Technical efficiency estimates of Cherokee agriculture: A pre- and post-removal analysis

Matthew T. Gregg
Roger Williams University, mgregg@rwu.edu

Follow this and additional works at: https://docs.rwu.edu/gsb_fp
Part of the Business Commons

Recommended Citation

This Article is brought to you for free and open access by the Gabelli School of Business at DOCS@RWU. It has been accepted for inclusion in Gabelli School of Business Faculty Publications by an authorized administrator of DOCS@RWU. For more information, please contact mwu@rwu.edu.
Technical efficiency estimates of Cherokee agriculture: A pre- and post-removal analysis

Matthew T. Gregg

Gabelli School of Business, Roger Williams University, Bristol, RI 02809, United States

1. Introduction

The objective of this paper is to answer one of the most important economic issues in American Indian history: how productive were Cherokee Indian farmers during the nineteenth century? As Indian removal's cause célèbre, Cherokee economic and agrarian progress was widely debated. This political debate focussed on conjectures regarding the productivity differentials between Cherokee and white farmers and between mixed-blooded and full-blooded Cherokee farm households. To date, Dunaway (1997) and Wishart (1995) have used simple partial productivity measures to promote the level of agrarian process by the Cherokees as a whole. In this paper, I exploit two detailed household-level data sets on nineteenth century Cherokee farm output to estimate the technical efficiency of individual Cherokee farmers. These estimates are then regressed on a variety of exogenous variables to identify the ceteris paribus impact of inter- and intra-ethnic differences on agricultural productivity.

The productivity of Cherokee agriculture before removal was largely characterized using racial constructs. While politicians like Andrew Jackson, Lewis Cass, and Wilson Lumpkin used the language of race for their own ends, historians such as Champagne (1992), McLoughlin (1986), Malone (1956), and Sturm (2002) all emphasize racial hierarchy as an important factor of political, social, and economic behavior in Cherokee history.1 Regarding Cherokee agriculture, McLoughlin and Conser (1977, 1984) and Wilms (1978) suggest that full-blooded Cherokees achieved lower yields than mixed-blooded Cherokees.2 Unlike nineteenth-century claims, McLoughlin and Conser (1984, p. 223) use rational, rather than essentialist, terms to explain the variation in productivity:

The smaller size of its [full-blooded North Carolina Cherokee] farms and the smaller yield of its crops (particularly its low wheat yield) are evidence both of the mountainous area, the poor soil, the lack of money to invest in slaves or plows, and con-

1 For example, Cass (1830) wrote that “[i]n any individuals among the Cherokees have acquired property … [a]nd we have as little doubt, that this change of opinion and condition is confined, in a great measure, to some of the half-breeds and their immediate connexions.” Even American Board missionaries, who lived among the Cherokees prior to removal and attempted to emphasize the view of tribal-wide assimilation, often used similar language: “The intermixture of white people with the Indians has undoubtedly been a considerable cause of the civilization of the latter” (qtd. in Perdue, 2004).

2 Mixed-blooded Cherokees are children from a Cherokee-white union. Scholars have used metis or mestizo to divorce these groupings from the racial interpretations developed in the early nineteenth century. This paper, like Robertson’s (2002) recent work on Oglala Lakota land use, adopts the term mixed-blood because it is still widely used in contemporary Indian communities and is more historically accurate than implying a French or Spanish connection.
sequently of the general poverty of the Cherokees in this region. Possibly isolation from white settlements, adherence to traditional life-style, and deliberate resistance to change contributed to these differences.

The adoption of an overlooked household-level Cherokee census taken in 1835 allows for at least two contributions to this literature. First, a more complete measure of farm productivity, an output-oriented measure of technical efficiency, can be estimated by incorporating multiple farm outputs and multiple factors of production. In the context of economic theory, Farrell (1957) lists managerial characteristics as the principal factors that influence productive efficiency. Therefore, if full-blooded households decided to use dated farming practices, then economic theory would suggest that full-blooded "traditionalists" would be less technically efficient than mixed-blooded "progressives." Second, the role of cultural persistence on efficiency can be isolated once control variables like farm size, soil quality, market orientation, and farm experience are included in the empirical model. Recent developments by Simar and Wilson (2007) provide a method to measure bias-corrected technical efficiency using non-parametric techniques and to model the determinants of those efficiency scores.

Within the larger ethnohistorical literature, this study introduces a quantitative analysis into the recent debate over the role of race on culture in Southeastern Indian communities. Relying on anthropological theories and mainly eighteenth-century observations, Perdue (2003, 2004) emphasizes how white men were strategically chosen and adopted into the nation on Cherokee terms. Given the Cherokee matrilineal kinship system, white males would have had a limited role within the household, since maternal uncles provided more influence on children. This kinship system also implies that only members of a Cherokee clan were considered Cherokee. Thus, the concept of racial hybridity may have been unrelated to Cherokee identity. Even when adopting racial categories, Perdue (2004) maintains that a closer look at acculturation reveals that “many ‘full bloods’, as well as ‘mixed bloods’, welcomed aspects of the U.S. ‘civilization’ program.” The fluidity within intra-ethnic groups is readily apparent when analyzing the leaders of the two factions over Cherokee removal. Major Ridge, a full-blooded Cherokee, led members of the mainly mixed-blooded, pro-removal Treaty Party, while the Anti-removal Party was comprised mostly of full-blooded Cherokees and was organized under the leadership of Principal Chief John Ross, who was one-eighth Cherokee.

On the other hand, Saunt et al. (2006) still maintain the traditional viewpoint that race was “essential to formations of power and resistance...[and] must be a central component of ethno-historical analysis.” Regarding economic behavior, historians have long maintained that the Cherokees who attempted to possess their remaining ancestral land. Removal advocates believed that the right to cultivate land was reserved to “civilized” men rather than to “savages” who used the land primarily to hunt (see de Vattel, 1758 for the earliest development of this philosophy). Given their constituents’ large demand for land disposal, politicians strongly believed that Indians were not using land up to its potential (Perdue and Green, 2007, p. 14). In particular, President Andrew Jackson used this rhetoric to promote his plan to remove the Cherokees and other Indian tribes west of the Mississippi River. According to Jackson (1833),

They have neither the intelligence, the industry, the moral habits, nor the desire of improvement which are essential to any favorable change in their condition. Established in the midst of another and a superior race, and without appreciating the causes of their inferiority or seeking to control them, they must necessarily yield to the force of circumstances and ere long disappear.

Pre-removal comparative farming data simply do not exist. The earliest household-level data on U.S. agriculture are located in the 1850 and 1860 manuscript censuses. Fortunately, the Cherokees living in Haywood and then Jackson County, North Carolina who avoided removal were enumerated alongside their white neighbors in the 1850 and 1860 population and agricultural manuscript censuses. By sampling white and Cherokee farm households from these two counties in 1850 and 1860, a richer database can be used to compute technical efficiency measures on both types of farms. The advantage of this data set is that the estimation problems from the data limitations in the 1835 Cherokee census can be resolved.

2. Methodology

The methodology used here largely follows from Färe et al. (1994) and Simar and Wilson (2007). When multiple inputs and outputs are used in production, the underlying production process and productive efficiency can be described with distance functions...

3. Henry Knox, the Secretary of War during Washington’s administration, devised the first federal Indian policy of “civilization.” The policy stated that treaty stipulations would furnish Indians with farm implements, livestock, and residential Indian agents to lead them in the adoption of Euro-American farming practices. The 1791 Treaty of Holston was the first Cherokee land cession to contain these features. See Royce (1975) for a list of nineteenth-century Cherokee treaty stipulations.

4. Historians also emphasize that wealth accumulation and world views varied within these groups. For example, Sturm (2002, p. 68) emphasizes that “these categories of identity were social constructions that had only a loose correspondence with racial ancestry, since Cherokees with white ancestry could also be poor, non-Christian or against black slavery.”

5. Thornton (1990) also uses the idiom of race to show the positive correlation between white ancestry and the number of slaves.
(Shephard, 1970). Given the existence of a best practice production frontier, an output distance function shows the amount actual output falls short of what can be feasibly produced given some level of inputs. The output distance function of a multiple-output and multiple-input production technology can be expressed as

\[ D_0(x, y) = \min(\theta : y/\theta \in P(x)) \]  

where \( \theta \) is a measure reflecting the amount by which the observed vector of outputs that can be radially increased and still remain feasible for a given input vector; and \( P(x) \) is the set of feasible pairs of outputs \((y, x)\) and inputs \((x, x)\), where both \( x \) and \( y \) are vectors. Under the typical assumptions, \( D_0(x, y) \) is nonincreasing in \( x \), and nondecreasing, homogenous of degree one, and convex in \( y \). If the household is fully efficient, then \( D_0(x, y) = 1 \).6 As the distance between the observed vectors of outputs and the frontier increases, \( D_0(x, y) \) decreases. Therefore, \( D_0(x, y) \) lies between zero and one.

This distance function can be estimated by using a linear programming method called Data Envelopment Analysis (DEA). This non-parametric approach builds a production frontier that envelops the data and measures the distance from observed output to the best practice frontier. The key advantages of DEA over the parametric stochastic frontier approach is that DEA can easily accommodate multiple outputs and zero values for some observations.7 This approach also does not require the adoption of market prices to aggregate output, which can confute revenue efficiency with technical efficiency (David and Temin, 1979).

With \( I \) households, \( M \) outputs, and \( N \) inputs, DEA computes technical efficiency measures by solving the following linear programming problem under the assumption of constant returns to scale:

\[
D_0(x, y) = \left[ \max_{\theta} \left\{ \sum_{i=1}^{I} \frac{z_i}{\theta_i} y_{mi} \right\} \right]^{-1}
\]

s.t. \( \theta_k y_{mi} \leq \sum_{i=1}^{I} z_i y_{mi}, \quad m = 1, \ldots, M \)

\[
\sum_{i=1}^{I} z_i x_{ni} \leq x_{ni}, \quad n = 1, \ldots, N
\]

\[
z_i \geq 0, \quad i = 1, \ldots, I
\]

where \( z_i \) reflects the intensity weights used to compute the piecewise production frontier (Färe et al., 1994). Since this problem is solved for each observation, \( \theta \) will reflect the efficiency score for each household. The technical efficiency term ranges from zero to one as one reflects 100% technical efficient.

In the second stage, these DEA efficiency scores are regressed on a set of explanatory variables. However, as illustrated by Simar and Wilson (2007), statistical inference from this second-stage regression is invalid since the DEA scores are by construction serially correlated. They do prove that consistent inference can be obtained by bootstrapping the DEA scores, correcting for sampling bias, and using maximum likelihood to estimate a truncated regression model of the DEA scores on a set of variables. Bootstrapping the efficiency measures simulates the underlying sampling distribution and the bias is subtracted from the original DEA score for each household to generate a bias-corrected technical efficiency measure. For this study, I follow Simar and Wilson’s (2007) first algorithm.

Although the term efficiency is used through the paper, deviations from the frontier may not in fact be “sub-optimal.” Unobserved constraints may lead some to optimally choose input and output quantities that appear to be inefficient. The measurements of efficiency may simply reflect differences in constraints as some households may farm on poorer soil, contain less farm experience, or suffer more from market imperfections. The goal is to control for these multiple factors of inefficiency in order to isolate the conditional impact of racial hierarchy on productivity.

3. The relative efficiency of Cherokee farming

3.1. Pre-removal data

The pre-removal data for this study come from a census of Cherokee farm households in 1835 collected by the War Department, who was in charge of Cherokee emigration throughout the 1820s and 1830s. The location of each household in the census was identified by the state of residence (i.e., Alabama, Georgia, North Carolina, or Tennessee), county within the state, and place-name within the county. Assisted by Cherokee interpreters, the census takers enumerated 36 different characteristics on 2670 households.

The category headings included information on the size of the household, number of slaves, various agricultural inputs and outputs, other wealth information like number of (grist and saw) mills and ferryboats, literacy information, and the racial characteristics of the household.

For the DEA analysis, two output and four input variables are used. Only bushels of corn and wheat grown were listed in the census. According to the census, 85% of the households grew corn while only 2% grew wheat. Other traditional crops such as beans, squash, peas, and potatoes were unfortunately omitted from the census (Hudson, 1976). However, the heavy reliance on corn within this Southern region after removal is shown in Weiman’s (1987) work on Upcountry farmers.9 Crop diversification estimates can be gleaned from the post-removal North Carolina Cherokee data on corn, peas, beans, and potato production. On North Carolina Cherokee farms that were roughly the same size as 1835 Cherokee farms, 95% of total crop output comprised solely of corn bushels. The pre-removal Cherokees were not self-sufficient, however, as Goodwin (1977), Perry (1974), and White (1975) all emphasize the reliance on “wild” foods in Cherokee diets. While the precise degree of diversification is unknown, these comparisons do suggest that it may have been limited. The chief limitation of these output variables is that it may bias the efficiency gains of larger slaveholders.10

The household labor variable treats adult males (over the age of 18), adult females (over the age of 16), and married whites (mainly males) as equals while counting half of the males under age.

---

6 If the vector of output for the ith household, \( y_i \), was located on the frontier, \( y_i = \theta y_i \), then \( y_i = \theta y_i \). In this case, for the ith observation, \( \theta_i = (y_i/\theta y_i) \) or \( \theta_i = 1 \). Therefore, the technical efficiency is equal to one when the vector of output cannot be radially expanded.

7 Flexible functional forms such as generalized Leontief, CES, and generalized McFadden all failed to hold regular conditions as they computed negative and non-diminishing marginal products. Data transformations, e.g. setting \( \ln(0) = 0 \), are not used since they reorder the observations in the range of [0, 1].

8 Regarding the accuracy of the data, there was only one account of a North Carolina census taker running into difficulty with data collection. However, Principal Chief John Ross helped thwart any backlash with an open letter read at a regional council meeting. In the letter, John Ross wrote, “I do not ask it as a favor but claim it as a right to proceed in taking your numbers and will view your refusal as a direct declaration that you have no friendship for the Government of the United States” (qtd. in Litton, 1940). There was no documented evidence of interference after this date.

9 Weiman (1987) shows that on Upcountry farms that contained at least 500 improved acres in 1860, corn still comprised 77% of total crop output.

10 Cotton would be the chief omitted variable on large plantations. McComb (1974) cites a letter from John Ridge, son of Major Ridge, to Albert Gallatin, Secretary of the Treasury under Jefferson’s administration, where Ridge claims that “cotton is generally raised for domestic consumption and a few have grown it for market and have realized very good profits.” Slaveholding households did account for only 7.8% of the total Cherokee households in 1835.
the age of 18 and females under the age of 16. These demarca-
tions were determined by the limitations of the census. How-
nevertheless, the assumed participation rate for young males is equivalent to
the four-state average during the antebellum period (Weiss, 1992). A
Cherokee household averaged six members while 7% of those
households also held a married white. Off-the-farm work was
limited as only 2% of the households generated income from
operating either a mill or a ferryboat. This suggests that the sub-
stitution between farm work and other work-related activities was
undoubtedly minimal by the third decade of the nineteenth
century. Lastly, adult females are added to the labor variable because historical evidence suggests that some females still kept their
role as chief cultivator of the household during the “civilization”
program.11
The slave variable assumes equal roles for male and female
slaves as age differences were not listed in the census. In their ana-
lyses of Cherokee slavery, Perdue (1979) and Halliburton (1977) do
not distinguish work roles by gender. Gang labor systems were not
employed by slaveholders in the Cherokee Nation as only 10% of the
slaveholders held over 15 slaves.
The land variable is defined as total acres cultivated. Rather than
constructing a variable reflecting land solely devoted to corn and
wheat, this variable is left unaltered. Weiman’s (1987) study on
antebellum Upcoming farming suggests that crop diversification
and farm size were unrelated: farms averaging 1–50 acres held a
corn share of 77.3%, while the largest farms in his sample (over
500 acres) held a corn share of 77.4%. Among Cherokee households
in 1835, there very little correlation between farm size and corn
yields.12
Another issue with the land variable is accounting for soil qual-
ity. The soil quality index constructed here is based around Field’s
(1988) suggestion. Specifically, households were split into groups
defined by three county-level soil types listed in Hilgard (1884).
Each category was further divided based on a household’s proxim-
ity to a waterway as defined in the 1835 census. The soil quality
index is computed by comparing the output per acre within these
six categories, using the category with the highest average yield
as the base category. Most technical efficiency studies include
soil quality and other environmental variables in the second-stage
regressions. Likewise, soil quality in this study is assumed to
directly affect efficiency and is thus incorporated into the second
stage.
Finally, the total number of buildings is included as a proxy for
the physical capital input since the value of farming imple-
ments and livestock were not enumerated. These buildings would
have reflected the number of dwellings, kitchens, slave quarters,
corn cribs, smokehouses, outhouses, and workshops consisted on
a farm. A farmer who invested heavily in structures probably also
invested in farm implements and other capital inputs. In practice,
the inclusion of this variable did not greatly alter the efficiency
measures.13
After removing unusable observations (i.e., households with
no acres cultivated and no corn produced), 2344 slaveholding
and non-slaveholding Cherokee families are used to estimate
household-level technical efficiency. Eventually all of the data con-
tained in the census is incorporated once the observations are
aggregated into communities.

11 Traditionally, Cherokee men helped only with breaking soil and harvesting crops. See Hatley (1988) for an overview of the continued role in women in agriculture during this period.
12 The sample correlation coefficient between corn yields and acres cultivated is −0.061 and is statistically insignificant at common significance levels.
13 The sample correlation coefficient between the technical efficiency computed with and without this capital input proxy is 0.963 and is significant at the 1% level.

### Table 1

<table>
<thead>
<tr>
<th>Variable Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Explanatory variables of pre-removal technical efficiency.</strong></td>
</tr>
<tr>
<td>Intra-ethnic categories</td>
</tr>
<tr>
<td>Farm size</td>
</tr>
<tr>
<td>Age</td>
</tr>
</tbody>
</table>

### 3.2. Pre-removal results

The definitions of the explanatory variables used in this stage are
located in Table 1. Intra-ethnic groups are eventually disaggregated
into four categories: households with all mixed-blooded Cherokees
but no married whites (221 observations); members with some mixed-
and full-blooded members (261 observations); households
that contained a married white (158 observations) and households
with all full-blooded members (1704 observations).14

Once the technical efficiency measures are computed using the
framework discussed in Section 2, truncated regression models
of the bias-corrected efficiency scores on a set of variables are
estimated using maximum likelihood and the results are located
in Table 2. Each regression controls for heteroskedasticity by imple-
menting the White method. Unless noted, each estimates in Table 2
is significant at the 1% level.

The unconditional regression model (1) implies that a 10% increase in the share of full bloods within a household decreased
technical efficiency by 1.7%. This initial finding is augmented in
the next five regressions by incorporating other factors (farm size, soil
quality, market orientation, and experience) in a step-wise fashion.
Each coefficient, except for the share of English readers, is signifi-
cant at the 1% level. Several notes need to be mentioned before
discussing the intra-ethnic variables. First, the inverse-U relation-
ship between efficiency and farm size and the efficiency gains from
market access and farming experiences are consistent with most
micro-level agricultural studies.15 Also, the impact of acres cul-

14 Using the census definitions, a mixed-blooded Cherokee is defined as either ¼ Cherokee, ¼ Cherokee, mixed-black, or mixed-Catawba. Less than 1% of the Cherokee
population in 1835 was mixed-black or mixed-Catawba.
15 Chayanov (1966) and Sen (1966) are the classic works on the inverse relationship
between farm size and (labor) productivity. Also, see Helfand and Levine (2004) and
Sherlund et al. (2002) for two recent examples of the positive influence of market
access and experience on efficiency.
Table 2
The determinants of pre-removal technical efficiency.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-ethnic categories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-blood share</td>
<td>−0.168</td>
<td>−0.119</td>
<td>−0.098</td>
<td>−0.084</td>
<td>−0.091</td>
<td></td>
</tr>
<tr>
<td>All mixed-blood</td>
<td>0.109</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial mixed</td>
<td>0.049</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married whites</td>
<td>0.113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farm size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.003</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Land²</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td></td>
</tr>
<tr>
<td>Spatial heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil quality index</td>
<td>0.079</td>
<td>0.067</td>
<td>0.068</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market orientation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market access</td>
<td>0.139</td>
<td>0.129</td>
<td>0.044¹</td>
<td>0.025¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English readers</td>
<td>0.044¹</td>
<td>0.025¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farm experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.129</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>2211.57</td>
<td>2247.66</td>
<td>2245.89</td>
<td>2271.79</td>
<td>2286.40</td>
<td>2296.37</td>
</tr>
<tr>
<td>χ²-Statistic</td>
<td>110.62</td>
<td>176.23</td>
<td>165.62</td>
<td>183.52</td>
<td>204.20</td>
<td>206.60</td>
</tr>
</tbody>
</table>

Notes: the standard errors are corrected for possible heteroskedasticity using the White method. All coefficients are statistically significant at the 1% level unless noted.

Table 3
Truncated regression results—omitting 125 most efficient observation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-ethnic categories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-blood share</td>
<td>−0.144</td>
<td>−0.110</td>
<td>−0.089</td>
<td>−0.076</td>
<td>−0.082</td>
<td></td>
</tr>
<tr>
<td>All mixed-blood</td>
<td>0.102</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial mixed</td>
<td>0.045</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married whites</td>
<td>0.101</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farm size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Land²</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td></td>
</tr>
<tr>
<td>Spatial heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil quality index</td>
<td>0.076</td>
<td>0.067</td>
<td>0.067</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market orientation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market access</td>
<td>0.131</td>
<td>0.122</td>
<td>0.051¹</td>
<td>0.033¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English readers</td>
<td>0.051¹</td>
<td>0.033¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farm experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.135</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>2057.35</td>
<td>2086.28</td>
<td>2081.49</td>
<td>2115.09</td>
<td>2130.77</td>
<td>2144.02</td>
</tr>
<tr>
<td>χ²-Statistic</td>
<td>110.75</td>
<td>166.95</td>
<td>154.10</td>
<td>178.64</td>
<td>199.31</td>
<td>202.53</td>
</tr>
</tbody>
</table>

Notes: the standard errors are corrected for possible heteroskedasticity using the White method. All coefficients are statistically significant at the 1% level unless noted.

A number of variations to these models are estimated to determine the robustness of these findings. First, since DEA is sensitive to outliers, 125 of the most efficient observations (5.3% of the sample) were omitted, and the DEA approach was re-run and the determinants re-estimated using the method described in Section 2. The findings are located in Table 3. The statistical and economic significance of the coefficients on married white households, all mixed-blooded households, partially-mixed households, and full-blood share remained. After omitting the most efficient producers, the technical efficiency on mixed-blooded households ranged from 7.6% to 11.0% above the efficiency of full-blooded households.

Second, the slaveholding households are omitted from the sample to determine if the shape of the frontier drastically changed.
Table 4
Truncated regression results—omitting slaveholding households.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-ethnic categories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-blood share</td>
<td>−0.157</td>
<td>−0.100</td>
<td>−0.077</td>
<td>−0.062</td>
<td>−0.069</td>
<td></td>
</tr>
<tr>
<td>All mixed-blood</td>
<td></td>
<td></td>
<td>0.096</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial mixed</td>
<td></td>
<td></td>
<td>0.029</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married whites</td>
<td></td>
<td></td>
<td>0.091</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farm size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land</td>
<td>0.012</td>
<td>0.016</td>
<td>0.017</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
</tr>
<tr>
<td>Land²</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
</tr>
<tr>
<td>Spatial heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil quality index</td>
<td>0.102</td>
<td>0.092</td>
<td>0.092</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market orientation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market access</td>
<td>0.108</td>
<td>0.097</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English readers</td>
<td>0.087</td>
<td>0.066</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farm experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>2157.96</td>
<td>2285.18</td>
<td>2284.72</td>
<td>2333.62</td>
<td>2345.86</td>
<td>2359.67</td>
</tr>
<tr>
<td>χ²-Statistic</td>
<td>63.17</td>
<td>201.71</td>
<td>200.10</td>
<td>225.21</td>
<td>246.63</td>
<td>252.08</td>
</tr>
</tbody>
</table>

Notes: the standard errors are corrected for possible heteroskedasticity using the White method. All coefficients are statistically significant at the 1% level unless noted.

After recomputing the technical efficiency measures, the regression results for the non-slaveholding households are contained in Table 4. The efficiency difference between mixed-blooded and full-blooded households is still statistically significant but the efficiency differential is now highest in all mixed-blooded households. However, the omission of slaveholders did not affect any estimated relationship between efficiency and the set of exploratory variables. For example, once the slaveholding households are removed, full-blooded households are estimated to be between 6.2% and 10.0% less efficient than mixed-blooded households.

Third, the unit of observation is adjusted from households to communities in order to incorporate all the data contained in the census. The aggregation method used here follows from McLoughlin and Conser (1977). Communities are defined by the place-names given in the census and using this technique, 103 communities are identified from the census data. Since each community contained at least one full-blooded member, intra-ethic decompositions cannot be used. Instead, the continuous measure (i.e., the full-blood share in a community) is used. Several models were considered and the results are located in Table 5. On average, communities with all full-blooded Cherokees were 7.8% less efficient than communities with all mixed-blooded members. Productivity differentials decreased once farm size, spatial heterogeneity, market orientation and farm experience were included. There are two possible reasons why the intra-racial variable become insignificant: (1) this ad hoc aggregating technique may not be an accurate way to combine observations; and (2) incorporating unusable observations with zero values for acres cultivated and corn cultivated, especially for slaveholders with 30 or more slaves, may imply heavier reliance on cash crops which would lower their measure of technical efficiency. Regardless, the variable of interest, the full-blood share, reflects a similar negative relationship with efficiency as shown in the earlier models. Therefore, all three tests suggest that it is unlikely that measurement error influenced these results.

3.3. Post-removal data

To conduct a direct test of the relative productivity of Cherokee farmers, a sample of Cherokee and white households are taken from the 1850 and 1860 population and agricultural manuscript.
The white population in Haywood County in 1850, which was later incorporated into Jackson County by 1860, was sampled in small blocks with an interval size of 30 households. These farm households were matched with information in the population schedule to determine the available farm labor supply. All Cherokee households in the agricultural schedules were collected from these 2 census years and matched to information in the population schedules. In sum, there are 271 total households in the sample, comprising of 145 white and 126 Cherokee households. Since the North Carolina Cherokees did not hold slaves, only non-slaveholding white households were left in the sample in order to make valid comparisons. Given the soil fertility and distance to markets, only 2% of the white sample contained slaveholders, none of whom held over five slaves.

The variables used to measure technical efficiency and to model its determinants are listed in Table 6. The most-commonly listed outputs are selected for the DEA analysis, specifically: corn, wheat, peas and beans, potatoes, oats, tobacco and wool. The census takers also enumerated the number of improved acres, value of farming implements and machinery, and value of livestock, all of which are considered inputs into agricultural production. Most importantly, Fogel and Engerman’s (1977) labor weights. In particular, the following weights were used for males: ages 10–14, 0.40; ages 15–19, 0.88; ages 20–54, 1.0; ages 55–59, 0.75. The weights for females were assumed to be 50% of the male participation rate for each corresponding age category.

The variables used to measure technical efficiency and to model its determinants are listed in Table 6. The most-commonly listed outputs are selected for the DEA analysis, specifically: corn, wheat, peas and beans, potatoes, oats, tobacco and wool. The census takers also enumerated the number of improved acres, value of farming implements and machinery, and value of livestock, all of which are considered inputs into agricultural production. Most importantly, Fogel and Engerman’s (1977) labor weights. In particular, the following weights were used for males: ages 10–14, 0.40; ages 15–19, 0.88; ages 20–54, 1.0; ages 55–59, 0.75. The weights for females were assumed to be 50% of the male participation rate for each corresponding age category.

On the other hand, post-removal Cherokee efficiency in North Carolina was slightly less or equal to the efficiency of their white non-slaveholding neighbors. While there is evidence of Cherokee farm inefficiency, white productivity premiums are not robust across model specifications. The differences between pre- and post-removal efficiency of full-blooded Cherokees suggest that cultural persistence may have been a bigger concern before removal.
For example, after Cherokee removal, the chief goal of the North Carolina Cherokees was to demonstrate their proclivity towards farming in order to obtain U.S. citizenship (Finger, 1984). This incentive did not exist among pre-removal full-blooded Cherokees who often refused to participate in the government’s “civilization” program. This empirical study corroborates with this aspect of Cherokee history.

Future research might determine the robustness of these findings by either adopting a parametric approach or uncovering comparative data on other Southeastern Indians. Census data also exist on postbellum Cherokee agriculture in present-day Oklahoma. These data could help extend the analysis of race and productivity well after removal. Regardless, through the adoption of these data sets, the determinants of Cherokee productive efficiency can at least shed more light on the ongoing debate over the role of race within Southeastern Indian communities.

References


