12-2018

The Power Law Distribution of Agricultural Land Size

Lauren Chamberlain
Utah State University

THE POWER LAW DISTRIBUTION OF AGRICULTURAL LAND SIZE

by

Lauren Chamberlain

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Economics and Statistics

Approved:

Sherzod Akhundjanov, Ph.D. Major Professor	Man-Keun Kim, Ph.D. Committee Member
Ryan Bosworth, Ph.D. Committee Member	Mark R. McLellan, Ph.D. Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2018

ABSTRACT

The Power Law Distribution of Agricultural Land Size

by

Lauren Chamberlain, Master of Science

Utah State University, 2018

Major Professor: Dr. Sherzod Akhundjanov
Department: Applied Economics

Power-law distributions explain a variety of natural and man-made processes spanning various disciplines including economics and finance. This paper demonstrates that the distribution of agricultural land size in the United States is best described by a power-law distribution. Maximum likelihood estimation is carried out using county-level data of over 3,000 observations gathered at five-year intervals by the USDA Census of Agriculture. Our analysis indicates that U.S. agricultural land size is heavy-tailed, that variance estimates generally do not converge, and that the top 5% of agricultural counties account for about 25% of agricultural land between 1997-2012. The goodness of fit of power-law distribution is evaluated using likelihood ratio tests and regression-based diagnostics. The power-law distribution of farm size has important implications for the design of more efficient regional and national agricultural policies as counties close to the mean account for little of the cumulative distribution of total agricultural land.

(31 pages)

PUBLIC ABSTRACT

The Power Law Distribution of Agricultural Land

Lauren Chamberlain

This paper demonstrates that the distribution of county level agricultural land size in the United States is best described by a power-law distribution, a distribution that displays extremely heavy tails. This indicates that the majority of farmland exists in the upper tail. Our analysis indicates that the top 5% of agricultural counties account for about 25% of agricultural land between 1997-2012. The power-law distribution of farm size has important implications for the design of more efficient regional and national agricultural policies as counties close to the mean account for little of the cumulative distribution of total agricultural land. This has consequences for more efficient management and government oversight as a disruption in one of the counties containing a large amount of farmland (due to natural disasters, for instance) could have nationwide consequences for agricultural production and prices. In particular, the policy makers and government agencies can monitor about 25% of total agricultural land by overseeing just 5% of counties.

.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. Introduction

Agricultural land size plays an important role in understanding U.S. agriculture productivity and wealth, yet it undergoes continuous shifts in distribution and utilization. First, the efficient use of capital and increasing use of production technologies in farming operations have caused a shift in the way the United States operates its agricultural land endowments. This results in larger regions of agricultural land being more productive than their smaller counterparts, a result that holds even after accounting for land scarcity, soil, geography, agrarian structure, and varying forms of agriculture (Adamopoulos and Restuccia, 2014). This has caused a polarization, making counties with larger areas of agricultural land more productive in terms of agricultural sales (MacDonald, Korbe, and Hoppe, 2013).

Second, the distribution of physical agricultural land has shifted over time because of population growth and urban sprawl. The total proportion of agricultural land in the United States has decreased from 44.7% total farmland in 1982 to 40.5% total farmland in 2012 (USDA, 2012). Many counties have very little agricultural land, and farm acreage varies significantly across U.S. counties. In 2012, some counties had less than 200 acres of agricultural land, while others had over 4 million acres (USDA, 2012).

The presence of very large agricultural land in certain counties, the very wide dispersion in farmland size, and the role of the agricultural sector in the U.S. economy make it crucial for policymakers to better understand farmland distribution for effective planning and policy design as well as efficient use of government subsidies and oversight. Despite the importance of quantitative analysis of the distribution of agricultural land area, there is little empirical work on this topic in the literature. In particular, previous studies have examined the distribution of production and sales among farms (Macdonald, Hoppe, and Newton, 2018), the determinants of farm size distribution in country-level wealth (Adamopoulos and Restuccia, 2014), and the growth process of farm size, specifically whether or not the growth process of county level agricultural land size obeys the Gibrat's law of proportionate effect (Mansfield, 1962; Shapiro, Bollman and Ehrensaft, 1987; Clark, Fulton, and Brown, 1992; Brenes Muñoz, Lakner and Brümmer, 2012).[1]

In this paper, we investigate the size distribution of U.S. county-level agricultural land for 1997, 2002, 2007 and 2012, and in each case find that agricultural land area is plausibly characterized by a power-law (Pareto) distribution, meaning the probability that a farm size is more than x acres is proportional to $1/x$. We rigorously examine the goodness of fit of the hypothesized Pareto distribution by employing new regression-based methods (Gabaix and Ibragimov, 2011), robust estimation of upper-tail power-law threshold (Clauset et al., 2009), and fitting alternative distributions. Our analysis provides evidence in favor of Pareto distribution, with estimates remaining robust across different periods, estimation methods, and diagnostic tests, and the distribution fitting the data as good or better than a series of alternative distributions.

The power-law distribution has been used to describe a variety of natural and man-made phenomena.[2] The omnipresence of power laws is derived partly because they are

[1] Gibrat's law of proportionate growth posits that the growth rate of a stochastic process does not depend on its size, but is proportionate to it (Gibrat, 1931). Further, Gibrat (1931) showed that the law of proportionate growth can generate the lognormal distribution for the size of the process. Later, Gabaix (1999) demonstrated that the proportionate growth process can also give rise to power law behavior at the upper tail of the process.

[2] Examples include firm size (Stanley et al., 1995; Axtell, 2001; Luttmer, 2007), city size (Gabaix, 1999, Krugman, 1996; Ioannides and Overman, 2003; Luckstead and Devadoss, 2014; Devadoss et al., 2016), frequency of words (Zipf, 1949; Irmay, 1997), income and wealth (Pareto, 1896; Champernowne, 1953; Wold and Whittle, 1957; Singh and Maddala, 1976; Klass et al., 2006;

preserved over an extensive array of mathematical transformations (Gabaix, 2009). Power law distributions are characterized by extreme kurtosis resulting in heavy (fat) tails, whereby the likelihood of an extreme (upper-tail) event occurring becomes more typical. Therefore, the robust Pareto fit to farmland indicates that U.S. agricultural land size is "heavy-tailed," with a handful of counties accounting for the majority of farmland.

The main contribution of this study is that we show that U.S. agricultural land size is well described by a power-law distribution across different periods (with power-law exponent of approximately 2). This finding is significant for two reasons. First, it becomes inconsequential to talk about average county agricultural land size as this statistic is not representative of the majority of counties; the total farmland is essentially determined by the largest counties (Gabaix, 2009). Focusing on quantile analysis and order statistics instead would be more appropriate in this case. According to our data, the top 5% of U.S. agricultural counties accounted for 25.75% of all agricultural land in 1997, 25.73% in 2002, 25.52% in 2007, and 25.42% in 2012. In contrast, the bottom 50% of counties only held 15.42% of land in 1997, 15.17% in 2002, 14.86% in 2007, and 14.71% in 2012. This heavy share of land concentrated in the top echelons of agricultural counties indicates most of the data is far from the mean, and that observations close to the mean account for little of the cumulative distribution of total farmland. This has consequences for more efficient management and government oversight as a disruption in one of the counties containing a large amount of farmland (due to natural disasters, for instance) could have nationwide consequences for agricultural production and prices. In particular, the policy makers and

---

Toda, 2012), consumption (Toda and Walsh, 2015; Toda, 2017), carbon dioxide emissions (Akhundjanov, Devadoss, and Luckstead, 2017), and natural gas and oil production (Balthrop, 2016), among others. See Gabaix (2009) for a review.

government agencies can monitor about 25% of total agricultural land by overseeing just 5% of counties.

Second, on a more technical concern, “fat tails” of agricultural land size—as suggested by power-law distribution—have significance for empirical research. Statistical analysis based on thin-tailed distributions (such as the normal) might dismiss extremely large agricultural land sizes as an outlier or improbable observation. However, when the distribution is appropriately characterized (as one with large kurtosis), it is apparent that the majority of farmland exists in the upper tail, which need not be discounted in empirical analysis. For example, when using ordinary least squares regression, all dependent variables are assumed to be normally distributed, an assumption also made when using maximum likelihood estimation and other methods. Because this research shows the power law is a fitting distribution for agricultural land size, it has implications that further research must make this consideration rather than assuming normality. Admittedly, given the near infinite number of distributions to choose from, it is likely there is a distribution that provides better fit to the data than a power law distribution. However, the power-law distribution analyzed here is able to capture the main features of the data (heavy tails) parsimoniously.

The remainder of the paper is organized as follows. The next section describes the data used in the analysis. Section 3 describes the methodology, briefly discussing maximum likelihood, regression-based estimation methods, and goodness of fit tests. In Section 4 we present the results and analysis for the size distribution. Section 5 provides concluding remarks.

II. Data

A considerable amount of data is required to identify a power-law-distributed process. This is because much of the power law behavior takes place in the tails of the distribution, where there are often the least number of observations. County-level agricultural land (in acres) data used in this study is obtained from the USDA Census of Agriculture (USDA, 1997, 2002, 2007, 2012). The National Agricultural Statistics Service (NASS) directs the census and obtains information from all U.S. farms. The census was obtained in inconsistent four, five, and ten-year intervals until 1982, after which constant five-year intervals began. The data used in this study is for 3,009 U.S. counties within 50 U.S. states, for four time periods: 1997, 2002, 2007 and 2012. We choose four periods to demonstrate the robustness of power law analysis to time period under consideration.

Table 1 provides summary statistics for our data set. Both the mean and median agricultural land size has decreased over time, with the former surpassing the latter in every year, indicating a heavy right tail. The sample estimates for skewness and kurtosis are large and positive, which suggests large weight in the tails of the distribution. This is also evident from the kurtosis of the kernel density of the data in Figure 1. It is apparent that the distribution is heavily left skewed with the upper tail of the distribution contributing to the majority of the data.

TABLE 1—SUMMARY STATISTICS

	1997	2002	2007	2012
Mean	308,874	303,463.4	298,412.1	295,848.1
Median	197,781	193,386	186,999	186,154
St. dev.	374,783.7	369,732.3	362,277.3	359,702.7
Skewness	3.49	3.62	3.52	3.52
Kurtosis	20.62	23.14	22.38	22.41
Min	90	40	187	143
Max	3,915,165	4,595,062	4,502,752	4,323,178
95% Quantile	953,098.6	935,806.8	946,962.4	958,518.8
Sample size	3,009	3,009	3,009	3,009

Notes: Only counties with observed agricultural land size in each time period were included in the analysis. The agricultural land size is in acres. Source: USDA Census of Agriculture 1997, 2002, 2007, 2012.

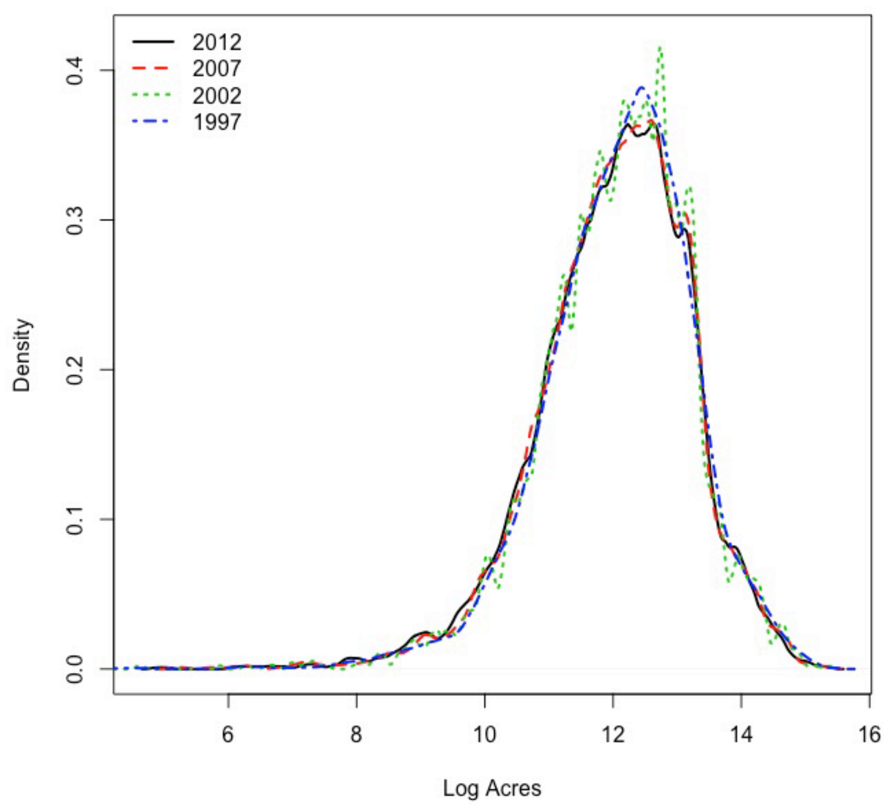

FIGURE 1. THE EMPICAL DISTRIBUTION OF U.S. COUNTY-LVEL AGRICULTURAL LAND SIZE.
THE STYLE IS NAMED FIGURE TITLE

Notes: The empirical distribution is obtained using kernel density with Epanechnikov kernel and the smoothing bandwidth based on unbiased cross-validation method.

III. Methodology

The methodology is split into three parts. First, we describe maximum likelihood and regression-based techniques to obtain power-law parameter estimates. Second, we discuss goodness-of-fit tests used to verify power-law behavior in farmland distribution. Third, we provide additional robustness checks, comparing the power law distribution fit to alternative distributions, including lognormal and exponential distributions.

3.1. Power Law Parameter Estimation

The probability distribution function (PDF) of a power law is given by

$$f(x) = \frac{\alpha - 1}{x_{min}} \left(\frac{x}{x_{min}} \right)^{-\alpha}$$

where x is an outcome of random variable (X) of interest (e.g., agricultural acres), x_{min} is the cutoff at which power law behavior takes hold, and α is the power-law exponent (the parameter of interest). Notice that the moment generating function for the power law distribution takes the following form

$$\langle x^k \rangle = \int_{x_{min}}^{\infty} x^k f(x) dx = \frac{\alpha - 1}{\alpha - 1 - k} x_{min}^k$$

for $\alpha > k + 1$. Thus, when $k < \alpha - 1$, only the first $\lfloor \alpha - 1 \rfloor$ moments of a power-law distribution exist. In Section 4, we show that $\hat{\alpha} \leq 3$ for agricultural land size across different periods, which implies that all the moments for x beyond variance are infinite. Clearly, it is possible to compute higher-order moments (e.g., skewness, kurtosis) for any finite sample. However, these sample estimates will not converge to any particular value as the sample size increases.

Given the observed sample $(x_1, \ldots, x_n)$, the joint log likelihood function for power-law distributed process is

$$\begin{aligned}\ln \mathcal{L}(\alpha; x) &= \sum_{i=1}^{n} \left[\ln(\alpha - 1) - \ln x_{min} - \alpha \ln \frac{x_i}{x_{min}} \right] \\ &= n \ln(\alpha - 1) - n \ln x_{min} - \alpha \sum_{i=1}^{n} \ln \frac{x_i}{x_{min}}\end{aligned}$$

Taking first-order condition with respect to α and solving the equation for α yield the maximum likelihood estimate (MLE) of

$$\alpha^{MLE} = 1 + n \left(\sum_{i=1}^{n} \ln \frac{x_i}{x_{min}} \right)^{-1} \quad (1)$$

The associated standard error of α^{MLE} is

$$\sigma = \frac{\alpha^{MLE} - 1}{\sqrt{n}}$$

In the literature, it is conventional to estimate the counter-cumulative parameter, given by $\gamma = \alpha^{MLE} - 1$, instead of equation (1). The estimator of γ is derived from equation (1), after making a small-sample adjustment, and is known as the Hill estimator:

$$\gamma^{Hill} = \frac{n - 2}{\sum_{i=1}^{n-1} (\ln x_i - \ln x_{min})}$$

The associated standard error of γ^{Hill} is $\gamma^{Hill}(n-3)^{-1/2}$.

The power-law exponent can similarly be obtained using regression-based techniques. In particular, the estimate of counter-cumulative parameter can be recovered through the following ordinary least squares (OLS) estimation:

$$\ln(i) = \beta_O - \gamma^{OLS} \ln x_i + \varepsilon_i$$

where (i) is the observation's rank in the distribution, γ^{OLS} is the parameter of interest, and ε_i is the error term. The associated standard error of γ^{OLS} is the asymptotic standard error of the form $\gamma^{OLS}(n/2)^{-1/2}$.

Both MLE and OLS methods rely on the correct specification of an upper-tail threshold parameter x_{min}. In this regard, there are two approaches in the literature. First, Gabaix (2009) suggests setting x_{min} at the 95th quantile of the data. Clearly, this approach is somewhat arbitrary and there is certain level of uncertainty about whether 95th quantile captures the true starting point of power-law behavior. Related to this approach, x_{min} has also traditionally been selected visually, whereby x_{min} is chosen as the point where the PDF or cumulative distribution function (CDF) becomes roughly straight on a log-log plot (Clauset et al., 2009). Both of these methods are evidently subjective and can be sensitive to the noise or fluctuation in the distribution tail. If the arbitrarily selected x_{min} is too small, the estimate of power-law exponent will be biased as it attempts to fit non-power law data to a power law model. If x_{min} exceeds the point where power law behavior begins, it results in discarding valuable data and causes the standard error to increase. Perline (2005) notes that sufficiently truncated Gumbel-type distributions (which include the lognormal) can generate a linear pattern on a log-log plot, thus mimicking the power law distribution.

The second, and preferred, approach is the data-driven procedure that removes the analyst from the selection process. This approach entails treating each data point as a potential candidate for x_{min}, and choosing the optimal x_{min} that minimizes some loss function (e.g., the mean squared error (MSE) of the power-law exponent). Here, we adopt the approach proposed by Clauset et al. (2009), where the best x_{min} is obtained by minimizing the Kolmogorov-Smirnov (KS) goodness-of-fit statistic. The KS statistic is specified as follows

$$KS = \max_{x \geq x_{min}} \left| E(x) - \hat{F}(x) \right|$$

where $E(x)$ is the empirical CDF and $\hat{F}(x)$ is the estimated power law CDF. Hence, the KS statistic measures the discrepancy between the empirical CDF and the estimated power-law CDF for the given candidate of x_{min}. The optimal x_{min} will thus minimize this distance, bringing the estimated distribution as close as possible to the empirical distribution. The algorithm is as follows:

Step 1: Set $x_{min} = x_1$;

Step 2: Estimate power-law exponent (γ^{Hill} and γ^{OLS}) using $x \geq x_{min}$;

Step 3: Calculate the KS statistic;

Step 4: Repeat steps 1-4 for all x_i for $i = 1, \ldots, n$;

Step 5: Choose x_{min} with the lowest KS statistic.

3.2. Goodness of Fit

Having significant parameter estimates alone does not provide sufficient evidence in favor of power-law fit to an empirical data. Additional goodness of fit tests and comparison with alternative distributions need to be carried out. Gabaix and Ibragimov (2011) suggest "rank – 1/2" test to verify the goodness of fit of power law model. First, define x^* as

$$x^* = \frac{\mathrm{Cov}\left[(\ln x_i)^2, \ln x_i\right]}{2\mathrm{Var}(\ln x_i)}$$

Then, estimate the following regression with OLS

$$\ln\left(i - \frac{1}{2}\right) = \alpha + \zeta \ln x_i + q(\ln x_i - x^*)^2 + \epsilon_i$$

The test statistic of interest is given by $\frac{q}{\zeta^2}$. The null hypothesis that agricultural land size is distributed according to a power law is rejected if $\frac{q}{\zeta^2} > 1.95(2n)^{-1/2}$.

3.3. Further Robustness Tests

Clauset et al. (2009) recommend comparing power law fit to the data with those of alternative heavy-tailed distributions, specifically the lognormal and exponential distributions. We adopt this here. The alternative distributions are fit to the upper tail data—similar to power law distribution—where x_{min} is obtained using the KS method discussed in the previous section. Using the fitted distributions, a log-log plot of the data can be constructed to visually evaluate the relative fit of competing distributions.[3]

The relative fit of alternative distributions can be compared more rigorously using the likelihood ratio test suggested by Clauset et al. (2009). The likelihood ratio statistic is given by

$$\mathcal{R} = \sum_{i=1}^{n} \left[\ln \hat{f}_1(x_i) - \ln \hat{f}_2(x_i) \right]$$

where $\hat{f}_1(x_i)$ and $\hat{f}_2(x_i)$ are the probabilities for x_i, $i = 1, \dots, n$, predicted by two competing distributions that are estimated via MLE. In our case, $\hat{f}_1(x_i)$ represents the power-law likelihood of x_i while $\hat{f}_2(x_i)$ represents the likelihood provided by an alternative distribution. The above statistic thus allows for a comparison of the power law

---

[3] The log-log plot is constructed by taking the logarithm of the rank of x in the data and the logarithm of x, and then plotting log rank of x against log x. Note that the counter-(complimentary-) CDF (also known as survival function) for power law distribution is $\text{Prob}(X > x) = k/x^{\gamma}$, where k is a constant. Now, taking the log of both sides of the counter-CDF of power law produces a linear relationship between log counter-cumulative probability and log data (i.e., $\ln x$), with the power-law parameter γ being the slope of the line.

distribution to the lognormal and exponential distributions. A positive value of the likelihood ratio statistic indicates the power-law distribution is the favored fit as it is more likely. In contrast, a negative value indicates the alternative distribution fits the data more closely. See Clauset et al. (2009) for asymptotic properties of and methods to obtain p-values for $\mathcal{R}$.

IV. Results

We implement maximum likelihood and regression-based techniques described in Section 3.1 to fit power law distribution to our data. We first determine the upper-tail threshold point x_{min} using a robust KS method, so that the selected x_{min} achieves the minimization of the KS goodness-of-fit statistic, and then fit power law distribution to the resulting upper-tail data (i.e., $x \geq x_{min}$). Estimation results are reported in Table 2.

The estimates of x_{min} lie far to the left of the 95th percentiles (from Table 1) for the corresponding periods, which indicates that power law behavior takes hold earlier than the 95th percentile. Thus, setting x_{min} arbitrarily at the 95th percentile would entail a loss of valuable data, inflating standard errors. As noted before, identification of a power-law-distributed process is data intensive. In particular, Clauset et al. (2009) recommend the sample size to be at least 50 observations for accurate power law analysis. From Table 2, it is clear that this requirement is far surpassed for all years in our case, with the upper-tail sample size ranging from 587 to 656 observations.

It is apparent that both the Hill and OLS estimates of the power law parameter γ are statistically significant for all years, with a slight difference between the two sets of estimates. The Hill estimates for county-level agricultural land size indicate that the top 5% of counties accounted for 24.54% of agricultural land in 1997, 23.31% in 2002, 22.82% in 2007, and 23.22% in 2012.[4] These parametric estimates are very close to empirical estimates obtained from the raw data,[5] which speaks about the model fit to the

[4] To identify fraction L of the total land held by the top fraction P of counties, use $L = P^{(\alpha-2)/(\alpha-1)}$, which is derived from the complementary-CDF of power law distribution.

[5] According to the raw data, the top 5% of counties accounted for 25.75% of all agricultural land in 1997, 25.73% in 2002, 25.52% in 2007, and 25.42% in 2012.

TABLE 2—POWER LAW PARAMETER ESTIMATES, ESTIMATED x_{min}

	1997	2002	2007	2012
γ^{Hill}	1.883	1.946	1.973	1.951
	(0.074)	(0.081)	(0.079)	(0.077)
γ^{OLS}	2.049	2.097	2.136	2.147
	(0.014)	(0.015)	(0.015)	(0.014)
X_{min}	424,121	449,671	440,462	426,329
Observations	656	587	614	643
Gabaix and Ibragimov (2011) test				
Goodness of fit test statistic	-0.139	-0.139	-0.136	-0.130
Goodness of fit threshold	0.054	0.057	0.056	0.054

Notes: Estimation is based on upper-tail observations ($x \geq x_{min}$), where x_{min} is determined based on the minimization of the KS statistic. Standard errors are in parentheses. For the Gabaix and Ibragimov (2011) test, the null hypothesis that agricultural land size is distributed according to a power law is rejected if test statistic > threshold. Clauset et al. (2009) recommend to have at least 50 observations for accurate power law analysis, a condition satisfied here

.

data. Further, the magnitude of the OLS estimates of γ implies that only the first two moments (mean and variance) of the power law distribution are finite, while the remaining moments are non-convergent. The magnitude of the Hill estimates, on the other hand, is just shy of 2, suggesting that only the first moment (mean) is finite.

The results from the Gabaix and Ibragimov (2011) goodness-of-fit test indicate that we fail to reject the null hypothesis of power-law-distributed agricultural land size across different periods. This provides strong evidence in favor of Pareto distribution fit for U.S. agricultural land size. For completeness, we also perform the analysis for x_{min} set at 95th percentile. See Appendix for the estimation and diagnostic results. Despite a drop in the sample size, our main findings remain qualitatively unaffected even with this approach. Given that x_{min} set at 95th percentile is arbitrary (and that the optimal x_{min} that minimizes the KS-statistic lies to the left of the 95th percentiles), we continue focusing on the estimated x_{min} in the remaining analysis. In Figure 2, we compare the power law fit to the upper-tail agricultural land data with those of alternative distributions using log-log plots. As noted in Section 3.3, the log counter-cumulative probability of power law distribution and log data produces a linear relationship, with the power-law parameter representing the slope of the line. Visually, power law distribution provides better overall fit to the data than competing distributions. In all periods the fitted power law is very accurate for observations located in the lower to mid quantiles of the upper tail, where the observed data forms a straight linear pattern, closely following the power-law fit. In the extreme upper-tail (after log land size of about 14.6), the tail of the Pareto distribution is heavy and decays more slowly, which results in overestimation of the frequency of the largest events

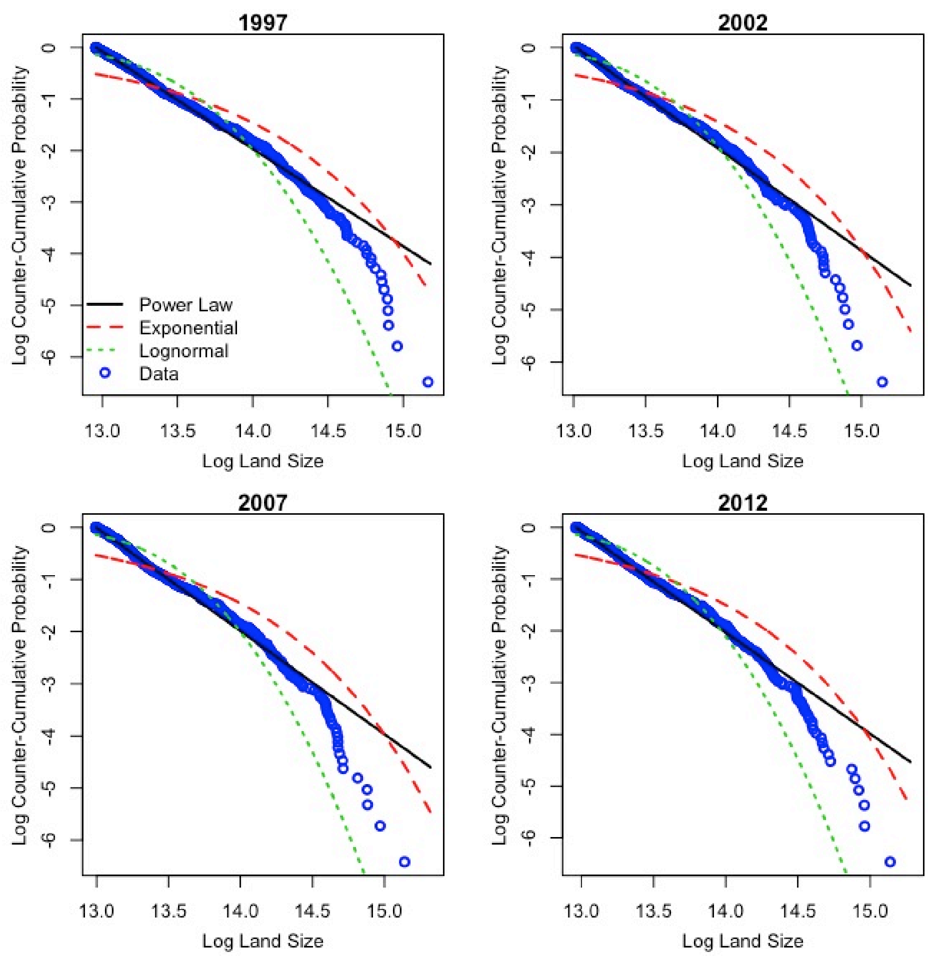

FIGURE 2. PLOT OF EMPIRICAL AND FITTED LOG COUNTER-CUMULATIVE PROBABILITY AND LOG AGRICULTURAL SIZE

TABLE 3—LIKELIHOOD RATIO TESTS OF COMPETING DISTRIBUTIONS

	1997	2002	2007	2012
Power law vs exponential				
Likelihood ratio statistic	503.284	466.060	493.321	509.232
P-value	0.000	0.000	0.000	0.000
Power law vs lognormal	*Power law vs*			
Likelihood ratio statistic	-5.631	-16.137	-30.479	-42.397
P-value	0.799	0.444	0.157	0.051

Notes: Estimation is based on upper-tail observations ($x \geq x_{min}$), where x_{min} is determined based on the minimization of the KS statistic. A positive value of the likelihood ratio statistic indicates that the power law is the better fitting distribution. A negative value indicates the alternative distribution fits the data more closely. P-values are calculated using the methods detailed in Clauset et al. (2009)..

Note that there are other, more flexible forms of the Pareto distribution (e.g., the tapered Pareto) that behave similarly to the Pareto in the lower quantiles, but decay more quickly in the extreme upper tail than the Pareto distribution. These modified Pareto distributions have an extra parameter, which makes these distributions more flexible, and capture the Pareto distribution as a special case. The literature has demonstrated that where the Pareto distribution overestimates the frequency of the largest events (in the extreme upper-tail), these modified Pareto distributions follow the data more closely (Patel and Schoenberg, 2011). Our main objective here is to show that the Pareto distribution generally provides a good approximation to agricultural land size, and not the exploration of different variations of the Pareto distribution.

In contrast, exponential distribution clearly does not fit the data well, neither in the early to mid-range of the upper tail nor in the extreme upper-tail. While lognormal distribution tends to noticeably deviate from the data in the early and medium range, its performance somewhat improves in the extreme upper tail, rivaling that of the power law fit.

In Table 3, we formally compare the power law distribution fit to the data with those of exponential and lognormal distributions using likelihood ratio tests. As noted in

Section 3.3, a positive value of the likelihood ratio statistic indicates that the power law distribution is the better fitting distribution, while a negative value implies that the alternative distribution fits the data more closely. It is evident that the power-law distribution provides significantly better fit than the exponential distribution for all years, with large positive likelihood ratio statistics. This finding corroborates our observations from Figure 2. With our comparison between power law and lognormal distributions, we do not find statistically significant difference between the two distributions for the time periods under consideration. The likelihood ratio statistics are negative but small, which proves to be insignificant at the standard significance levels. This further clarifies our observations from Figure 2, where we noted that power law distribution fits the data better than lognormal in the early and medium range of the upper-tail data, whereas the lognormal tends to capture the extreme upper tail more closely. For both power law and lognormal distributions, the goodness of fit in one part of the data seems to be offset by a lack of fit in another part.

TABLE 3— LIKELIHOOD RATIO TESTS OF COMPETING DISTRIBUTIONS

	1997	2002	2007	2012
Power law vs exponential				
Likelihood ratio statistic	503.284	466.060	493.321	509.232

Taken together, given that (a) power law parameter estimates from both maximum-likelihood and regression-based methods are statistically significant and robust, (b) power law fit passes goodness-of-fit tests, and (c) power law provides a fit that is at least as good as or better than a series of alternative distributions, this provides evidence that power law is the suitable fit for the U.S. agricultural land size.

5. Conclusions

Agricultural land plays a key role in the economy of a country, and understanding the size distribution of agricultural land is fundamental to policy design. This paper presents evidence for a power law distribution for the upper tail of U.S. county-level agricultural land size. Our analyses demonstrate that the power law distribution passes extensive diagnostics tests, is robust across different periods, and fits the data at least as good as or better than a series of alternative distributions. This finding is significant because it implies that U.S. agricultural land size is heavy-tailed, that variance estimates generally do not converge, and that the top 5% of agricultural counties account for about 25% of agricultural land between 1997-2012. The power-law distributedness has implications for both more efficient management and agricultural policy design and empirical research. Understanding mechanisms that give rise to the emergence of such regularity for agricultural land size represents an avenue for further research.

References

Adamopoulos, T. and Restuccia D. 2014. “The Size Distribution of Farms and International Productivity Differences”. *American Economic Review*, 104(6): 1667-1697.

Ahearn, M.C., Yee J., and Korb P. 2005. “Effects of Differing Farm Policies on Farm Structure and Dynamics.” *American Journal of Agricultural Economics*, 87(5), 1182-1189.

Akhundjanov, S.B., Devadoss S., and Luckstead J. 2017. “Size Distribution of National CO_2 Emissions.” *Energy Economics*, 66: 182-193.

Axtell, R. 2001. “Zipf Distribution of U.S. Firm Sizes.” *Science*, 293: 1818-1820.

Balthrop, A. 2016. “Power Laws in Oil and Natural Gas Production.” *Empirical Economics*, 51(4): 1521-1539.

Brenes Muñoz, T., Lakner S., and Brümmer B. 2012. “Economic Growth of Farms: An Empirical Analysis on Organic Farming.” *In 2012 Conference, August 18-24, 2012, Foz do Iguacu, Brazil* (No. 126756). International Association of Agricultural Economists.

Champernowne, D. 1953. “A model of income distribution.” *Economic Journal*, 83: 318-351.

Clark, J.S., Fulton M., and Brown D.J. 1992. “Gibrat’s Law and Farm Growth in Canada.” *Canadian Journal of Agricultural Economics*, 40: 55-70.

Clauset, A., Shalizi R., and Newman M. 2009. “Power Law Distributions in Empirical Data.” *SIAM Review*, 51: 661-703.

Devadoss, S., Luckstead J., Danforth D., and Akhundjanov S. 2016. “The Power Law Distribution for Lower Tail Cities in India.” *Physica A: Statistical Mechanics and its Applications*, 442: 193-196.

Gabaix, X. 1999. “Zipf’s Law for Cities: An Explanation.” *Quarterly Journal of Economics*, 114(3): 739-767.

Gabaix, X. 2009. “Power Laws in Economics and Finance.” *Annual Review of Economics*, 1: 225-293.

Gabaix, X. and Ibragimov R. 2011. “Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents.” *Journal of Business and Economic Statistics*, 29(1): 24-39.

Gibrat, R. 1931. *Les Inegalites Economiques*. Librairie du Recueil Sirey, Paris.

Ioannides, Y.M. and Overman H.G. 2003. “Zipf’s Law for Cities: An Empirical Examination.” *Regional Science and Urban Economics*, 33(2): 127-137.

Irmay, S. 1997. “The relationship between Zipf’s law and the distribution of first digits.” *Journal of Applied Statistics*, 24(4): 383-394.

Key, N. and Roberts M.J. 2006. “Government Payments and Farm Business Survival.” *American Journal of Agricultural Economics*, 88(2): 382-392.

Key, N. and Roberts M.J. 2007. “Do Government Payments Influence Farm Size and Survival?” *Journal of Agricultural and Resource Economics*, 32(2): 330-348.

Klass, O.S., Biham O., Levy M., Malcai O., and Solomon S. 2006. “The Forbes 400 and the Pareto wealth distribution.” *Economics Letters*, 90(2): 290-295.

Krugman, P. 1996. *The Self-Organizing Economy*. Blackwell Publishers, Cambridge, Massachusetts.

Luckstead, J., and Devadoss, S. 2014. “Do the World’s Largest Cities Follow Zipf’s and Gibrat’s Laws?” *Economics Letters*, 125(2): 182-186.

Luttmer, E.G.J. 2007. “Selection, Growth, and the Size Distribution of Firms.” *Quarterly Journal of Economics*, 122: 1103-1144.

MacDonald, J.M., Hoppe R., and Newton D. 2018. “Three Decades of Consolidation in U.S. Agriculture.” EIB-189, U.S. Department of Agriculture, Economic Research Service.

MacDonald, J.M., Korbe P., and Hoppe R. 2013. “Farm Size and the Organization of U.S. Crop Farming.” ERR-152, U.S. Department of Agriculture, Economic Research Service.

Mansfield, E. 1962. “Entry, Gibrat’s Law, Innovation, and the Growth of Firms.” *American Economic Review*, 52(5): 1023-1051.

Pareto, V. 1896. *Cours d’Économie Politique Professé a l’Université de Lausanne*. Vol. I.

Patel, R.D. and Schoenberg F.P. 2011. “A graphical test for local self-similarity in univariate data.” *Journal of Applied Statistics*, 38(11): 2547-2562.

Perline, R. 2005. “Strong, Weak and False Inverse Power Laws.” *Statistical Science*, 20(1): 68-88.

Richardson, J.W. and Condra G.D. 1981. “Farm Size Evaluation in the El Paso Valley: A Survival/Success Approach.” *American Journal of Agricultural Economics*, 63(3): 430-437.

Shapiro, D., Bollman R.D., and Ehrensaft P. 1987. “Farm Size and Growth in Canada.” *American Journal of Agricultural Economics*, 69(2): 477-483.

Singh, S.K. and Maddala G.S. 1976. “A function for size distribution of incomes.” *Econometrica*, 44(5): 963-970.

Stanley, M.H., Buldyrev S.V., Havlin S., Mantegna R.N., Salinger M.A., and Stanley E.H., 1995. “Zipf Plots and the Size Distribution of Firms.” *Economics Letters*, 49(4): 453-457.

Sumner, D.A. and Leiby J.D. 1987. “An Econometric Analysis of the Effects of Human Capital on Size and Growth Among Dairy Farms.” *American Journal of Agricultural Economics*, 69(2): 465-470.

Toda, A.A. 2012. “The double power law in income distribution: explanations and evidence.” *Journal of Economic Behavior and Organization*, 84(1): 364-381.

Toda, A.A. 2017. “A note on the size distribution of consumption: more double Pareto than lognormal.” *Macroeconomic Dynamics*, 21(6): 1508-1518.

Toda, A.A. and Walsh K. 2015. “The double power law in consumption and implications for testing Euler equations.” *Journal of Political Economy*, 123(5): 1177-1200.

USDA. 1997. Census of Agriculture: http://agcensus.mannlib.cornell.edu/AgCensus/censusParts.do?year=1997

USDA. 2002. Census of Agriculture: https://www.agcensus.usda.gov/Publications/2002/

USDA. 2007. Census of Agriculture: https://www.agcensus.usda.gov/Publications/2007/

USDA. 2012. Census of Agriculture: https://www.agcensus.usda.gov/Publications/2012/

Wold, H.O.A. and Whittle, P. 1957. “A model explaining the Pareto distribution of wealth.” *Econometrica*, 25: 591-595.

Zipf, G.K. 1949. *Human Behavior and the Principle of Least Effort*. Addison-Wesley.

Appendix

A. Estimated Parameters Using x_{min}=95th percentile

We apply the maximum likelihood and regression-based methods to the 5% upper-tail data for each year. That is, we set x_{min} at the 95th percentile of the data (Gabaix, 2009). Table A1 presents the main results. As can be seen, both the Hill and OLS estimates of the power law parameter remain statistically significant for all years. Now, the magnitude of the OLS estimates of γ for the first three years indicates that only the first two moments (mean and variance) of the power law distribution are finite, while the remaining moments are non-convergent. Similarly, the magnitude of the Hill estimates suggests that only the first two moments are finite. Our results from the Gabaix and Ibragimov (2011) goodness-of-fit test suggest that we again fail to reject the null hypothesis of power-law-distributed agricultural land size for all years. This corroborates our main findings in the paper and provides further support for Pareto distribution fit for U.S. agricultural land size.

TABLE A1— POWER LAW PARAMETER ESTIMATES, x_{min} =95TH PERCENTILE

	1997	2002	2007	2012
γ^{Hill}	2.175	2.183	2.363	2.516
	(0.179)	(0.179)	(0.194)	(0.207)
γ^{OLS}	2.913	2.894	2.979	3.003
	(0.052)	(0.047)	(0.046)	(0.037)
x_{min}	953,098.6	935,806.8	946,962.4	958,518.8
Observations	151	151	151	151
The Gabaix and Ibragimov (2011) test				
Goodness of fit test statistic	-0.202	-0.173	-0.156	-0.134
Goodness of fit threshold	0.112	0.112	0.112	0.112

Notes: Estimation is based on 5% upper-tail observations ($x \geq x_{min}$=95th percentile). Standard errors are in parentheses. For the Gabaix and Ibragimov (2011) test, the null hypothesis that agricultural land size is distributed according to a power law is rejected if test statistic > threshold. Clauset et al. (2009) recommend to have at least 50 observations for accurate power law analysis, a condition satisfied here.