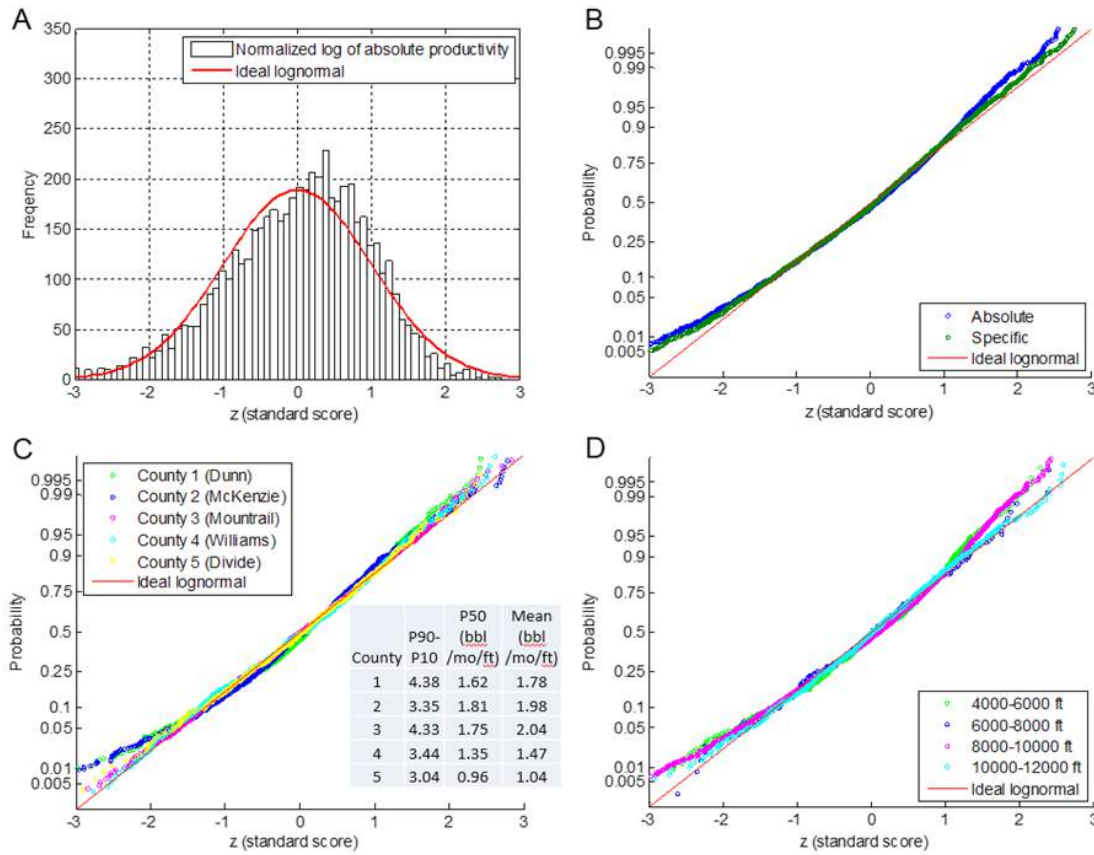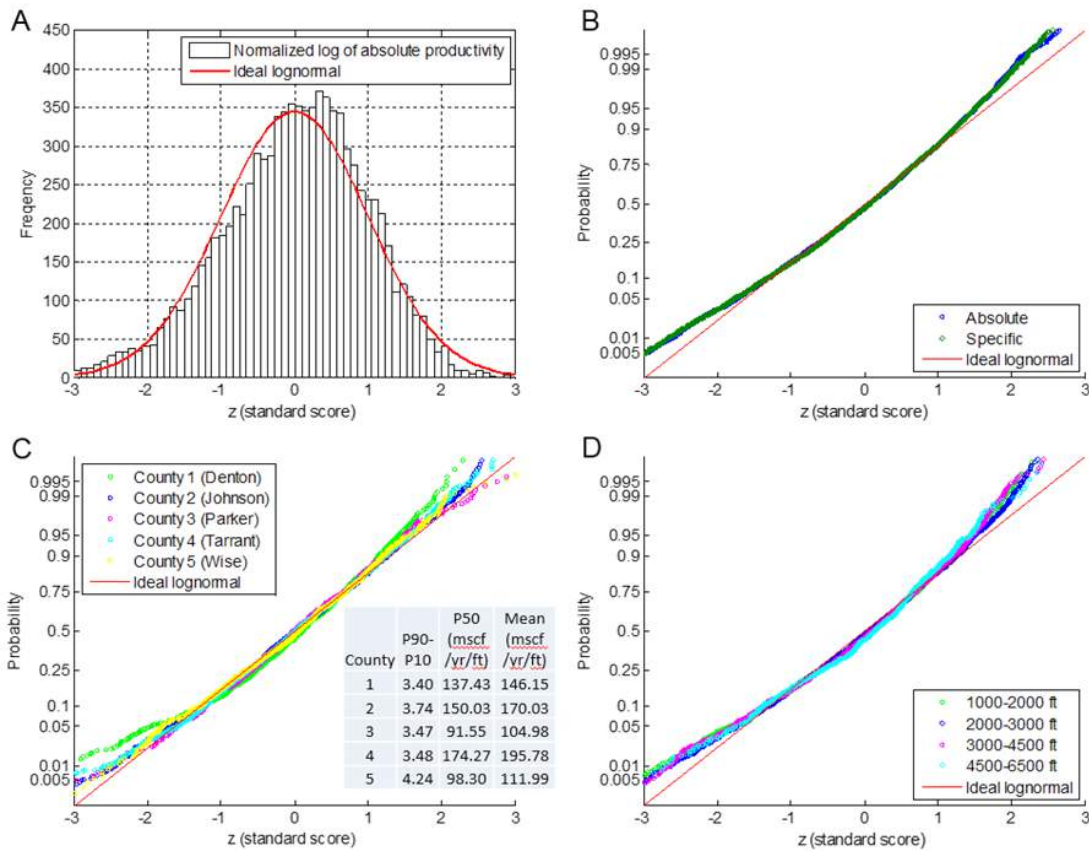
S2 Fig. Lognormal distribution of absolute and specific productivity in the Bakken shale. (A) Lognormal histogram of well productivity in the Bakken. **(B)** Probability plot comparing normalized absolute and specific productivity in the Bakken to an ideal lognormal distribution. **(C)** Probability plot comparing normalized specific productivity in different counties of the Bakken to an ideal lognormal distribution. Legend includes P90-P10 ratio, P50 (i.e. median), and mean production rates for each county. Units of mscf/mo/ft are 1000 standard cubic feet per month per foot. **(D)** Probability plot comparing normalized absolute productivity for Bakken wells with different perforation lengths to an ideal lognormal distribution.
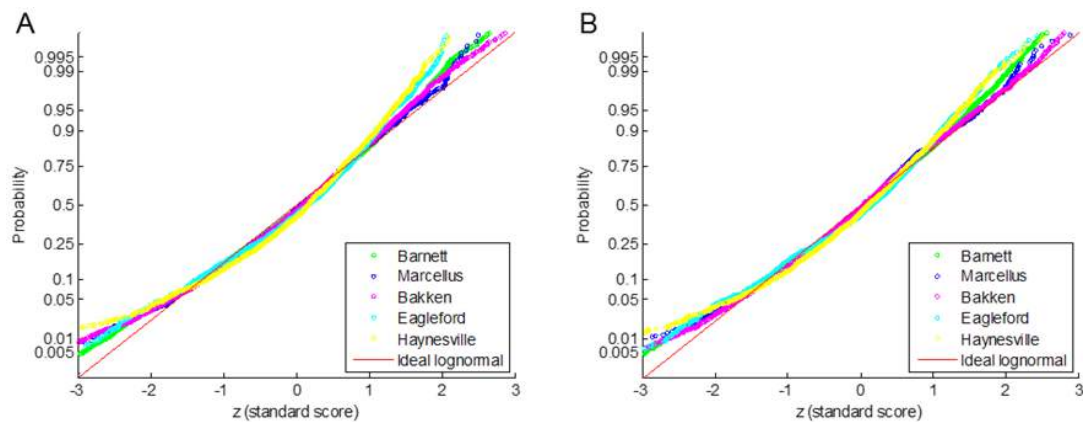
S3 Fig. Lognormal distribution of absolute and specific productivity in the Eagle Ford shale. (A) Lognormal histogram of well productivity in the Eagle Ford. **(B)** Probability plot comparing normalized absolute and specific productivity in the Eagle Ford to an ideal lognormal distribution. **(C)** Probability plot comparing normalized specific productivity in different counties of the Eagle Ford to an ideal lognormal distribution. Legend includes P90-P10 ratio, P50 (i.e. median), and mean production rates for each county. Units of mscf/mo/ft are 1000 standard cubic feet per month per foot. **(D)** Probability plot comparing normalized absolute productivity for Eagle Ford wells with different perforation lengths to an ideal lognormal distribution.
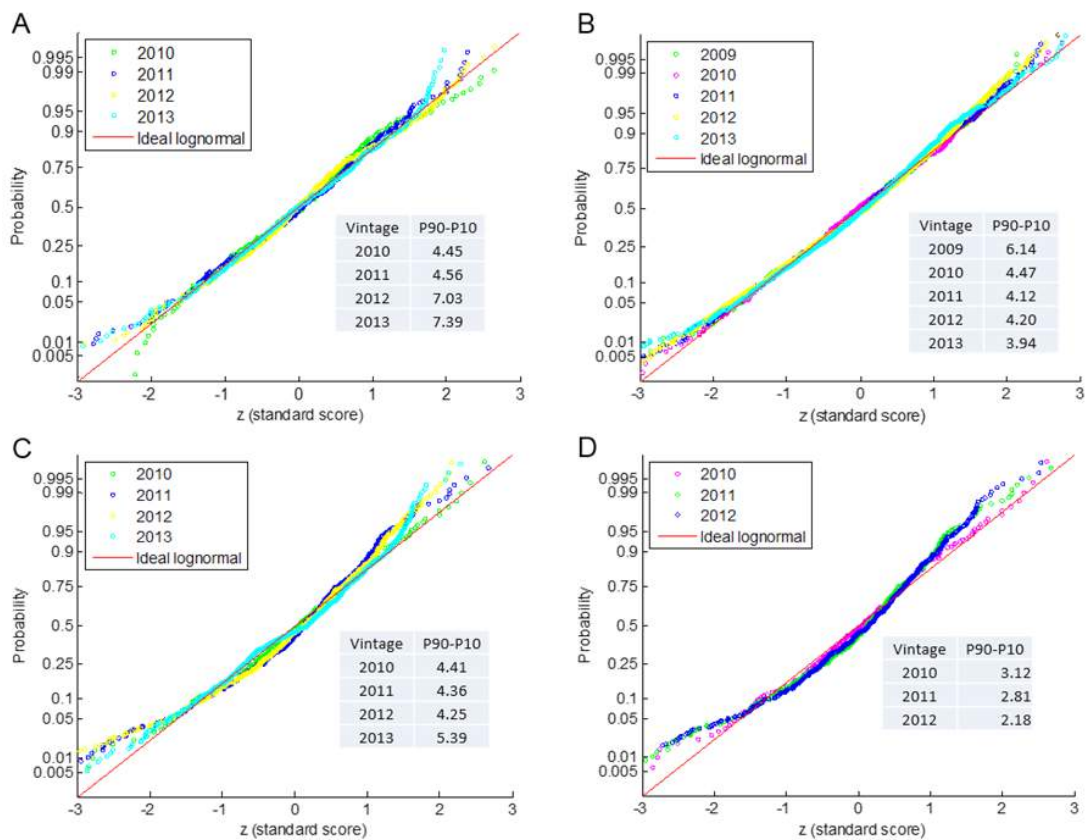
S4 Fig. Lognormal distribution of absolute and specific productivity in the Haynesville shale. (A) Lognormal histogram of well productivity in the Haynesville. **(B)** Probability plot comparing normalized absolute and specific productivity in the Haynesville to an ideal lognormal distribution. **(C)** Probability plot comparing normalized specific productivity in different counties of the Haynesville to an ideal lognormal distribution. Legend includes P90-P10 ratio, P50 (i.e. median), and mean production rates for each county. Units of mscf/mo/ft are 1000 standard cubic feet per month per foot. **(D)** Probability plot comparing normalized absolute productivity for Haynesville wells with different perforation lengths to an ideal lognormal distribution.
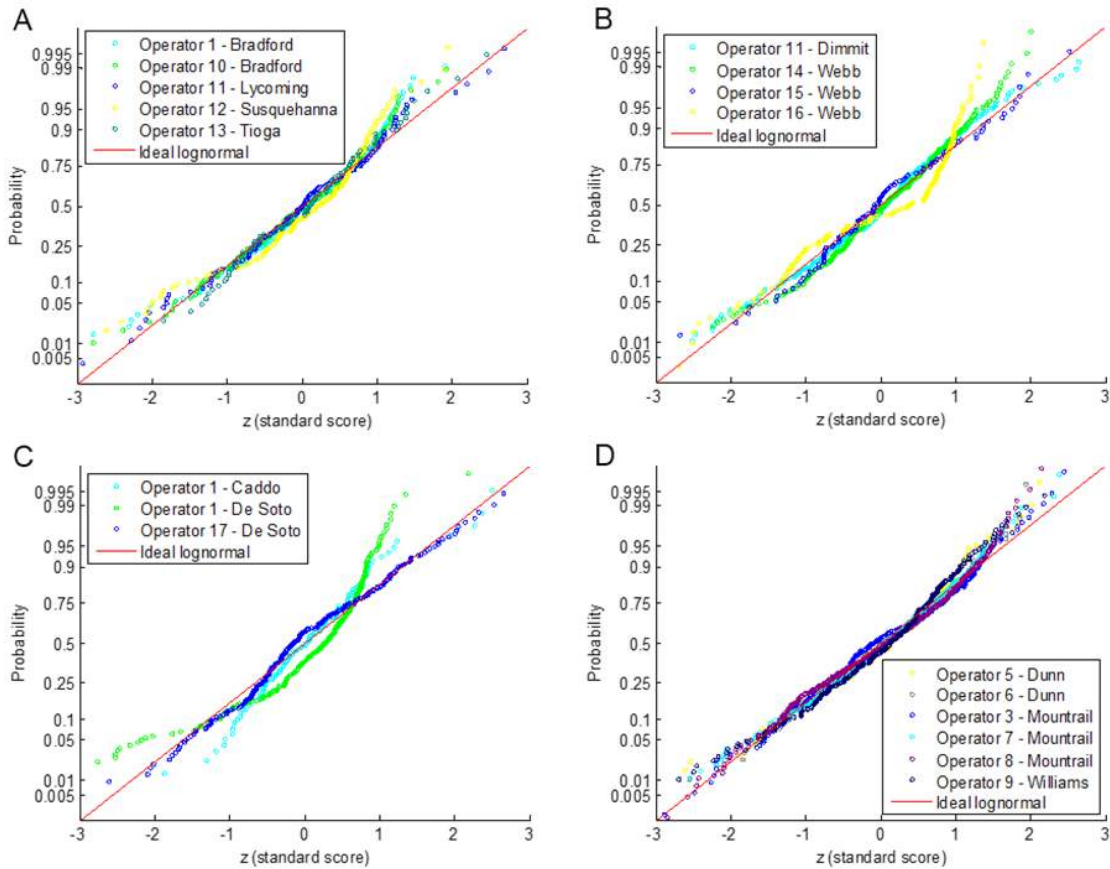
S5 Fig. Lognormal distribution of absolute and specific productivity in the Barnett shale using 12 month gas production instead of peak production rate. (A) Lognormal histogram of well productivity in the Barnett. **(B)** Probability plot comparing normalized absolute and specific productivity in the Barnett to an ideal lognormal distribution. **(C)** Probability plot comparing normalized specific productivity in different counties of the Barnett to an ideal lognormal distribution. Legend includes P90-P10 ratio, P50 (i.e. median), and mean production rates for each county. Units of mscf/mo/ft are 1000 standard cubic feet per month per foot. **(D)** Probability plot comparing normalized absolute productivity for Barnett wells with different perforation lengths to an ideal lognormal distribution.

S6 Fig. Lognormal distribution of productivity across major shale plays using 12 month production instead of peak production rate. (A) Probability plot comparing normalized absolute productivity of five major shale plays to an ideal lognormal distribution. **(B)** Probability plot comparing normalized specific productivity of five major shale plays to an ideal lognormal distribution.

S7 Fig. Lognormal distribution of specific productivity in different vintages of major shale plays. **(A)** Probability plot comparing normalized specific productivity of different vintages in the Marcellus to an ideal lognormal distribution. Legend includes the P90-P10 ratio. **(B)** Probability plot comparing normalized absolute and specific productivity in the Barnett to an ideal lognormal distribution. **(C)** Probability plot comparing normalized absolute productivity of five major shale plays to an ideal lognormal distribution. **(D)** Probability plot comparing normalized specific productivity of five major shale plays to an ideal lognormal distribution.

S8 Fig. Lognormal distribution of specific productivity for different operators in major shale plays. (A) Probability plot comparing normalized specific productivity for different operators in the Marcellus to an ideal lognormal distribution. **(B)** Probability plot comparing normalized specific productivity for different operators in the Haynesville to an ideal lognormal distribution. **(C)** Probability plot comparing normalized specific productivity for different operators in the Eagle Ford to an ideal lognormal distribution. **(D)** Probability plot comparing normalized specific productivity for different operators in the Bakken to an ideal lognormal distribution.

Tables and Data

S1 Table. The plays used in the study include Barnett, Bakken, Marcellus, Eagle Ford, and Haynesville. In order to generate the database of wells in each play for the study (S1 Database) the criteria in this table were used for each play.

Play	Barnett	Bakken	Marcellus	Eagle Ford (Gas)	Haynesville
State(s)	Texas	North Dakota	Pennsylvania	Texas	Texas, Louisiana
Number of wells	9138	5483	1584	2159	2221
Production type	Gas	Oil; Oil and Gas	Gas	Gas	Gas
Well type	Horizontal	Horizontal	Horizontal	Horizontal	Horizontal
Status	Active	Active	Active	Active	Active
Earliest start date	2005	2009	2009	2010	2009
Depths	At least 3,000 feet	At least 5,000 feet	At least 4,000 feet	At least 4,500 feet	At least 10,500 feet
Perforation length	At least 1,000 feet	At least 1,000 feet	At least 1,500 feet	At least 1,000 feet	At least 1,000 feet
Peak rate	Highest rate in first 3 months	Highest rate in first 6 months	Highest rate in first 3 months	Highest rate in first 3 months	Highest rate in first 3 months
Production history	At least 3 months of production	At least 6 months of production	At least 3 months of production	At least 3 months of production	At least 3 months of production

S2 Table. General details about the plays included in the study are summarized here, including the location, geological age, extent, depth, average thickness, total organic carbon, porosity, and technically recoverable resources. This information has been assimilated from a range of references and is only intended to be descriptive of the plays. Many of the values still have a high degree of uncertainty associated with them. References for these plays have been included as well [40-50].

Play	Barnett	Bakken	Marcellus	Eagle Ford	Haynesville
Location	North Texas [49]	North Dakota, South Dakota, and Montana [46]	Ohio, Pennsylvania, New York, and West Virginia [49]	South Texas [49]	East Texas and northwest Louisiana [49]
Geological age	Mississippian [49]	Devonian-Mississippian [50]	Devonian [49]	Cretaceous [49]	Jurassic [49]
Extent, sq mi	6500 [41]	6500 [41]	95000 [41]	3300 [41]	9000 [41]
Depth, ft	7500 [41]	4500-7500 [41]	4000-8500 [41]	7000 [41]	10500-13500 [41]
Average Thickness, ft	300 [41]	22 [41]	125 [41]	200 [41]	250 [41]
Total organic carbon (TOC), %	4.2 [49]	11 [43]	4.8 [49]	3.8 [49]	2.4 [49]
Porosity, %	5 [40]	8 [41]	8 [41]	9 [41]	8.5 [41]
Technically recoverable resources (MMBO, million barrels of oil. BCFG, billion cubic feet of gas. MMBNGL, million barrels of natural gas liquids.)	100 MMBO, 26,700 BCFG, 1,100 MMBNGL [45]	7,400 MMBO, 6,700 BCFG, 500 MMBNGL [46]	84,200 BCFG, 3,400 MMBNGL [47]	900 MMBO, 51,900 BCFG, 2,000 MMBNGL [48]	60,700 BCFG [48]

Supporting information references

40. Montgomery, S.L., et al., *Mississippian Barnett Shale, Fort Worth basin, north-central Texas: Gas-shale play with multi–trillion cubic foot potential.* AAPG bulletin, 2005. 89(2): p. 155-175.

41. EIA, *Review of emerging resources: US Shale gas and shale oil plays.* 2011, US Energy Inf. Admin., Dep. Energy.

42. Bowker, K.A., *Barnett Shale gas production, Fort Worth Basin: Issues and discussion.* AAPG bulletin, 2007. 91(4): p. 523-533.

43. Sonnenberg, S.A., *TOC and Pyrolysis Data for the Bakken Shales, Williston Basin, North Dakota and Montana, in The Bakken-Three Forks Petroleum System in the Williston Basin,* R.M.A.o. Geologists, Editor. 2011. p. 308-331.

44. Hammes, U., H.S. Hamlin, and T.E. Ewing, *Geologic analysis of the Upper Jurassic Haynesville Shale in east Texas and west Louisiana.* AAPG bulletin, 2011. 95(10): p. 1643-1666.

45. USGS, *Assessment of Undiscovered Oil and Gas Resources of the Bend Arch-Fort Worth Basin Province of North-Central Texas and Southwestern Oklahoma, 2003.* 2004, U.S. Department of the Interior, U.S. Geological Survey.

46. USGS, *Assessment of Undiscovered Oil Resources in the Bakken and Three Forks Formations, Williston Basin Province, Montana, North Dakota, and South Dakota, 2013.* 2013, U.S. Department of Interior, U.S. Geological Survey.

47. USGS, *Assessment of Undiscovered Oil and Gas Resources of the Devonian Marcellus Shale of the Appalachian Basin Province, 2011.* 2011, U.S. Department of Interior, U.S. Geological Survey.

48. USGS, *Assessment of Undiscovered Oil and Gas Resources in Jurassic and Cretaceous Strata of the Gulf Coast, 2010.* 2011, U.S. Department of Interior, U.S. Geological Survey.

49. Curtis, M.E., et al., *Microstructural investigation of gas shales in two and three dimensions using nanometer-scale resolution imaging.* AAPG bulletin, 2012. 96(4): p. 665-677.

50. Webster, R.L., *Petroleum Source Rocks and Stratigraphy of the Bakken Formation in North Dakota,* in RMAG Guidebook, Williston Basin, Anatomy of a Cratonic Oil Province. 1987.