Journal of Institutional Economics (2020), **16**, 127–144
doi:10.1017/S1744137419000110

Rationing by racing and the Oklahoma land rushes

Douglas W. Allen* and Bryan Leonard

Department of Economics, Simon Fraser University, Burnaby Canada and School of Sustainability, Arizona State University, Phoenix USA
*Corresponding author. Email: allen@sfu.ca

Abstract

Yoram Barzel was always aware that competition is ubiquitous and takes many forms, and he was among the first to analyze settings where individuals compete on the basis of time, rather than price. This paper applies his insights to study the Oklahoma land rushes, where thousands of individuals raced to establish property rights to land. A simple modification of Barzel's analysis generates a model of rationing by racing, and we test its predictions using new data on the timing and location of over 73,000 homestead claims within the five distinct land rushes and one lottery. We find that increases in land quality or decreases in the cost of racing generate corresponding increases in the equilibrium speed, implying that potential rents are dissipated by investments in speed. The analysis highlights the lasting significance of Barzel's insights regarding non-price competition.

Keywords: homesteading; Oklahoma land rush; racing

1. Introduction

Although Yoram Barzel is best known for his work on property rights, transaction costs, and institutional economics, he was trained as (and in many ways, always remained at heart) a Chicago price theorist. In the late 1960s and early 1970s, his price theory foundation started to commingle with an interest in matters of rent dissipation, and just like the mixing of warm and cold ocean currents off of the Grand Banks of Newfoundland, the fertile result was several seminal papers that laid the groundwork for his later work on property rights.[1] Of these works, his theory on rationing by waiting (Barzel, 1974) endures because it was theoretically sound, widely applicable, and a critical component in the theory of the establishment of property rights.[2]

Here, after a brief review of Barzel's rationing insights, the model is slightly adapted and then used to examine an interesting and unique episode of US history: the Oklahoma land rushes.[3] Starting in 1862 the US federal government allocated some of its federal lands through homesteading. Overall, about one-third of all federal lands (equivalent in size to the states of California and Texas) were given out through homesteads, and almost always in a similar fashion. In theory, regions of states and territories would be opened to homesteading once the allotted federal lands had been surveyed by the federal government. Individuals who then arrived first to the local land office, paid a

---

[1] These works include his paper on rent dissipation in innovations (Barzel, 1968a), the introduction of quality in price indices (Barzel, 1968b, 1969), dissipating entry into production (Barzel, 1971), quality changes in the face of taxation (Barzel, 1976), and his analysis of slavery (Barzel, 1977).

[2] As Lueck (1995) Lueck and Miceli (2007), and Suen (1989) point out, efforts to capture stocks left in the public domain led to dissipations based on Barzel (1974). The Barzel waiting model has led to a large theoretical literature, particularly in the fields of development and health (see, for example, Gravelle and Siciliani (2009), Kulshreshtha (2007) or Bose (1997)). However, few papers have tested the model directly, and no one has empirically examined racing explicitly. See Yeung *et al.* (2004) for a study of time costs and doctor choice, or Deacon and Sonstelie (1985) for estimates on the value of time based on waiting for gasoline.

[3] Although we adapt the waiting model, the idea of racing is found in Barzel (1968a).

registration fee, and then settled, occupied, and made improvements, could receive title to the land after five years. In practice, settlers often went ahead of the federal surveyors, and often settled outside of areas planned for homesteading. *Ex post*, and particularly in the early days of settlement in any given state or territory, these preemptive settlements were recognized as legitimate homesteaders and given title – even in areas where homesteading was legally either prohibited or restricted.[4]

The Oklahoma situation was different. Known at the time as "Indian Territory," about half of the state was purchased by the US government from the existing tribes at the end of the 19th century, and then six large tracts of non-uniform land were opened for settlement at fixed starting dates. Unlike the other parts of the frontier, efforts were made to keep settlers out until the actual start time, and as a result settlers lined up on the borders and waited for the race to start (Billington and Ridge, 2001). Thus, unlike other areas subject to homesteading where settlers arrived early and waited for civilization to catch up, in Oklahoma there is a definitive start date and set of bounded territories that induced a race to establish ownership. Furthermore, the last territory to be given away was done so as a lottery for the right to participate in a limited race.[5] All of these features make Oklahoma excellent testing grounds for the rationing by waiting model.

The Oklahoma natural experiment has several advantages compared to previous studies of the "free" allocation of land under the Homestead Act. A core theoretical prediction in the literature is that allocating land at a zero money price induced a rush to settle land before it was economically optimal to do so (Anderson and Hill, 1990). Empirical verification of this prediction has proved elusive due to a lack of detailed data on homesteading and, more fundamentally, due to the lack of a clear benchmark for the efficient rate of settlement. Our study of Oklahoma overcomes both limitations. The fixed start date of the rushes provides a clear benchmark from which to measure speed. Moreover, because of the fixed starting date there are no missing initial observation (no left-side censoring), and no attrition over time because the physical plots of land never leave the sample.

2. Rationing by waiting

The Barzel model of rationing is ostensibly a model of waiting mechanics.[6] That is, the model starts with the assumption that something of value is in the public domain (either by nature or placed there by an original owner), and the first to possess the thing becomes the new legal and economic owner.[7]

Barzel's insight into queues starts by recognizing that the forces of competition remain even when time, rather than nominal price, is used to allocate. As a result, competition forces potential owners to compete on the time dimension, arriving earlier to ensure access to the good. As in the case of allocation by price, the marginal person who is successful in acquiring the good by time is just indifferent between acquisition and never having bothered. That is, for the marginal person, the costs of waiting just offsets the value of the good being acquired, and therefore the marginal consumer's surplus is zero. Since the time costs are not transferred to another, the marginal person(s) completely dissipates the value of the good acquired because the act of using time is all-or-nothing. On the other hand, intra-marginal individuals do not completely exhaust their surplus, and they gain from time allocation because it is the marginal person's valuation that determines the "time price" and therefore, the mechanics of the line.

The conclusion that an equilibrium amount of time must be spent when allocation is by waiting is counter-intuitive. For example, it is commonly thought that if the mechanical logistics of giving something away are made easier that the time spent in line will be reduced. However, improving the processing time of a give-away only increases the number in the line until, in equilibrium, the value of time for the marginal person spent in line equals the value of what is given away.

---

[4] For example, generally homesteading was either not allowed within the checkerboard swath of lands given to land grant railroads, or the amount of land that could be homesteaded was restricted in size. However, settlement on 160-acre parcels was common (see Allen, 2019 for details).

[5] That is, the lottery was not for land parcels directly, but for the right to enter the race to claim a parcel.

[6] Barzel's great seminal idea is that all forms of rent dissipation are mitigated by the parties involved. In analyzing waiting mechanics, he ignored this issue. Likewise, in this paper, we ignore efforts to mitigate the cost of racing.

[7] Why the thing of value is in the public domain is left unaddressed.

The equilibrium time price in the Barzel model depends on the distribution of individual demands for the good (which depend, in part, on time costs), the amount given away per person, and the total amount being given away. Although there is complete dissipation for the marginal individual, the total dissipation depends on the heterogeneity of those competing by waiting.[8] With identical time costs, the entire surplus of the stock being given away is dissipated.

3. Rationing by racing

A race with a specific start time has a very similar mechanism to rationing by waiting.[9] When land is captured by a race, occupation is to the swiftest. Rather than waiting in a queue, individuals take possession upon arrival and incur costs to arrive first. Assume there are N quarter-sections of land available in a given race.[10] Each parcel has attribute $Q \in (0,1)$, where $Q = 1$ means that a given quarter-section of land was reached first by someone. The probability of being first in a race conditional on speed, s, is given by:

$$\begin{aligned} p(Q = 1|s) &= 1, \text{ if } s \geq s^*; \\ &= 0, \text{ if } s < s^*, \end{aligned}$$

where s^* is the necessary speed for success. Speed is acquired through effort and capital (e.g. a horse); let the costs of effort and capital be w and v respectively. Optimal acquisition of speed implies a cost function $c_i(w,v,s^*)$, which varies across K individuals indexed by i. An individual's willingness to pay for $Q = 1$ is the difference between the total value of the land and the cost of racing. Define this willingness to pay as:

$$Z_i(w, v, s) = \Omega_i(w) - c_i(w, v, s),$$

where $\Omega_i(w)$ is the individual's total value of the land, and which assumes there is no utility in racing. Individuals are willing to enter the race as long as $Z_i > 0$, or the individual's value of the land is greater than their cost of racing.

Given that one plot is available per person, the quantity demanded for land is equal to the number of people for whom $Z > 0$ for a given speed. As shown in Figure 1, the equilibrium speed s^* is determined by equating the total amount of land available to this market demand.

In equilibrium $N = K$, all racers go at the same speed, and all are successful in obtaining a plot of land. Costs of speed are different, however, and so low-cost racers receive surplus from the land, but the marginal racer fully dissipates the value of the land. If the value of land (Ω) increases or the costs of racing (c) fall then the demand for land quarters increases as does the equilibrium speed – leading to every quarter taken, to be taken sooner. Obviously, other things being equal, an increase in the number of land quarters available lowers the equilibrium speed.

When land quality is not homogeneous the model is easily extended. Suppose there are two levels of land quality: high and low. Let N^h be the number of homogeneous land plots of high quality q^h available for homesteading and N^l be the set of low quality homogeneous land plots ($q^l < q^h$). For both N^h and N^l there are separate race markets, with separate equilibrium speeds s^l and s^h, which also depend on the relative demands for the different quality lands. Adding more qualities of land simply increases the number of races.

Figure 2 shows the separate high and low quality markets, where it is assumed that the low quality land is not scarce. Individuals race for the high quality lands, but there is no race for the low quality

---

[8] Rent dissipation is reduced when individuals have heterogeneous time costs because the marginal successful person needs only arrive early enough to beat the next potential claimant, who has higher time costs (Barzel, 1968a; Lueck, 1995; Suen, 1989).

[9] This model used here is a minor variation of Suen's (1989) rationing by waiting model.

[10] A "section" of land is one square mile, and contains 640 acres. "Quarter-sections" are created by dividing a section into four equal parts, each containing 160 acres.

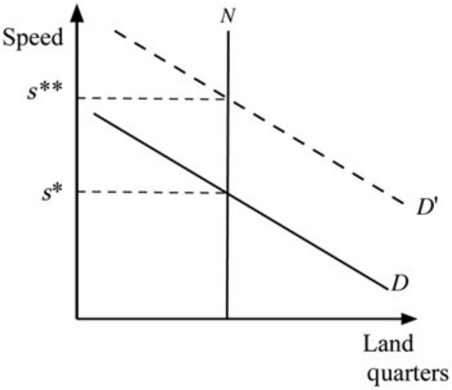

Figure 1. Equilibrium speed allocation

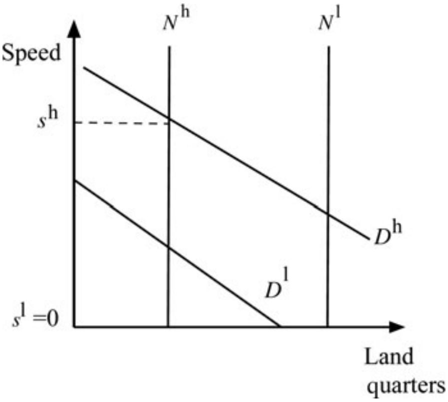

Figure 2. Heterogeneous land qualities

lands, and some remain untaken. Observationally, for a given starting date, the plots of land will be taken up sequentially, with the highest quality going first, and subsequent lower qualities following. Clearly it is possible that not all "free land" is taken.

The recognition that land has different qualities leads to an interesting prediction regarding how much land remains untaken over time. One way of describing how much land is untaken is to use what is called a "survival function."[11] A survival function reports the probability that a parcel of land remains untaken (survives) at some time t after the race begins. Such a function can be conditioned on the quality of land within an area. Hence, when there is an increase in the average quality of land within a given territory, holding constant the variance in land qualities, the probability of a given plot being taken increases at every moment t. This is shown in panel (a) of Figure 3, where the survival function of the higher quality territory, $S(h)$, lies everywhere below the lower quality territory survival function, $S(l)$. It lies below because the probability of remaining *untaken* has fallen.

On the other hand, if there is a mean-preserving increase in the spread of land qualities within a territory, the survival functions cross as in panel (b) of Figure 3, where it is assumed var(2) > var(1). Here, in the territory with the higher variance, the very best land is taken early, leading to a steep survival function. That is, initially there is a faster race for the best lands.

As the best lands are taken, the remaining lands are of lower quality and so are mostly left untaken. That is, the race slows down and this results in a flatter survival function. Hence the relative shapes of

---

[11] Such functions were first used in medical studies where the data contained described how long patients lived after a treatment.

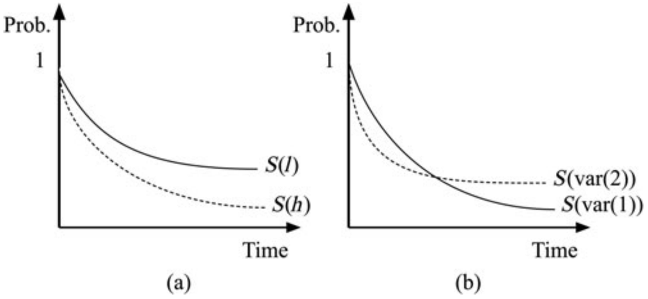

Figure 3. Changes in land quality mean and variance

the survival functions of land within a given territory will depend on the average quality, the variance in quality, and the cost of racing.

4. The Oklahoma land rushes and lottery

The area known today as Oklahoma was acquired by the US as part of the Louisiana Purchase in 1803. Over the next 30 years it was explored, but remained of little consequence until Andrew Jackson started negotiations with the Five Civilized Tribes of the South. Through treaty, payments, and force these five tribes marched along the "Trail of Tears" to the eastern regions of Oklahoma and established the "Indian Territory," a massive block of land that included all of modern Oklahoma except for the panhandle. By treaty, the tribes were quasi-sovereign states, and several sided with the South during the Civil War. After the war the territory came under reconstruction and those tribes who had signed treaties with the Confederacy (the Choctaw, Cherokee, Chickasaw, Creek, and Seminole) were forced transfer lands to the US, who then entered new treaties for land with the other western tribes (the Cheyenne, Arapaho, Kiowa, Comanche, and Apache) in 1866 (Kidwell, 2018). Throughout the 1860s and 1870s other tribes of the plains were relocated to the territory.

By the mid-1870s there were just 17,000 natives legally occupying 28 million acres, but living mostly in the eastern half of the territory (Wickett, 2000: 45). At the same time, and certainly by the mid-1880s, white settlement of the states surrounding the Indian Territory had become dense enough that settlers began to lobby and encroach on the Indian Territory lands. In response, the federal government created a long strip of neutral "no man's land" along the Oklahoma–Kansas border for protection, and "[t]hroughout the 1870s and into the early 1880s the United States government diligently tried to remove as many illegal white squatters from Indian lands as possible" (Wickett, 2000: 48). By the later 1880s the situation was untenable, and the General Allotment Act (known as the Dawes Act, 1887), the Curtis Act (1898) and other subsequent treaties were used to break up the tribal authority, end native self-government in the Indian Territory, and create private title by allocating (usually) 80–160 acres to each tribal member. The federal government then purchased back the surplus lands left over (under legislated terms), mostly in the western portion of the state. Prices ranged from 30 cents per acre to $1.46 per acre, and by 1893 15,100,538 acres had been acquired by the federal government for settlement (Kidwell, 2018).

Then, between 1889 and 1901 six large areas were established for homesteading. The first territory, opened up for homesteading on April 22, 1889, was purchased from the Creek and Seminole and known as the "Unassigned Lands." This rush was followed by the Deep Fork Rush (made up of lands purchased from the Iowa, Sac and Fox, and Shawnee–Pottawatomie lands) on September 22, 1891; the Cheyenne–Arapaho Rush on April 19, 1892; the Cherokee Outlet lands on September 16, 1893; and the Kickapoo Rush on May 23, 1895. For each race there was no fee, other restrictions, or registration to enter, and hence the number of racers was determined by individual choice.

Finally, between July 9 and August 6, 1901 the Caddo, Comanche, Kiowa, Apache, and Wichita lands were given away through a lottery. Each day a group of lottery winners was selected from names in a barrel, and then these winners would race to their preferred site and make a claim. The

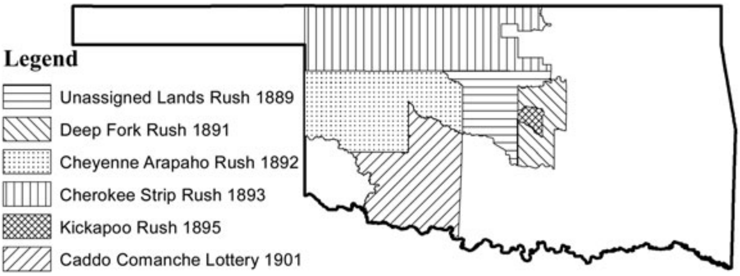

Figure 4. Oklahoma Land Rush/ Lottery Territories

next day the process was repeated. Figure 4 shows the six different territories involved in the land give-away, and shows that they made up almost one-half of the entire state.

Empirical work

Data

Interestingly, there is actually little statistically known about the Oklahoma land rushes, mostly due to the fact that until recently land settlement data for the entire western part of the US were limited and of low quality. Economic historian Richard Edwards notes that, as of 2008:

> Data available to scholars on homesteading are of very poor quality – inconsistent, unreliable, inaccessible, incomplete … When scholars look for data on homesteading, there are seven basic sources. Unfortunately, each is seriously flawed. (Edwards, 2008: 181–182)

However, since that time, the almost two million historical patent files have been digitized, and made publicly available.[12] Among other things, each digitized patent record provides the location of the land patent, and the year the settler received title. Here this information is georeferenced and mapped.

For this project the land patents for each of the six territories were selected from the BLM files, starting in 1889 with the Unassigned Lands rush and terminating with 1911.[13] During this period and over this area there were 70,215 homestead land patents issued. Interestingly, not all homesteads were 160 acres. Of the total, 10,646 were “quarter-quarters” (40 acres). Different sized homesteads usually reflected relative proximity to (planned) townsites. Due to some measurement error in the BLM data, there are 36,817 plots that can only be located to a specific section (640 acres), and not a quarter. Therefore, in order to use all plots regardless of size, we conduct our analysis at the section level.[14]

The BLM land patent files contain information on the actual plots of land that were *taken*. It is important for our analysis to know which plots of land were not taken, and to know the actual latitudes and longitudes of all plots of land. This was accomplished by taking a geocoded Public Land Survey System (PLSS) shapefile of *every section* within the territory, provided by the state of Oklahoma, and merging it with the BLM land patent data. Those sections that were not matched with a land patent were considered not taken (empty).[15]

Table 1 shows the average values of our land quality variables with standard deviations reported in parentheses. The standard deviation of elevation is a measure of land roughness (Ascione *et al.*, 2008), and for farming and most other uses a smaller variation is more valuable. Soil quality, average rainfall

---

[12] See the data appendix for details on these files.

[13] The end date of 1911 was selected only because this is when the railroad GIS shapefile information ceases. By this time every one of the rushes was well over.

[14] This means that some of our distance measures will contain some random errors. Of the non-quarter plots, the average size was 150 acres. See the data appendix for a description of the digitizing error.

[15] An empty section is not matched with *any* BLM observation, which means it was not homesteaded, sold, allotted, or otherwise transferred to private ownership. See the data appendix for more information on the PLSS file.

Table 1. Summary statistics by region

	Kickapoo (Rush)	Unassigned Lands (Rush)	Cherokee Strip (Rush)	Cheyenne–Arapaho (Rush)	Deep Fork (Rush)	Caddo Comanche (Lottery)	All
Parcel acres	136.62 (37.82)	151.09 (25.84)	151.85 (24.22)	151.45 (24.86)	140.55 (49.86)	153.51 (21.71)	150.85 (27.45)
Standard elevation	7.14	5.82	6.23	8.41	8.00	7.10	7.04
	(2.74)	(2.78)	(3.64)	(3.40)	(2.22)	(6.29)	(3.87)
Average soil	10.48	12.08	10.22	10.05	10.16	12.78	10.75
	(2.35)	(2.75)	(3.67)	(3.40)	(1.92)	(3.21)	(3.36)
Average precipitation	887.61	774.90	593.43	593.60	904.89	692.52	659.9
	(42.01)	(51.01)	(74.41)	(58.08)	(59.44)	(49.12)	(117.75)
Average 1889 rail	37.73	18.90	22.73	67.04	40.67	109.87	45.74
distance (km)	(7.85)	(12.80)	(16.38)	(29.51)	(15.67)	(34.57)	(37.41)
Average stream	1.86	1.85	1.96	2.00	2.09	1.82	1.94
distance (km)	(1.78)	(1.48)	(1.52)	(1.52)	(1.46)	(1.36)	(1.49)
Average border	1.85	11.53	18.19	19.00	6.55	13.86	15.88
distance (km)	(1.78)	(8.54)	(12.84)	(13.56)	(4.75)	(1.38)	(12.43)
Mean factor score	0.291	0.335	−0.008	−0.236	0.252	−0.039	0.001
	(0.167)	(0.195)	(0.355)	(0.282)	(0.179)	(0.367)	(0.363)
N sections	470	16,610	33,346	28,411	7,182	10,100	96,119
(for analysis)							
Available parcels	300	12,000	40,000	25,000+	7,000	13,000	NA
Entrants	3,000+	50,000	100,000	25,000	20,000	170,000	NA
Days' notice	5	30	28	7	4	17	NA

Notes: Stream distance is in kilometers, rail distance is in tens of kilometers. Values based on lands homesteaded or left empty. Data on available parcels, entrants, and the number of days between the announcement of the race and the start date are from Bohanon and Coelho (1998). Standard deviations in parentheses.

in 1895, and distances to streams and rail are our other measures of land quality. All of these measures are exogenous, including the two rail lines, which were in place well before the territories were opened up to settlement.[16] The variable "Border Distance," which gives the distance of a given land section to the race starting line on the territory border, is one measure of the cost of racing.[17]

Table 1 also reports the mean factor score and its standard deviation, which is a measure of the latent value "land quality," based on the five measures of land quality in the table.[18] Looking at the mean factor score reveals some interesting differences across the various territories. Not surprisingly, the Unassigned Lands have the highest mean score (these lands are almost the highest quality on every dimension).[19] The Deep Fork and Kickapoo territories are also high in average quality, and have similar standard deviations in mean scores. On the other hand, the Cherokee Strip, Caddo Comanche, and Cheyenne–Arapaho territories are all below the average quality measure. However, the Caddo Comanche territory has the highest variance in land qualities.

The final three rows of Table 1 report the number of available parcels, the total number of entrants, and the time that elapsed between the public announcement and actual start of each race based on Bohanon and Coelho (1998), who relied on secondary historical accounts for the numbers of racers. Although these numbers are likely exaggerated (see Leonard and Allen, 2018), they provide some context for conceptualizing the demand for land in each area in Figure 1. Evidently, demand was greatest relative to the fixed supply of land in the rushes where there was more advance notice – the Unassigned Lands, the Cherokee Strip, and the Caddo Comanche Lottery. This is consistent with our modeling assumption that individuals make investments in speed prior to the start of the race.

Testable predictions

The rationing by racing model makes a number of predictions that can be tested with the available data. We consider two sets of predictions: the first concerns racing speeds within any given race and the second focuses on differences in speeds between races.

The first prediction of the model is that higher quality land should be taken sooner. This implies that for any given race, there should be a positive relationship between the racing speed and each of the various measures of land quality such as terrain ruggedness, soil quality, and rainfall. Second, the lower the cost of racing, the sooner the lands should have been taken. The distance of a given land plot to the starting line is a good proxy for costs, since the farther the distance, the farther the race. We expect these predictions to hold broadly across all of the races and within any particular race, because they concern the effect of parcel-level land characteristics on claimants' behavior.

The model also implies several predictions about how equilibrium speed should vary across races. First, races with a higher average land quality should be associated with earlier claims, on average (see Figure 2). Second, races with more participants relative to the supply of parcels should have faster speeds (see Figure 1). Third, races with lower racing costs lead to lands being taken sooner. Fourth, a mean-preserving spread in land quality should lead to a crossing point between the survival functions for two races.

Our data on individual claims allow us to test the first set of hypothesis directly. We also provide suggestive evidence that differences across the races are consistent with the model, but formally testing our race-level predictions is complicated by the fact that we only observe six distinct races. We begin by analyzing differences in the survival and hazard functions across races before proceeding to a more rigorous analysis of individuals' racing behavior.

---

[16] See Data Appendix for a map of the 1889 railroad locations.

[17] The starting line was the entire border of the territory.

[18] The factor analysis shows there are three relevant factors: rail distance, stream distance, and land qualities.

[19] The unfortunately ambiguous term "Unassigned Lands" is simply the name given to the first territory opened up (see Figure 4).

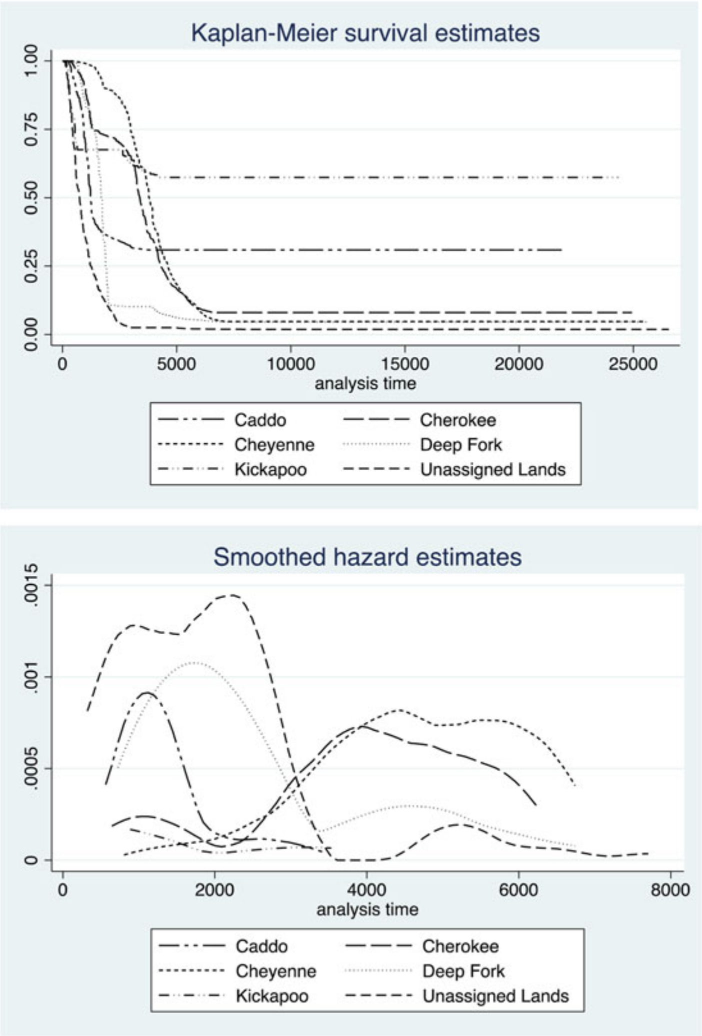

Figure 5. Hazard and survival rates

Non-parametric survival and hazard rates

Figure 5 shows the survival and smoothed hazard estimates for each region for lands that were either homesteaded or left untaken. The Kaplan and Meier non-parametric estimator of the survival function $S(t)$ gives the probability of a plot of land not being taken by moment t. The same non-parametric estimator of the hazard function $h(t)$ provides the instantaneous rate of take-up (homesteading) at a given time t, which is simply a different way of representing the information in the survival function. Analysis time is measured here in days.[20] These non-parametric measures can be used to broadly test the model.

Both graphs show that there was a great deal of variance across the races in terms of how long each rush took to complete and the percentage of lands taken. In terms of the length of time, the land rushes ranged from 5 to 16 years. In terms of the fraction of land ultimately taken in each region,

---

[20] The BLM land patent data identify the date the claim was filed. This is unlikely to be the date that the homesteader arrived on the plot of land and literally claimed the site. For various reasons, a homesteader may have stayed on site, built a shelter for his family, and then made his way to the land office. It is unlikely the homesteader waited months, but at the day level, arrival time has considerable measurement error.

in the Caddo Comanche lottery and the last land rush in Kickapoo, not all of the "free" land was taken. Both of these facts suggest that the historical narratives and estimated numbers of home-seekers were likely greatly exaggerated. Had 100,000 people shown up on day one to race for 40,000 plots, the race would have been over in days, all land taken, and many would have been disappointed.[21]

From Figure 5 we see that three rushes (Unassigned Lands, Deep Fork, and Caddo Comanche) had lands taken earlier than the other three (Cheyenne–Arapaho, Cherokee Strip, and Kickapoo).

For some intuition, consider the two extreme cases: the Unassigned Lands territory and the Kickapoo territory. For the Unassigned Lands the survival function rapidly comes close to zero, and for any given t, the probability of survival is lowest for this territory (similarly, the hazard rates are highest for the Unassigned Lands). The rush here was swift (about 7,000 claims within a year and most claims within four years), and by 1911 only 1.6% of lands were unclaimed.[22] In contrast, for the Kickapoo territory there was a small initial rush (accounting for about 75% of homesteads) within the first two years followed by some claims made eight years later, but with much of the land remained untaken. Indeed, by 1911 53% of the sections in Kickapoo remained unclaimed.

Using the mean scores for land quality found in Table 1, and grouping the races into high and low quality lands, we can test the model prediction that higher quality land should be taken sooner. Figure 6 shows the survival functions where the three highest quality races are grouped together and contrasted with the two lower quality races.[23] The results confirm the racing hypothesis, since the survival functions of the higher quality lands always lie everywhere to the left of the low quality lands – meaning these lands were taken first.[24]

Figure 7 shows the survival and smoothed hazard estimates for the Caddo Comanche (lottery) territory compared to the other five land rush territories in order to test another model prediction. One significant difference for the Caddo Comanche territory was that it allocated lands based on a lottery. The lottery took place over several months, with winners announced every day. Unlike the other races, claimants in the lottery made a claim in the order in which their name was drawn and competed with only the other winners. This spot in the queue could be forfeited if a claim was not made within a certain time (Bohanon and Coelho, 1998).

The model suggests that the lottery can be thought of as a series of races. On the first day a set of racers are chosen, and they can select from all plots within the territory. On the second day, a new set of racers are chosen, and they select from the plots remaining untaken. Each subsequent lottery has a decreasing number of high quality parcels of land left. Thus, the supply of high quality land falls over time. On the demand side, the lottery initially lowers the market demand for land because there are fewer racers in the market on the first day. However, lottery winners who do not claim land on the day they win join the pool of racers the following day. Thus, over time, the demand for high quality land increases. Therefore, speed should have increased over the first few days of the race, but within a short period of time the actual effect of the lottery should have been small relative to the other races.

The factor that really distinguished the Caddo Comanche territory was the large variance in the quality of its lands (from Table 1) – terrain ruggedness, soil quality, and the mean factor score all have a higher standard deviation on the Caddo Comanche lands than most other territories, despite having similar means. The model depicted in Figure 2 predicts that the two survival functions should cross because the higher quality lands will be taken more rapidly, while the lower quality lands will be left untaken.

From Figure 7 the survival graph clearly shows that the Caddo Comanche lottery allotted lands sooner than the other land rushes, and at the end of the day not all of the land was allotted. Likewise with the hazard graph on the right we see that the hazard peak was earlier for the Caddo

---

[21] See Allen and Leonard (2018) for a full description of this narrative, and other evidence for why it is inconsistent with the land patent data.

[22] This does not mean that 98.4% of lands were homesteaded. There were plots allotted to natives, some were sold, or otherwise transferred in other ways. Of all lands in the Unassigned Lands Territory, about 85% were homesteaded.

[23] The Caddo Comanche territory is not included because it was a lottery.

[24] Unfortunately, in Table 1 there seems to be no case of a mean preserving increase in spread across the various territories.

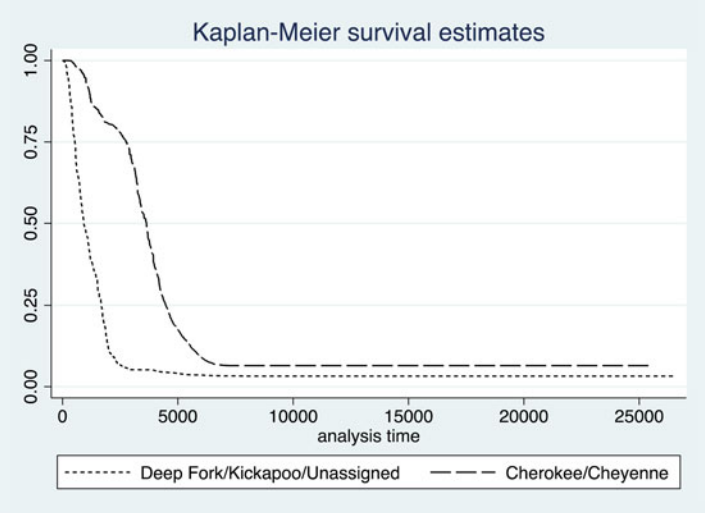

Figure 6. Survival of high *versus* low quality lands

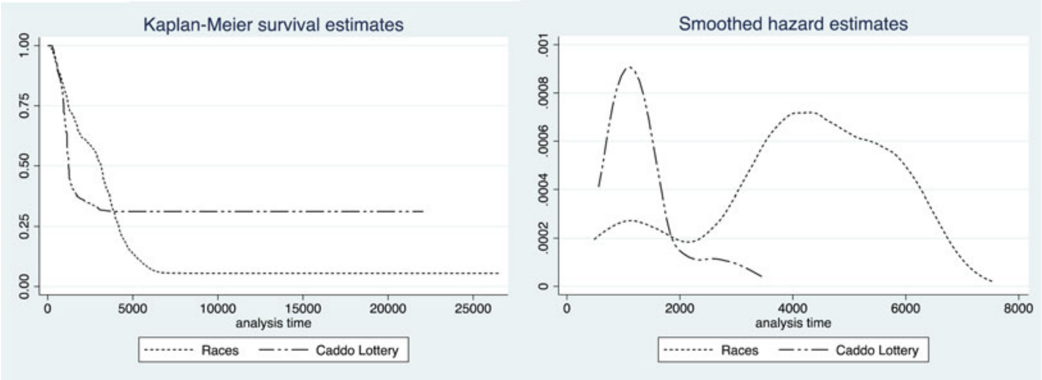

Figure 7. Caddo Comanche Lottery *versus* other rushes

Comanche lands, and stops sooner, both consistent with the racing model. In addition, the Caddo Comanche survival function crosses the average survival function for the other races, as predicted. Unfortunately there are no obvious mean-preserving spreads in quality among the five races that would allow us to hold constant the cost of racing for testing the predictions of Figure 2.

The Kaplan–Meier survival estimates are broadly consistent with the model of rationing by racing. Comparisons of aggregate survival and hazard functions across races make it difficult to assess the effect of our various measure of land quality on speed. However, we next examine the issue of land acquisition through a Cox proportional hazard estimation in order to control for the various dimensions of land quality.

Cox semiparametric estimation

The Cox proportional hazards model assumes that the hazard rate for the i^{th} data subject takes the following form:

$$h(t|x_i) = h_0(t)e(x_i\text{ß}_x),$$

where $h_0(t)$ is a constant baseline hazard, and ß_x are the regression coefficients to be estimated. The model places no parametric assumptions on the baseline hazard, which is reasonable in this context where there is no *a priori* knowledge about what it might look like across the six territories. The model has no constant term because it is absorbed by the baseline hazard function (which is not estimated).

The values reported in the regression tables below are the *hazard ratio*, $e(x_i\beta_x)$. The hazard ratio is the Cox hazard rate evaluated for a one unit change in a variable x, divided by the Cox hazard rate evaluated at the original value of x. Hence, a hazard ratio of 1 means that the change in variable x had no change in the hazard rate. A hazard ratio less than 1 means the change in x lowered the hazard rate (failure was less likely), which in the context here would mean a plot of land was less likely to be homesteaded. Likewise, a hazard ratio greater than 1 means that the change in x increased the likelihood of a plot of land being taken/homesteaded.

The advantages of the Cox model, aside from the lack of assumptions regarding the baseline hazard, is that it allows for an estimation of multiple covariate effects. A disadvantage of hazard models in general is dealing with unobserved heterogeneity across groups (or races, in our case). The leading approach for accommodating unobserved heterogeneity is to estimate a shared frailty model, which specifies a group-level random effect that shifts the hazard function multiplicatively. In our case, we are also interested in the mean difference in arrival times across races, which is not identified if a group-level random effect is included.

Our approach, therefore, is to estimate a linear model with $ln(days\ elapsed)$ as the dependent variable to capture mean differences between races, and then to focus on shared frailty Cox proportional hazards model to study the effect of other covariates. Table 2 reports the results of the linear model (column a) and three different Cox estimations (columns b through d). Columns (a) and (b) include all six rushes. Column (b) includes all races and stratifies estimates by race. Column (c) omits the Caddo Comanche lottery, whereas Column (d) includes only the lottery.

$$Y = ln(days) \text{ in Column (a). Columns (b)} - \text{(d) fit a Cox model.}$$

Across all of the regressions, the plot quality and cost measure coefficients/hazard ratios are consistent with the rationing by racing model. Flatter lands, higher quality soils, moist environments, and larger parcel size all induce faster races to claim the land (shorter durations in column (a) and hazard ratios greater than 1 in columns (b) through (d)). Likewise, lands closer to railroads and the starting lines are also taken sooner. Interpretation of the hazard ratios is straightforward. For example, in the case of terrain roughness, a one unit increase in the standard deviation of elevation lowers the hazard *rate* of lands taken by about 1.5%.

The total number of acres variable squared is included because smaller plots are mostly high valued townsite lands, while very large plots greater than the standard 160 acres are mostly low valued quasi-desert lands in the western portions of the state. The signs on these variables are consistent with this interpretation.

The fixed effects coefficients in column (a) indicate that plots were taken later in all of the races other than in the lottery, with the exception of the Unassigned Lands, which had lands taken sooner (recall these lands had the highest average quality). Comparison of columns (c) and (d) indicates that the various measures of land quality had similar effects on speed, and therefore time to being taken, within the races *versus* the lottery. Two variables show different effects (precipitation and stream distance), and this likely results from the importance of railroad distance. The Caddo Commanche area was the most remote in terms of rail, making this feature exceptionally important. The general similarity implies that claimants responded to parcel-level land characteristics similarly in the lottery *versus* the races, with the important difference that the initial speed was greater in the lottery. Taken together, these two findings are consistent with our interpretation that the lottery *per se* had little effect on the racing speed, and that the variance in land quality is what drives the difference.

Table 3 provides robustness checks and falsification tests on the pooled hazard estimates from Table 2. Column (a) uses the standard Cox model, whereas column (b) is a shared frailty model.

Table 2. Linear and Cox proportional hazard model estimates

	All races	All races	Non-Caddo	Caddo
	(linear)	(Cox)	(Cox)	(Cox)
	(a)	(b)	(c)	(d)
Standard elevation	0.0139	0.9888	0.9807	0.9949
	(13.50)	(−10.07)	(−15.64)	(−2.09)
Average soil	−0.0256	1.0317	1.0341	1.0183
	(−21.86)	(25.00)	(25.89)	(4.04)
Average precipitation	−0.0007	1.0004	1.0001	0.9978
	(−7.84)	(4.55)	(1.04)	(−3.37)
1889 rail	0.0031	0.9987	0.9993	0.9462
distance	(20.08)	(−7.41)	(−3.61)	(−6.97)
Stream	0.0111	0.9813	0.9686	1.0337
distance	(5.04)	(−7.35)	(−12.02)	(3.30)
Border distance	0.0031	0.9999	1.0000	0.9999
	(11.34)	(−1.65)	(6.87)	(−14.25)
Total acres	−0.0177	1.0247	1.0244	1.0329
	(−20.98)	(18.32)	(17.82)	(5.39)
Acres squared	0.00007	0.9999	0.9999	0.9998
	(20.64)	(−17.35)	(−16.06)	(−6.59)
Unassigned Lands	−0.7892 (−31.16)			
Deep Fork	0.0444			
	(1.56)			
Cheyenne	0.2072			
	(10.36)			
Cherokee	0.2231			
	(9.92)			
Kickapoo	1.1239			
	(10.65)			
N	74,578	74,578	66,889	7,689

Notes: z-scores in parentheses. Values either homesteaded or left empty.

Column (c) of Table 3 reports the same regression, but where the factors of the quality variables are used rather than the land quality variables themselves. Table 2A in the Data Appendix reports the scoring coefficients of the factor analysis, which is based on the correlations of the land quality variables. From that table, factor 1 can be interpreted as an index variable that captures the land attributes roughness, soil quality, and precipitation. Factor 2 is an index for roughness and distance to a stream, and factor 3 is an index mostly influenced by distance to the 1889 railroad. Column (c) is quite striking, and shows that a unit move in “land quality” improvement increases the hazard rate of land taken by almost 40%. Speed allocation is very sensitive to values of Ω, the individual’s value of the land in our theoretical model.

Columns (d) and (e) represent falsification tests. There were three major ways in which lands were allocated in these six territories. The first, and most significant was through homesteading. The second

Table 3. Robustness hazard model estimates

	All races (a)	All races (b)	All races (c)	Indian allotments (d)	Cash sales (e)
Standard elevation	0.9830	0.9874		1.0116	0.9867
	(−15.41)	(−11.44)		(1.64)	(−6.13)
Average soil	1.0285	1.0287		0.9299	1.0067
	(23.24)	(22.63)		(−9.14)	(2.71)
Average precipitation	1.0025	1.0007		1.0040	0.9969
	(36.85)	(7.04)		(7.56)	(−14.73)
1889 rail	0.9930	0.9970		0.9913	0.9882
distance	(−59.12)	(−17.52)		(−8.06)	(−5.42)
Stream	0.9704	0.9831		0.9098	0.9632
distance	(−11.70)	(−6.33)		(−5.81)	(−6.91)
Border distance	0.9964	0.9997		0.9997	0.9959
	(−10.82)	(−0.96)		(−0.07)	(−5.89)
Total acres	1.0246	1.0249	0.9889	0.989	1.0278
	(17.77)	(18.40)	(−8.32)	(−8.31)	(15.97)
Acres squared	0.9999	0.9998	1.0000	1.000	0.9998
	(−17.39)	(−17.70)	(9.15)	(−9.21)	(−18.23)
Factor 1			1.399		
			(32.43)		
Factor 2			0.770		
			(−11.798)		
Factor 3			0.774		
			(−8.88)		
N	74,578	74,578	74,578	8,351	22,963
Controls	Yes	Yes	Yes	Yes	Yes
Territory FE	No	Yes	Yes	Yes	Yes

Notes: *z*-scores in parentheses. $Y = 1$ if homesteaded for (a)–(c). $Y = 1$ if Indian allotment for (d), $Y = 1$ if cash sale for (e).

was through cash sales, and finally, the third was through "Indian allotments." Prior to any lands being allocated by homesteading, the Jerome Commission negotiated settlements with the various tribes involved. These settlements included individual cash payments, and allotments of lands to the resident Native Americans. These allotments were called Indian allotments. Little is known about how these allotments were made, but presumably the process was similar to the broader allotment of reservation lands under the Dawes Act of 1887. In that context, land was set aside for allotments and then assigned to individual Native Americans at the *discretion* of the local Bureau of Indian Affairs agent (Carlson, 1981). Most notably, there was no race or price mechanism used.

Column (d) shows how the various land quality measures relate to the timing of the Indian allotments. Of these allotments, 71% were given out in 1872, and 60% were given in the Deep Fork territory. Deep Fork is the easternmost territory of the various rush territories and was relatively close to the 1889 railroad going through the Unassigned Lands. Thus, the hazard ratios for precipitation and rail distance are not too surprising. From the elevation/roughness and the soil quality variables, however, it would appear that the early allotments were of lower quality, which is an opposite result from allocation with speed. This is consistent with the hypothesis that agents of the Bureau of Indian Affairs (BIA) strategically reserved better land for whites, allocating the lowest quality lands to natives first.[25]

Column (e) runs a final Cox estimation on the patent timing of cash sales for land. When land is allocated by market price, higher quality land sells for more. In an open market the marginal buyer should be indifferent between various land qualities (all Cox coefficients should be 1). Although land sales used price to allocate, a competitive open market was not ubiquitous. For most plots the price was set by legislation, and only parcels worth more than this price would have been chosen. Hence, land sales should have responded to land quality in a similar fashion to homesteaders; that is, higher quality lands should have sold first. This is consistent with all of the land quality variables except precipitation.[26] Again, different allocation mechanism, different relationship between land qualities and patent timing.

5. Conclusion

The term "racing," when used in economics, is almost always a metaphor or a reference to a sporting event, and is not used to indicate a method of allocation. On the other hand, the term "rush" is almost always used in the context of a gold, land, or other resource rent-seeking rush. In this latter case the question of allocation is on the table, but to our knowledge there has been no attempt to model the rush *per se*. Rather, the interest is in the amount of wealth dissipation caused by the rushing.

Here we have mostly ignored the question of wealth dissipation in order to examine the mechanics of an actual race. In our analysis of the Oklahoma land rushes through the lens of Yoram Barzel's model of rationing by waiting, we demonstrate three things. First, the analysis shows the general applicability of Barzel's model. At the time of writing, Barzel used his model to discuss waiting for gasoline in an era of price controls. The case of Oklahoma is about as "out of sample" a test as one can imagine. Yet despite the different century and different context, the Barzel analysis stands up.

Second, despite a few attempts to directly test the Barzel model, this application probably comes closest to testing the actual predictions of the model, even though the tests are conducted through an analysis of speed rather than waiting times. Finally, as with most of Barzel's work, once an allocation mechanism like racing is understood, it naturally begs the question: why was such a dramatic and costly race used to allocate so much land? We leave this question for future consideration.

Acknowledgments. **Thanks to Yoram Barzel, D. Bruce Johnsen, Dean Lueck,** and seminar participants at Canterbury, Auckland, and Loughborough Universities for comments, and to Ian Broad (ian.broad@outlook.com) for his assistance in geo-referencing the Oklahoma data.

References

Allen, D. W. (forthcoming, 2019), 'Establishing Economic Property Rights by Giving Away an Empire', *Journal of Law and Economics*.

---

[25] Running a logit regression with Indian Allotment as the dichotomous dependent variable on the same covariates, provides correlations consistent with the Cox hazard ratios: Native Americans were given poorer quality lands. This finding contrasts somewhat with Leonard *et al.* (2018), who find that higher quality lands were more likely to be allotted in a sample of reservations in the Northern Great Plains. These allotments were made at the discretion of a different BIA agent facing different settlement pressure than in Oklahoma. The difference may also be due to the fact that Leonard *et al.* focus on the extensive margin and not on the timing of claims.

[26] Precipitation is strongly correlated with location of territory, and so this variable measures the sales across the different territories. If a similar estimation is done to (c) on the three factors, the results are the opposite for cash sales. Factor 1 has a coefficient/*z*-score of 0.933/−3.31; factor 2, 1.180/3.74; and factor 3, 1.45/6.75.

Allen, D. W. and B. Leonard (2018), 'Did Hundreds of Thousands Rush During the Oklahoma Land Openings?' Mimeo, SFU.

Anderson, T. and P. J. Hill (1990), 'The Race for Property Rights', *Journal of Law and Economics*, **33**(1): 177–197.

Ascione, A., A. Cinque, E. Miccadei, F. Villani, and C. Berti (2008), 'The Plio-Quaternary uplift of the Apennine Chain: New Data from the Analysis of Topography and River Valleys in Central Italy', *Geography*, **102**(1): 105–118.

Barzel, Y. (1968a), 'Optimal Timing of Innovations', *Review of Economics and Statistics*, **50**(3): 348–355.

Barzel, Y. (1968b), 'Costs of Medical Treatment: Comment', *American Economic Review*, **58**(4): 936–939.

Barzel, Y. (1969), 'Productivity and the Price of Medical Services', *Journal of Political Economy*, **77**(6): 1014–1027.

Barzel, Y. (1971), 'Investment, Scale and Growth', *Journal of Political Economy*, **79**(2): 214–231.

Barzel, Y. (1974), 'A Theory of Rationing by Waiting', *Journal of Law and Economics*, **17**(1): 73–95.

Barzel, Y. (1976), 'An Alternative Approach to the Analysis of Taxation', *Journal of Political Economy*, **84**(6): 1177–1197.

Barzel, Y. (1977), 'An Economic Analysis of Slavery', *Journal of Law and Economics*, **20**(1): 87–110.

Billington, R. A. and M. Ridge (2001), *Westward Expansion: A History of the American Frontier*, Albuquerque, NM: University of New Mexico Press.

Bohanon, C. E. and P. Coelho (1998), 'The Costs of Free Land: The Oklahoma Land Rushes', *Journal of Real Estate Finance and Economics*, **16**(2): 205–221.

Bose, P. (1997), 'Adverse Selection, Waiting Lists and Restaurant-Rationing', *Journal of Industrial Organization*, **15**(3): 335–347.

Carlson, L. A. (1981), 'Land Allotment and the Decline of American Indian Farming', *Explorations in Economic History*, **18**(2): 128–154.

Deacon, R. and J. Sonstelie (1985), 'Rationing by Waiting and the Value of Time: Results from a Natural Experiment', *Journal of Political Economy*, **93**(4): 627–647.

Edwards, R. (2008), 'Why the Homesteading Data Are So Poor (And What Can Be Done about It)', *Great Plains Quarterly*, Summer: 181–190, available at http://digitalcommons.unl.edu/greatplainsquarterly/1374/?utm_source=digitalcommons.unl.edu%2Fgreatplainsquarterly%2F1374&utm_medium=PDF&utm_campaign=PDFCoverPages (accessed 28 February 2019).

Gravelle, H. and L. Siciliani (2009), 'Third Degree Waiting Time Discrimination: Optimal Allocation of a Public Sector Healthcare Treatment under Rationing by Waiting', *Health Economics*, **18**(8): 977–986.

Kidwell, C. S. (2018), 'Allotment', *Encyclopedia of Oklahoma History and Culture*, available at www.okhistory.org (accessed 25 May 2018).

Kulshreshtha, P. (2007), 'An Efficiency and Welfare Classification of Rationing by Waiting in the Presence of Bribery', *Journal of Development Economics*, **83**(2): 530–548.

Leonard, B. and D. W. Allen (2018), 'Property Rights and Path Dependence: 19th Century Land Policy and Modern Economic Outcomes', mimeo, ASU.

Leonard, B., D. Parker and T. L. Anderson (2018), 'Poverty from Incomplete Property Rights: Evidence from American Indian Reservations', working paper, ASU.

Lueck, D. (1995), 'The Rule of First Possession and the Design of the Law', *Journal of Law and Economics*, **38**(2): 393–436.

Lueck, D. and T. Miceli (2007), 'Property Law', *Handbook of Law and Economics*, A. M. Polinsky and S. Shavell (eds), North Holland: Elsevier.

Suen, W. (1989), 'Rationing and Rent Dissipation in the Presence of Heterogeneous Individuals', *Journal of Political Economy*, **97**(6): 1384–1394.

Wickett, M. (2000), *Contested Territory: Whites, Native Americans, and African Americans in Oklahoma 1865–1907*, Baton Rouge, Louisiana State University Press.

Yeung, R., G. Leung, S. McGhee, and J. Johnston. (2004), 'Waiting Time and Doctor Shopping in a Mixed Medical Economy', *Health Economics*, **13**(11): 1137–1144.

Data appendix

Land patent files

The Bureau of Land Management, Eastern States Office (BLM), provides bulk data from federal land patents at www.glorecords.blm.gov/BulkData/default.aspx. These multiple files contain the Public Land Survey System (PLSS) locations for each land patent. The PLSS system identifies a plot of land through aliquot part, section, township, and range numbers. Each township is 36 sections, and is identified by a township and range designation. For example, a township might be 8 North, Range 3 East, which means it is in the eighth tier of townships north of the baseline, and in the second column of townships east of the reference meridian. A section is one square mile (640 acres), and is often divided into quarter-sections (160 acres). The "aliquot parts" identify the quarter-section (e.g. NW is the northwest quarter-section). As lands are subdivided, the aliquot parts become very complicated. Hence, in the BLM records, one specific Nebraska plot is

Table A1. Definitions, Means

Variable	Definition	Mean	Min.	Max.
Standard elevation	= Standard deviation in elevation of a given section	7.04	0	75.07
Average elevation	= Mean elevation of a given section	484.96	229.2	781.0
Stream distance	= Nearest distance (km) between major stream and section	1.94	0	10.9
Average soil	= Mean of soil index for given section	10.75	0	18
Standard soil	= Standard deviation in soil index of a given section	0.70	0	8.97
Average precipitation	= Mean rainfall (mm) for given section measured in 1895	659.9	478.4	1,087.9
Standard precipitation	= Standard deviation of precipitation by section	1.54	0	16.38
Rail distance	= Nearest distance (10s of km) from section to 1889 rail	4.57	0	18.64
Total acres	= The size of the plot of land	150.85	1	463
Caddo	= 1 if land was in the Caddo Comanche lottery territory	0.105	0	1
Cherokee	= 1 if land was in the Cherokee Strip territory	0.347	0	1
Cheyenne	= 1 if land was in the Cheyenne Arapaho territory	0.295	0	1
Deep Fork	= 1 if land was in the Deep Fork territory	0.074	0	1
Kickapoo	= 1 if land was in the Kickapoo territory	0.004	0	1
Unassigned Lands	= 1 if land was in the Unassigned Lands territory	0.172	0	1

described as SW1/4NW1/4 26 T3S R11W, which is the southwest 40 acres, in the northwest quarter-section, of section 26 in township 3 South, Range 11 West. For each section the geographic center is used for its location, and this center touches each quarter-section, which is the unit of analysis.

The PLSS location variable is not statistically useable. Rather GIS files for Oklahoma were downloaded from the Oklahoma GIS Data Clearinghouse (https://okmaps.org), which included the PLSS information along with the latitudes and longitudes of every section. The sections located in the six territories were extracted, and the BLM information merged in. Each section of land is a "polygon," and can be thought of as a "cookie cutter" for other GIS files.

We used the geographic mapping program ArcGIS to build our data files. Using the section polygons we then assembled multiple shapefile data and added it to the data set. An important shapefile used was the railroad shapefile created by Jeremy Atack, which is located at https://my.vanderbilt.edu/jeremyatack/data-downloads/.

From this file we were able to calculate the closest distance of every section from the railroads that existed in 1889. Likewise, the stream shapefile (from the Oklahoma clearinghouse) was used to calculate the nearest distance from a section to a major stream.

Information on our land quality variables came from various "raster" files. These are pixel maps that contain the relevant information. The section polygons were used to cut out the relevant pixels, and from these the information was then extracted into a flat data file and merged with the other information.

The elevation data are from: https://datagateway.nrcs.usda.gov/.

The soil quality data are from: www.geographer-miller.com/productivity-index-grid/.

The precipitation data are from: PRISM historic climate data (http://prism.oregonstate.edu/historical/).

We used the Contiguous USA Albers projection system to generate distance measures between the centers of each section to the nearest object specified (a railroad or stream).

The definitions of variables and some summary statistics are provided in Table A1.

Factor analysis

The factor scoring coefficients were calculated using the regression method where the factors were correlated (oblique rotation). The coefficients are given in Table A2.

Table A2. Factor coefficients

Variable	Factor 1	Factor 2	Factor 3
Rail distance	0.010	0.067	0.229
Standard elevation	0.211	0.214	0.146
Average soil	0.368	0.098	0.008
Average precipitation	0.285	0.041	0.014
Stream distance	0.004	0.173	0.040

1889 railroads

Figure A1 shows the railroads that existed in Oklahoma in 1889, just as the land rushes were about to start. For the calculation of nearest distance, the railroads in the eastern portion of the state are irrelevant. Most of the territory lands were a long distance from existing railroads.

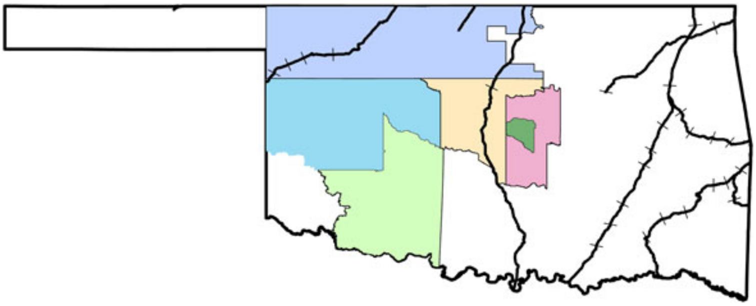

Figure A1. 1889 railroads

Cite this article: Allen DW, Leonard B (2020). Rationing by racing and the Oklahoma land rushes. *Journal of Institutional Economics* **16**, 127–144. https://doi.org/10.1017/S1744137419000110