Migrating to Riches?
Evidence from the California Gold Rush

Karen Clay
Heinz School of Public Policy and Management
Carnegie Mellon University
Pittsburgh, PA 15213
kclay@andrew.cmu.edu

Randall Jones
Carnegie Mellon University
Pittsburgh, PA 15213
rcjones@andrew.cmu.edu

This Version: January 14, 2008

The discovery of gold in California in January 1848 set off a very large and extremely rapid migration to California. This paper uses newly collected data from the 1850 and 1852 Censuses of Population for California together with the 1 percent public-use sample of the 1850 Census of Population to examine who went to California and how they did economically. We have four main findings. First, the propensity to migrate was affected by the individual's age and literacy, distance of the state from California, and average state latitude. The effects of distance and latitude are consistent with higher cost of the trip leading to lower propensity to migrate. Second, in 1850 miners earned more in absolute terms, but less in relative terms than day laborers in other parts of the United States. Third, individual characteristics including, notably, whether the individual was foreign-born had little ability to predict average daily product from mining. Fourth, the short run effect of having migrated to California on real state wealth was small or even zero for miners and positive and large for non-miners, which is consistent with contemporary historical reports.

We are grateful to Jeremy Atack, Maria Ferreyra, David Gerard, Joel Tarr, Werner Troesken, two anonymous referees and seminar participants at Northwestern University for helpful comments. We gratefully acknowledge the research assistance of Mercy Chang, Kerem Aksoy, Marie Yetsin, and Erin Deasy and funding provided by the Heinz School and the Carnegie Mellon undergraduate research summer fellowships.

1. Introduction

The discovery of gold in California in January 1848 set off a very large and extremely rapid migration of young men to California. By the end of 1850, 1.9 percent of native-born men ages 20-40 were already in California, and due to the undercount and loss of census records for some counties, this number was probably closer to 3.1 percent. The magnitude of the rush is all the more impressive given that very few people took the gold rush seriously until the end of 1848. In comparison, the Great Black Migration out of the South involved 26 percent of the black population representing 2.8 percent of the total population over 40 years. Only military-related migrations would induce a more rapid migration of young men in a comparably short period of time.

Given the magnitude of the rush and its national importance, we know remarkably little about who went or the factors that affected their decision to go. Further, we know very little about what happened to participants in the gold rush once they got to California. The historical literature has drawn extensively on a few hundred published and unpublished diaries to present detailed information on various aspects of life in California.[1] The writers of these materials were, however, almost certainly not typical, and so they have a limited amount to say about the many others who went.[2] Economic historians have largely focused on mining district constitutions or on the development of mining-related industries.[3] Strikingly, few historians or economic historians have drawn

[1] Some prominent examples include Holliday (1981), Levy (1990), Rohrbough (1997), Roberts (2000).
[2] The fact that nearly all of the diarists are American already suggests a bias, since 28 percent of the population in California was foreigners. Historians have advanced some further hypotheses about the nature of the bias including the fact that writers were middle class, largely Northeastern, and often married or engaged, which gave them a reason to write. (See Owens (2002) and Roberts (2000) on the source bias inherent in most works.)
[3] See, for example, Umbeck (1981), Zerbe and Anderson (2001), and Clay and Wright (2005).

on the 1850 or 1852 Censuses of Population for California, despite their potential to shed light on various dimensions of the gold rush.[4]

This paper uses newly collected data from the 1850 and 1852 Censuses of Population for California together with the 1 percent public-use sample of the 1850 Census of Population to examine who went to California and how they did economically. The newly collected data includes 100-percent samples of the 1850 Census of Population from El Dorado, Sutter, and Yuba Counties covering 31,345 individuals. In addition to the standard census questions, the enumerator in El Dorado County asked individuals to report their average daily product from mining, thus shedding new light on earnings from mining in California. We also use the 1852 Census of Population. It asked fewer questions than the 1850 Census, but it asked a key question that the 1850 Census did not: Where was the individual last resident before coming to California? This information allows us to impute states of last residence based on the birthplaces listed in the 1850 Census and examine factors related to migration.

Our results contribute both to the California history literature and to the economics literature on migration. Specifically, they challenge historians' estimates of wages and the extent of discrimination against foreigners in the mines.[5] Our results, however, support the contention of miners and historians that merchants and other service providers reaped most of the profits from mining. With respect to migration, much of the research thus far has examined migration later in the nineteenth century.[6] This paper

[4] Mann (1982) is a notable exception, but he examines just two cities – Grass Valley and Nevada City.
[5] See Paul (1947), Bancroft (1888), Rohrbough (1997), Johnson (2000).
[6] See Collins (1997), Hatton and Williamson (1998), Ferrie (1997), Herscovici (1998), Walker (2000), Stewart (2006).

breaks new ground by presenting evidence on the factors that drove migration to the gold rush and the short-run economic gains that miners realized.

The paper is organized as follows. The next section provides some historical background on the gold rush. Section 3 reviews the main results from theoretical models of migration and from models of the dynamic response of wages to a large demand shock. Section 4 examines where individuals migrated from and the factors that underlay their decision to migrate. Sections 5 and 6 examine the average daily product from mining and the real estate holdings of miners and non-miners. Section 7 concludes.

2. The Gold Rush

On January 24th, 1848 James Marshall discovered gold at John Sutter's mill in Coloma in what would later be El Dorado County. Information about the discovery of gold took most of 1848 to disseminate, however. In March, *The Californian*, a San Francisco newspaper printed a story about the discovery of gold. But the streets of San Francisco did not immediately empty. In May, Sam Brannan arrived in San Francisco and began to advertise the arrival of the gold rush. A store owner at Sutter's Fort and the publisher of the *California Star*, Brannan stood to gain from any increase in gold mining activity. In June, an estimated 4,000 to 5,000 miners were at work in the gold district. This represented a very large share of the adult male population in California.[7] By the end of July, 2,000 copies of a special edition of the *California Star* had reached Missouri. In August, the *New York Herald* printed a story on the discovery of gold. And in December, President Polk confirmed the rumors in his address to Congress.

[7] It is important to note that fewer than five hundred native-born men ages 20-40 were in California in 1846 (Langum 1987). The population of native-born men ages 20-40 did increase by the end of 1847, but at most the number was 2,500. The remaining miners were Mexicans, Native Americans, newly arrived Americans, mostly military men, and the young and the old.

By late 1848, the whole American nation and many foreign countries knew about the California gold rush. By December 1849, the number of gold miners had risen to 40,000. The number of gold miners would peak at 100,000 in 1852.[8] As Figure 1 shows, gold production rose rapidly, peaking around 1853.

Many headed for California in the spring of 1849 either overland or by ship. Estimates suggest that between 5,000 and 6,000 wagons left Missouri in the spring of 1849. Others took overland routes that began further south or even in Mexico. The hundreds of surviving diaries suggest that those who made the overland journey were young men, typically traveling in the company of brothers, cousins, in-laws and friends from their town or village. Frequently they elected one of their members to be the leader and some signed compacts that bound them to travel together to California. For the period 1848-1850, lower-bound estimates of overland migration are more than 101,000.[9] Ships had also begun to leave New York and other cities on the Atlantic seaboard for California. The ships either took the long route around Cape Horn to California or left the passengers in Panama. In the latter case, the passengers then traveled overland across Panama and took a second ship from Panama to San Francisco. For the period, 1849-1850, arrivals by sea are conservatively estimated at 75,462.

Migrating to California from within the United States typically required cash outlays of $100-300.[10] For a migrant who paid $200 in 1850, this represented more than six-

---

[8] These figures come from contemporary estimates, summarized in Paul (1947), p. 43.

[9] Estimates of overland migration are summarized in Wright (1940) p. 342. Passenger arrivals by sea may be found on p. 341. Note that a substantially higher figure for 1849 immigration is presented in the *State Register and Book of Facts* (San Francisco, 1857). Total arrivals were as much as 50 percent larger than the recorded 1852 non-Indian population of 223,856, suggesting that many newcomers had already left.

[10] Rohrbough (1997), p. 40. Bancroft (1888), pp. 134, 144. Gordon (1983), p. 21. Rydell (1948), p. 154. It is worth noting that the cost of migrating to California was far more expensive in both time and money than migrating from Western Europe to the United States. For example, the Atlantic trip typically took less

months wages for an unskilled laborer.[11] The cost varied depending on when and how one travelled. The overland route was generally thought to be cheaper, but was more rigorous and took four to six months to traverse. Further, it was less convenient for those located near the Atlantic coast than the sea routes via Panama or around Cape Horn. The sea route was physically less demanding and could be somewhat shorter, but this depended on a number of factors, such as the amount of time spent in Panama waiting for a ship. Migrants on both routes probably experienced higher mortality both en-route and once in California than they would have had they stayed home.[12]

Once they arrived in California, most individuals tried their hand at mining. In 1849 and 1850, all that was required was a pick or shovel and a pan to sort out the gold from the rock, sand, and other debris.[13] Quite quickly miners began to organize themselves to build or operate cradles or long toms, basically slanted boxes with bars on the bottom to catch the gold. The main requirement for being a miner during this early period was the ability to endure long hours of physical labor. Those who found the physical labor too difficult and those with skills – often acquired from their previous occupation – moved on to other jobs.[14] They became clerks, carpenters, merchants, doctors, lawyers, or cooks.

3. Theory

Theoretical models of migration share a common structure. Migrants compare the present value of lifetime earnings for alternatives including remaining in the same location. They then move if the expected net gain is positive. In a 2000 review article,

---

than two months. The cost of passage to the United States from Great Britain in 1850 was $17-25. Passage from Sweden in 1860 cost about $25 (Galenson 1984).

[11] $200 is worth approximately $5,000 in 2005 dollars.

[12] The increased risk of mortality came from the fact that the average migrant experienced less sanitary conditions, a poorer quality diet, and a greater risk of accidents than if he had stayed home.

[13] This paragraph draws on Paul (1947), pp. 50-66.

[14] For more discussion of occupations, see Margo (1997).

Borjas notes that three empirically testable propositions follow directly from this framework.

- "An improvement in the economic opportunities available at the destination increases the net gains to migration, and raises the likelihood that the worker moves."[15]
- "An improvement in the economic opportunities at the current location decreases the net gains to migration, and lowers the probability that the worker moves."
- "An increase in migration costs lowers the net gains to migration, and reduces the likelihood of a move."

Borjas discusses the related question of who will move. Depending on the setting, migration can induce either positive selection – migration of high skill workers – or negative selection – migration of low skill workers. Which will occur depends on whether the rate of return to skills in the new location is high or low.

In the California context, the first of the testable hypotheses has little content. The information environment was noisy, and individual priors regarding the economic opportunities in California may have varied, but individuals were all considering migrating to the same location – California. If we had the years in which individuals migrated, then we might be able to investigate the effect of the variation in expected payoff to migrating to California over time. The data we use to examine migration are, however, from the 1850 Census of Population, which does not contain information on when the individual arrived.

[15] This quote and the two that follow are from Borjas (2000), p. 4.

The second and third testable hypotheses are relevant, however. We would expect states with more economic opportunities to have experienced lower migration. Similarly we would expect states with higher costs of migration to have experienced lower migration.

It is unclear whether positive selection or negative selection was at work in this setting. One constraint on negative migration was the cost of the trip. Migration models implicitly assume perfect capital markets, which would allow individuals to borrow against future earnings. In practice, many migrants had to borrow the cash necessary to make the trip from friends or family. Not every prospective migrant had friends and family who were willing or able to supply the funds.

Some positive selection may have been occurring. We know more about those who wrote diaries and letters than the many other individuals who went. In particular, we usually know their previous occupation and something about their social class. Many diarists were from middle-class backgrounds, which indicates that there could have been positive selection. The diarists themselves were, however, unlikely to have been representative of all migrants, so we are limited in the inferences we can make.

Both Margo (1997) and Carrington (1996) provide dynamic models of labor market adjustments to a large positive shock to labor demand. The basic intuition is fairly straightforward. Nominal wages rise initially as the full impact of the shock is felt and then fall as more individuals migrate to the region. In both models, the existing stock of labor will initially supply additional labor through some combination of reduced leisure and reduced labor in other sectors. Over time, additional labor will be supplied through migration. Margo (1997) notes "The comparative static path followed by real

wages may be more complex, but it is likely, too, that real wages rise initially and then fall."[16] We do not analyze the primary implication of the models, because we do not have data on wages over time. The models do, however, provide a backdrop against which to understand our main results on productivity in mining and real estate wealth.

4. Migration

To examine where individuals were migrating from, we use data from three sources – the 1 percent public-use national sample of the 1850 Census of Population, a 100 percent sample of the 1850 Census of Population for three mining counties in California, and a 100 percent sample of the 1852 Census of Population for one mining county in California.[17] A significant problem with using the 1 percent sample to understand the gold rush is its low density. [18] It includes approximately 900 people in California, roughly 600 of which were gold miners. To address this problem, a data set was collected by the authors and undergraduate research assistants from the manuscript records of the 1850 Census of Population. A 100-percent sample was transcribed from the original handwritten records for three gold mining counties in California – El Dorado, Sutter, and Yuba counties.[19] The three counties were located in the northern part of the gold mining region which is sometimes referred to as the "Northern Mines", and we will refer to them as the northern mining counties. This sample includes 31,345 individuals who were living in 9,892 households. Eighty percent of the individuals in these counties were gold

[16] Margo (1997), p. 11.

[17] There have been some concerns about data quality for both the 1850 and 1852 Censuses of Population in California. We discuss these concerns in the Appendix.

[18] This sample is downloadable in electronic form from usa.ipums.org.

[19] These data have not been previously used, because of the labor intensive nature of the data-collection process. We are in the process of collecting 100-percent samples for additional mining counties.

Because of population pressures, these counties were subdivided during the nineteenth century. This sample includes the modern day counties of El Dorado, Sutter, Yuba, Sierra, Nevada, Placer and parts of Amador and Alpine counties. We are in the process of collecting 100-percent samples for additional mining counties.

miners. The total number of California gold miners recorded in the 1850 census was about 60,000, so this sample represents 42 percent of all miners. This much larger sample allows us to conduct a much more detailed analysis of migration and outcomes than would otherwise be possible.

Table 1 compares the characteristics of the northern mining counties, California, and the nation as a whole. The population of California and of the northern mining counties differed from the national average in several dimensions. The average age was older, in large part because there were very few children in California. Unlike the rest of the country, where 52 percent of the population was male, 94 percent of the population in California and 97 percent of the population of the northern mining counties was male. Self-reported literacy rates were slightly higher in California. This is consistent with the view that literate individuals were more likely to migrate.[20] The percentage of the population that was foreign-born was higher in California than in the northern mining counties.

The first question we are interested in is the distribution of birthplaces of men ages 20-40 who were in California.[21] Table 2 lists the nine states of birth with the highest percentages (>5 percent) and absolute numbers of men in California. Figure 2 shows for

[20] See Zerbe and Anderson (2001) for more discussion on this point.

[21] To calculate this, we use the public-use sample, but drop our three counties and use our 100-percent sample, appropriately weighted. We adjust for the undercount by multiplying by 1.6. The adjustment was arrived at by summing the arrivals by year found in Wright (1940), pp. 341-342. The total arrivals in 1849 and 1850 were conservatively estimated to be slightly more than 175,000, which is 1.91 times the official census total. There is no data on return migration in 1849. Return migration in 1849 was unlikely due to the unavailability of crews for ships. Although it was possible in theory to return overland, few ever took this route. In 1850, return migration was 26,600 according to Wright. Depending on when they left, some of these individuals would have been captured in the census. Similarly, some of those who arrived in 1850 would have arrived too late for enumeration – although vast majority of the enumeration was done in the fall and some of it continued into 1851. Based on these numbers, we conservatively estimate that the undercount was on the order of 60 percent of the official census total. This agrees with Joseph Kennedy's contemporary estimate of California's population as 165,000 (Anderson 1988, p. 46). His estimate suggests that the undercount was about 80 percent.

all states of birth the percentage of men who were in California. As expected, men who were born on or near the western frontier are heavily represented in California. In absolute numbers, very populous states such as New York and Massachusetts were heavily represented in California.

In addition to the states of birth of native-born men, Table 2 lists the nine countries other than the United States with the greatest numbers of men in California. Mexico had the largest number of men in California. This is not entirely surprising, since Mexico controlled California until 1846. Further, because of geographic proximity, many Mexican miners were able to move rapidly to the gold rush. These numbers agree fairly closely with contemporary estimates, which suggested that there were 8,000-10,000 individuals from Sonora and Lower California in California in 1849.[22] In 1850, a memorial requesting a customs house in Los Angeles stated "at least ten thousand Sonorans pass through Los Angeles on their way to the mines each spring, generally returning to Mexico in the autumn."[23] The United Kingdom was well represented, with large numbers of individuals from England, Ireland, and Scotland. Were these three aggregated with Canada, they would have edged out Mexico for the most men in California. China had relatively few men in California in 1850, although the numbers would grow significantly over the next few years.[24]

We draw on a 100-percent sample from the 1852 Census of Population for El Dorado County. Like the 100-percent sample from the 1850 Census, the data set was collected by

[22] Moerenhout (1849) and Stillman (1849) quoted in Doris Marion Wright (1940), p. 325.
[23] Quoted in Doris Marion Wright (1940), p. 325.
[24] The census counts agree closely with contemporary estimates of emigration from China, which indicate that the number was "nearly a thousand". See Chen (2000), p. 38. On the Chinese and anti-Chinese sentiment later in the gold rush, see Kanazawa (2005).

the authors and undergraduate research assistants.[25] The sample includes observations for 11,936 individuals in El Dorado County. The 1852 Census asked fewer questions than the 1850 Census, and so on the whole is less valuable than the 1850 Census. It did, however, ask one important question that the 1850 Census did not: Where was the individual last resident before coming to California? We use this information to create probability distributions by birthplace of the state of most recent residence prior to California. For example, 46 percent of the men who were born in New York, had migrated to California, and were captured by the 1852 Census came directly from New York. The remaining men who were born in New York had most recently resided in frontier states such as Michigan (16 percent), Illinois (10 percent), and Wisconsin (8 percent).

We can then apply the probability distributions to the birthplaces listed in the 1850 Census to impute where the men were living prior to coming to California. This procedure will be reasonably accurate as long as two things hold. First, we need for the migration patterns from individual birthplaces not to have changed too significantly between 1849/1850 and 1852.[26] Second, we need the migration paths conditional on birthplace for native-born men in El Dorado County in 1852 to be reasonably representative of the migration paths of the population of California as a whole. As we will see later in this section, using the imputed state of last residence instead of state of

[25] These data have not been previously used, because of the labor intensive nature of the data-collection process. Relative to the 1850 Census, the transcription process was slightly faster, because the records were transcribed into typewritten form by a WPA project in the 1930s. Unfortunately, the quality of the typescript was too low to use computer transcription (optical character recognition software).

[26] To highlight the issue of changing migration patterns, suppose that all men born in New York who came to California in 1849 and 1850 had previously migrated to frontier states, and all men born in New York who came to California in 1851 and 1852 came directly from New York. We would be imputing too few individuals from the frontier and too many from New York in 1850. In general, we expect this approach will over-impute people from their birth state, particularly for birthplaces that are more distant from California.

birth improves our ability to predict who migrated. This suggests that imputation using the 1852 Census captures additional information not captured by birthplace alone.

Our imputed last residence highlights the extent to which miners are being drawn from the western frontier. Table 3 lists the eight states that had lost 5 percent or more of their men ages 20-40. Many of the states are overlap with the states in Table 2. These states lost both a large share of men born in the state and a large share of men who were in the state prior to the gold rush. The fact that many of these states were losing such a large share of their prime age males to the gold rush is nonetheless striking. Minnesota and Utah each lost 16 percent of their prime age men to the gold rush. The absolute numbers of men are also striking. Missouri lost nearly 7,300 prime age men, and Illinois and New York each lost more than 6,000 men. Figure 3 shows for all states of last residence the percentage of men who were in California.

In Table 4, we examine the factors associated with the migration of native-born men ages 20-40 to California. We estimate a regression of the following form:

$$y_i = X_i\beta + Z_i\delta + \varepsilon_i \qquad (1)$$

The dependent variable is whether individual i was in California (1) or not (0). This is regressed on a vector of the individual's exogenous personal characteristics (X_i) and a vector of the characteristics of the individual's state of birth or last residence (Z_i). The personal characteristics include age, race, and literacy.[27]

The characteristics of the individual's state of birth or last residence include proxies for both the state labor market and the average cost of migrating from that state to California. Recall that the second prediction from the theoretical models of migration

[27] Because people learn to read before age 12, we consider literacy to be exogenous with respect to the decision to migrate to California.

was that individuals in locations with greater economic opportunities (i.e., stronger labor markets) will be less likely to migrate. We use the average real wages of laborers in the state in 1850 (the average daily wages of a laborer without board from the 1850 Social Statistics multiplied by six and divided by the weekly cost of board from the 1850 Social Statistics) and the share of men ages 20-40 in the state who were foreign-born as somewhat crude proxies for the state labor market.[28] Higher shares of foreign born men are believed to negatively affect labor market outcomes and so induce migration of native-born men. This is sometimes referred to in the literature as the foreign-born push factor.

The third prediction from the theoretical models of migration was that individuals with higher costs of migration will be less likely to migrate. We use distance from the state to San Francisco, the availability of water transport, average state latitude and average state latitude squared as crude proxies for the cost of travelling from the state to California.[29] Of the four measures, average state latitude and average state latitude squared are least obviously related to cost. What we are trying to capture here is the well-documented fact that during the nineteenth century individuals were most likely to move East-West (Steckel 1983). One reason individuals moved East-West was that their human capital in farming tended to transfer best to locations with a similar climate. Thus,

[28] Real wages of the state are the average daily wages of a laborer in the state without board from the 1850 Social Statistics multiplied by six and divided by the weekly cost of board. The share of men ages 20-40 that were foreign-born in 1850 was computed from the 1-100 national sample of the 1850 Census of Population. The real wages are conditional on the population actually present in 1850, as is the share of the population that were foreigners. While one could correct for the latter by imputing the population from the state who were in California, it is not possible to impute the wage effect. The availability of water transport is the share of counties in a state having access to navigable rivers, lakes, or the ocean from Rappaport and Sachs (2003). While these three categories are not mutually exclusive, the number of counties having lakes and navigable rivers or oceans and navigable rivers or lakes and oceans is very small.

[29] To the extent that water transport is correlated with commercial development, it may also be picking up the effects of state labor market. Unfortunately, the two effects are of different signs, so it is not clear a priori which one will dominate.

it was less costly to migrate East-West than, for example, North-South. Most migrants to California probably did not intend to farm, although many would later do so. In this case, the propensity to move East-West may also reflect other costs such as the fact that the major overland migration routes ran through Missouri and the major water migration routes left from New York. Thus, the costs of reaching these routes would be lowest for those in middle-tier states. All regressions have standard errors clustered on birthplace and are weighted to adjust for differences between our 100-percent sample and the public-use 1-percent sample.

What we find in column 1 is that individual characteristics explain a relatively small fraction of the variance. The coefficients on age, age squared, white, and literate are, however, all statistically significant and of the expected sign. Individuals who were literate and in their mid-twenties were more likely to be in California than individuals who were illiterate or in the younger or older parts of the age distribution.

In columns 2 and 3, when we include characteristics of the individual's state of birth, the fit improves relative to column 1. In column 2, the coefficient on distance is negative and significant, indicating that individuals were more likely to migrate from closer birthplaces. This confirms what we observed in Figure 2. The coefficient on transport is positive and significant, indicating that individuals were migrating from states where there was greater access to water transportation. The coefficients on latitude and latitude and latitude squared are both statistically significant. They indicate that individuals are most likely to migrate from latitudes of 37 degrees, which is roughly the latitude of Kentucky and Virginia. Someone from Vermont, which is at 44 degrees, would be 0.025

less likely to migrate than someone at 37 degrees. The coefficients on all three variables are consistent with individuals being less likely to migrate as the cost of doing so rise.

In column 3, when we add real wages and the share of foreign-born in the state of birth in 1850, only age, literacy, distance, latitude, and latitude squared are significant. The coefficients on both average real wages for common laborers and the share foreign born are, however, of the expected sign. Individuals in states with higher average real wages for common laborers were less likely to migrate. Individuals in states with greater shares of foreign born men are also more likely to migrate. This is consistent with a greater share of foreign born in a state negatively affecting labor markets and so playing a role in the decision to migrate.[30] While the evidence on labor markets' effects on the propensity to migrate is somewhat weaker than the evidence on the cost of migration, the coefficients are consistent with this prediction.

In columns 4 and 5, we include the weighted average of the characteristics of the state of last residence conditional on birthplace. For example, recall that for individuals born in New York, 46 percent came directly from New York, 16 percent came from Michigan, 10 percent came from Illinois, 8 percent came from Wisconsin and the remaining 20 percent came via other states. The characteristics of the state of last residence would be the weighted average of New York, Michigan, Illinois, Wisconsin, and other states. The fit is improved relative to columns 2 and 3, suggesting that the characteristics of the state of last residence have more predictive power than the characteristics of birthplace do.

The most striking differences between columns 2 and 4 are that in column 4, the coefficients on latitude are of opposite sign and water transport is no longer statistically

[30] See Collins (1997) on the Great Migration of Blacks out of the South, Hatton and Williamson (1998) on internal migration within the United States 1880-1910, Card (1990) on migration from Miami after the arrival of large numbers of Cubans, and Filer (1992) on migration within the United States 1975-1980.

significant. Thus conditional on distance from state of last residence, migrants are now slightly more likely to come from the periphery. This may reflect the fact that migrants are being disproportionately drawn from the upper Midwest and to a lesser degree from states like Louisiana (see Figure 3). If we compare columns 3 and 5, we see that coefficients on latitude are also of opposite sign and the coefficient on foreign born is statistically significant. The latter is consistent with a foreign-born push factor playing a role in the decision to migrate internally.[31]

Having examined who migrated, we now turn to how they did economically.

5. Average Daily Product in the Gold Rush

The enumerator in El Dorado County recorded "average value of each miners daily product" on the 1850 Census of Population.[32] For miners who were mining on their own, average daily product would have corresponded to their earnings. For miners who were mining for others, their wage on average would have been less than the average daily product. Fifty-five percent of all men ages 20-40 reported positive average daily product from mining. Men who did not report values may have done so for at least four reasons. First, they may not have been mining. Second, they may have been unwilling to report values, if, for example, they thought that the information would be used for tax or other purposes. Third, they may have just arrived and so had no average daily product to report. This was particularly likely in the fall, when most overland immigrants arrived. Fourth, they may not have been present at the time the household was enumerated, and so value

[31] In unreported specification checks in which various state variable such as wages and transportation were dropped, foreign-born remained marginally significant.

[32] This is what he wrote at the top of each column in which he recorded the average daily product from mining. Why he recorded this information is not entirely clear.

may not have been reported. We will investigate the issue of who reports average daily product in more detail shortly.

Table 5 presents the mean, standard deviation, and minimum and maximum reported average value of daily product for a number of groups, including all men, miners, native-born miners and foreign-born miners. Although the range was relatively large, the median value was $4 and the 25th and 75th percentiles were $3 and $5 per day. These numbers are in line with data from a variety of contemporary sources.[33] Margo (1997) drawing on data from wages paid to civilian employees at U.S. Army forts in California during the gold rush, finds average daily wages of $3.78 per day in 1850. Data from the 1850 Census of Manufacturing for California on wages paid in mining also suggest that the average was close to $4 per day. Contemporary estimates were printed in the *Sacramento Placer Times* on Oct. 26, 1850. The editor of that newspaper estimated that average daily earnings were: Feather River, $6; Yuba River, $4; Bear River, $4; American River $5.[34] Another estimate from 1850 covering fourteen companies engaged in river damming operations indicated that the average daily product was $3.16.[35]

In Table 6, we present average daily product and the cost of board from the 1850 Social Statistics for El Dorado County, three other mining counties, a number of states and the nation as a whole.[36] These numbers indicate that the average miner in El Dorado who reported average daily product was, if he was working on his own account, earning somewhat less than what they would have earned as a day laborer in El Dorado County.

---

[33] These numbers are considerably lower than $10 per day that Rodman Paul reported for 1850 in his appendix on wages in the mines (Paul 1947, p. 349). He provides very little evidence as to how he arrived at this number.

[34] Cited in Quaife (1949), p xxxv.

[35] Data from Woods (1851) cited in Quaife (1949), p xxxv.

[36] Not all enumerators reported wages in Social Statistics and those that did often noted that the wages were fluctuating greatly.

This is not at all surprising, since miners were willing to accept a lower average payoff in the hopes that they would strike it rich.

To the extent that the weekly price of board reflected the actual cost of food, miners needed to spend roughly half their daily earnings on food. Some savings could probably be had by buying and preparing the food themselves, but the main point is that expenditures on food were a substantial share of earnings.

Was the average native-born miner earning more mining in 1850 than he would have had he not migrated? The answer will depend on his most recent state of residence and his previous occupation. Because it is useful for understanding the following results, and because it is important for results in the next section, in Table 7 we show the distribution of occupations for men ages 20-40 nationally and for selected states where many miners had last been resident. Nationally, the most common occupation for this age group was farmer, with 36 percent of men reporting that they were farmers. The next two most common occupations were craftsmen and laborers. Each had a 17 percent share.

Because uniform statistics on the wages paid to day laborers were collected for all states as part of the Census of Social Statistics, we compare miners to day laborers. In absolute terms, miners who had previously been day laborers were certainly doing better. Table 6 showed the average state wages for a number of occupations nationally and for selected states where many miners had last been resident. If we multiply the United States average daily wage without board by six and subtract the weekly price of board, we find that the average day laborer would have made $3.26 per week after board. If we do the same exercise for El Dorado County (outside the cities), we find that the average day laborer made $22.00 per week after board. This amount is sufficiently large that it

probably exceeded the earnings of most individuals in most occupations in the rest of the United States.

In relative terms, however, miners were also doing worse than a day laborer. If we use the cost of board as a price index, the average day laborer in the United States made 2.89 times the cost of board, whereas the average day laborer in El Dorado County made 2.57 times the cost of board. In comparison, the average miner in El Dorado County made 1.80 times the cost of board, which is substantially less than what the average day laborer in the United States made. To the extent that miners were drawn from more lucrative occupations than day laborer, they would have been doing even worse in relative terms.

In the first two columns of Table 8, we examine the question – Who reported positive average daily product from mining? We estimate a regression of the following form:

$$y_i = X_i\beta + \varepsilon_i \qquad (2)$$

The dependent variable is whether individual i reported positive daily average daily product from mining (1) or not (0). This is regressed on a vector of the individual's personal characteristics (X_i). In column 1, we only include the characteristics that are clearly exogenous such as age, race, literacy, and whether the individual was foreign-born and foreign-born but from an English speaking area. The explanatory power in column 1 is effectively zero, although the coefficients on literacy, foreign-born, and foreign-born English speaking are all statistically significant.

In column 2 we add two additional characteristics – when the individual was enumerated and whether the individual listed their occupation as miner. The effects of both variables are substantial. Individuals who were enumerated later were more likely

than individuals enumerated earlier to report average daily product from mining. This is consistent with the enumerator starting the census in more populous locations, which had proportionately fewer miners, and moving on to more remote locations, where the proportion of miners was higher. Non-miners were 64 percent less likely than individuals who listed other occupations to report average daily product from mining.

In the next two columns of Table 8, we examine the question – Conditional on having reported average daily product, what were its determinants? We estimate a regression similar to specification (2), where now the dependent variable is ln (average daily product + 1). As before, the dependent variable is regressed on a vector of the individual's personal characteristics. In column 3, as in column 1, the exogenous characteristics have virtually no explanatory power. This is what we would expect if average daily product were effectively random. Individuals who were literate reported somewhat higher average daily product, and whites and foreign-born individuals who were from English speaking nations reported slightly somewhat lower average daily product, but the effects are small. Unless there was variation in hours worked or certain miners were confined to poorer quality locations, we would expect all miners to have had a similar chance at success.

In column 4, we add when the individual was enumerated and whether the individual listed their occupation as miner. The explanatory power remains low, although we do find that individuals who were enumerated later reported higher average daily product, as did miners. The higher average daily product reported by individuals who were enumerated later is consistent with the enumerator starting the census in more populous (lower average daily product) locations then moving to more remote (higher average

daily product) locations. Individuals who listed their occupation as miner were probably devoting more of their time to mining and therefore had higher average daily product.

Strikingly, average daily product reported by foreign-born non-English-speaking men (the coefficient on foreign born) was no different than those of native-born men. Of course it is possible that this is a selection story, whereby high earning foreign-born non-English-speaking miners, most of whom were Mexican, are more likely than high earning native-born miners to report their earnings. Also, there is some discussion in the historical literature that the Mexican miners were previously miners in Mexico and so had greater experience and skill than other groups. We are unable to evaluate these possibilities, beyond the analysis we already presented on who reported earnings.

If selection was not a significant factor, and this is a big if, this result is striking for two reasons. First, it casts some doubt on claims by historians that foreign miners, particularly non-English-speaking foreign miners, were uniformly marginalized or otherwise discriminated against during the gold rush.[37] Second, both wage data from California during the gold rush and wage data from later in the nineteenth century show that foreign-born men – particularly newly-arrived foreign-born non-English-speaking men – had lower wages than native-born men.[38] One would infer that this is at least in part due to lower productivity. Yet we see no evidence that foreign-born non-English-speaking men had lower productivity in gold mining in El Dorado County in 1850. The two observations are not necessarily inconsistent – Mexican miners may have been as

[37] See Bancroft (1888), Paul (1947), Rohrbough (1997), Johnson (2000).

[38] Most of the data on the earnings of foreign born men relative to native born men comes from locations outside of California later in the nineteenth century (Hannon 1982a, 1982b; Eichengreen and Gemery 1986; Hanes 1996; Hatton 1997; Hatton and Williamson 1998). The exception is Margo 1997, who finds in Table 2, column 1 that Mexicans were paid less than non-Mexicans as common laborers/teamsters at forts in California around this time. There is some question as to whether this is really discrimination or just a reflection of the jobs within this category that Mexican's were doing.

productive as American miners, while Mexican teamsters may have been less productive than American teamsters. What the observations highlight is how little we know about wages and productivity during the nineteenth century.

6. Real Estate Wealth in the Gold Rush

To understand the economic success of individuals in California, it is helpful to consider wealth in addition to average daily product from mining. 1850 Census of Population asked individuals to report the gross value of their real estate holdings. Unfortunately, the gross value of real estate holdings is a less than perfect proxy for real estate wealth, because of the possibility that individuals had real estate-related debt. Largely for convenience, and following in the tradition of other scholars, we will refer to the reported value as real estate wealth. Interpreting the values reported in the census is complicated for a number of other reasons. First, it was almost certainly true that individuals who owned real estate were less likely to migrate. Second, those who owned real estate may have sold it to finance the trip. Third, if the holding was in California, prices were fluctuating daily, so the accuracy of the reported value was likely to be lower than in other parts of the United States. Fourth, much of the real estate that individuals reported owning in California may have represented ownership of use rights to mining claims. Thus any differences between California and the rest of the United States have to be interpreted cautiously.

Given that we will be comparing men in California with men who stayed behind, it is worth discussing what we expect to find. It is entirely possible that we would find that California was a bonanza for those who went, or more likely that it was a bust. These outcomes will depend on the extent to which expectations induced too many or too few

men to migrate. Suppose, however, that on average expectations were correct. The outcomes in California were highly skewed. Thus, measured wealth may be lower for those who migrated, particularly if we use a natural log transformation, as is common with wealth data. It may also be lower, because of the cost of the trip or because individuals in California chose to invest their wealth in something other than real estate.

Table 9 estimates the effects of being in California on real estate wealth. We estimate a regression similar to specification (2), where now the dependent variable is ln (real estate wealth + 1). The results are sensitive to what one takes as the counter-factual – the set of occupations that one views miners as likely to have been drawn from. Based on the coding of occupations in the public-use sample and our coding using the same criteria, individuals can be classified into one of ten occupational categories. In column 1, we use the full sample and include occupational fixed effects. The coefficient on the variable in California is negative and significant. What is measures is the average wealth difference between individuals in California and individuals elsewhere holding occupation constant. In particular, operatives, the occupational grouping that includes miners, in California are being compared to operatives elsewhere.

In columns 2 and 3, we split the sample to separately consider individuals who were i) operatives, ii) service workers, iii) laborers, iv) farm laborers or v) farmers and individuals who were vi) professionals, vii) managers, viii) clerks, ix) craftsmen, and x) salesmen. Thus the sample in column 2, contains, miners, other operatives, and occupations that miners in California were likely to have been drawn from. These occupations all typically required substantial amounts of manual labor, and so men in these categories would be relatively well suited to the physical demands of mining. We

do not occupational fixed effects, because historical evidence suggests miners were drawn from these groups, and so they collectively represent the relevant comparison group. The coefficient on the variable in California remains negative and significant.

In column 3, we consider professionals, managers, clerks, craftsmen, and salesmen. In this specification we include occupational dummies, because individuals in California in these occupations were likely to have engaged in these occupations previously. So, for example, the relevant comparison group for professionals in California was professionals elsewhere. The coefficient on the variable in California is negative but smaller and statistically insignificant, suggesting that any negative effect on wealth was largely confined to the occupations in column 2.

One possible concern with what we have done so far concerns the inclusion of farmers in column 2. There are two important issues with respect to farmers. First, the individuals most likely to migrate to California were relatively unlikely to be farm owners. Instead, they may have been the children or relatives of farmers who were listed as farmers but did not actually own farms.[39] These individuals were arguably more similar to farm labors than they were to farmers. Evidence from diaries indicates that a relatively small fraction of migrants were previously farmers (Holliday 1981, Clay and Chang 2007). Second, farmers in California were much wealthier than their counterparts in the United States. Further a disproportionate number of the individuals classified as farmers in California arrived prior to the gold rush.

In column 4, we run the same regression as in column 2, but now omit farmers from the comparison group for miners. The coefficient on the variable in California is

[39] More than half of all men who listed their occupation as farmers did not report owning real property in the 1850 Census of Population , which suggests that they did not actually own a farm.

positive, small, and statistically insignificant. In column 5, we run the same regression as in column 3, but now add farmers to this category. As was the case in column 3, the coefficient on the variable in California is negative and statistically insignificant.

Thus far, we have not discussed selection effects, although it is clear that different types of men were probably migrating from Missouri than from Connecticut. In the bottom panel of Table 9, we instrument for the decision to migrate using the distance from the state of last residence to California. As we saw in Table 4, where we used a logit to examine the determinants of migration to California, other characteristics of the state of last residence were sometimes statistically significant. Their inclusion tended, however, to depress the F-statistic in the first stage regression

Once we control for selection in the bottom panel of Table 9, the effect of being in California is now generally positive and ranges from very large positive and significant to small and insignificant. In column 1, the average effect of being in California on those who went is positive, statistically significant, and large. In columns 2 and 4, where we focus on miners and occupations that they were likely to have come from, we see that the results are very sensitive to the inclusion of farmers. When we include farmers in the set of occupations that miners would be drawn from, the effect is positive, statistically significant, and large. When we exclude farmers from the set of occupations that miners would be drawn from, the effect is negative, statistically insignificant, and small. We believe the latter specification is preferable given the problems with farmers we outlined above. In columns 3 and 5, we examine the success of non-mining occupations in California, first with farmers excluded and then with farmers included. The effect is large and statistically significant in column 3, when farmers are excluded, and large but

not statistically significant in column 5, when farmers are included. Because of the problems with farmers, we prefer the former specification.

Did the average native-born man in California in 1850 have greater real estate holdings than he would have had he not migrated? The answer appears to have depended very much on his occupation both before and after migrating. To the extent that miners in California were drawn from the ranks of operatives, service workers, laborers and farm laborers (column 4), the effect of migrating on their wealth was effectively zero. To the extent that non-miners in California were previously professionals, managers, clerks, craftsmen, and salesmen, the effect of migrating on their wealth was positive and large (column 3). Interestingly, these results are consistent with commentary during the gold rush and with later historical analysis. Both have indicated that the bulk of the wealth earned during the gold rush was earned by merchants, proprietors of hotels, doctors and other service providers. The returns to mining were low, because of the large numbers of migrants and the ease of entry into mining.

7. Conclusion

This paper presented new evidence on who migrated to California during the gold rush and how migration affected earnings and wealth. Economists and economic historians have long been interested in the factors that determine migration. The propensity to migrate to California was related to an individual's personal characteristics, proxies for the cost of the trip from the state of birth or state of last residence, and proxies for the strength local labor market in the state of birth or state of last residence. With respect to the cost variables, distance to California was negative and the most consistently significant across specifications. In some specifications latitude, as a proxy for distance

to the major overland and water routes, which ran through and departed from middle tier states, and access to water transport were also significant. Proxies for the local labor market, which included average real wages of common laborers and the share of foreign born, had the correct sign but were not always significant. Our findings are consistent with what Collins (1997) and Hatton and Williamson (1998) find for the late nineteenth and early twentieth centuries.

One question in the historical literature has been how well miners did economically. We brought new evidence to bear on this question from the 1850 Census. Migration to California increased earnings in absolute terms for all or nearly all occupations. For miners, however, migrating may well have lowered earnings in relative terms. Day laborers were doing less well in California in relative terms than day laborers elsewhere in the United States, because of the high cost of food. And in El Dorado County, the average daily product from mining was somewhat below those of day laborers – $4.19 versus $5.00. These estimates are lower than the most commonly cited estimates (Paul 1947) by more than a factor of two. So in relative terms, miners were on average no better off, and in all likelihood worse off, than day laborers in the rest of the United States.

Another question in the economics and historical literatures has focused on differences in earnings between foreign-born and native-born men. California historians have emphasized that foreign-born miners were consistently marginalized (Bancroft 1888, Paul 1947, Rohrbough 1997, Johnson 2000). Using data from California during the gold rush and data from other locations and sectors in the late nineteenth century United States, economic historians have found consistent differences in the wages of

foreign-born and native-born men engaged in similar activities (Hannon 1982a, 1982b; Eichengreen and Gemery 1986; Hanes 1996; Hatton 1997; Hatton and Williamson 1998). The differences are often attributed in part to differences in productivity. We drew on newly assembled data from 1850 in El Dorado County to show that the average daily product from mining were the same for foreign-born non-English-speaking men and native-born men. This result is striking, because it demonstrates that marginalization was not consistent across time and place.

A final question in both the California history literature and the economics literature has been how migration affected migrants' wealth. Using instrumental variables to correct for selection, we show that the short-run effect of migration on the real estate wealth of the migrant was occupation specific. Miners in California did relatively poorly on average because of free entry, whereas individuals who pursued other occupations in California did much better on average as measured by real estate wealth than if they had pursued these occupations at home. This is consistent with what historians have contended – that most of the economic rents in the gold rush accrued to merchants and other service providers (Bancroft 1888, Holliday 1981, Rohrbough 1997). It also adds to the small but growing literature on how migration to the frontier during the nineteenth century affected migrants' wealth (Ferrie 1997, Herscovici 1998, Walker 2000, Stewart 2006).

8. References

Anderson, M. J., 1988, The American Census: A Social History. New Haven: Yale University Press.

Bancroft, H. H., 1888, History of California, Volume VI. Santa Barbara: W. Hebberd.

Berry, T. S., 1984, Early California Gold, Prices, Trade. The Bostwick Press, Richmond.

Borjas, G., 2000, "Economics of Migration", International Encyclopedia of the Social and Behavioral Sciences, Section No. 3.4, Article No. 8.

Card, D., 1990, "The Impact of the Mariel Boatlift on the Miami Labor Market", Industrial and Labor Relations Review, 43: 245-57.

Chen, Y., 2000, Chinese San Francisco, 1850-1943: A Trans-Pacific Community. Stanford: Stanford University Press.

Clay, K. and G. Wright, 2005, "Order Without Law: Property Rights during the California Gold Rush", Explorations in Economic History, 42: 155-83.

Clay, K. and M. Chang, 2007, "Migration, Risk, and Insurance: Evidence from the California Gold Rush," Working Paper.

Collins, W. J., 1997, "When the Tide Turned: Immigration and the Delay of the Great Black Migration", Journal of Economic History, 3: 607-32.

Eichengreen, B. J. and H. A. Gemery, 1986, "The Earnings of Skilled and Unskilled Immigrants at the End of the Nineteenth Century," Journal of Economic History 46: 441-54.

Ferrie, J., 1997, "Migration to the Frontier in Mid-Nineteenth Century America: A Re-Examination of Turner's 'Safety Valve'," Working Paper.

Filer, R. K., 1992, "The Effect of Immigrant Arrivals on Migratory Patterns of Native Workers," in G. J. Borjas and R. B. Freeman (eds.), Immigration and the Workforce: Economic Consequences for the United States and Source Areas. Chicago: University of Chicago Press.

Fishback, P., W. Horace, and S. Kantor, 2006, "The Impact of New Deal Expenditures on Mobility During the Great Depression," Explorations in Economic History, 43: 179-222.

Galenson, D. W., 1984, "The Rise and Fall of Indentured Servitude in the Americas: An Economic Analysis", Journal of Economic History, 44: 1-26.

Gordon, M. M., 1983, "Overland to California in 1849: A Neglected Commercial Enterprise", Pacific Historical Review, 52: 17-36.

Hanes, C., 1996, "Immigrant's Relative Rate of Wage Growth in the Late Nineteenth Century," Explorations in Economic History, 33: 35-64.

Hannon, J. U. 1982a, "Ethnic Discrimination in a Nineteenth Century Mining District: Michigan Copper Mines, 1888," Explorations in Economic History, 19: 28-50.

Hannon, J. U. 1982b, "City Size and Ethnic Discrimination: Michigan Agricultural Implements and Iron Working Industries, 1890," Journal of Economic History, 42: 825-845.

Hatton, T. J., 1997, "The Immigrant Assimilation Puzzle in Late Nineteenth Century America," Journal of Economic History, 57: 34-62.

Hatton, T. J. and J. G. Williamson, 1998, The Age of Mass Migration: Causes and Economic Impact. Oxford: Oxford University Press.

Herscovici, S., 1998, "Migration and Economic Mobility: Wealth Accumulation and Occupational Change among Antebellum Migrants and Persisters," Journal of Economic History, 58: 927-956.

Holliday, J. S., 1981, The World Rushed In: The California Gold Rush Experience. New York: Simon and Schuster.

Johnson, S. L., 2000, Roaring Camp: The Social World of the California Gold Rush. W. W. Norton & Co.

Kanazawa, M., 2005, Immigration, Exclusion, and Taxation: Anti-Chinese Legislation in Gold Rush California. Journal of Economic History, 65: 779-805.

Lainhart, A., 1992, State Census Records. Genealogical Publishing Co.

Langum, D., 1987, Law and Community on the Mexican California Frontier. Norman: University of Oklahoma.

Levy, J. A., 1990, They Saw the Elephant: Women in the California Gold Rush. Hamden, CT: Archon Books.

Mann, R., 1982, After the Gold Rush: Society in Grass Valley and Nevada City, California, 1849-1870. Stanford: Stanford University Press.

Margo, R. A., 1997, "Wages in California during the Gold Rush", NBER Working Paper, #h0101.

Owens, K. N., 2002, Riches for All: The California Gold Rush and the World. Lincoln: University of Nebraska.

Paul, R., 1947, California Gold. Cambridge: Harvard University Press.

Quaife, M. M., 1949, Pictures of Gold Rush California, Chicago: Lakeside Press.

Roberts, B., 2000, American Alchemy: The California Gold Rush and Middle-Class Culture. Chapel Hill: University of North Carolina.

Rohrbough, M., 1997, Days of Gold: The California Gold Rush and the American Nation. Berkeley: University of California Press.

Rydell, R. A., 1948, The Cape Horn Route to California, 1849, Pacific Historical Review, 17: 149-163.

Rappaport, J. and J. Sachs, 2003, "The United States as a Coastal Nation," Journal of Economic Growth, 8: 5-46.

Southern California Genealogical Society and Family Research Library, 2005, Index to the 1852 California Census: The "Gold Rush" Census. CD 388.

Steckel, R. H., 1983, "The Economic Foundations of East-West Migration during the 19th Century", Explorations in Economic History, 20: 14-36.

Stewart, J., 2006, "Migration to the Agricultural Frontier and Wealth Accumulation, 1860-1870," Explorations in Economic History, 43: 547-577.

Umbeck, J. R., 1981, Theory of Property Rights with Applications to the California Gold Rush. Ames: Iowa State University Press.

Walker, T., 2000, "Economic Opportunity on the Urban Frontier: Wealth and Nativity in Early San Francisco," Explorations in Economic History 37, 258-277.

Wright, D. M., 1940, "The Making of Cosmopolitan California: An Analysis of Immigration, 1848-1870," California Historical Society Quarterly, 19: 323-343.

Zerbe, R. O. and C. L. Anderson, 2001, "Culture and Fairness in the Development of Institutions in the California Gold Fields," Journal of Economic History, 61:114-43.

Appendix: Data Quality Issues for the 1850 and 1852 Censuses of Population in California

The 1850 Census of Population in California has suffered from a number of criticisms. Before addressing these criticisms, it is important to correct a widespread misconception about the 1852 Census. Some historians believe, incorrectly, that the 1852 Census was done to correct deficiencies in the 1850 Census. The California state constitution of 1849 provided for a state census in 1852 and every five years thereafter. The later censuses were never taken.[40] When the 1850 Census of Population returns for Contra Costa, Santa Clara, and San Francisco Counties were lost, the 1852 Census was used to establish the population of California so that seats in the United States House of Representatives could be apportioned.[41]

The first criticism of the 1850 Census is that data were lost. This criticism is correct -- data were lost for three important counties including San Francisco. The loss of the returns for these counties, while unfortunate, does not, however, invalidate the returns for other counties.

The second criticism of the 1850 Census is that not everyone in California was enumerated. Given the rapidity with which people were arriving, this criticism is almost certainly accurate. Specific locations were only enumerated once and so provide information on who was there at the time of enumeration. Because enumeration occurred over the course of a year, later arrivals would have been missed. Individuals in more remote locations may have been missed as well.[42]

---

[40] Lainhart (1992), p. 2.
[41] Anderson (1988), p. 46.
[42] We indirectly control for this issue by including timing of enumeration in some regressions.

The nature and extent of the bias that under-enumeration would induce is not entirely clear. For example, it is not clear whether foreign-born miners would have selected more remote locations to escape violence from native-born miners or whether their property rights were actually more secure in more densely populated areas. Similarly, most overland migrants had arrived by early October. Much of the enumeration was done in the fall, so it is not clear how many of these migrants would have escaped enumeration. We have less information on the timing of the arrival of ships, although many left New York and other Atlantic ports in the spring, so these ships would have reached San Francisco by the early fall as well.

The third criticism is duplicate enumeration. For an individual to be enumerated twice in our sample, they would have to be present in multiple locations where the census was being taken. To check this, we examined every 28 year old man who was born in Massachusetts who showed up either in our data set or in the public-use sample in California. Twenty-eight was the most common age and Massachusetts was the second most common birthplace. Of the 95 men who listed their age as 28 and their birthplace as Massachusetts, there was only one possible first name last name match, B. Wright and B. Y. Wright. They each lived with one other man and those men had different names, suggesting that these were probably – although not definitely – different men. This by no means exhausts the possibility of duplicate enumeration, but it does suggest that it probably was not extremely prevalent.

The fourth criticism is falsification of the census. The quality of the enumerators appears to have varied. The enumerator for El Dorado County was unusually diligent in that his returns are complete and even have additional data on average daily product of

miners. The enumerator for Yuba County appears to have been less diligent judging from the lower density of information on his returns. Although there is no way to directly check for falsification, all of the returns showed substantial variation both within and across pages in names, birthplaces, ages, occupations and the number of individuals per dwelling and household. This gives us greater confidence that the enumerators were actually out in the field conducting the enumeration.

As long as the 1850 Census provides an approximately random sample of people who were in California at the time, it can shed new light on the gold rush. Because most of the focus is on native-born men, we only need the 1850 Census to be an approximately random sample of native-born men who were in California. This appears to be the case.

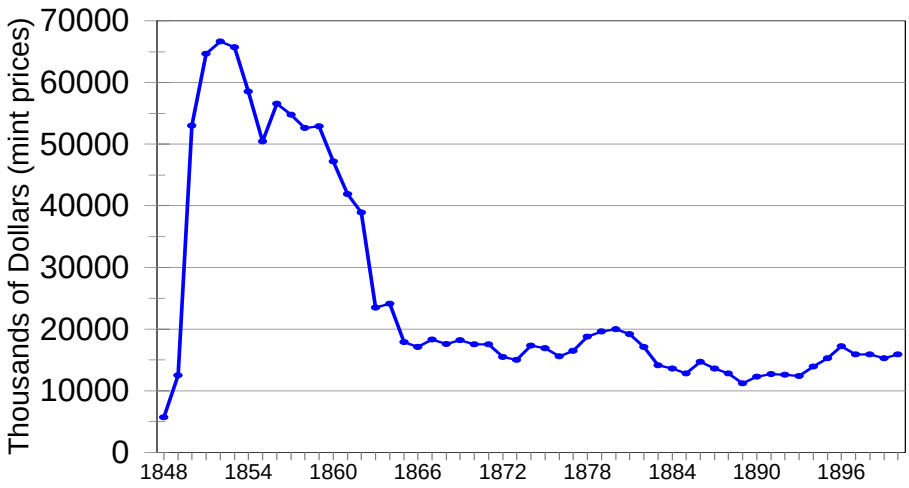
The 1852 Census of Population suffers from a number of deficits including under-enumeration of the foreign-born and the destruction of the original census returns after they were transcribed by a WPA project.[43] As with the 1850 Census, because we are not focusing on the counts, all that we need is that the transcribed records were approximately a random sample of native-born men who were in a particular location (El Dorado County). To the extent that they were, we can use the information to impute states of last residence based on birthplace for the richer 1850 data.

One question is whether the 1850 and 1852 Censuses could profitably be matched. In principle the two censuses can be matched, but in practice we found the match rate to be below 5 percent. This is attributable to several factors including i) the loss of census records for three counties including San Francisco in 1850, ii) the transience of the population – people would come for a year or two and then return home, iii) difficulties

[43] See Southern California Genealogical Society (2005), Introduction for specifics on the problems with the 1852 Census.

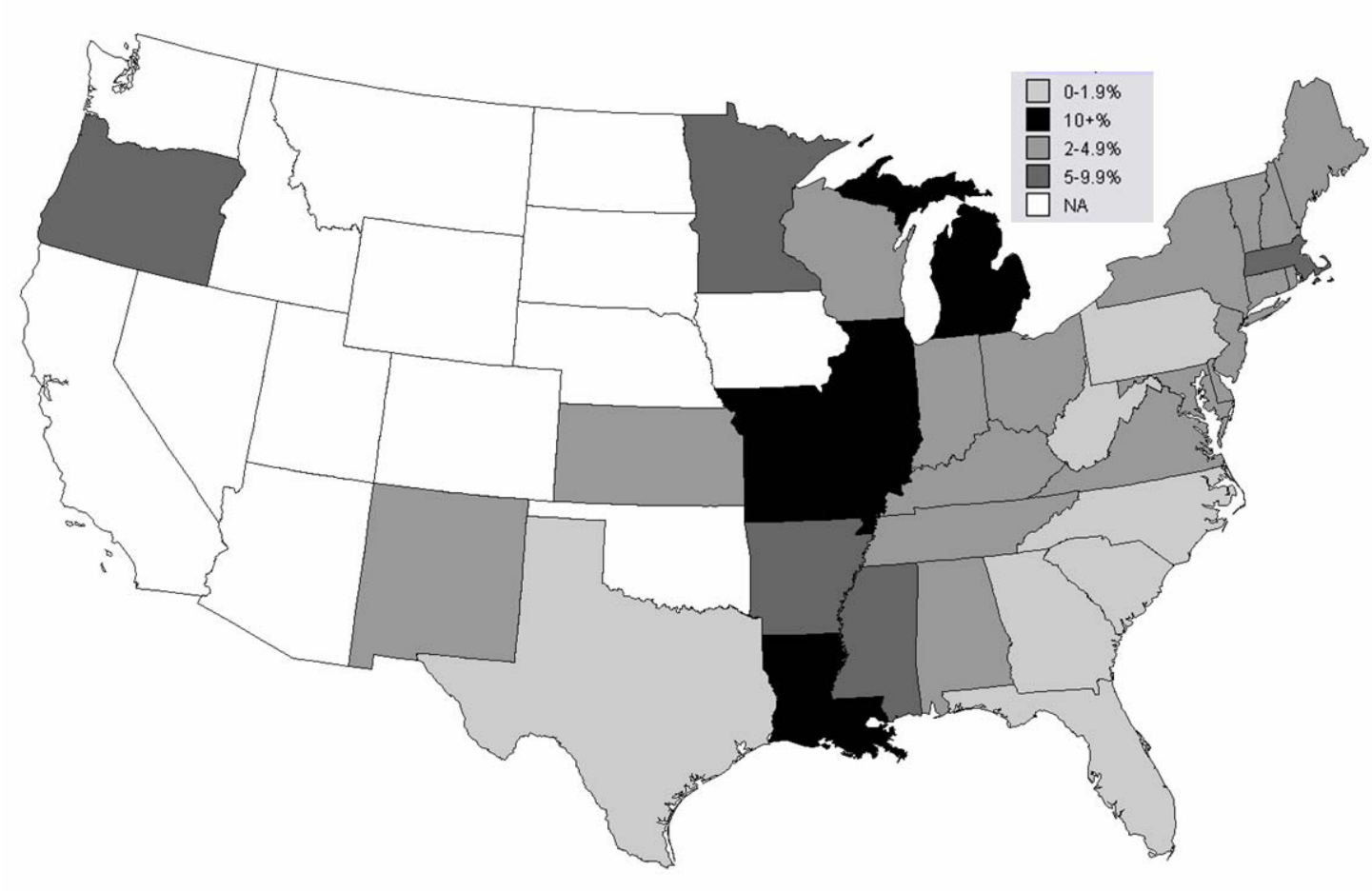
in transcribing names – Robert Smith might be written down by the 1850 and 1852 enumerators, but the transcribers both have to have accurately read the handwritten name as Robert Smith, iv) difficulties in achieving unique matches – there may be 5 Robert Smiths in 1850 who were born in New York and were 26 and 3 Robert Smiths in 1852 who were born in New York and were 28. The problem is that we don't know whether these are the same Robert Smiths or differ ones. Using unique first name, last name combinations is possible, but requires throwing out most of the data.

Figure 1: Gold Production in California over Time



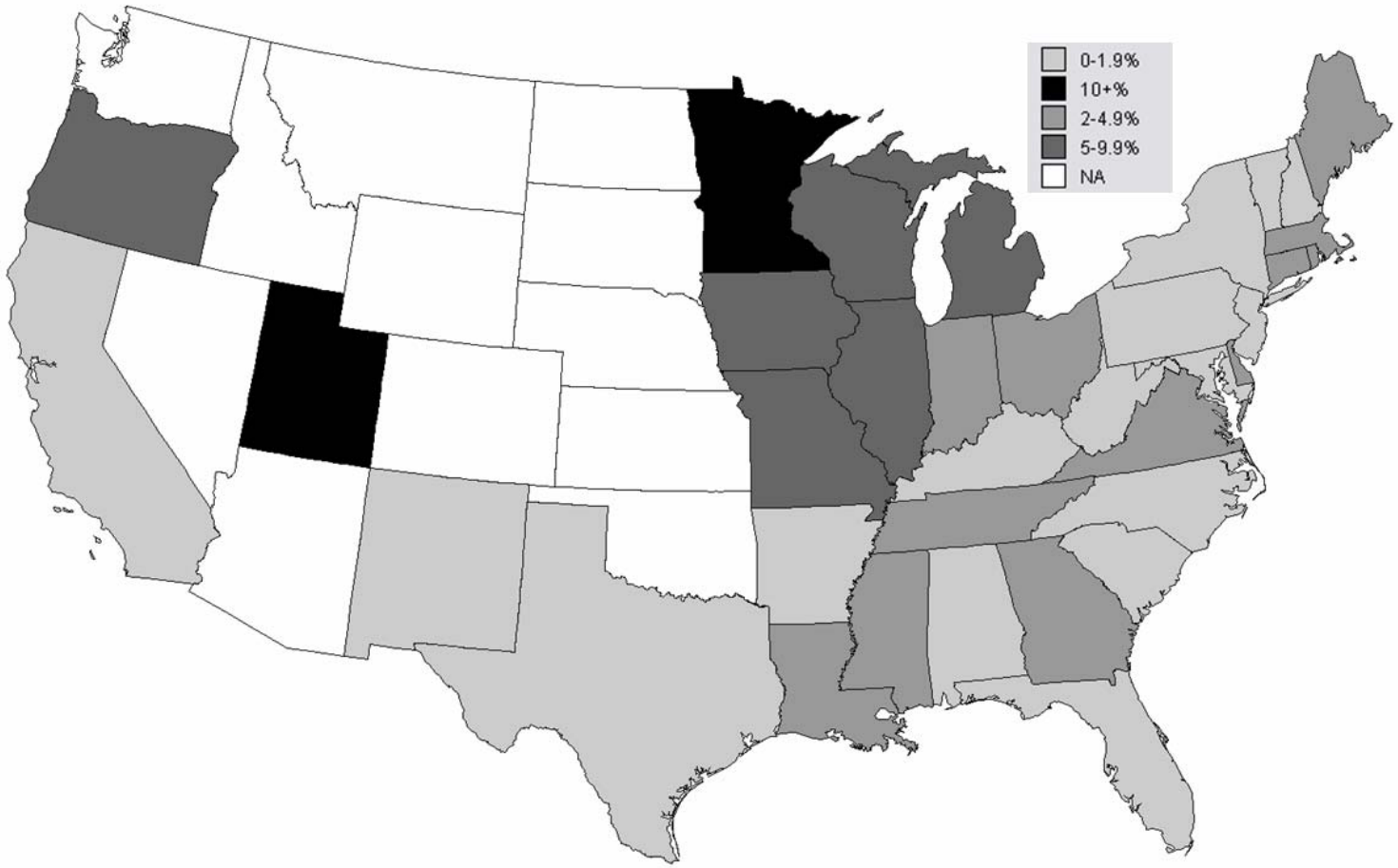
Source: Berry (1984), pp. 74, 76, 78.

Figure 2: Percentage of Men Ages 20-40 from Birthplace Who Were in California in 1850



Notes: Based on Table 2. NA are states that were not included in the 1850 Census.

Figure 3: Percentage of Men Ages 20-40 from Last Residence Who Were in California in 1850



Notes: Based on Table 3

Table 1: Summary Statistics in 1850

	Northern Mining Counties		California		National	
	Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation
Age	28.00	8.84	28.07	9.87	22.31	17.59
Sex	1.03	0.18	1.06	0.24	1.49	0.50
White	0.98	0.12	0.98	0.11	0.98	0.15
Literate	0.97	0.17	0.94	0.23	0.95	0.22
Foreign-born	0.15	0.36	0.28	0.44	0.11	0.32
Foreign-born English speaker	0.06	0.24	0.07	0.24	0.07	0.25
Obs.	31,345		31,959		196,792	

Notes: The observations for the sample and for the nation are from the authors' data and public-use sample respectively. The observations for California are the sum of the observations for our three counties (31,345) and the observations from the 1-100 public use sample for counties in California, excluding our three counties, (614). Taking into account the differences in weighting, these two samples represent the approximately 92,700 people in California.

Table 2: Birthplaces of Men Ages 20-40 in California in 1850

State	Share from birthplace in CA (>5%)	State	Number from birthplace in CA	Country	Number from birthplace in CA
Minnesota	0.051	Maine	3,582 (0.031)	Chile	810 (0.007)
Mass.	0.053	Virginia	4,469 (0.039)	China	842 (0.007)
Oregon	0.077	Tennessee	4,474 (0.039)	Scotland	998 (0.009)
Mississippi	0.091	Missouri	4,765 (0.042)	Canada	1,293 (0.011)
Arkansas	0.097	Penn.	5,485 (0.048)	England	3,010 (0.026)
Louisiana	0.105	Kentucky	5,701 (0.050)	France	3,336 (0.029)
Illinois	0.143	Ohio	7,492 (0.066)	Germany	3,691 (0.032)
Michigan	0.174	Mass.	7,793 (0.068)	Ireland	4,064 (0.036)
Missouri	0.218	New York	12,712 (0.112)	Mexico	8,480 (0.074)

Notes: All statistics are for free men ages 20-40. All counts were multiplied by 1.6 to address the 1850 undercount. The percentage of the California population that various birthplaces represent is listed in parentheses.

Table 3: Last Residence of Men Ages 20-40 in California in 1850

State	Share of men in imputed last residence (>5%)	State/Country	Number from imputed last residence in CA
Illinois	0.061	Michigan	2,775 (0.040)
Michigan	0.062	Wisconsin	3,110 (0.045)
Wisconsin	0.072	Massachusetts	4,438 (0.064)
Oregon	0.073	Mexico	5,057 (0.073)
Iowa	0.089	Ohio	5,296 (0.076)
Missouri	0.094	Illinois	6,010 (0.087)
Utah	0.161	New York	6,243 (0.090)
Minnesota	0.161	Missouri	7,267 (0.105)

Notes: All statistics are for free men ages 20-40. All counts were multiplied by 1.6 to address the 1850 undercount. The percentage of the California population that various birthplaces represent is listed in parentheses.

Table 4: Logit of Determinants of Participation in Gold Rush for Native-Born Men 20-40 (Marginal effects reported)

	In California	In California (birthplace)	In California (birthplace)	In California (last residence)	In California (last residence)
Age	0.009***	0.009***	0.009***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age squared	-0.00017***	-0.00016***	-0.00016***	-0.00015***	-0.00014***
	(0.00004)	(0.00003)	(0.00003)	(0.00003)	(0.00003)
Non-white	-0.007	-0.003	-0.003	-0.003	-0.003
	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)
Illiterate	-0.023***	-0.023***	-0.023***	-0.020***	-0.019***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Distance in 1000s miles		-0.048***	-0.049***	-0.038***	-0.030***
		(0.012)	(0.011)	(0.006)	(0.009)
Latitude		-0.022*	-0.025*	0.036*	0.037*
		(0.012)	(0.013)	(0.018)	(0.021)
Latitude squared		0.0003**	0.0003**	-0.0004*	-0.0005*
		(0.0002)	(0.0002)	(0.0002)	(0.0003)
Water Transport		0.013*	0.009	0.013	0.003
		(0.007)	(0.011)	(0.010)	(0.011)
Real wages in 1850			-0.009		0.059
			(0.012)		(0.081)
Share foreign-born in 1850			0.028		0.057*
			(0.031)		(0.034)
Observations	47804	47804	47516	47521	47141
R-Squared	0.0151	0.0410	0.0396	0.0547	0.0593

Notes: All regressions are weighed by the person weights (100 or 101 for the public-use sample outside of California, 160 or 161.6 for the public-use sample inside California and 1.6 for El Dorado, Sutter, and Yuba Counties). The number of observations is lower in column 3 than in the previous columns, because some relatively uncommon birthplaces are not represented in the 1852 Census, and thus we do not have transition probabilities to construct the weighted average of the state temperatures, distance, and transportation. The number of observations in columns 4 and 5 are lower than in columns 2 and 3 because real wages in 1850 and share foreign-born in 1850 do not exist for some states/birthplaces that are on the frontier (e.g. Kansas, Nebraska). Standard errors are in parentheses. All standard errors are clustered on birthplace. Using robust standard errors did not change the levels of significance. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 5: Summary Statistics on Average Daily Product from Mining in El Dorado County in 1850

	Mean daily product	Std. Deviation	Min	Max	Observations
All	4.19	2.57	0.10	40.00	8,597
Miners	4.19	2.57	0.10	40.00	8,530
Foreign-born Miners	4.03	2.22	0.10	30.00	1,915
Native-born Miners	4.24	2.66	0.10	40.00	6,615

Table 6: Earnings from Selected Occupations in 1850

	Day labor w/board	Day labor w/o board	Carpenter w/o board	Weekly price of board
County-Level Observations from Mining Areas in California				
El Dorado (cities)	4.50	5.00	6.00	12.00-14.00
El Dorado (outside cities)	5.00	6.00	8.00	14.00
Calaveras	4.00	6.00	8.00	14.00
Sacramento	4.00	6.00	10.00	12.00
San Joaquin	5.00	8.00	25.00	12.00
State-Level Observations for Selected States				
California	4.00	5.00	7.60	11.00
Missouri	0.55	0.75	1.48	1.31
New York	0.67	0.90	1.38	1.78
Illinois	0.62	0.85	1.47	1.47
Massachusetts	0.84	1.09	1.45	2.12
U.S.	0.61	0.84	1.43	1.72

Notes: The county level observations for California are from the manuscript schedules for Social Statistics for 1850, which are available through interlibrary loan. The state level observations are from ICPSR 2896: Historical, Demographic, Economic, and Social Data: The United States, 1790-2000. The United States average is population weighted.

Table 7: Distribution of Occupations in 1850 for Males ages 20-40

	U.S.	Missouri	New York	Illinois	Massachusetts
Professional	0.031	0.029	0.030	0.027	0.029
Farmer	0.359	0.464	0.255	0.558	0.118
Manager, Official, Proprietor	0.047	0.048	0.054	0.035	0.077
Clerical, Sales	0.024	0.023	0.038	0.013	0.044
Craftsmen	0.174	0.126	0.215	0.132	0.283
Operatives	0.099	0.086	0.101	0.047	0.168
Service Workers	0.013	0.019	0.022	0.004	0.018
Laborers	0.171	0.129	0.201	0.096	0.207
Non-occupational response	0.079	0.070	0.080	0.085	0.052

Notes: Author's computations from the public-use sample of the 1850 Census of Population.

Table 8: Determinants of Average Daily Product from Mining in El Dorado County in 1850

	Positive product Logit, marginal effects	Positive product Logit, marginal effects	Ln product OLS	Ln product OLS
Age	-0.011 (0.008)	-0.004 (0.009)	0.015 (0.012)	0.011 (0.011)
Age squared	0.00012 (0.00013)	0.00004 (0.00015)	-0.0002 (0.0002)	-0.0002 (0.0002)
Nonwhite	-0.069 (0.053)	0.002 (0.070)	0.146** (0.061)	0.071 (0.061)
Illiterate	0.081*** (0.017)	0.071*** (0.020)	-0.197*** (0.028)	-0.179*** 0.029
Foreign-born	0.112*** (0.012)	0.080*** (0.015)	0.015 (0.017)	-0.019 (0.017)
Foreign-born English speak.	-0.041*** (0.018)	-0.057*** (0.021)	-0.076*** (0.025)	-0.058*** (0.025)
Enumeration date		0.006*** (0.0001)		0.002*** (0.0002)
Non-Miner		-0.644*** (0.006)		-0.247*** 0.081
Constant			1.010*** (0.180)	90.977*** 6.306
Observations	15,643	15,643	8,597	8,597
R-squared	0.0064	0.2731	0.0081	0.0326

Notes: Standard errors are in parentheses. All standard errors are robust. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 9: The Effect of Being in California on Real Property Holdings in 1850

	OLS Estimates				
Dependant variable	Ln Real Prop	Ln Real Prop	Ln Real Prop	Ln Real Prop	Ln Real Prop
Sample	All	Miners and equivalent	All other	Miners and equiv. excl. farmers	All other including farmers
Age	0.576***	0.695***	0.353***	0.196***	0.671***
	(0.035)	(0.043)	(0.068)	(0.043)	(0.036)
Age squared	-0.0066***	-0.0083***	-0.0034***	-0.0017**	-0.0080***
	(0.0005)	(0.0007)	(0.0011)	(0.0008)	(0.0005)
Non-white	-1.403***	-1.652***	-0.747***	-0.430***	-0.833***
	(0.120)	(0.104)	(0.180)	0.090	0.195
Illiterate	-0.840***	-0.913***	-0.433***	-0.412***	-1.072***
	(0.064)	(0.082)	(0.107)	(0.093)	(0.076)
In California	-1.066***	-1.212***	-0.443	0.118	-0.462
	(0.134)	(0.145)	(0.323)	0.122	0.317
Occupational fixed effects	Yes	No	Yes	No	Yes
R-squared	0.1712	0.1737	0.1490	0.0741	0.2107
Observations	47,824	36,046	11,778	24,868	22,956
	IV Estimates				
Dependent variable	Ln Real Prop	Ln Real Prop	Ln Real Prop	Ln Real Prop	Ln Real Prop
Sample	All	Miners and equivalent	All other	Miners and equiv. excl. farmers	All other including farmers
Age	0.481***	0.659***	0.402***	0.216***	0.656***
	(0.049)	(0.055)	0.087	0.038	0.044
Age squared	-0.0048***	-0.0063***	-0.0040***	-0.0019***	-0.0077***
	(0.0008)	(0.0009)	0.0015	0.0007	0.0007
Non-white	-1.320***	-1.571***	-0.987***	-0.434***	-0.760***
	(0.110)	(0.102)	0.248	0.104	0.205
Illiterate	-0.520***	-0.615***	0.435***	0.430***	1.052***
	(0.144)	(0.140)	0.153	0.098	0.077
In California	10.973***	8.931***	15.019***	-0.190	11.171
	(4.439)	(3.603)	7.241	0.481	8.782
Occupational fixed effects	Yes	No	Yes	No	Yes
F-test of excluded inst.	11.55	9.86	10.18	23.33	7.44
Observations	47,541	35,789	11,752	24,613	22,928

Notes: All regressions are weighed by the person weights (100 or 101 for the public-use sample outside of California, 160 or 161.6 for the public-use sample inside California and 1.6 for El Dorado, Sutter, and Yuba Counties). All standard errors are clustered on birthplace. Distance in the first stage regression is

distance from place of last residence. Only native-born men are included in the sample. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.