NBER WORKING PAPER SERIES

SHOCKING BEHAVIOR : RANDOM WEALTH IN ANTEBELLUM GEORGIA AND HUMAN CAPITAL ACROSS GENERATIONS

Hoyt Bleakley
Joseph P. Ferrie

Working Paper 19348
http://www.nber.org/papers/w19348

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2013

For comments on early versions, we thank Lou Cain, Greg Clark, Robert Pollack, Chris Roudiez, Rachel Soloveichik, and seminar and conference participants at the NBER Cohort Studies meeting, the Stanford SITE meeting, the University of Washington, the University of Michigan, the University of Chicago, and the Minnesota Population Center and for helpful comments, to Steven Ruggles for access to a preliminary version of the 1850 full-count file. Bleakley gratefully acknowledges research support from the Center on Aging, the Center for Population Economics, the Stigler Center, and the Deans' Office of the Booth School of Business, all at the University of Chicago. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Shocking Behavior : Random Wealth in Antebellum Georgia and Human Capital Across Generations
Hoyt Bleakley and Joseph P. Ferrie
NBER Working Paper No. 19348
August 2013
JEL No. J01,J13,J24,J62,N11,N31,O12,O15

ABSTRACT

Does the lack of wealth constrain parents' investments in the human capital of their descendants? We conduct a fifty-year followup of an episode in which such constraints would have been plausibly relaxed by a random allocation of wealth to families. We track descendants of those eligible to win in Georgia's Cherokee Land Lottery of 1832, which had nearly universal participation among adult white males. Winners received close to the median level of wealth – a large financial windfall orthogonal to parents' underlying characteristics that might have also affected their children's human capital. Although winners had slightly more children than non-winners, they did not send them to school more. Sons of winners have no better adult outcomes (wealth, income, literacy) than the sons of non-winners, and winners' grandchildren do not have higher literacy or school attendance than non-winners' grandchildren. This suggests only a limited role for family financial resources in the transmission of human capital across generations and a potentially more important role for other factors that persist through family lines.

Hoyt Bleakley
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
bleakley@chicagobooth.edu

Joseph P. Ferrie
Department of Economics
Northwestern University
Evanston, IL 60208-2600
and NBER
ferrie@northwestern.edu

1. Introduction

The role of parents' resources in shaping the human capital of their children has been a central concern of economists since the work of Schumpeter (1951) and of social scientists more generally since Dewey (1889).[1] The extent to which parents are constrained in this process is of particular concern when capital markets are imperfect and parents cannot fully borrow against the future labor earnings of their children. In such situations, much of the productive potential of the children of resource-constrained parents may go unrealized, leaving both them and society as a whole poorer as a result. Nevertheless, the importance of such constraints is difficult to assess in the absence of randomized perturbations to family wealth.

In this study, we assess the impact over several generations of a large, exogenous random shock to family wealth that had the potential to relax the constraint faced by parents and allow them to make investments in the human capital of their children that they would not have otherwise been able to afford. Both the size of the windfall we study (close to the median wealth at the time) and the near-universal participation in the system that provided this windfall to roughly 15% of adult white males makes this setting unique until the advent of large-scale income-based social experiments in the late twentieth century.

Parental wealth might predict child outcomes for reasons other than borrowing constraints. Though more advantaged parents may simply have more assets (e.g. financial or human capital) that they can directly transfer to their children, their advantages might result as well from underlying characteristics (e.g. ability, ambition, or access to superior investment opportunities) that have led to their accumulation of more assets. If these underlying characteristics are passed on to their children, they would exhibit superior outcomes regardless of their parents' direct investments in them.

---

[1] See the summary in Becker and Tomes (1986).

The design of effective interventions depends on what plays a greater role in human-capital formation: financial constraints or the household's underlying characteristics. For example, if investment in the next generation is governed primarily by the resources parents have available, policies that relax that constraint will lead to greater investment.[2] But if children's human capital is instead the result of particular characteristics transmitted to them by their parents, policies that merely relax the financial resource constraint will prove ineffective in generating additional investment in children. The challenge is determining whether parents' financial and human capital are crucial inputs themselves in producing children's human capital, or simply of these resources are themselves merely indicators of higher underlying productivity which can be transmitted across generations.

Previous research seeking to disentangle the various intergenerational determinants of human capital (Chevalier 2004, Black *et al.* 2005, Oreopoulos *et al.* 2006) has focused on identifying plausibly exogenous policy changes that forced parents to invest more in their children's human capital than otherwise (e.g. the imposition of compulsory school attendance laws or raising school leaving ages). In the present study, we focus instead on a random shock to financial resources to separate these effects. If the household's resources are expanded through a random wealth shock but children's human capital is nonetheless unaffected, this is evidence for the role of underlying characteristics rather than parents' resources in shaping children's outcomes, as posited by Clark and Cummins (2013) in the British context.

We examine the results of a large-scale lottery in the mid-nineteenth century in the U.S. state of Georgia (the 1832 Cherokee Land Lottery).[3] At this time in Georgia, it was already clear that parents'

---

[2] In an extreme case, a "poverty trap" (Galor and Zeira 1993), parents cannot make even small investments in their children's human capital because of a fixed cost, and they cannot make large investments because of the binding financing constraint and their inability to borrow against their children's future earnings. In this situation, large transfers may be necessary to move families past the threshold at which it becomes feasible to begin investing in their children.

[3] See Figure 1 for the location of Old Cherokee County, the area settled through this lottery.

resources were linked to the human capital outcomes of their children and even their grandchildren: correlations were both substantively large and statistically significant between parental resources and the school attendance, literacy, wealth, and occupations of their children and grandchildren.[4] In this setting, we analyze the effects of random disbursements of wealth on fertility and human-capital investments over a long horizon.

The lottery generated a shock to an individual's wealth that we can plausibly expect was exogenous to his characteristics. Registration in this lottery was cheap and widespread (over 97% of those eligible were registered, by our calculations), unlike many studies of lotteries whose participants are a selective subset of the population. The prize in this lottery was a claim on a parcel of land with an average value close to the median wealth of the period. Lottery winnings were essentially a pure wealth shock – there was no homesteading requirement and the claim could be readily liquidated without even setting foot on the land. In addition, we are able to undertake a long-run followup on the effects of this wealth shock: we are able to examine family outcomes over nearly fifty years after the wealth disbursement.

In the first three decades of the nineteenth century, Georgia used lotteries to allocate over two-thirds of its area to white settlers. Following several large corruption scandals in Georgia in the 1790s, this peculiar manner of opening land was chosen because of its relative incorruptibility. (Section 2 presents background details on the Cherokee Land Lottery and discusses related work in the literature on intergenerational transmission.) We conduct a follow-up on these random wealth shocks using a sample (drawn from the 1850 Census) of over 14,000 men eligible to win land in the 1832 Land Lottery. From this sample of eligibles, we identify winners using a list published by the state of Georgia (Smith, 1838). Those found in the list comprise the treatment group, and the lottery eligibles not found in list

[4] See Section 8.2 below.

serve as a control group. Note that not we cannot verify that all of the men in our sample of eligibles did in fact register for the lottery, but our calculations (seen in Section 3) indicate that this was a minor subset. Further, in our sample, lottery losers look similar to lottery winners in a series of balancing tests using outcomes determined prior to the lottery and placebo regressions using a sample drawn from South Carolina instead. (These results are found in Sections 3 and 8.3, respectively.) We compute that the net-present-value of land won was at least several hundred dollars and perhaps as high as $1,000 in 1850 dollars. Further, we estimate with 1850 data that winners were, on average, $700 richer than the controls, almost two decades after the lottery.[5] (We present the estimation strategy, main regression equations, and estimates for 1850 wealth in Section 4.)

To measure the long-term effect of wealth on investments in children, we collect information from various years of the Census manuscripts. (Further description of the data are found in Section 3.) Existing indices of the census contain information on net fertility, residence, spousal age, etc. Adding to this, we transcribed information on wealth, literacy, and school attendance from the 1850 Census manuscripts, where children present in the household are recorded with their lottery-eligible fathers. We also link a subsample of these children to the 1870 and 1880 Census manuscripts to observe their outcomes as adults, as well as the outcomes of any grandchildren present by 1880.

When compared to a similar population that did not win the lottery, winners had only slightly higher (post-lottery) fertility but were no more likely to send their children to school. The fertility estimates, however, do not imply a particularly steep Engel curve for the number of children. The result for school attendance, on the other hand, suggests that the lack of paternal wealth was not a significant impediment to investing in a child's education. We show that the increase in (post-lottery) fertility took place along both intensive and extensive margins. In contrast, we do not find a significant result for

[5] An unskilled laborer in the South Atlantic region earned $0.74 per day in 1849 (Margo and Villaflor 1987, p. 880), so $700 in 1850 represented 945 days of work by an unskilled worker.

schooling when decomposing the result by age or gender of the children. We also show that these results (and all others of the paper) are robust to controlling for various factors, including characteristics of the person's name. The latter strengthens our conclusions in that, although we used the name to link to the list of winners, this does not appear to bias our estimate of the treatment effect. Further, while contextual influences on fertility and school attendance are no doubt important, they are not apparently an important mechanism for these results: effects are not sensitive to controlling for county of residence, nor do we find evidence that lottery winners move to counties with unusual fertility, schooling, land value, slave intensity, farm sizes, land improvement, urbanization, or transport access.

Effects of a random shock to paternal wealth on sons' human capital do not manifest themselves as we follow these children into adulthood in 1870 and 1880. School attendance is an imperfect proxy for human-capital investments and some aspects of human capital may not appear until adulthood. Linkage forward also allows us to determine if any fertility effect persists across generations. The sons' 1870 wealth (in real estate and/or personal property) is not statistically distinguishable between control and treatment groups. Nevertheless, a mechanical split of the 'extra' 1850 paternal wealth among his children would suggest a treatment effect in 1870 of $140 (in 1850 dollars), which we can reject for reasonable discount rates. Further, in 1880, we do not find differences in occupational standing or literacy as a function of their father's lottery status. In 1880, the grandchildren themselves do not have significantly greater literacy or school attendance if their grandfather was a lottery winner. If anything, the grandchildren of the treated are less likely to be in school in 1880, some five decades after the lottery. Treated families have fewer grandchildren per son in 1880, which roughly offsets the small fertility effect in the previous generation, leaving a statistically similar number of grandchildren by lottery status. In other words, the additional wealth causes a one-generation blip to the size of that dynasty.

The failure of lottery winners to invest more in their children's human capital is not the result of a lack of a substantial return on such investment. Cross-sectional comparisons show signs of returns to skill in nineteenth-century Georgia. This evidence is found in Section 8.1. We show that literacy and childhood school attendance predict adult wealth and sons with more siblings tend to have worse adult outcomes. Whether this reflects a causal effect is uncertain, but the standard methodologies for measuring these relationships indicate their presence in one form or another in the context of our study. The presence of positive returns and the absence of an effect of lottery winning indicate that parents did not use the wealth to relax a financial resource constraint. The results are instead consistent with the presence of deeper, underlying characteristics that persist through family lines and are associated with superior outcomes, like those posited by Clark and Crimmins (2013). Section 8.2 provides evidence on correlations between parental resources and children's human capital outcomes. It also provides suggestive evidence for the presence of linkages that persist through family lines in the form of regressions on individual-level outcomes (wealth, fertility, human capital) using the average values of these variables for all individuals (except the individual in our sample) with the same surname. The results are also not the product of biases introduced in the process of generating our data: a placebo exercise in Section 8.3 uses the same procedure we employed but applied to South Carolina, for which any data linkages generated should be entirely spurious, and finds no effect of lottery winning. Finally, we perform a simulation in Section 8.4 using our control sample, and find that the cross-sectional relationships between parental wealth and human capital are *not* consistent with our estimates derived from the lottery treatment.[6]

Section 9 concludes the study.

[6] The cross-sectional relationships for human capital and wealth are present in our control sample and similar to modern intergenerational correlations despite the Civil War's occurrence between the 1832 lottery and when the children of 1832 lottery participants are observed as adults.

2. Background

2.1. The Cherokee Land Lottery in Northwest Georgia, 1832

Georgia, unlike most U.S. states, placed most of its land in the hands of the public through a series of land lotteries.[7] At its origin, the colony of Georgia was located along the Savannah River, where most of the land was distributed through the headright system, in which arriving settlers were given land in proportion to the number of individuals they brought with them. The impetus for the lottery system as a new means of distributing land was a widely-reported corruption scandal, the Yazoo Land Fraud of the 1790s. The legislature, in response to the public uproar brought on by the scandal, introduced lotteries as an ostensibly fair and transparent system to distribute the state's land, beginning in 1805. As new land was acquired by the state in treaties concluded with the indigenous population, new lotteries were conducted, the last following the eviction of the Cherokee from northwest Georgia beginning in the early 1830s.

It is this last Georgia lottery in 1832 on which we focus our attention in this study. A list of the lottery's winners was easy to obtain, and the late date of the lottery makes it possible for us to identify the lottery's winners and losers in the first U.S. Census (1850) with complete names, ages, and birthplaces for all household members, as well as measures of literacy, school attendance, and real estate and slave wealth. Once individual lottery winners and losers were located in 1850, it was straightforward to locate their children in 1850, 1860, 1870, and 1880, and their grandchildren in 1880.

The lottery's rules were simple: every male age 18 and older who had resided in Georgia for the three years prior to the 1832 drawing was eligible to one draw (Cadle 1991, pp. 267-283). Some others (widows, orphans, and military veterans) were entitled to two draws. Our inability to identify these groups in our treatment and control groups, as well as the small numbers represented by these groups,

---

[7] This summary of early Georgia land policy is drawn from Cadle (1991, pp. 60-108 and 267-283).

led us to exclude them from the subsequent analysis. Finally, members of a band of outlaws known collectively as "The Pony Club" were excluded from participation in the lottery and are ignored in our analysis. In theory, individuals who had won land in prior lotteries conducted by the state were also excluded. A nominal fee of 12.5¢ was imposed on registrants. In view of the near-universal participation rate (see below), we doubt that these latter two constraints were rigorously enforced.

We can estimate the participation rate in the 1832 lottery. The state's 1830 census returns report 77,968 white males age 15+.[8] Using data in Cadle (1991), we estimate that just under 76,000 males registered for single draws, a number only 2.8% lower than the population of males in the eligible age range by 1832 based on the 1830 census. The small discrepancy can be accounted for by out-migration and mortality between 1830 and 1832 , combined with those aged 15 at the time of the 1830 census who had not reached 18 by the time of the 1832 lottery. Thus, the participation rate among the eligible population was extremely high – in fact, close to universal (from 97.2 percent to 99.5 percent depending on the assumptions we make). Smith (1838) reports that there were 15,000 winners in the 1832 lottery, excluding widows and orphans, which corresponds to a winning rate of roughly 19%, only slightly higher than rates observed in Columbia County (16.0%) and Oglethorpe County (16.8%) where the full lists of lottery participants and winners have survived.

As Cherokee County was being surveyed in preparation for the lottery, lists of those eligible to participate were forwarded to the state capital at Milledgeville. The survey divided what was to become the 10 northwestern counties of present-day Georgia into four sections, each of which was then divided into districts (generally square in shape, though less regularly-shaped right along the border between Cherokee County and both the older-settled region to the east and the Chattahoochee River to the county's southeast). The districts were then further divided into parcels of 160 acres, with 324 parcels

[8] Calculated from ICPSR Study 2896 (Haines, 2010). These figures also appear in Bleakley and Ferrie (2013).

in each of the square districts.

The lottery itself was conducted to ensure the greatest possible transparency, with a slip of paper for each participant placed in one large barrel and a slip of paper for each plot placed in a second barrel, which also contained enough additional blank slips so the number of slips was the same in each barrel. Names and parcels (or blank slips) were then drawn simultaneously from the two barrels until both were empty. As a result, winning was random and so was the quality of the parcel won among winners.

Once the lottery was completed, winners could immediately sell their winning draw. Unlike land distributions in many Midwestern states, there was no requirement that the recipient spend any time on the land or make any improvements whatsoever. The only requirement imposed was that winners register their claim and pay an $18 registration fee to the state.[9] The land could not be immediately occupied, however, as the Cherokee Nation was engaged in legal action to fight their eviction, and the final ruling in favor of the state did not come until 1838. As a result, some lottery winners may have exercised their option of immediately "flipping" their property.

We estimate that the value of a winning draw was perhaps as high as $700 in 1850 for a 160-acre parcel. This $700 figure is based on the value of farmland in the ten counties of northwestern Georgia in 1850, minus the value of implements and machinery.[10] In Table 2 below, we find that winners were in fact $700 wealthier than losers by 1850 – the equivalent of more than 900 days of earnings for an unskilled laborer in the South at this time. Even if they sold their parcel between 1832 and 1850 and bought land that rose in value at a similar rate, we would also expect them to be wealthier in 1850 than lottery losers. Those who sold out before the uncertainty over the timing of the expulsion of the Cherokee might have received somewhat less than this, but the timing was the only source of

[9] The registration fee need not have been an obstacle to liquidity-constrained winners in that there were many who simply sold the claim itself.

[10] See Bleakley and Ferrie (2013) for the details of this calculation.

uncertainty in this process, as the Indian Removal Act of 1830 under which the eviction was conducted had already been applied elsewhere.[11]

2.2. Related literature

The literature on the effect of parental resources on child outcomes is so large that we cannot possibly do it justice here. One could start with the claims of Malthus (1806). As presented by Becker (1992), the simple Malthusian idea was that income was a "preventive check" that was the main constraint on fertility. Post-Malthus, however, many societies experienced the Demographic Transition in which fertility declined while human-capital investments took off. Becker argued that a simple trade-off between child quantity and child quality was at play during this transition.

Once opportunities for investing in children's human capital become available, it is possible to imagine how parents' circumstances affect the outcomes of each of their children (rather than merely the total *number* of their children). Becker and Tomes (1986) modeled the decision made by parents to invest in their children, subject to a budget constraint and the presence of "a family's cultural and genetic 'infrastructure.'" (1986, p. S6). In this setting, wealthier and better-educated parents face a different budget constraint than poorer and less-educated parents, resulting in a correlation in outcomes across generations even if all families possess the same "infrastructure." Conversely, parents facing identical budget constraints might also see different outcomes for their children if their "infrastructures" are different. Some of the advantages enjoyed by certain parents might be dissipated (*i.e.,* not exclusively generate better per-child outcomes) to the extent that they result in greater fertility. Becker and Tomes (1986) predict on the basis of studies available in 1986 that any earnings advantage would be erased within three generations.

[11] The bottom third of Cherokee County was thought to contain gold, and was distributed in a separate lottery in smaller, 40-acre parcels. We focus only on the main lottery of 160 acre parcels.

The existence of such intergenerational correlations in outcomes in the nineteenth century U.S. is clear: for example, Long and Ferrie (2013) show the links between the occupations of fathers and sons in the U.S., 1850-80. Sacerdote (2005) examines father-son links after the Civil War and finds that it took roughly two generations for the descendants of those born into slavery in the U.S. (and faced severely limited opportunities for human or financial capital accumulation) to converge to the human capital outcomes of blacks who were born free. In Britain, Clark and Crimmins (2013) use evidence from rare surnames to show how advantages in educational opportunities (attendance at Oxford and Cambridge) persist for eight centuries.

A large number of contemporary studies have examined correlations in human capital across generations (see Black *et al.* 2005 for a summary of several such studies). Oreopoulos *et al.* (2006) use a change in compulsory schooling laws in the U.S. and show that parents' education has a causal impact on children's education. Black *et al.* (2005) use a large sample of twins and a change in schooling policy that was rolled out only gradually across Norway to isolate the effect of parents' own human capital on that of their children from unobserved family effects, finding that the latter (including the genetic inheritance of ability) were most important. What all of these studies have in common is an interest in separating the effect of parents' outcomes (e.g. higher educational attainment) from the effect of their ability and its effect on their own education and then on their children's. The focus in these exercises has largely been upon finding plausible exogenous variation in educational attainment that is not the product of variation in ability or other unobserved family characteristics.

We are not, however, acquainted with any study that uses random variation in wealth to study human capital transmission at the start of the demographic transition, nor are we aware of one that follows up on random wealth shocks over such a long horizon (five decades). We examine just such an "experiment" – outcomes for families that participated in a large-scale lottery with a significant prize awarded to a large number of winners. Since the lottery took place well after parents had completed their own schooling, there was little opportunity for the outcome to alter their own human capital stock. Instead, it should have relaxed the budget constraint faced by poorer households and allowed them to invest more in the human capital of their children. If human capital was unaffected in the next generations, this is evidence in favor of the view (recently advanced by Clark and Cummins 2013) that a substantial portion of the intergenerational correlation in outcomes is driven by fundamental, family-specific effects (the "family's cultural and genetic 'infrastructure'" in the Becker and Tomes 1986 model) and that the latitude to improve mobility across generations through interventions that address only the outcomes themselves (e.g. improving parents' or children's education) is severely limited.

Our results also relate to a literature analyzing so-called Conditional Cash Transfer (CCT) programs that have become quite popular in the contemporary developing world. While the 1832 lottery was a wealth shock with no strings attached, a CCT is an ongoing payment conditioned on certain behaviors, such as sending one's children to school. (See Das, Do, and Özler, 2005 for a review.) There is mixed evidence on whether the transfer itself promotes school attendance if such conditionality is removed (Baird, McIntosh, and Özler, 2011; Akresh, de Walque, and Kazianga, 2013). Nevertheless, work in Brazil by Bursztyn and Coffman (2012) is not consistent with the idea that school investments are held back by imperfectly altruistic parents who cannot borrow against the future earnings of their children versus some other within-household bargaining problem.

3. Data

3.1 Data Sources and Construction

The present study follows up on the outcomes of lottery winners and losers and their children and grandchildren. In order to do this, we first need to identify who was eligible, and who then won. We find these individuals, their children, and their childrens' children in later, publicly available data sources, and ascertain their outcomes. We initially search for these individuals in the Census manuscripts of

1850-1880 using a preliminary version of the full-count file for the 1850 census from the IPUMS project, the full-count file for the 1880 census from the NAPP project, and also indexes to the 1860 and 1870 censuses searchable on Ancestry.com.

The names of winners in the 1832 Georgia land lottery were published in Smith (1838), who lists each parcel in the 1832 lottery area and the name of the winner of that parcel, as well as the county and minor civil division where the winner resided in 1832. A version of this list was obtained on-line from on accessgenealogy.com. It was compared against a copy of the Smith book that was converted to a computer file using an OCR program and any discrepancies were resolved.

Although we possess a list of winners, there is no surviving state-wide list of all participants from which we could construct a control population to compare to those treated by winning the lottery. To create the control population, we exploited the lottery's eligibility requirements and information available in the 1850 Census of Population, which identifies all household members by name, age, and state of birth. The bulk of those eligible to participate in the lottery had to have been males age 18+ in 1832 who had been present in Georgia continuously over the preceding three years. Using the full-count file created by the transcription of the 1850 census, we identified all white males who would have been age 18+ in 1832 and who had at least one child who was born in Georgia 1829-1832 and who had no children born outside Georgia in the same interval. There were 14,306 individuals who met these criteria, 1,758 of whom were subsequently located – using their surname and given name – in the Smith (1838) list of lottery winners.[12] For both the control and treated populations, information reported in the 1850 census (county of residence, marital status, number and ages of all children, occupation,

[12] An individual was considered to have been uniquely linked if exactly one individual in the Smith list appeared in the 1850 census group of eligibles with the correct given name and a surname that differed by no more than 15 units in the SPEDIS "phonetic distance" function in SAS (which assigns points to different sorts of transcription errors such as omitting a letter, sums the points, and adjusts for the name length). If several individuals were matched by given name and all had exactly the same SPEDIS value (below 15), the individual was considered to have been multiply matched.

literacy, school attendance, and real estate value) was then transcribed and combined with information on the number and age and gender of slaves owned as reported in the slave schedule that was created by census marshals concurrently with the population schedule. The value of slave wealth was estimated using slaves' age and gender and contemporaneous slave prices disaggregated by these characteristics.

Linkage to later censuses was then performed to generate multigenerational outcomes for the control and treated populations. The male sons of the 1850 male household heads previously identified as the control and treatment groups were sought in the 1880 U.S. Census of Population in two ways: (1) the characteristics of 1850 sons (name, year of birth, birthplace, and parents' birthplaces) were used to locate them in the 1880 U.S. Census 100% File[13]; and (2) individuals not successfully linked 1850-80 were located in the Ancestry.com on-line 1850 U.S. Census index, where any hints to their 1880 record were followed (these hints are generated by Ancestry.com on the basis of both actual links among individuals made by genealogists in the construction of their family trees, and links generated by Ancestry.com through a machine learning process in which actual genealogist-generated links were used as training data and the system then generated links automatically for individuals not previously linked by genealogists). When 1850 sons were identified as 1880 household heads through either of these mechanisms, the 1880 information on their entire 1880 family was transcribed (occupation, literacy, school attendance).

The male sons of the 1850 male household heads we previously identified as the control and treatment groups were sought in the 1870 U.S. Census of Population in two ways: (1) individuals linked to the 1880 100% file in the manner described above were located in the Ancestry.com on-line 1880 U.S. Census index, from which any hints to their 1870 record were followed; and (2) individuals not successfully linked 1850-80 in the manner described above had their hinted links forward from their

---

[13] When multiple matches were found in 1880 for the same 1850 individual, the match that minimized the SPEDIS "phonetic distance" between the 1850 individual sought and the 1880 individual located was chosen; if multiple 1880 individuals minimized this distance, the observations were rejected.

1850 census record on Ancestry.com followed. When 1850 sons were identified as 1870 household heads through either of these mechanisms, their 1870 real estate and personal estate were transcribed.

The initial sample drawn from the 1850 census yielded 47,749 children age 5-17 whose schooling and literacy were observed (as the lottery occurred in 1832, the number of children under 18 years of age was also an outcome that we observed for families of winners and losers). The linkage to 1880 yielded 14,963 male children of lottery winners and losers whose outcomes could be observed in both 1850 and 1880, together with 40,658 grandchildren in 1880 of the original lottery winners and losers. Finally, the linkage to 1870 yielded 24,510 male children of lottery winners and losers whose 1870 outcomes were observed in 1870; of these 6,823 were adults in 1870, so their 1870 real and personal wealth was observed.

3.2 Summary Statistics and Balancing Tests

Table 1 presents the sample's summary statistics. Each variable appears in its own row, and each panel contains similar variables.[14] Column 1 displays values (means and, in parentheses, standard deviations) for the entire sample, while Columns 2 and 3 report, respectively, the corresponding values for lottery losers and winners. Column 4 reports *p*-values for a test of the null hypothesis that the means in Columns 2 and 3 are identical (where the test is a simple bivariate regression on a dummy for lottery winner). Clustered standard errors are calculated throughout the analysis when the data have a grouped structure. Sample sizes are reported in square brackets.

We use two measures of winning land in the 1832 Cherokee Land Lottery (Panel A, Table 1). If an individual was uniquely matched to the list of winners (Smith, 1838), the first measure is coded to one; otherwise this measure is coded to zero. By this measure, 12.4% of our observations are lottery

[14] Portions of Tables 1 and 2 also appear in Bleakley and Ferrie (2013).

winners. This measure has a mean of zero for losers (Column 2) and a mean of one for winners (Column 3) by construction. The second measure is designed to account for the few cases where more than one individual is matched to the list of winners. If *n* individuals are matched to the same winner, the match variable is recorded to 1/*n*. Our maintained assumption in constructing this measure is that one of the "tied" individuals in fact won a parcel, but in the absence of additional information, we can do no better than assigning equal probabilities of this even to all *n* individuals in the "tied" set. The mean value for this measure of the probability of winning is 15.5%, which is 3% higher than the original measure but similar to the winning rates in Columbia and Oglethorpe counties where we have actual lists of both lottery participants and lottery winners.

The second measure is higher than the first, as some individuals who were multiply matched have a zero for the first measure but 1/*n* for the second. In 9 cases, there was one unique match to 1850 but several similar quality matches (e.g. the additional multiple matches had full given names but lacked middle initials, while the unique match had a middle initial and full given name in the winners list and in the 1850 census). In these cases, there is one observation with the value one and n-1 with the value zero by the first measure, and *n* observations with the value 1/*n* by the second. Overall, these two lottery winning indicators have a very high correlation.

Panel B of Table 1 presents outcomes determined before the 1832 lottery, which should not be affected by whether the individual was a lottery winner or loser. The comparisons between Columns 2 and 3 here represent a balancing test – an analysis of how well the treated group compares to the control group prior to the application of the treatment. Average age, the fraction born in Georgia, the number of co-resident children present in 1850 and born in the three-year pre-lottery window, and the fraction of adults who could not read or write are similar in the control and treatment groups.

We then examine characteristics associated with the surname of each individual. Since surname was determined at birth and could not have been affected by the lottery, we would also expect no differences between the control and treatment groups in these measures. We account for minor variation in spelling by using the Soundex code for each. Surnames (prior to Soundex coding) are 6.2 characters in length on average, though this measure is slightly lower for lottery winners. On average, each individual's surname occurs 36 times in the sample, with no difference between winners and losers. Surnames began with the letter "M" or "O" (a rough indicator of Celtic origin) in 10% of all cases, with no difference between winners and losers.

Finally, we constructed average characteristics from other males resident in Georgia in 1850 with the same surname.[15] Mean real estate wealth of those people with the same surname as the sample individuals is $1,200, while the median wealth is below $300. The surname average rate for illiteracy is 22%. Neither measure differs between the winners and losers.[16]

Panel C summarizes our measures of 1850 wealth. Outcomes measured here and in the rest of Table 1 are no longer expected to be the same between winners and losers and thus, unlike Panel B, are not useful as a balancing test. We report real estate wealth, slave wealth, and the sum of these. Although we label the latter "total wealth," there are other forms in which wealth could be held that were not recorded in the 1850 census (a personal wealth question was added in 1860 and 1870). Mean wealth in all 3 measures (real estate, slave, and "total") are all several hundred dollars higher for lottery winners than for lottery losers. The economically large magnitude is similar to the value of winning a parcel that we calculated previously.

---

[15] Although some lottery winners and losers no doubt migrated out of Georgia after the 1832 lottery, we have limited our attention to Georgia in constructing these surname-average characteristics because roughly half of the counties in Georgia have already been completely transcribed. Individuals in the lottery sample are themselves excluded from these surname averages.

[16] Although this is a weak test due to noise in the surname averages, surname averages (in results not shown) are statistically significant predictors of individual-level behavior even controlling for a variety of other covariates.

Summary statistics for fertility and school attendance among the children of winners and losers are shown in Panel D. Lottery winners had, on average, 0.2 more children born after 1832 who survived to 1850 than did lottery losers. By contrast, the fraction of school-eligible children who attended school at any time during the 12 months prior to the census reference date (June 1, 1850) did not differ between the winners and losers.

Additional characteristics for spouses and 1850 locations are compared in Panel E. Slightly more winners than losers still had a spouse present in 1850, while among those winners and losers with spouses present in 1850 the winners had spouses 6 months younger on average than the losers. Roughly equal percentages of spouses were illiterate among the winners and losers. Most of the sample still resided in Georgia in 1850 (Figure 1), while most of the balance outside Georgia was in Alabama. Although the fractions residing in Georgia and Alabama do not differ between the winners and losers, the equality of the distributions of winners and losers across counties is strongly rejected by a simple χ^2 test. As we will see below, lottery winners were slightly more likely than losers to reside in 1850 in Old Cherokee County (the counties settled through the 1832 Cherokee Land Lottery).

4. Estimation strategy

Our data allow us to analyze outcomes for lottery winners themselves, their children, and their grandchildren – a span of roughly 50 years from the date of the lottery. We were able to construct treated (winners) and control (losers) groups based on the list of winners and the criteria for participation, where the latter allowed us to identify all individuals likely to have been eligible to win and the former allowed us to identify winners in that eligible population. In this sample, the treatment effect of winning a parcel in the lottery can be assessed directly by comparing mean outcomes for winners and losers (and their descendants), or by estimating a simple bivariate regression with a relevant outcome on the left-hand side and a dummy variable for winning a parcel on the right-hand side. We adopt the regression-based approach to permit both the inclusion of additional control variables and the continuous $1/n$ lottery status indicator. Although the random assignment of parcels among participants reduces the omitted-variable problem and thereby diminishes the need to introduce additional controls, such controls can improve the precision of our estimated treatment effect and reduce the residual variation. These controls can also reduce any biases resulting from our process for imputing lottery status, although the inability of lottery status to predict pre-determined outcomes reduces this concern.[17]

We estimate OLS regressions of the following form:

$$Y_{ijk} = gT_j + BX_{ijk} + d_a + d_k + e_{ijk} \quad (1)$$

in which where i is the individual, j indexes the lottery-eligible person, T_j, (a binary variable) denotes treatment—winning a parcel in the lottery—and control variables are: d_a (a set of age dummies); d_k (a set of county × state location dummies to allow for differences between control and treated in settlement patterns; and X_{ijk} (a vector of other control variables specified below). The error term is allowed to vary by both i and j. When we examine outcomes for the original lottery participants, $i=j$. But many of the regressions below use instead samples of children or grandchildren of the lottery participants, generating potentially numerous observations (i) for each lottery participant (j). In these regressions, standard errors will be clustered at the lottery participant (j) level. The estimate of g that we recover should be uncontaminated by omitted-variable or endogeneity problems, as a result of the random assignment of treatment by the lottery.

We also employ an additional specification that incorporates characteristics measured at the level of surnames, in the simplest case adding a fixed effect for each surname. Such a specification controls for numerous differences that might be constant in family lines (patrilineal lines here, as we only have information on surnames), allowing the impact of winning a parcel in the lottery to persist within

---

[17] We also address this issue directly with a placebo test in Section 8.3.

extended patrilineal families. Clark and Cummins (2013) and Güell *et al*. (2012) both highlight striking persistence in a variety of outcomes across family lines, an effect that surname fixed effects would absorb. At the same time, our imputation process for determining lottery status relies on matching by surname, so noise introduced in this process can be absorbed by surname fixed effects.

5. Analysis of wealth differences for original lottery participants

As a first step in assessing the impact of lottery winnings on family outcomes across generations, we estimate the direct effect on lottery participants of winning a parcel on both real estate and total wealth (real estate plus slave wealth) levels in Table 2. The effect is large: the baseline estimate in Column 1 is an impact of $750 on 1850 wealth, similar to the unconditional difference in Table 1 and our estimate of the value of a parcel of land in northwest Georgia by 1850. Although winnings could in theory have been invested in a variety of instruments other than land and slaves, such alternative investment opportunities were rare in the Deep South in the antebellum period.[18] The baseline estimates suggest that the effect of winning a parcel in the lottery persisted for at least the two decades following the drawing.[19]

In the rest of Table 2, we employ specifications with different fixed effects: surname characteristics (initial letter, length, and frequency in the sample) in Columns 2-4, which yield results within a third of a standard error of the baseline; dummies for each surname (by Soundex code) in Column 5, which reduces the effect of winning by half a standard error, although the effect remains

---

[18] Ransom and Sutch (1988, Table A.1, pp. 150-1) report that the total value of slaves in the U.S. in 1860 (the first time the census reported both real and personal wealth) was $3.1 billion. In that year, total real estate and personal estate in the South were $3.4 billion and $4.7 billion, respectively (IPUMS 1860 1% Sample: Ruggles *et al.*, 2010). Thus, slaves accounted for 2/3 of all personal wealth in 1860, and land plus slaves accounted for 80% of total wealth in 1860.

[19] Bleakley and Ferrie (2013) present a more detailed analysis of the 1850 wealth of the lottery winners and losers.

large ($600) even 18 years after the lottery; dummies for given name in Column 6, which raises the estimated effect by half a standard error; and both given name and surname fixed effects in Column 7, which also yields a substantial impact of lottery winning, slightly below the baseline but higher than with surname controls alone.[20]

Finally, estimates of the effect of winning are similar whether we use the binary or the $1/n$ match versions of our lottery status. We present only the binary variable results in what follows. The Table 2 results reveal that including surname fixed effects has a larger impact, so we provide that specification as an alternative throughout.

6. Effects on Child Quantity versus Child Quality

Lottery winners tended to have slightly more children, but did not send them to school more. In Becker's (1982) terminology, they invested in child quantity but not child quality. These results are found in Table 3, where, as above, we estimate equation (1). We report the coefficients on the binary measure of lottery winning. Column 1 reports results when the dependent variable is the number of children born after 1832 (the year of the lottery) who were still present in the household in 1850. (Recall that the number of children born in the three years prior to the 1832 Lottery was not significantly related to lottery status.) In the basic specification, we estimate lottery winners have 0.13 more children on average and, in the specification augmented with surname fixed effects, we estimate instead a coefficient of 0.19. These numbers are consistent with the unconditional difference seen in Table 1 of 0.2. When considered over the entire set of children still in the household, this represents a 3% increase in fertility, as seen in Column 2.

[20] The specification in Column 7 uses two sets of fixed effects: one for each surname and one for each given name. Dummies for each given-name × surname cell would entirely absorb the lottery-status variable, as lottery status in the eligible population was determined by linkage to the Smith (1838) winners list using surname and given name.

The remaining two columns of Table 3 examine school attendance by children in the household aged at least five years old but not more than 17 years old. (Note that this age range excludes children born prior to the lottery.) These children are linked to the lottery status of their father, and the standard errors are adjusted for clustering at the level of the father. Column 3 uses the OLS estimator and therefore this regression is a linear probability model, while Column 4 uses the logit estimator, with marginal effects evaluated at the mean of observables and assuming the surname fixed effects are all zero. The resulting coefficients imply an effect of winning the lottery of close to zero, and we can rule out effects of more than a few percentage points.

In Table 4, we consider some decompositions and possible mechanisms for the quantity/quality result. One hypothesis for these results is that richer husbands might be able to remarry more easily (and/or to a younger spouse) if his first wife had died in childbirth (which was not an uncommon occurrence in this period). This higher remarriage probability could result in higher fertility in families headed by lottery winners. But we see in Columns 1 and 2 that there is not a statistically significant difference by lottery status in the wife being present or in the wife's age, if she is present. Next we consider the extensive margin of post-lottery fertility in Column 3, where we see that lottery winners are more likely to have children after 1832 than the lottery losers. Indeed, the entire CDF of the number of post-lottery children (N_{kp}) is shifted out for lottery winners, although such differences are strongest for when assessing whether winners were more likely than losers to have had one or two additional post-lottery children, as seen in Columns 3 and 4. In results not shown, we find evidence of differential stopping behavior: the average age of children in the household or the age of the youngest child is about 0.2 years lower for winners, although this result is only marginally significant (p=13%). In Column 5, we find essentially no effect on the gender composition of children, suggesting that the fertility effect is not due to the differential survival of one gender or the other. Finally, in Columns 6-9, we obtain similar school-attendance results when decomposing the sample by gender or by broad age groups.

Locational choice, at least at the county level, does not appear to be a central mechanism in driving these results. First note that results are quite similar whether we include fixed effects for county × state of residence (in 1850) (in Tables 2 and 3) or not (in Table 1). We further investigate this mechanism by examining characteristics of the 1850 county of residence in Table 5. We begin by noting that lottery winners are slightly more likely to end up in Old Cherokee County in 1850 (Column 1), although this difference in probabilities is quite small (2.2%). The lack of a homesteading requirement implies that there is no mechanical reason why the lottery winners should have higher rates of residence in Old Cherokee County than the lottery losers. Nonetheless, some of them may have chosen to settle on their parcel rather than flip it, and this decision apparently stuck for a small fraction. However, the treatment group shows no differential probability of residing in the state of Georgia (Column 2) or residing in a county that is farther east (Column 3). But lottery winners do, on average, live somewhat farther south when compared to lottery losers. This may be because land in the Upcountry frontier was cheaper and therefore more attractive to the poorer lottery losers. Alternatively, this may be because someone with enough capital to buy a slave preferred to stay farther south where slave agriculture was more productive.

The remainder of Table 5 (Columns 5-16) uses county-level data to construct left-hand-side variables describing the local economic and demographic conditions in the 1850 county of residence. Because of the repeated data within counties, we now cluster the standard errors on county of residence. Most important for the quantity/quality results, we do not see differences by lottery status in the average school-enrollment or fertility rates (Columns 5-7). This suggests that lottery winners were not differentially moving to areas that were more conducive to higher fertility or school attendance (the latter being perhaps because of the provision of school infrastructure). Additionally, being a lottery winner does not predict differences in county-of-residence farm values, farm sizes, land improvement, or slave density (Columns 8-11). While this might suggest that lottery winners bought more acreage instead of moving to counties with more valuable land, we cannot rule out that they bought the land that was more valuable within a county. Finally, we do not find statistically significant differences for county-of-residence urbanization or access to transport (Columns 12–15). Additional analyses in Panels C and D add a dummy variable for residence in Old Cherokee County and controls for surname fixed effects and find no difference from the results in Panel B.

7. Outcomes of the next generations in 1870 and 1880

We now follow up on the outcomes in 1870 and in 1880 of children observed in the 1850 households. Someone who was a child in 1850 will have advanced to adulthood in those later years, thus giving us an opportunity to observe the adult outcomes of children whose parents were eligible to win in the Cherokee Land Lottery. Many of those in the second generation following the lottery had formed households by 1880, which also allows us to observe the childhood outcomes of the grandchildren of those who were eligible to participate in the 1832 Lottery. Note that here we are examining outcomes that are almost 50 years after the lottery took place.

We next track this sample by taking the children under 19 in the 1850 households and looking for likely matches in the 1870 and 1880 censuses. We use the 100% file from NAPPdata.org for 1880, as well as indexes for 1870 and 1880 that are searchable on Ancestry.com. The conditions under which lottery winning predicts linkage are discussed below. It should be noted that we only attempt to link male children across censuses, because female children would almost certainly change their surname at marriage. Linkage rates to 1870 and 1880 are somewhat low (28% and 35%, respectively). The lower linkage rate for 1870 results from the exclusive reliance on the hints generated by Ancestry.com to perform the matching – for 1880, we were able to use both the hints and the 1880 100% file from NAPP. Approximately 59% of the lottery eligible men have at least one child in the 1880 sample.

The relationship between having a father win the 1832 Lottery and various outcomes for these children as adults is presented in Table 6. As above, we present results from a basic specification that includes dummies for age and place of residence, and for an augmented specification that controls for surname fixed effects as well (Panels A and B, respectively). For our purposes, the bulk of the outcomes of interest are drawn from the 1880 census, so we focus on those first. Note in Columns 1 and 2 that having a lottery-winning father is a significant predictor for the child being linked to the 1870 or 1880 census. This might induce a bias in the coefficients for other outcomes in 1880, although the fact that the 1880 outcomes are all either binary or have limited ranges puts an upper bound on the magnitude of such bias. In any event, this differential linkage seems to result from differences in the characteristics of given (first) names. Accordingly, if we condition on a variety of characteristics of the given name, lottery winning no longer significantly predicts differential linkage. Therefore, to our standard set of specifications, we add a Panel C in which we also control for the number of letters in the given name.[21]

Next, we turn to outcomes of the children of winners and losers. Having a lottery-winning father predicts linkage to 1870 or 1880, but this correlation dissipates when controlling for characteristics of the given name. The first outcome variable examined is illiteracy in Column 3, measured as whether the lottery participant's son is unable to read and unable to write. In Column 4, the outcome variable is the occupational score, in adulthood (1880), of the children of the lottery participants. Neither of these outcomes is significantly different when comparing the children of lottery winners versus those of lottery losers. In Columns 5 through 7, we consider outcomes in 1870. It is perhaps too early in 1870 to reliably measure the outcomes of grandchildren simply because many of the children of lottery

[21] We find similar results if we use other characteristics of the given name.

eligibles would just be starting families. But 1870 is the last nineteenth-century census in which wealth is reported, which we use as an outcome variable here. We transcribed both real-estate and personal wealth, and the results here are for the sum of these two variables. To compare with the estimates above, we deflate the 1870 wealth to 1850 dollars using the consumer price index from measuringworth.com (Williamson, 2013). Results are similar using other deflators.

The 1870 total wealth is statistically and economically similar between control and treatment groups. An alternative point of comparison is a mechanical split of the lottery winnings among the average number of children. This would suggest a treatment effect in 1870 of $140 (in 1850 dollars), which we also cannot reject at conventional levels of confidence. On the other hand, we can reject values larger than that. Note that the deflator adjusts for inflation only and does not convert the 1870 wealth into its present-value equivalent; results for 1870 wealth would drop by a factor of 2 to 5 for annual interest rates of 3% to 8%. At standard confidence levels, we could handily reject a lottery-winning effect of $140 in 1850 dollars for interest rates much above 3% *per annum*. These estimates for the wealth of the sons are inconsistent with a claim of supernormal returns in intergenerational transmission. Finally, we show in Appendix Figure 1 using a quantile regression of 1870 wealth on treatment that effects are similar across the distribution of 1870 wealth.

In Columns 8 through 10, we turn to outcomes in the third generation (the grandchildren of lottery participants). Note that children in the household in 1880 are, by construction, the grandchildren of lottery eligibles. There are two principal outcomes we consider: illiteracy (Column 8) and school enrollment (Column 9). Differences in illiteracy between the grandchildren by lottery status of their paternal grandfather are not statistically significant. In contrast, the grandchildren of lottery winners have a 2 to 3 percentage point lower probability of attending school. (These two columns use restricted ages corresponding to the age ranges over which the variables are measured and/or meaningful.) The result for schooling is the opposite of what one would expect if wealth were relaxing a constraint on human-capital investment. Nor are these results consistent with moving along or relaxing a quantity/quality tradeoff in that there is less of both variables. Men whose fathers had won the lottery had fewer children by 1880 (Column 10), although this effect is only significant at the 10% level. Nevertheless, the magnitude of this effect is approximately the same as it was for the previous generation. A regression at the grandfather level winning on the number of grandchildren cannot reject equality between winners and losers in the 1832 lottery. Furthermore, the fact the fertility effects in the first and second generations roughly cancel out suggests that the wealth shock induced only a one-generation blip in the size of the dynasty.

8. Discussion

In this section, we address four distinct questions: (1) is there evidence in these data of a return to skill? (Yes.); (2) is there evidence of intergenerational correlations in outcomes at this time, and are such correlations consistent with the presence of characteristics that might be passed along family lines? (Yes.); (3) do we obtain similar results using a placebo sample constructed using children born in South Carolina rather than Georgia in the years prior to 1832? (No.); and (4) are these results consistent with the relationship between parental wealth and children's human capital observed in our control population? (No.)

8.1. Was there a return to skill in nineteenth-century Georgia?

A possible response to the results above is that antebellum Georgia is not the right environment to observe parents investing in skills or facing a quantity/quality trade-off, perhaps because it was too early in that region's path of economic development. But was this indeed the case? It is possible that the contemporary reader might be unduly influenced by the seemingly moribund state of education in the South after the Civil War. Nevertheless, Bleakley and Hong (2013) show that antebellum rates of school enrollment among white children in the South were considerably higher than rates postwar, and, indeed, that the South would have caught up to the North by circa 1890 if the antebellum trends in school enrollment had continued after the Civil War.

Schultz (1975) has emphasized the importance of returns to education in agriculture once farming has passed out its "traditional" phase (in which prices are stable, long-used production techniques can be employed year after year, and there are no new technological or financial innovations that need to be dealt with). In light of the non-traditional nature of farming in Georgia from the 1830s forward (with new crops like new cotton varieties being introduced, and increasingly national and international markets for the state's products with wide year-to-year swings in prices), it would not be surprising to find a substantial value for education in this environment.

In any event, lacking an intervention or instrument that specifically manipulates time in school or the price of fertility control or some such, we cannot provide causal evidence on the return to school or the technological rate of substitution between quantity and quality in this context. Instead, we apply standard methods using observational data to get a first-pass estimate of these effects. We use the lottery-eligible sample from above because (i) it allows for estimates that are most internally comparable to the results above, and (ii) it can be used 'off the shelf' without the need for further linkage, transcription, or data description. There are a variety of outcomes in these data that are suitable for estimating such models, including wealth, occupational score, illiteracy, school attendance, and family size.

We find substantial (correlational) evidence of returns to skill in our sample. These results are found in Panel A of Table 7, which display estimates of equation (2):

$$(2) \qquad Y_{ijks} = wQ_{ijks} + d_a + d_k + e_{ijk}$$

where i is the individual, j indexes the lottery-eligible person, Q ("quality") is the skill variable, d_a are dummies for age, and d_k are dummies for state/county of residence. To this baseline specification, we add, in some cases, fixed effects for surname/Soundex. Column 1 of Panel A regresses the 1850 wealth of the lottery-eligible man on whether he is illiterate. Men in our sample had substantially lower wealth in 1850 if they could not read and write. The remaining columns of Panel A consider 1870-80 outcomes for the children of the lottery-eligible men. As seen in Columns 2 and 3, if these men were illiterate in 1880, they had lower income (using the occupational income score as a proxy) and wealth. We see also in Columns 4-6 that attending school in 1850 is associated with lower illiteracy and higher income in 1880 and higher wealth in 1870.

Though the absence of an effect of lottery winning on children's (and grandchildren's) human capital might reflect a quantity/quality trade-off (winners increased their family size but did not invest more in their children's education), this is implausible in light of the very small effect of winning on the number of children in winners' households (and the negative sign on the number of children in winners' sons' households) and the associated high value for children this would imply: for example, winners had between 0.13 and 0.19 children more than losers (Table 2), despite their having won as much as $700 in the lottery. This implies that an additional child was worth between $3,600 and $5,300. Kotlikoff (1979) reports that a prime-age male slave could be purchased at auction in 1832 in New Orleans for $701. Given high antebellum infant and child mortality rates, parents would have had to place an implausibly high premium on their own children's labor (and any non-pecuniary benefits from populating their households with their own children rather than slaves) for the measured effect of lottery winning on fertility to be consistent with investment at the extensive margin (quantity) in lieu of investment at the intensive margin (quality).

In sum, it appears that in antebellum Georgia there were indeed returns to human-capital investment. To reiterate, this table departs from previous ones in that we are not estimating the treatment effect of winning the lottery, but rather estimating the relationship (not necessarily a causal one) between human capital variables and other outcomes.

8.2. Intergenerational correlations

Next, we consider the extent to which outcomes are in fact correlated across generations at this time and whether outcomes are related to characteristics of other people who share the same surname and are therefore likely related along patrilineal lines of descent. As an example of correlation across generations, we examine the 1850 outcomes as predictors of the son's own 1870 wealth. There is evidence of a strong relationship between the log wealth of fathers and sons in Columns 6 and 7. In the control sample, the elasticity of son's 1870 wealth with respect to the father's 1850 wealth is 0.23 (and the correlation is 0.57) and statistically significant.[22] This linkage persists despite the intervening Civil War that destroyed much of the South's physical capital (though most had been restored by 1870), but – perhaps more importantly – resulted in emancipation and the disappearance from slaveowners' balance sheets of a significant quantity of capital. Despite the loss of this capital, the link between fathers' and sons' wealth remains strong. This is perhaps not surprising as the war would not have destroyed human capital acquired before it took place. The correlations in the control sample are also substantively large and statistically significant at the one percent level between father's 1850 log total wealth and his children's 1850 school attendance (0.25), 1880 literacy (0.11) and 1880 occupational score (0.11), and between the father's 1850 log total wealth and his grandchildren's 1880 school attendance (0.04).

We now show that this linkage comes through characteristics that are common across patrilineal lines, using surname-specific averages of fertility, school attendance, and wealth as possible proxies for differences across extended families in either preferences or prices. We used the 1850 100% census file to construct the average fertility, school attendance, and real-estate wealth among Georgia-resident

[22] Charles and Hurst (2001, 2003) report an intergenerational wealth elasticity of 0.37 and an intergenerational correlation of 0.23 to 0.50 for the modern U.S.

households for each (Soundex) surname.[23] Those individuals that appear in our lottery-eligible sample are excluded from the construction of the averages. We first check for the statistical power of these proxies by regressing the individual-level outcome on the surname average:

$$(3) \qquad Y_{ijks} = aY_s + d_a + d_k + e_{ijk}$$

where s denotes the surname for each observation, Y_s is the surname-average of the Y variable, and each regression contains dummies for age and for state/county of residence. Furthermore, due to the group-level nature of the regressor, we adjust the standard errors for clustering at the surname level. The base sample for these regressions is the same as for analogous estimates of equation (1) displayed in earlier tables, with the exception that some households are omitted if there were no other households in Georgia with the same surname and therefore no one from whom to form the surname-level averages. Estimates of this equation are found in Panel B of Table 7.

The surname-averaged variable is indeed a strong and statistically significant predictor of the individual-level outcome. The coefficient of zero is rejected in all three cases for conventional confidence intervals. A mechanistic model in which the patrilineal dynasty (proxied by surname) predicts outcomes one-for-one is even more strongly rejected; the coefficients are closer to 1/5th or 1/8th. (Note that we are not arguing that this is a causal effect of the behavior of their relatives on the individuals' choices, but rather a proxy for some shifter that is common within the group.) The surname has a little effect on the son's 1870 log total wealth, however, when both the father's log 1850 total wealth and the surname-average log 1850 total wealth are included together in Column 7 – the father's wealth dominates the surname effect.

[23] On the use of surnames in this way, see Clark and Crimmins (2013) and Güell *et al.* (2012).

8.3. Falsification exercise using a placebo sample from South Carolina

We perform a falsification exercise using South Carolina rather than Georgia and do not find statistically significant results. One of the challenges in identifying the treatment effect associated with winning the 1832 Lottery is that our method of imputing lottery status via name matching may introduce biases through sample selection. To check for this possibility, we construct a placebo sample using households with children born only in South Carolina (rather than Georgia) during the same pre-lottery window (the three years prior to the Cherokee Land Lottery of 1832).[24] We use the names among this South Carolina sample to impute a pseudo-lottery-status by linking to the Smith (1838) list. As above, we use both a dummy for a unique match to the Smith list and a variable that allows for probabilistic matches, deflated to $1/n$ case of ties. By the eligibility rules of the Cherokee Land Lottery, any matches from the South Carolina sample to this list must be spurious. It is then reassuring that the fraction of unique matches in the placebo sample derived from South Carolina is only one quarter of the fraction in the Georgia sample.

In Table 8, we estimate equation (1) using this placebo sample, for the different variables indicating lottery status, and using both the basic specification and the one that includes surname/Soundex fixed effects. These results are found in Panels A and B, with analogous results from the Georgia sample provided for reference in Panel C. The first four columns of Table 8 show outcomes that were determined prior to the 1832 Lottery, and there are no statistically significant results. (Note that a series of falsifications checks using pre-lottery variables was also performed for the Georgia sample, as shown in Table 1, Panel B.) The remaining columns show post-lottery outcomes such as residing in old Cherokee County and fertility by 1850. There is no statistically significant

[24] This exercise required the transcription of an additional 55,739 observations.

pseudo-treatment effect for the South Carolina sample, in contrast to what we find for Georgia.[25] Nor is there a statistically significant effect of treatment for 1850 real-estate wealth, along either intensive or extensive margins (Columns 7 and 8, respectively). In Column 9, there is some evidence of a positive relationship between the pseudo-treatment and school attendance. If we choose to subtract this estimate from the Georgia estimates, it would make the above estimates even less supportive of the idea that wealth allowed families to buy their way around credit constraints to invest more in their children's schooling.

8.4. Simulated results using cross-sectional relationships in the control group

Based on the cross-sectional relationship between paternal wealth and sons' outcomes, we would have expected much larger effects of winning the lottery on sons' human capital but not on fertility. We come to this conclusion by conducting a simple shift/share analysis using the expected change in the wealth distribution interacted with the relationship between wealth and various outcomes in the control group. We use the control group to conduct this calculation because we wish to compare the results from the randomized wealth with those for wealth in a sample that did not receive a random wealth disbursement. Some readers might ask why we did not instead set this up as a two-stage-least-squares (2SLS) problem with 1850 wealth as the endogenous regressor. This is inappropriate in that lottery winners may have spent some of their wealth precisely on the human-capital formation of children. This would violate the 2SLS exclusion restriction in that lottery treatment has an effect on child outcomes via a channel *other* than measured 1850 wealth. Such transitional dynamics of wealth would not be present in the control group, which did not receive the extra wealth. However, we would not argue that

[25] We generally limited this falsification test to variables that were already available in the 1850 census index. We also transcribed wealth and school attendance in these households from the 1850 Census manuscripts. Our efforts to link to the slave schedule were considerably more skilled-labor intensive, so we did not duplicate these efforts for the placebo sample.

the relationship between child outcomes and wealth in the control group is necessarily causal, but rather is a useful benchmark. One additional complication that motivates our use of the shift/share analysis (versus a more common comparison of 2SLS and OLS estimates) is that the relationship between wealth and various outcomes might not be linear.

The specifics of the shift/share calculation are as follows. We use 100 grid points, evenly spaced across the distribution of log 1850 total wealth, to discretize the 1850 wealth distribution. Within each cell j there is an estimated average x_j of some outcome. Let the vector of these averages be **x** and the probability of being in each cell summarized by the vector $\mathbf{p}=\{p_j\}$. The expected value of this outcome variable across the whole sample is therefore the dot product of **p** and **x**. Suppose the distribution of wealth is perturbed to be **q**. The change in the expected value of the outcome variable would be $\Delta = (\mathbf{q}-\mathbf{p})\cdot\mathbf{x}$. For a given perturbation of the wealth distribution, we compute the distribution of Δ with 500 bootstraps from the control sample. In the case of child or grandchild outcomes, we use a block bootstrap grouped by the lottery-eligible father.

Results from this exercise are shown in Table 9. The outcome measures and the year in which they are measured are displayed on the leftmost columns of the table. Each row and column group displays the mean and, in square brackets, the 95% confidence interval from a different simulation. The rightmost columns display estimates from the lottery-based design above. A dagger denotes that the confidence interval for that simulation does not overlap with the confidence interval estimated from the lottery treatment. For each simulation, we specify an expected value of winning, discounted to 1850 and denoted in the "EV" column-group headings. The first expected value we use for the simulation is $700, corresponding to our estimate in Section 2.1 of the value of land won. We also consider expected values $200 above and $200 and $400 below $700. We focus mostly on the $700 case, but discuss the robustness to alternate assumptions.

We also allow for heterogeneity in the value of land winnings by using a simple, two-point distribution including zero as a possible "prize." For each outcome and expected value of winning, we conduct three simulations with varying degrees of heterogeneity. These are denoted in the column "Fraction with zero," and indicate the fraction t of the simulated winners that receive zero change in wealth. In other words, we use the control sample to construct $\mathbf{p}$ using 1850 total wealth. We then define a perturbed-wealth variable equal to measured wealth for t of the sample and equal to measured wealth plus EV/(1- t) for the remaining (1- t). (Receiving zero wealth is randomly assigned separately for each bootstrapped sample.) In the end, these alternate assumptions do not make much difference. For a given outcome and expected value of the lottery prize, the simulated and estimated confidence intervals tend to overlap either in all three cases or in none at all. (Of the 44 blocks of cells in Table 9, 38 have either zero or three daggers.) In practice, this relative insensitivity arises from the approximate linearity of the relationship between most outcomes and 1850 wealth, at least across the densest part of the wealth distribution.

The simulations are generally consistent with our estimates of lottery treatment on fertility above. These are seen in row-groups A, B, and D, in Table 9. The fertility/wealth relationship in the control sample (not shown) is approximately flat in terms of economic significance. (A smoothed plot of 1850 fertility versus 1850 wealth displays an inverse ∪-shape. However, the range on the y-axis is quite small and only a minor change in fertility is associated with large changes in wealth.) In the simulations, wealth shocks of various sizes change the fertility rate by only a few children per hundred. The 95% confidence intervals for the simulation typically do not contain the point estimates from above, but they do overlap with the estimated confidence interval.

With a few exceptions noted below, the simulations are generally *not* consistent with our estimates of lottery treatment on human-capital variables, particularly at the low end. These results are found in Table 9, row-groups C and D-J. First, consider 1850 school attendance in row-group C. By this simulation, a homogeneous $700 wealth shock would increase school attendance by approximately 5.4%. This is different in both statistical and economic terms from the lottery-based estimate of -0.001. The simulation delivers larger (smaller) effects for larger (smaller) wealth shocks. In contrast, while we find positive rather than negative simulated effects for the occupational income score of the sons in 1880, there is generally a substantial overlap between the confidence intervals of the simulation and the estimate. (This outcome is complicated by the fact that essentially the entire sample was involved in farming, thus narrowing the occupational range.) The results for the sons' literacy, however, show economically and statistically significant differences between the simulation and the estimates. In the simulation, a positive wealth shock should have reduced the rate of illiteracy. However, the relationship between 1850 wealth and human capital of the descendants is weaker for the grandchildren than for the sons. Accordingly, there is substantial overlap in the confidence intervals for grandchildren's human capital, except for the $900 wealth shock.

Finally, we examine the 1870 wealth of sons, which the simulations suggest would have been markedly different at the low end (row-groups H-J). By the simulation, we would have expected an increase in the proportion of the sons with positive wealth in 1870, rather than a decrease as was estimated above. Relatedly, the simulations imply a large increase in the natural log of the sons' 1870 wealth, while we observed essentially no change using the lottery-based estimates. While these latter two outcomes are weighted towards changes in the lower tail of the sons' wealth distribution, we also examine the level of wealth in row-group H. In each simulation for 1870 wealth levels, the estimated and simulated confidence intervals have substantial overlap. These results taken together indicate the strongest effect of winning the lottery on the low end of the sons' wealth distribution. We might expect this pattern of results on *a priori* grounds as well in that high-wealth families were presumably less likely to be liquidity constrained.

9. Conclusions

The state of Georgia allocated most of its land to the public through a system of lotteries. These episodes provide a unique opportunity to assess the impact of shocks to wealth, in that the random assignment implied that the wealth shock was uncorrelated with individual characteristics. We focus on the 1832 Cherokee Land Lottery. We assess the impact on the winners themselves and their families into the third generation. Using 1850 Census microdata, we draw a sample of male household heads that likely were eligible for the lottery. The rate of registration for this eligible population was very high. We identify the lottery winners using Georgia state records and define them as our treatment group. We cannot reject that the treatment variable was randomly assigned in several balancing and placebo tests. We estimate that lottery winners won some $700 – close to median wealth in 1850 and the equivalent of nearly two and a half years of wages for an unskilled laborer in the South.

We focus on child outcomes in response to this wealth shock. Lottery winners slightly increased their family size after the lottery more than non-winners, but were not more likely to send their children to school. Children of lottery winners did not have more wealth, literacy, or income as adults. Further, the grandchildren of winners were not more likely to be literate or attend school. Indeed, the sons of lottery winners actually have fewer children and, if anything, send their children to school less that the control-group sons. This reduction of treated fertility in the second generation actually leaves the estimated number of grandchildren similar between control and treatment groups, effectively nullifying any fertility effect from treatment in the long run.

Despite the substantial size of the financial windfall received by lottery winners and the presence of returns to human capital, it does not appear that lottery winners invested more in their children (or that winners' children in turn invested more in their own children) than did losers (or losers' children). The random nature of the lottery assures us that winning was orthogonal to parents' wealth or their underlying characteristics. Taken together, these findings are inconsistent with parents' financial resources being a significant constraint on their shaping of children's outcomes. The results are also inconsistent with a wealth-based "poverty trap" for human capital. The observed intergenerational links are consistent instead with the presence of underlying characteristics that are passed down along family lines and are associated with better outcomes.[26]

10. References

accessgenealogy.com. 1832 Cherokee Country Georgia Land Lottery. Web site. http://www.accessgenealogy.com/georgia/landlottery/index.htm [Accessed May 29, 2009.]

Akresh, Richard, Damien de Walque and Harounan Kazianga. 2013. Cash Transfers and Child Schooling: Evidence from a Randomized Evaluation of the Role of Conditionality. World Bank Policy Research Working Paper 6340. January.

Baird, Sarah, Craig McIntosh, and Berk Özler. 2011. Cash or Condition? Evidence from a Cash Transfer Experiment. *Quarterly Journal of Economics* 126 (4): 1709?53.

Becker, Gary. 1982. *A Treatise on the Family*. Chicago: University of Chicago Press.

Becker, Gary. 1992. Fertility and the Economy. *Journal of Population Economics*, Vol. 5, No. 3: 185-201.

Becker, Gary S., and Nigel Tomes. 1986. Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, Vol. 4, No. 3 (July): S1-39.

Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes. 2005. Why The Apple Doesn't Fall Far: Understanding Intergenerational Transmission Of Human Capital. *American Economic Review*, Vol. 95, No. 1 (Mar.), pp. 437-449.

Bursztyn, Leonardo and Lucas C. Coffman. 2012. The Schooling Decision: Family Preferences, Intergenerational Conflict, and Moral Hazard in the Brazilian Favelas. *Journal of Political Economy*, 120(3): 359-397.

Bleakley, Hoyt and Joseph Ferrie. 2013. Up from Poverty? The 1832 Cherokee Land Lottery and the Long-run Distribution of Wealth. NBER Working Paper No. w19175.

---

[26] Clark (2007, p. 8) describes a similar process by which characteristics associated with better economic outcomes persisted and spread within family lines in England as the fertility of the affluent exceeded that of the poor, setting the stage for the Industrial Revolution: "The attributes that would ensure later economic dynamism – patience, hard work, ingenuity, innovativeness, education – were thus spreading *biologically* throughout the population." [emphasis added] It is easy to imagine how attributes like these and the attitudes toward human capital accumulation they inform might be transmitted through social as well as biological channels, generating the intergenerational correlations and their apparent immunity to wealth shocks that we have shown.

Bleakley, Hoyt and Sok Chul Hong. 2013. When the Race between Education and Technology Goes Backwards: The Postbellum Decline of White School Attendance in the Southern U.S. Unpublished manuscript, University of Chicago. April.

Cadle, Farris W. 1991. *Georgia Land Surveying History and Law*. Athens, Georgia: Univ. Of Georgia Press.

Charles, Kerwin Kofi, and Erik Hurst. 2001. The Correlation of Wealth across Generations. Working Paper (June). University of Chicago, Booth School of Business.

Charles, Kerwin Kofi, and Erik Hurst. 2003. The Correlation of Wealth across Generations. *Journal of Political Economy*, Vol. 111, No. 6 (December): 1155-1182.

Chevalier, Arnaud. 2004. Parental Education and Child's Education: A Natural Experiment. IZA Discussion Paper No. 1153.

Clark, Gregory. 2007. *A Farewell to Alms*. Princeton: Princeton Univ. Press.

Clark, Gregory and Neil Cummins. 2013. What is the True Rate of Social Mobility? Surnames and Social Mobility, England 1800-2012. Unpublished manuscript.

Das Jishnu, Quy-Toan Do, and Berk Özler. 2005. Reassessing Conditional Cash Transfer Programs. *The World Bank Research Observer*, 20(1):57-80. http://wbro.oxfordjournals.org/cgi/reprint/20/1/57

Dewey, John. 1889. Galton's Statistica M ethods. *Publications of the American Statistical Association*, Vol 1,No.7: 331-34.

Galor, Oded and Jospeh Zeira. 1993. Income Distribution and Macroeconomics. *Review of Economic Studies*, Vol. 60(Jan.): 35-52.

Graham, Paul K. 2010. *Georgia Land Lottery Research*. Atlanta: Georgia Genealogical Society.

Güell, Maia, Jose V. Rodriguez, and Christopher I. Telmer. 2012. Intergenerational Mobility and the Informational Content of Surnames. Unpublished manuscript. January.

Haines, Michael R., and Inter-university Consortium for Political and Social Research, 2010. *Historical, Demographic, Economic, and Social Data: The United States, 1790–2002*, Inter-university Consortium for Political and Social Research, icpsr.org.

Kotlikoff, Laurence J, 1979. The Structure of Slave Prices in New Orleans, 1804 to 1862. *Economic Inquiry*, Vol. 17, No. 4 (Oct.): 496-518.

Long, Jason, and Joseph Ferrie (2013). Intergenerational Occupational Mobility in Great Britain and the United States since 1850. *American Economic Review*, Vol. 103, No. 4 (June): 1109-37.

Malthus, Thomas Robert. 1806. *Essay on the principle of population; or, A view of its past and present effects on Human happiness*. London: Johnson.

Margo, Robert A., and Georgia C. Villaflor. 1987. The Growth of Wages in Antebellum America: New Evidence. *The Journal of Economic History*, Vol. 47, No. 4 (Dec.): 873-895.

North Atlantic Population Project. 2004. NAPP: Complete Count Microdata, Preliminary Version 0.2. Computer File. Distributed by University of Minnesota Minnesota Population Center: Minneapolis, MN. www.nappdata.org.

Oreopoulos, Philip, Marianne E. Page, and Ann Huff Stevens. 2006. The Intergenerational Effects of Compulsory Schooling. *Journal of Labor Economics*, Vol. 24, No. 4 (Oct.): 729-760.

Ransom, Roger, and Richard Sutch. 1988. Capitalists without Capital: The Burden of Slavery and the Impact of Emancipation. *Agricultural History*, Vol. 62, No. 3, Quantitative Studies in Agrarian History (Summer): 133-160.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor].

Sacerdote, Bruce. 2005. Slavery and the Intergenerational Transmission of Human Capital. *The Review of Economics and Statistics,* Vol. 87, No. 2 (May): 217-234.

Schultz, T.W. (1975). The Value of the Ability to Deal with Disequilibria. *Journal of Economic Literature*, Vol. 13, No. 3 (Sep.): 827-846.

Schumpeter, J.A. (1951). *Imperialism and Social Classes*, translated by Heinz Norden. New York: Augustus M. Kelley.

Smith, James F. 1838. *The Cherokee land lottery, containing a numerical list of the names of the fortunate drawers in said lottery, with an engraved map of each district.* New York: Harper and Brothers.

Williams, H. David. 1989. Gambling Away the Inheritance: The Cherokee Nation and Georgia's Gold and Land Lotteries of 1832-33. *The Georgia Historical Quarterly*, Vol. 73, No. 3, Special Issue Commemorating the Sesquicentennial of Cherokee Removal 1838-1839 (Fall): 519-539.

Williamson, Samuel H. 2013. Seven Ways to Compute the Relative Value of a U.S. Dollar Amount, 1774 to present. MeasuringWorth.com, accessed March 15, 2013.

Table 1: Summary Statistics

	(1) Whole Sample	(2) Lottery "Losers"	(3) Lottery "Winners"	(4) p-value, mean difference [N]
	Panel A: Lottery Winner or Loser			
Dummy for unique match to Smith (1838) list	0.124 (0.329)	0	1	---
Dummy for match to Smith (1838), deflated to 1/n in case of ties	0.155 (0.335)	0.037 (0.121)	0.995 (0.053)	0.000 [14375]
	Panel B: Predetermined Outcomes			
Age, in years	51.2 (8.5)	51.3 (8.5)	50.9 (8.6)	0.122 [14375]
Born in Georgia	0.497 (0.500)	0.497 (0.500)	0.498 (0.500)	0.889 [14375]
Born in South Carolina	0.212 (0.408)	0.210 (0.407)	0.222 (0.416)	0.263 [14375]
Born in North Carolina	0.180 (0.384)	0.180 (0.384)	0.178 (0.383)	0.804 [14375]
Number of Georgia-born children in the three years prior to the lottery	1.333 (0.542)	1.333 (0.541)	1.332 (0.542)	0.910 [14375]
Cannot read and write	0.147 (0.354)	0.147 (0.354)	0.142 (0.350)	0.593 [14340]
Number of letters in surname	6.19 (1.61)	6.20 (1.62)	6.13 (1.51)	0.072 [14375]
Frequency with which surname appears in sample	36.2 (46.3)	36.3 (46.9)	35.3 (41.9)	0.380 [14375]
Surname begins with "M" or "O"	0.101 (0.302)	0.101 (0.301)	0.104 (0.305)	0.740 [14375]
Mean wealth of families in Georgia with same surname	1186.3 (1257.8)	1185.4 (1288.4)	1192.3 (1021.8)	0.811 [13848]
Median wealth of families in Georgia with same surname	289.1 (716.6)	290.0 (717.6)	282.7 (709.9)	0.686 [13848]
Mean illiteracy of adults in Georgia with same surname	0.219 (0.107)	0.219 (0.108)	0.218 (0.098)	0.648 [13848]

Notes: Table continues on next page.

Table 1 (continued): Summary Statistics

	(1) Whole Sample	(2) Lottery "Losers"	(3) Lottery "Winners"	(4) p-value, mean difference [N]
	Panel C: Measures of Wealth in 1850			
Real-estate wealth	1999.0 (4694.2)	1970.8 (4422.0)	2198.2 (6290.1)	0.068 [13094]
Slave weath	1339.1 (5761.0)	1297.3 (5329.7)	1635.3 (8189.0)	0.021 [14375]
Total wealth (sum of wealth in real estate and slaves)	3323.7 (8691.0)	3245.5 (7952.9)	3876.5 (12734.4)	0.006 [13094]
	Panel D: Child Quantity versus Quality			
Number of children in household born after the 1832 lottery	3.955 (2.546)	3.930 (2.539)	4.135 (2.586)	0.002 [14375]
School attendance among children aged 5-18, inclusive	0.342 (0.474)	0.342 (0.475)	0.341 (0.474)	0.799 [47749]
	Panel E: Other Outcomes			
Spouse present in household	0.806 (0.395)	0.804 (0.397)	0.820 (0.384)	0.109 [14375]
Spouse age, in years	45.9 (7.8)	46.0 (7.8)	45.5 (7.8)	0.037 [11591]
Spouse cannot read and write	0.235 (0.424)	0.236 (0.424)	0.231 (0.421)	0.676 [11563]
Resides in Georgia	0.723 (0.447)	0.722 (0.448)	0.729 (0.445)	0.548 [14375]
Resides in Alabama	0.144 (0.351)	0.144 (0.351)	0.145 (0.352)	0.935 [14375]

Notes: This table displays summary statistics for the main data used in the present study. The sample consists of all household heads in the 1850 census with children born in Georgia during the three years prior to the Cherokee Land Lottery of 1832 and no children born outside of Georgia during the same period. Column (1) presents means and standard deviations (in parentheses) of variables for this entire sample. We use two measures of whether the person won land in the drawing for the Cherokee Land Lottery of 1832. The first measure is coded to 1 if that person is a unique match to a name found on the list of winners published by Smith (1838); anyone else in the sample is coded to zero. The second measure takes individuals that "tie" for a match to the Smith list with (n-1) other observations and recodes them to 1/n. These variables are summarized in Panel A. Columns (2) and (3) present means and standard deviations of variables for the subsamples of, respectively, lottery losers and winners (decomposed using the first measure). Column (4) presents the p-value on the test of zero difference in means between the subsamples of losers and winners. In square brackets, we report the sample size used for this test, although the test involving children or surnames adjust for the clustering of errors. With the exception of the measure of surname length, we use the Soundex version of each name to account for minor spelling differences. For the variables that are means by surname, we use the 1850 100% census file to construct average fertility, school attendance, and real-estate wealth among Georgia-resident households for each (soundex) surname. (Those individuals that appear in our lottery-eligible sample are excluded from the construction of these indices.) Real-estate wealth is as reported on and transcribed from the manuscript pages of the 1850 Census of Population. Slave wealth was estimated by linking the household to the 1850 Slave Schedule and imputing a market value of slave holdings adjusting for the reported ages and gender of slaves on the Schedule. Numbers in curly brackets in Panel C are the 25th, 50th, and 75th percentiles of the respective wealth measures. Data sources and additional variable and sample definitions are found in the text.

Table 2: Lottery Status versus Total Wealth in 1850

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Binary Match to Smith (1838)							
	723.4	714.4	710.1	632.4	593.6	855.1	677.8	723.6
	(325.3) **	(319.5) **	(325.4) **	(311.2) **	(352.3) *	(348.9) **	(385.6) *	(325.2) **
	Panel B: Allow for 1/n Matching to Smith (1838)							
	777.7	749.8	762.5	660.2	572.0	922.7	645.6	777.6
	(310.7) **	(303.0) **	(310.5) **	(300.2) **	(335.6) *	(331.3) ***	(332.6) **	(310.6) **
Additional Fixed-Effect Controls:	None	First letter of surname	Number of letters in surname	Freq. of surname in sample	Surname	Given name	Surname; Given name	None; Adjust truncated lower tail

Notes: This table displays OLS estimates of equation (1) in the text. Each cell presents results from a separate regression, and only the coefficient on winning the lottery is reported. The sample consists of all household heads in the 1850 census with children born in Georgia during the three years prior to the Cherokee Land Lottery of 1832 and no children born outside of Georgia during the same lapse of time. The dependent variable in this table is total measured wealth. This variable is the sum of real-estate wealth, which was reported to enumerators on the population schedule, and slave wealth, which was computed from the slave schedule. The sample size is 13,094. The baseline specification also includes dummies for age and for (state x county) of residence. Additional sets of fixed effects are included in columns 2-7, as reported in the bottom row. In columns 4-7, we use the Soundex version of each name to account for minor spelling differences. Two variables are constructed to measure whether the person was a lottery winner. The first measure, used in Panel A, is coded to 1 if that person is a unique match to a name found on the list of winners published by Smith (1838); anyone else in the sample is coded to zero. The second measure, which is used in Panel B, takes individuals that "tie" for a match to the Smith list with (n-1) other observations and recodes them to 1/n. A single asterisk denotes statistical significance at the 90% confidence level; double 95% and triple 99%. Data sources and additional variable and sample definitions are found in the text.

Table 3: Effects on Child Quantity versus Quality

	(1)	(2)	(3)	(4)
	Number children born post 1832	Natural log of total children	Attended school	Attended school (Logit)
1. Estimates of the effect of winning the lottery				
Panel A: Basic Specification				
	0.134	0.032	-0.001	-0.005
	(0.059) **	(0.015) **	(0.011)	(0.051)
Panel B: Control for Surname Fixed Effects				
	0.193	0.030	-0.003	-0.010
	(0.073) ***	(0.014) **	(0.011)	(0.033)
2. Estimation sample				
	Lottery-eligible person, linked to household characteristics [N=14375]		Children aged 5-17, inclusive [N=47749]	

Notes: This table displays estimates of equation (1) in the text. Each cell presents results from a separate regression, and only the coefficient on winning the lottery is reported. Estimates are computed using OLS, except in Column 4, which uses logit. The basic specification (shown in Panel A) also includes dummies for age and for (state x county) of residence. The specification used in Panel B includes fixed effects for surname (soundex). The base sample consists of all households in the 1850 census with children born in Georgia during the three years prior to the Cherokee Land Lottery of 1832 and no children born outside of Georgia during the same period. The sample for Columns 1-2 consists of household heads, while the sample for Columns 3-4 consist of their children at least 5 but not more than 17 years of age. The dependent variables are indicated in the column headings. A household is coded as a lottery winner if the head is a unique match to a name found on the list of winners published by Smith (1838); anyone else in the sample is coded to zero. A single asterisk denotes statistical significance at the 90% confidence level; double 95% and triple 99%. All standard errors are heteroskedasticity robust and clustered on the lottery-eligible man if there are multiple observations per household. Data sources and additional variable and sample definitions are found in the text.

Table 4: Mechanisms and Decompositions, Fertility and School Attendance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Spouse present	Spouse age, if present	Post-1832 children > 0	Post-1832 children > 3	Child gender is male	Attended school in past year			
1. Estimates of the effect of winning the lottery									
Panel A: Basic Specification									
	0.013	-0.104	0.014	0.021	0.000	0.003	-0.006	-0.007	0.007
	(0.010)	(0.138)	(0.007) **	(0.011) *	(0.008)	(0.013)	(0.013)	(0.012)	(0.013)
Panel B: Control for Surname Fixed Effects									
	0.009	-0.038	0.015	0.018	0.000	0.005	-0.008	-0.013	0.009
	(0.011)	(0.153)	(0.007) **	(0.010) *	(0.009)	(0.014)	(0.013)	(0.013)	(0.014)
2. Estimation sample									
	Lottery-eligible person, linked to household characteristics [N=14375]				Children under 18 [N=47749]	Males, age 5-17 [N=24510]	Females, age 5-17 [N=23239]	Children, age 5-12 [N=26756]	Children, age 13-17 [N=20993]

Notes: This table displays OLS estimates of equation (1) in the text. Each cell presents results from a separate regression, and only the coefficient on winning the lottery is reported. The basic specification (shown in Panel A) also includes dummies for age and for (state x county) of residence. The specification used in Panel B includes fixed effects for surname (soundex). The base sample consists of all households in the 1850 census with children born in Georgia during the three years prior to the Cherokee Land Lottery of 1832 and no children born outside of Georgia during the same period. The sample for Columns 1-4 consists of household heads, while the sample for Columns 5-8 consist of their children at least 5 but not more than 17 years of age, with subsamples noted in the last row. The dependent variables are indicated in the column headings. A household is coded as a lottery winner if the head is a unique match to a name found on the list of winners published by Smith (1838); anyone else in the sample is coded to zero. A single asterisk denotes statistical significance at the 90% confidence level; double 95% and triple 99%. All standard errors are heteroskedasticity robust and clustered on the lottery-eligible man if there are multiple observations per household. Data sources and additional variable and sample definitions are found in the text.

Table 5: Differences in 1850-County-of-Residence Characteristics by Lottery Status

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Resides in Old Cherokee County	Resides in Georgia	Miles East	Miles North	School Enroll. Rate	Total Fertility Rate (TFR5)	Total Fertility Rate (TFR19)	Log of Farm Value per Acre	Log of Average Farm Size	Log of Improved Land Ratio	Log Slaves per Area	Log Pop. Density in 1850	Log Pop. Density in 1830	Log Fraction Urban	Access to Water Transport	Access to Railroads
Panel A: Basic Specification															
0.022	0.005	4.320	-4.026	-0.003	0.006	0.011	-0.007	-0.014	-0.017	-0.001	-0.045	-0.111	0.072	-0.007	0.018
(0.008) ***	(0.011)	(3.643)	(2.211) *	(0.003)	(0.004)	(0.012)	(0.021)	(0.017)	(0.017)	(0.026)	(0.029)	(0.052) **	(0.072)	(0.011)	(0.016)
Panel B: Control for Surname Fixed Effects															
0.022	0.004	4.265	-4.661	-0.004	0.005	0.009	-0.011	-0.005	-0.024	0.000	-0.057	-0.117	0.066	-0.001	0.015
(0.008) ***	(0.013)	(3.997)	(2.306) **	(0.003)	(0.004)	(0.012)	(0.022)	(0.017)	(0.018)	(0.026)	(0.030) *	(0.054) **	(0.062)	(0.011)	(0.016)
Panel C: Basic Specification, Control for Residence in Old Cherokee County															
---	---	4.654	-5.924	-0.004	0.005	0.007	-0.012	-0.004	-0.015	0.008	-0.048	-0.101	0.072	-0.003	0.018
		(4.343)	(2.781) **	(0.003)	(0.004)	(0.012)	(0.022)	(0.016)	(0.017)	(0.025)	(0.029)	(0.052) *	(0.072)	(0.011)	(0.015)
Panel D: Control for Surname Fixed Effects, Control for Residence in Old Cherokee County															
---	---	4.560	-6.569	-0.005	0.004	0.005	-0.016	0.004	-0.022	0.009	-0.058	-0.108	0.066	0.002	0.014
		(4.727)	(2.761) **	(0.003) *	(0.004)	(0.012)	(0.022)	(0.016)	(0.019)	(0.026)	(0.030) *	(0.054) **	(0.062)	(0.011)	(0.015)

Notes: This table displays OLS estimates of equation (1) in the text. Each cell presents results from a separate regression, and only the coefficient on winning the lottery is reported. The basic specification (shown in Panel A) also includes dummies for age. The specification used in Panel B includes fixed effects for surname (soundex). Panels C and D repeat specifications from Panels A and B, respectively, but also include a dummy variable for residence in Old Cherokee County. The sample consists of all household heads in the 1850 census with children born in Georgia during the three years prior to the Cherokee Land Lottery of 1832 and no children born outside of Georgia during the same period. The dependent variables are the locational county-specific characteristics denoted in the column headings. Location data used in Columns 3 and 4 are county centroids computed from NHGIS data, and are converted into miles east or north of the NAD83 reference point in central Oklahoma. County data used in Columns 5-14 are drawn from ICPSR study #2896. The number of observations for Columns 1-4 is 14375 and for Columns 5-14 is 14237 because of missing data for some (mostly unorganized) counties. A household is coded as a lottery winner if the head is a unique match to a name found on the list of winners published by Smith (1838); anyone else in the sample is coded to zero. A single asterisk denotes statistical significance at the 90% confidence level; double 95% and triple 99%. All standard errors are heteroskedasticity robust and, in Columns 3-14, clustered at the (state x county) level to account for multiple observations per county. Data sources and additional variable and sample definitions are found in the text.

Table 6: Outcomes of Next Generation(s) in 1870-80

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Linked to 1880 census	Linked to 1870 census	Unable to read and write	Occup. Score	Total Wealth ($)	Wealth Positive	Natural log of Wealth	Unable to read and write	Enrolled in school	Number children under 18
1. Estimates of the effect of father or grandfather winning the lottery										
Panel A: Basic Specification										
	0.037	0.018	0.004	-0.124	57.6	-0.031	-0.038	-0.002	-0.020	-0.092
	(0.010) ***	(0.008) **	(0.008)	(0.285)	(86.8)	(0.016) *	(0.048)	(0.013)	(0.012) *	(0.056) *
Panel B: Control for Surname Fixed Effects										
	0.027	0.011	0.005	-0.087	112.4	-0.017	0.023	0.001	-0.028	-0.092
	(0.011) **	(0.008)	(0.009)	(0.326)	(100.9)	(0.019)	(0.060)	(0.014)	(0.012) **	(0.056) *
Panel C: Control for Surname Effects and Length of Given Name										
	0.017	0.003	0.011	0.328	115.7	-0.008	0.055	-0.003	-0.035	-0.105
	(0.013)	(0.010)	(0.011)	(0.398)	(119.0)	(0.025)	(0.078)	(0.015)	(0.014) **	(0.060) *
2. Estimation sample										
	Children in 1850	Children in 1850	1850 children as adults in 1880	1850 children as adults in 1880	1850 children as adults in 1870	1850 children as adults in 1870	1850 children as adults in 1870	Children in 1880, 10-19 years old	Children in 1880, ages 5-19	1850 children as adults in 1880
	[N=40024]	[N=24510]	[N=14963]	[N=14956]	[N=6823]	[N=6823]	[N=6823]	[N=23544]	[N=40658]	[N=14963]

Notes: This table displays OLS estimates of equation (1) in the text. Each cell presents results from a separate regression, and only the coefficient on winning the lottery is reported. The basic specification (shown in Panel A) includes dummies for age and for (state x county) of residence. The specification used in Panel B also includes fixed effects for surname (soundex), and the specification in Panel C adds to this dummies for the length (number of letters) of the given name. The base sample of children in 1850 is as described in prior tables, and this sample is used in Columns 1 and 8 to estimate the differential probability of linkage to 1870 and 1880 censuses. The samples in the remaining columns are drawn from the 1870 or 1880 households of those male children linked from 1850. The dependent variables are indicated in the column headings. A household is coded as a lottery winner if the head is a unique match to a name found on the list of winners published by Smith (1838); anyone else in the sample is coded to zero. A single asterisk denotes statistical significance at the 90% confidence level; double 95% and triple 99%. All standard errors are heteroskedasticity robust and clustered on the lottery-eligible man if there are multiple observations per household. Data sources and additional variable and sample definitions are found in the text.

Table 7: Estimated Returns to Skill from the Cross Section

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Estimated returns to human capital						
Dependent variable:	Total wealth in 1850	Occup. score in 1880	Total wealth in 1870	Literacy in 1880	Occup. score in 1880	Total wealth in 1870	
Measure of human capital:	Cannot read and write, 1850	Cannot read and write, 1880	Cannot read and write, 1880	Attend school in 1850	Attend school in 1850	Attend school in 1850	
Estimates from basic specification:	-2737 (103) *** [-770]	-3.107 (0.188) ***	-661 (100) *** [-300]	-0.038 (0.008) *** [-300]	2.004 (0.325) ***	633 (97) *** [205]	
Estimates using surname fixed effects:	-2828 (162) ***	-2.994 (0.218) ***	-648 *** (149)	-0.027 (0.010) ***	1.586 (1.586) ***	548 *** (115)	
Sample:	Lotto-eligible	Children in 1850 as adults in 1870-1880 (ages [5,18] for col 4-6)					
Sample size:	13063	14956	6501	7524	7524	5380	
	Panel B: Does surname average predict own level?						
Dependent variable:	Number children born post 1832	Attended school in 1850	Total wealth in 1850, levels	Total wealth in 1850, logs	Son's total wealth in 1870, logs		
Average of outcome for surname in 1850 Georgia	0.169 (0.063) ***	0.130 (0.015) ***	0.225 (0.088) **	0.164 (0.022) ***	0.076 (0.033) **		0.023 (0.032)
Father's total wealth in 1850, logs						0.226 (0.012) ***	0.225 (0.012) ***
Number of observations:	14213	45688	12661	12553	5080	5080	5080

Notes: This table display OLS estimates of equations (2) and (3). (Terms in square brackets for wealth outcomes are from a quantile regression at the median.) This table departs from previous ones in that we are not estimating a treatment effect of winning the lottery, but rather estimating the relationship (not necessarily a causal one) between human-capital variables and other outcomes. Each cell presents results from a separate regression. In addition to the displayed coefficients, regressions include dummies for age and state/county of residence. The second specification in Panel A includes fixed effects for surname (soundex). The samples are as described in the previous tables. In the regressions with school attendance, children under 5 and over 18 years of age in the year in which that variable is measured are excluded from the sample. The dependent and main independent variables for each column are indicated in the first two rows of each table. For Panel B, we use the 1850 100% census file to construct average fertility, school attendance, and real-estate wealth among Georgia-resident households for each (soundex) surname. (Those individuals that appear in our lottery-eligible sample are excluded from the construction of these indices.) A single asterisk denotes statistical significance at the 90% confidence level; double 95% and triple 99%. All standard errors (shown in parentheses) are heteroskedasticity robust and clustered at the level of the original lottery-eligible person. Data sources and additional variable and sample definitions are found in the text.

Table 8: Falsification test using South Carolina instead of Georgia to construct sample

Dependent variables: / Lottery-status variables:	(1) Born in Georgia	(2) Born in South Carolina	(3) Number Ga.-born children, pre-lottery	(4) Number SC-born children, pre-lottery	(5) Resides in Old Cherokee County	(6) Number children born post 1832	(7) Real-estate Wealth ($)	(8) Real-estate Wealth >$100	(9) In school, children ages [5,18]
Panel A: South Carolina, basic specification									
Dummy for unique match to Smith (1838) list	0.001	-0.017		-0.019	0.005	0.019	-41.1	0.001	0.025
	(0.004)	(0.013)		(0.019)	(0.007)	(0.077)	(236.2)	(0.016)	(0.021)
Dummy for match to Smith (1838), deflated to 1/n in case of ties	0.004	-0.016		-0.002	0.010	0.001	-15.6	0.001	0.025
	(0.004)	(0.012)		(0.018)	(0.007)	(0.074)	(232.0)	(0.015)	(0.020)
Panel B: South Carolina, including surname fixed effects									
Dummy for unique match to Smith (1838) list	0.000	-0.003		-0.016	0.006	0.016	-93.3	0.005	0.053
	(0.004)	(0.014)		(0.021)	(0.008)	(0.096)	(229.4)	(0.016)	(0.023) **
Dummy for match to Smith (1838), deflated to 1/n in case of ties	0.003	-0.004		-0.004	0.009	-0.029	-72.6	0.015	0.055
	(0.004)	(0.014)		(0.020)	(0.008)	(0.093)	-(72.6)	(0.015)	(0.022) **
Panel C: Analogous results for Georgia, dummy for unique match to Smith list									
Basic specification	-0.004	0.014	0.002		0.022	0.134	295.2	0.002	-0.001
	(0.012)	(0.011)	(0.014)		(0.008) ***	(0.058) **	(154.4) *	(0.011)	(0.011)
Control for surname fixed effects	0.001	0.012	0.009		0.023	0.193	315.8	0.002	-0.003
	(0.014)	(0.012)	(0.016)		(0.008) ***	(0.073) ***	(146.8) ***	(0.011)	(0.011)

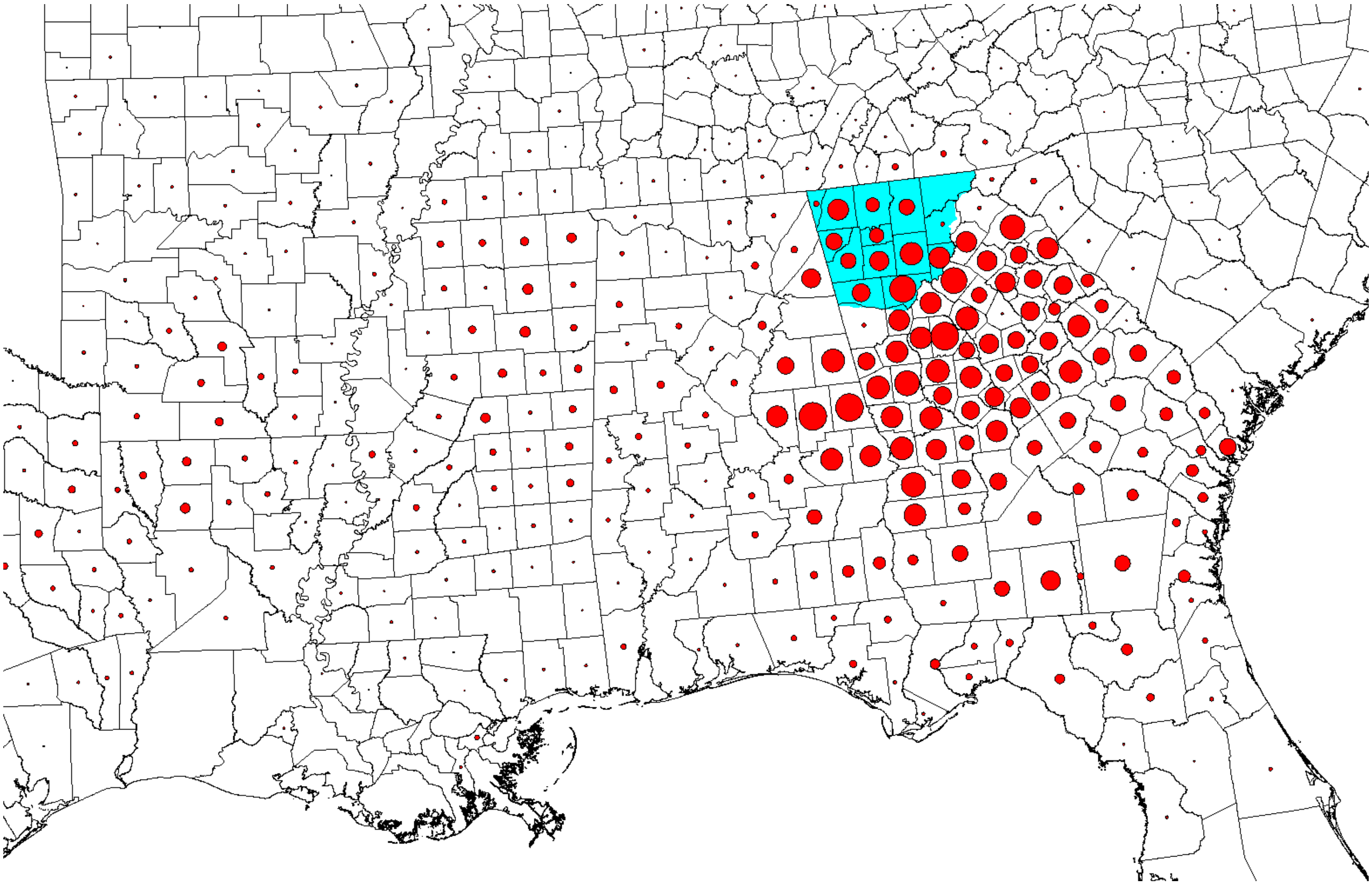
Notes: This table displays estimates of equation (1) in the text. Each cell presents results from a separate regression, and only the coefficient on "winning the lottery" is reported. The sample for Panels A and B consists of all households in the 1850 census with children born in South Carolina during the three years prior to the Cherokee Land Lottery of 1832 and no children born outside of Georgia during the same period. The sample for Panel C, which repeats some results from earlier tables, uses households with Georgia-born children in this same window. We use two measures of whether the person won land in the drawing for the Cherokee Land Lottery of 1832. The first measure is coded to 1 if that person is a unique match to a name found on the list of winners published by Smith (1838); anyone else in the sample is coded to zero. The second measure takes individuals that "tie" for a match to the Smith list with (n-1) other observations and recodes them to 1/n. Note that these are spurious measures for the South-Carolina samples because the birthplace of their children implies that they lived outside of Georgia at some point during the three years prior to the lottery, and were therefore ineligible. The basic specification also includes dummies for age. The other specification used includes fixed effects for surname (soundex). The dependent variables are indicated in the column headings. A single asterisk denotes statistical significance at the 90% confidence level; double 95% and triple 99%. All standard errors (shown in parentheses) are heteroskedasticity robust and clustered on the lottery-eligible man if there are multiple observations per household. Data sources and additional variable and sample definitions are found in the text.

Table 9: Simulated effects of wealth and comparison with estimates from the lottery

		Outcome measure:	Fraction w/ zero	Results from simulations or estimates; Mean and [95% Confidence Interval] EV $700	EV $500	EV $300	EV $900	Estimates from above	
(A)	1850	Number children born post 1832	0	.029 [-.008, .065]	.038 [.003, .073]	.045 [.010, .077]	.018 [-.022, .056]		
			.25	.012 [-.023, .041]	.023 [-.006, .047]	.031 [.005, .054]	.001 [-.034, .033]	.134 [.018, .250]	3A1
			.5	-.004 [-.027, .017] †	.006 [-.013, .028]	.017 [-.003, .036]	-.014 [-.039, .011] †		
(B)		Natural log of total children	0	.005 [-.005, .013]	.007 [-.001, .015]	.009 [.001, .017]	.003 [-.008, .012]		
			.25	.002 [-.007, .009]	.004 [-.003, .010]	.006 [.000, .012]	-.001 [-.009, .007]	.032 [.003, .061]	3A2
			.5	-.002 [-.007, .004]	.001 [-.005, .006]	.003 [-.002, .007]	-.004 [-.011, .003]		
(C)		Attended school (children)	0	.054 [.050, .057] †	.042 [.038, .045] †	.028 [.025, .031] †	.064 [.060, .068] †		
			.25	.049 [.046, .053] †	.039 [.036, .042] †	.026 [.024, .029] †	.058 [.055, .062] †	-.001 [-.023, .021]	3A3
			.5	.043 [.040, .045] †	.034 [.032, .037] †	.024 [.022, .026] †	.049 [.047, .053] †		
(D)	1880 (sons and grandchildren)	Number grandchildren (per son) under 18	0	-.045 [-.073, -.012]	-.034 [-.057, -.003]	-.021 [-.043, .009]	-.055 [-.082, -.018]		
			.25	-.042 [-.065, -.015]	-.033 [-.051, -.005]	-.021 [-.039, .001]	-.049 [-.075, -.026]	-.092 [-.202, .018]	6A2
			.5	-.035 [-.061, -.020]	-.029 [-.046, -.011]	-.020 [-.032, -.002]	-.040 [-.070, -.030]		
(E)		Occupational score (sons)	0	.347 [.207, .482]	.229 [.086, .351]	.112 [-.013, .233]	.458 [.319, .598]		
			.25	.354 [.243, .477]	.247 [.134, .338]	.126 [.025, .216]	.459 [.331, .609]	-.124 [-.683, .435]	6A4
			.5	.346 [.260, .442]	.252 [.174, .322]	.142 [.076, .209]	.435 [.311, .550]		
(F)		Unable to read and write (sons)	0	-.021 [-.025, -.015] †	-.018 [-.022, -.013] †	-.015 [-.019, -.010]	-.023 [-.028, -.018] †		
			.25	-.018 [-.022, -.014] †	-.016 [-.019, -.011]	-.012 [-.016, -.009]	-.020 [-.024, -.015] †	.004 [-.012, .020]	6A3
			.5	-.015 [-.017, -.011]	-.012 [-.015, -.009]	-.010 [-.012, -.007]	-.016 [-.019, -.013] †		
(G)		Unable to read and write (grandchildren)	0	-.030 [-.035, -.025]	-.026 [-.031, -.022]	-.020 [-.025, -.016]	-.033 [-.039, -.028] †		
			.25	-.025 [-.030, -.022]	-.022 [-.026, -.019]	-.018 [-.021, -.014]	-.028 [-.032, -.024]	-.002 [-.027, .023]	6A5
			.5	-.020 [-.023, -.017]	-.018 [-.020, -.015]	-.014 [-.016, -.012]	-.023 [-.026, -.019]		
(H)		Attended school (grandchildren)	0	.005 [.001, .009]	.001 [-.003, .004]	-.002 [-.005, .001]	.008 [.004, .013] †		
			.25	.007 [.003, .010]	.003 [.000, .006]	-.001 [-.003, .002]	.010 [.006, .014] †	-.020 [-.044, .004]	6A6
			.5	.008 [.006, .011] †	.005 [.003, .008]	.001 [-.001, .003]	.011 [.008, .013] †		
(I)	1870 (sons only)	Total Wealth ($)	0	70 [-199, 227]	20 [-252, 177]	-27 [-261, 115]	116 [-170, 281]		
			.25	93 [-123, 218]	46 [-169, 163]	-5 [-213, 110]	135 [-82, 270]	58 [-113, 228]	6A8
			.5	111 [-43, 208]	68 [-78, 151]	22 [-131, 99]	146 [-22, 262]		
(J)		Wealth is positive	0	.035 [.028, .043] †	.032 [.024, .039] †	.026 [.018, .033] †	.038 [.029, .048] †		
			.25	.029 [.023, .036] †	.026 [.020, .033] †	.022 [.016, .027] †	.031 [.025, .040] †	-.031 [-.062, .000]	6A9
			.5	.022 [.017, .027] †	.020 [.015, .025] †	.017 [.013, .021] †	.023 [.019, .030] †		
(K)		Natural log of total wealth	0	.320 [.298, .339] †	.239 [.221, .259] †	.129 [.111, .148] †	.382 [.359, .403] †		
			.25	.292 [.277, .310] †	.231 [.217, .247] †	.144 [.131, .157] †	.345 [.323, .362] †	-.038 [-.132, .056]	6A10
			.5	.249 [.238, .266] †	.208 [.195, .219] †	.142 [.133, .152] †	.290 [.274, .303] †		

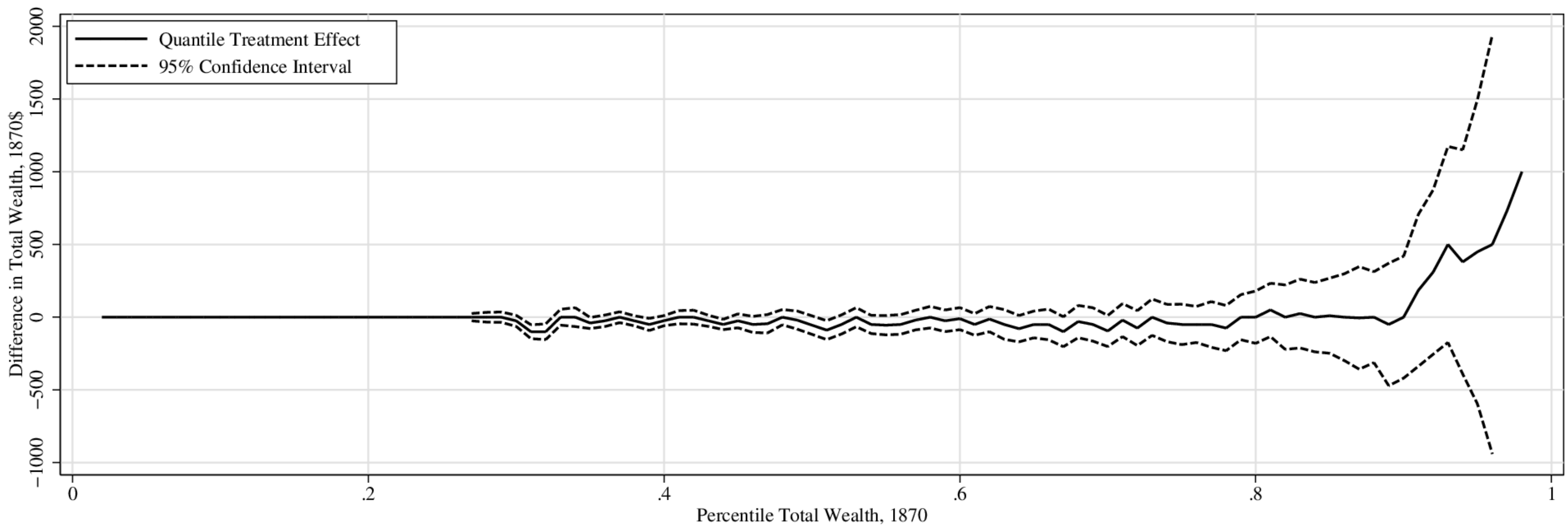
Notes: This table provides a shift-share analysis with the differences in probability generated by various perturbations of the wealth distribution and the relationship between each outcome and 1850 wealth in the control group. The outcome measures and the year in which they are measured are displayed on the leftmost columns of the table. We use a discretized distribution of 1850 wealth using 100 grid points evenly spaced across 1850 log wealth. For each simulation, we specify the expected value of winning (in 1850$), as denoted in the "EV" column-group headings. For each outcome and expected value of winning, we conduct three simulations with varying degrees of heterogeneity in the value of land winnings. These are denoted in the column "Fraction w/ zero", and indicate the fraction of the simulated winners that receive zero change in wealth. The rightmost columns display estimates from the treatment/control comparisons above. The final column on the right (a number-letter-number sequence) denotes the Table, Panel, and Column from which the estimate is drawn. Each row and column group displays the mean and, in square brackets, the 95% confidence interval from a different simulation. A dagger denotes that the confidence interval for that simulation does not overlap with the confidence interval estimated from the lottery treatment. The data for the simulation are the lottery losers, defined as those with no match to the Smith (1838) list. The statistics for each simulation come from 500 bootstrapped samples of the control group, with the lottery-eligible man being the block for the bootstrap when the outcomes are for their descendants.

Figure 1: Old Cherokee County and the 1850 Locations of the Sample



Notes: This figure displays a map of the southeastern United States with information on the location (by county) in 1850 of the lottery-eligible households in our main sample. Black lines indicate the 1850 county boundaries, drawn from the NHGIS database. The area shaded in blue in northwest Georgia denotes old Cherokee County, which was allocated by the Cherokee Lottery of 1832. The sample consists of all household heads in the 1850 census with children born in Georgia during the three years prior to the Cherokee Land Lottery of 1832 and no children born outside of Georgia during the same period. If households in our sample are resident in a county in 1850, we place a red dot at the county centroid. The area of a dot is proportional to the number of sample households resident in that county. A minor fraction of sampled households resides in counties outside the frame of this map. Such households are included in the econometric analysis, but we chose to zoom in on this region to make the feature is legible in this figure. Data sources and additional variable and sample definitions are found in the text.

Appendix Figure 1: Quantile Regression Estimates of Treatment Effects on Childrens' 1870 Wealth



Notes: This figure displays quantile-regression estimates of equation (1) in the text. The coefficient on winning the lottery is reported for various quantiles. Data sources and additional variable and sample definitions are found in the text.