
Migration in the 1930s: Beyond the Dust Bowl

Myron P. Gutmann, Daniel Brown, Angela R. Cunningham,
James Dykes, Susan Hautaniemi Leonard, Jani Little,
Jeremy Mikecz, Paul W. Rhode, Seth Spielman
and Kenneth M. Sylvester

This paper analyzes in detail the role of environmental and economic shocks in the migration of the 1930s. The 1940 US Census of Population asked every inhabitant where they lived five years earlier, a unique source for understanding migration flows and networks. Earlier research documented migrant origins and destinations, but we will show how short-term and annual weather conditions at sending locations in the 1930s explain those flows, and how they operated through agricultural success. Beyond demographic data, we use data about temperature and precipitation, plus data about agricultural production from the agricultural census. The widely known migration literature for the 1930s describes an era of relatively low migration, with much of the migration that did occur radiating outward from the Dust Bowl region and the cotton South. Our work about the complete United States will provide a fuller examination of migration in this socially and economically important era.

Introduction

People on the move are an enduring image of the United States in the 1930s, from photographs (Agee and Evans 1941), literature (Steinbeck 1939), and history (Egan 2006; Gregory 1989). People *did* move in the 1930s, spurred by the economic difficulties of the Great Depression, by heat and drought, and by a multitude of other pressures. The scale of migration in the 1930s is visible in [figure 1](#), which shows the rate of out-migration from counties between 1935 and 1940, and in [table 1](#), which shows out-migration rates for states, divided into moves across county boundaries within states, and those that crossed state boundaries.¹ [Figure 1](#) also outlines the area usually recognized as the extent of the dust storm activity in the 1930s and 1940, which provides a sense of ways that weather and agricultural stress acted on the lives of US residents in this era.²

Despite the lore of the Dust Bowl and the “Great Migration” from the South to the North, the volume of internal US state-to-state migration was not all that great from the early twentieth century until after World War II (US Bureau of the Census 1946). Modest interstate migration rates belie a continuing mobility made up of streams of people moving relatively short distances from one type of community to another (Bogue et al. 1957; Ferrie 2006; Hall and Ruggles 2004; US Bureau of the Census

1. These figures only include adults aged 20 years and older and not living in group quarter in 1940, with a known county of origin based on our county definitions. See Appendix C.

2. These extents are digitized from Cunfer (2005).

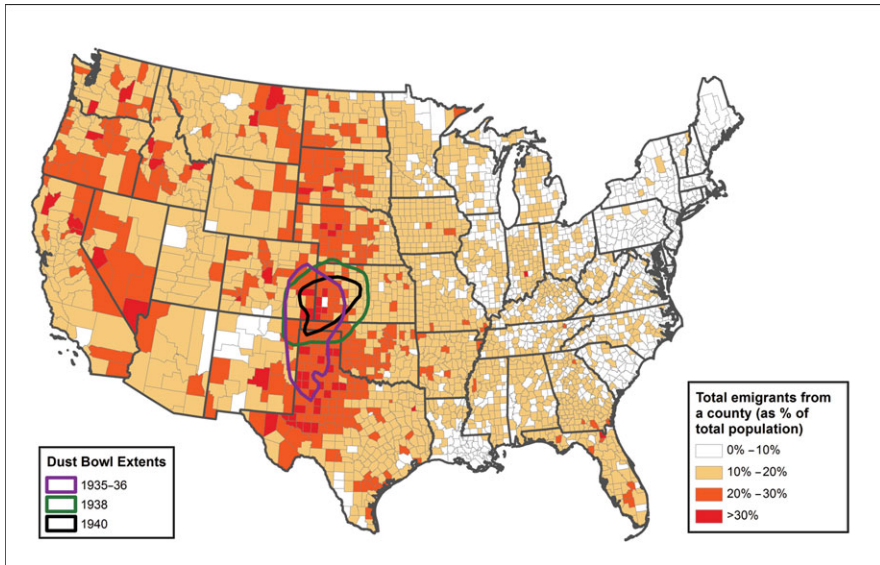


FIGURE 1. *Total emigrants from county, 1935–1940, as a percentage of total estimated 1935 county population.*

1946). In addition, as [figure 1](#) shows, areas with substantial out-migration were not limited to the Dust Bowl or the Deep South. Nonetheless, migration was a frequent subject of discussion, important enough that the US Census tracked migration for the first time in 1940, asking exactly where every person enumerated had lived five years earlier (US Bureau of the Census 2002). With this wealth of data, research about migration in the 1930s has explored many questions but left even more unanswered (Bogue and Hagood 1953; Bogue et al. 1957; Boustan et al. 2010, 2012; Fishback et al. 2006; Hornbeck 2012; Lively and Taeuber 1939; Long and Siu 2013; McLeman 2013; McLeman and Smit 2006; Tolnay et al. 2005; US Bureau of the Census 1946; White 2005; White et al. 2005).

How much do we really know about the causes of migration in the 1930s? To what extent do we know that the big forces that are supposed to have been important actually were significant? This is an important question, and the availability of census data for 1940 that shows where virtually every American lived in 1935 makes it possible to think about the factors that drove migration at a very refined scale. Moreover, the availability of data about weather, agriculture, and employment—among other factors—with comparable refinement make it possible to add real insight to our understanding. Unlike previously utilized sample or tabulated population data they show the migration experience of almost every American resident at the county level, and tell us where they went (although for this article we are only interested in whether they moved across county or state lines). Managing these detailed and complex

TABLE 1. *Estimated adult domestic migration by 1935 state of residence*

State	Migration type (% of adults in migration type)			Total
	Stayed in county	Intercounty/ Intrastate move	Interstate move	
Alabama	1,359,598 89.2	96,101 6.31	68,491 4.49	1,524,190
Arizona	196,343 83.5	14,777 6.28	24,019 10.21	235,139
Arkansas	904,154 85.21	77,218 7.28	79,705 7.51	1,061,077
California	3,560,988 88.06	367,511 9.09	115,219 2.85	4,043,718
Colorado	531,664 82.47	57,350 8.9	55,670 8.64	644,684
Connecticut	1,021,383 95.42	19,813 1.85	29,197 2.73	1,070,393
Delaware	149,723 95.22	1,742 1.11	5,774 3.67	157,239
Dist. of Columbia	341,208 91.29	0 0	32,536 8.71	373,744
Florida	885,954 88.14	73,582 7.32	45,629 4.54	1,005,165
Georgia	1,455,648 88.46	126,115 7.66	63,783 3.88	1,645,546
Idaho	223,837 82.47	23,408 8.62	24,170 8.91	271,415
Illinois	4,715,569 92.24	177,628 3.47	218,927 4.28	5,112,124
Indiana	1,906,166 90.33	125,113 5.93	78,931 3.74	2,110,210
Iowa	1,339,286 86.47	118,863 7.67	90,633 5.85	1,548,782
Kansas	952,265 82.99	85,748 7.47	109,383 9.53	1,147,396
Kentucky	1,438,411 90.44	80,828 5.08	71,302 4.48	1,590,541
Louisiana	1,200,842 91.84	71,176 5.44	35,566 2.72	1,307,584
Maine	463,464 92.96	21,564 4.33	13,541 2.72	498,569
Maryland	995,452 94.15	29,203 2.76	32,617 3.09	1,057,272
Massachusetts	2,532,894 94.2	84,696 3.15	71,260 2.65	2,688,850
Michigan	2,886,555 91.22	190,914 6.03	86,849 2.74	3,164,318
Minnesota	1,518,284 89.34	110,082 6.48	71,136 4.19	1,699,502
Mississippi	989,627 87.97	95,532 8.49	39,849 3.54	1,125,008
Missouri	2,108,622 88.15	127,600 5.33	155,807 6.51	2,392,029
Montana	269,676 82.69	29,275 8.98	27,170 8.33	326,121
Nebraska	702,010 81.8	67,731 7.89	88,426 10.3	858,167
Nevada	45,020 78.71	4,092 7.15	8,082 14.13	57,194

TABLE 1. *Continued*

<i>State</i>	<i>Migration type</i> (% of adults in migration type)			<i>Total</i>
	<i>Stayed in county</i>	<i>Intercounty/ Intrastate move</i>	<i>Interstate move</i>	
New Hampshire	270,340 <i>94.17</i>	5,937 <i>2.07</i>	10,794 <i>3.76</i>	287,071
New Jersey	2,425,833 <i>93.15</i>	98,628 <i>3.79</i>	79,633 <i>3.06</i>	2,604,094
New Mexico	204,125 <i>85.77</i>	14,723 <i>6.19</i>	19,136 <i>8.04</i>	237,984
New York	8,114,585 <i>94.7</i>	225,695 <i>2.63</i>	228,054 <i>2.66</i>	8,568,334
North Carolina	1,639,261 <i>91.25</i>	112,437 <i>6.26</i>	44,755 <i>2.49</i>	1,796,453
North Dakota	313,584 <i>82.79</i>	23,175 <i>6.12</i>	42,022 <i>11.09</i>	378,781
Ohio	4,018,762 <i>91.95</i>	209,068 <i>4.78</i>	142,849 <i>3.27</i>	4,370,679
Oklahoma	1,065,432 <i>79.87</i>	127,046 <i>9.52</i>	141,546 <i>10.61</i>	1,334,024
Oregon	528,513 <i>82.27</i>	68,990 <i>10.74</i>	44,908 <i>6.99</i>	642,411
Pennsylvania	5,789,001 <i>94.17</i>	197,106 <i>3.21</i>	161,113 <i>2.62</i>	6,147,220
Rhode Island	421,421 <i>96.09</i>	4,666 <i>1.06</i>	12,498 <i>2.85</i>	438,585
South Carolina	871,526 <i>91.5</i>	48,602 <i>5.1</i>	32,331 <i>3.39</i>	952,459
South Dakota	317,760 <i>81.26</i>	31,049 <i>7.94</i>	42,232 <i>10.8</i>	391,041
Tennessee	1,480,037 <i>90.09</i>	86,456 <i>5.26</i>	76,406 <i>4.65</i>	1,642,899
Texas	3,075,063 <i>83.69</i>	464,807 <i>12.65</i>	134,326 <i>3.66</i>	3,674,196
Utah	265,103 <i>86.92</i>	17,161 <i>5.63</i>	22,743 <i>7.46</i>	305,007
Vermont	187,575 <i>91.55</i>	8,509 <i>4.15</i>	8,812 <i>4.3</i>	204,896
Virginia	1,326,306 <i>92.3</i>	64,306 <i>4.48</i>	46,343 <i>3.23</i>	1,436,955
Washington	904,744 <i>86.05</i>	86,388 <i>8.22</i>	60,278 <i>5.73</i>	1,051,410
West Virginia	932,077 <i>90.86</i>	54,051 <i>5.27</i>	39,748 <i>3.87</i>	1,025,876
Wisconsin	1,780,784 <i>91.59</i>	101,204 <i>5.2</i>	62,391 <i>3.21</i>	1,944,379
Wyoming	113,452 <i>83.06</i>	8,660 <i>6.34</i>	14,477 <i>10.6</i>	136,589
Total	70,739,927 <i>90.36</i>	4,336,326 <i>5.54</i>	3,211,067 <i>4.10</i>	78,287,320

data—there were more than 130 million American inhabitants in 1940—created significant challenges, and describing how we overcame those challenges constitutes a significant part of this article. However, our findings are important as well.

To a great extent, the conventional story of the 1930s is right but too limiting: People left areas where the weather was challenging for agriculture in the mid-1930s, but that experience was not limited to the Dust Bowl region. The weather was hot and dry in a much larger part of the United States, and migrants escaped those areas as well. Moreover, much more was happening. There were regional processes that intensified the agriculture and weather effects, and broader economic processes that are predictable. Our research confirms much that we should have known all along, but with much more detailed data that give us confidence that we understand what was happening.

Theoretical Background

A large theoretical literature (Greenwood 1997; Lee 1966; Massey 1990; Massey et al. 1998; Ravenstein 1885, 1889; Roy 1951) has sought to explain migration. Though much of it focuses on international migration, its core elements have also been applied to internal migration. Neoclassical economics provides the dominant perspective, informed by contributions from the New Economics of Migration, by Massey's notion of Cumulative Causation Theory and by the descriptive richness of Migration Systems Theory (Bakewell 2013; Fussell et al. 2014; Massey 1990; Massey, et al. 1993). In these approaches, a person's likelihood of migrating is a function of not only his or her individual characteristics but also of the characteristics of their places of origin and destination, including the distribution of income, land, and human capital; the organization of agriculture and industry; public policy; and cultural frameworks, reflected in local ethnic, religious, and racial conditions (Fishback et al. 2006; Massey et al. 1993). The structure of social networks shapes migration. Because these processes are not independent of spatial context, we see migration not merely as people moving in unconnected and location-free contexts but as interactions between people and locations following specific pathways.

Most research sees migration as a tension between pushes and pulls, and theorizes that migration serves as a mechanism to restore equilibrium between competing forces. Disaster-related demographic theory builds on these patterns of movement to identify the migratory systems that existed prior to the disaster and to gauge whether a given shock transforms the preexisting migration system (Black et al. 2011; Fussell et al. 2014; McLeman 2013). The disaster literature also adds the concept of vulnerability to the determinants of migration (Adger 2006; McLeman 2013; McLeman and Smit 2006). The vulnerability paradigm focuses on the exposure of people to stress over time, prior to the period of crisis, based on the condition of the economic and ecological systems they inhabit. Assessments of vulnerability have focused on both short-term moves and long-term reorganizations of human-environment systems (Adger 2006; Berkes et al. 1998; Black et al. 2013), asking whether a specific "demographic signature of disaster" exists (DeWaard et al. 2014). In the United States, the Dust Bowl story has driven research about environmental migration in the 1930s, yet drought and land degradation were not limited to the southern plains. Extreme heat

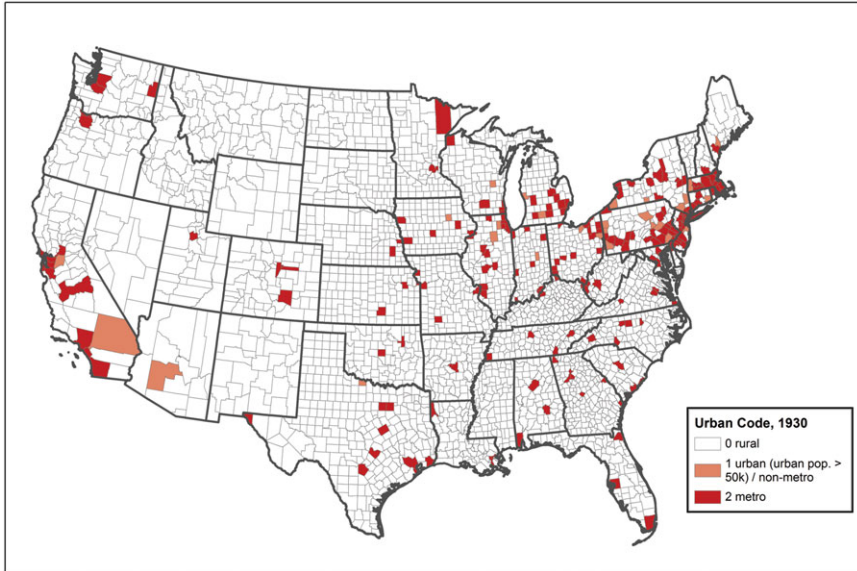


FIGURE 2. Counties by urban code, 1930.

also led to agricultural stress elsewhere (Giesen 2011; Gregory 2005; McEwan et al. 2014; Olmstead and Rhode 2008).

Our analysis is not limited to environmentally driven agricultural shocks. The United States was in the midst of a depression in the 1930s, with high unemployment, low wages, and a weak economic recovery. We also ask whether people left places with comparatively poor economic conditions, or stayed in places with relatively good economic conditions, and did so in a way that reveals the characteristics of places as drivers of migration. The United States still had a significant rural population in the 1930s, with more than 40 percent of its inhabitants living in places of less than 2,500 residents in both 1930 and 1940 (US Bureau of the Census 2012). The West, South, and Midwest were significantly more rural, while the Northeast was significantly less rural. Farming was still a major industry in many parts of the country, and weather had a major impact on agriculture and, consequently, the rural economy. The limits of urbanization are visible in figure 2, which shows where all the counties that had an urban population of 50,000 or more were located. The map is sparse.

For the analysis reported here, we focus on the attributes of counties rather than the attributes of individuals. We are interested in how likely it was for someone living in a given US county in 1935 to migrate to another county by the time of the 1940 Census, and the attributes of counties that made them more or less likely to send their inhabitants elsewhere by 1940. We realize that this is a substantial simplification, because the migration literature is heavily focused on the idea first raised by Roy (1951) that migrants self-select for upward mobility. Moreover, Borjas

(1987) expanded selection theory by arguing that it is the more highly skilled who migrate, and Kanbur and Rapoport (2005) argue that selectivity by education is key. While we recognize that individual attributes are as important as those of context in determining who migrates, from which communities, and where they go, our approach is an important first step in understanding how migration operated in this era, and how county characteristics shaped out-migration flows.

Understanding Relative Levels of Out-migration

The main questions we explore revolve around the role of the environment and economy in encouraging or discouraging out-migration from counties between 1935 and 1940. We discuss the data that we rely on, and questions raised by the nature of those data, in a later section. We begin by discussing the factors that may have led one county to experience more out-migration than another. Our focus is mostly on processes that are important for the less urbanized parts of the United States, where natural phenomena, such as precipitation and temperature, may have had a strong effect, either acting on their own or acting through agriculture. We also include measures related to employment by industry and unemployment, which we hypothesize play a role in determining migration even in urban areas. While other researchers have examined the role of New Deal support programs in explaining migration (Fishback et al. 2006), our preliminary analysis suggested that they were less important than other factors, and they are not included in our statistical models.

Given the severe drought of the mid-1930s, which has been described recently as the worst drought of the last millennium (Cook et al. 2014), we begin with an examination of the role of weather in influencing levels of out-migration from US counties. We measure weather by looking at annual total precipitation for various time periods, and average daily maximum temperature as a percentage of a longer term (1920–40) average. The drought was worst in 1934 (and to some extent similarly severe in 1933 and 1935), so we have created relative weather measures for 1934, 1933–35, and 1933–39. We show the scale of the drought for these time periods in figures 3 and 4. Both temperature (higher than normal) and precipitation (lower than normal) diverged most significantly from expected patterns in 1934, somewhat less so in 1933–35, and came closer to normal for the seven-year period from 1933 to 1939. What is also clear is that the spatial patterns were rather different, with the highest sustained temperatures in the front range of the Rocky Mountains and the northwestern Great Plains (see especially figure 4.b), and the lowest sustained precipitation in the classic Dust Bowl areas in Oklahoma, Texas, Kansas, New Mexico, and Colorado, plus parts of the northern plains (figure 3.c), and to a lesser extent in the intermountain west (California, Nevada, Utah, Wyoming, and Idaho; figure 3.b).

We utilize a variety of ways to measure changes in agriculture during the 1930s, making use of data from the 1930, 1935, and 1940 censuses of agriculture, which represent agricultural results in 1929, 1934, and 1939, respectively (Haines et al. 2014). The agricultural census includes one direct measure, the percent of land with failed

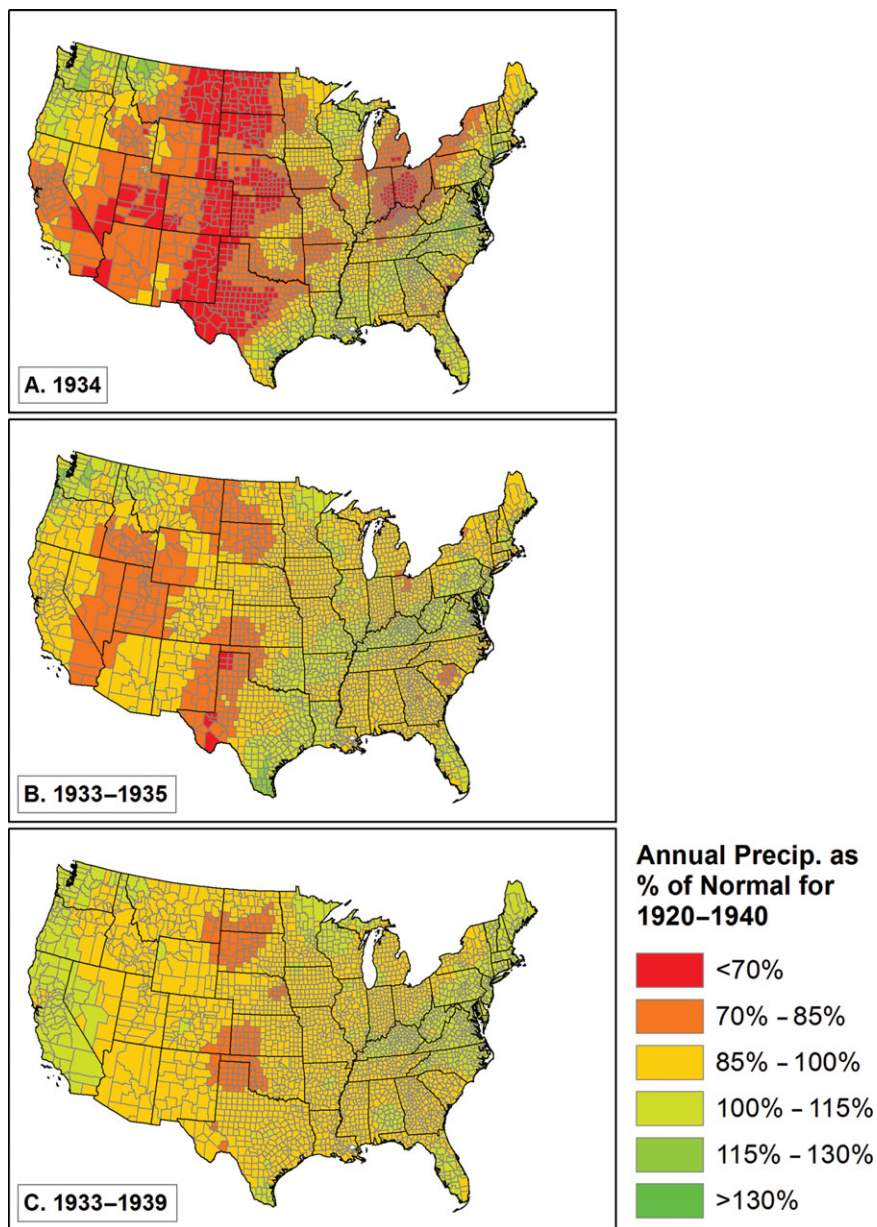


FIGURE 3. Annual precipitation as a percentage of the 1920 to 1940 average.

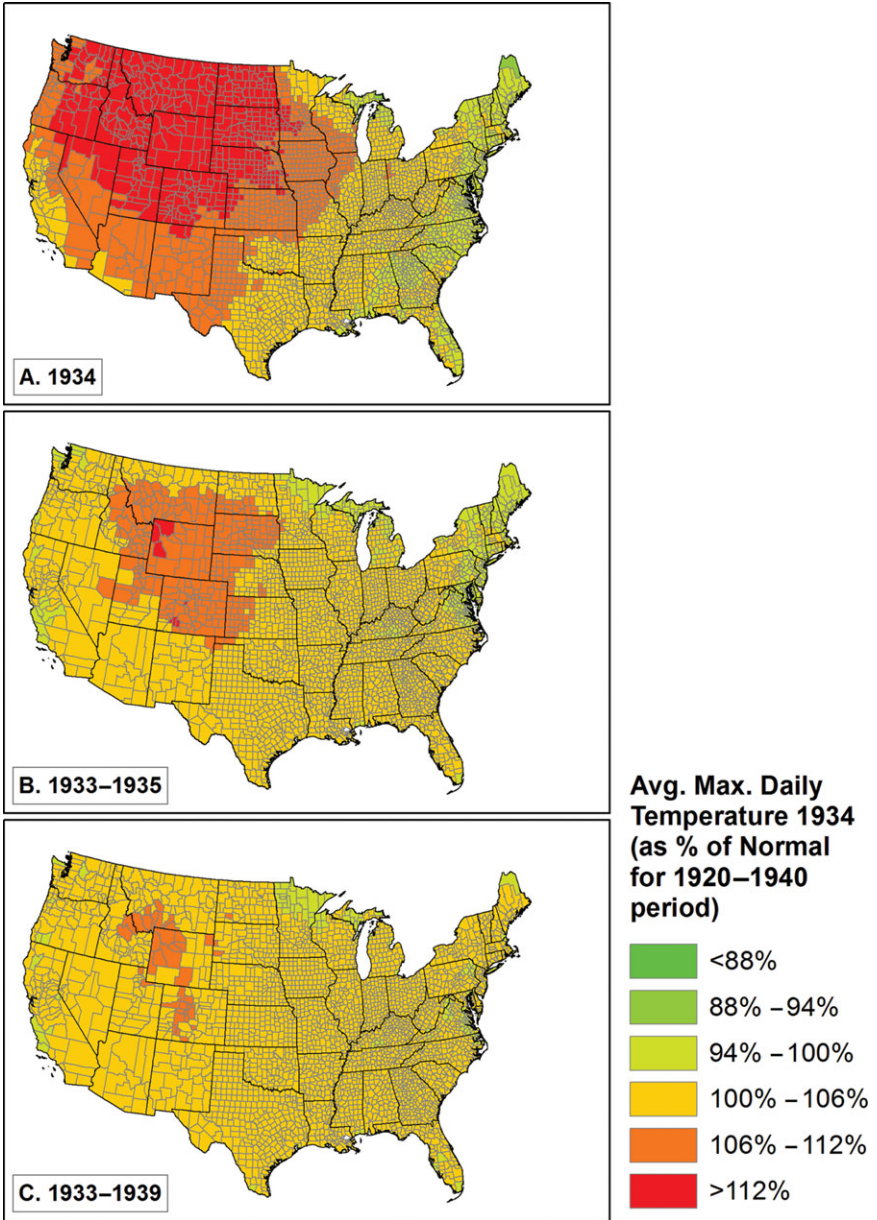


FIGURE 4. Annual daily maximum temperature as a percentage of the 1920 to 1940 average.

crops. We display this in [figure 5](#). The two panels of the figure show the consequences of the most severe weather, with significant failure levels throughout the central United States in 1934, and in areas of Kansas, Colorado, Arizona, New Mexico, and South Dakota (and spots elsewhere) in 1939.

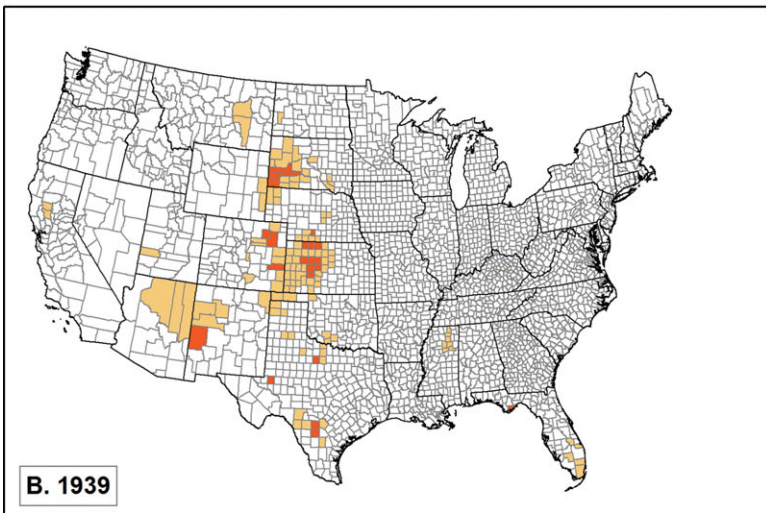
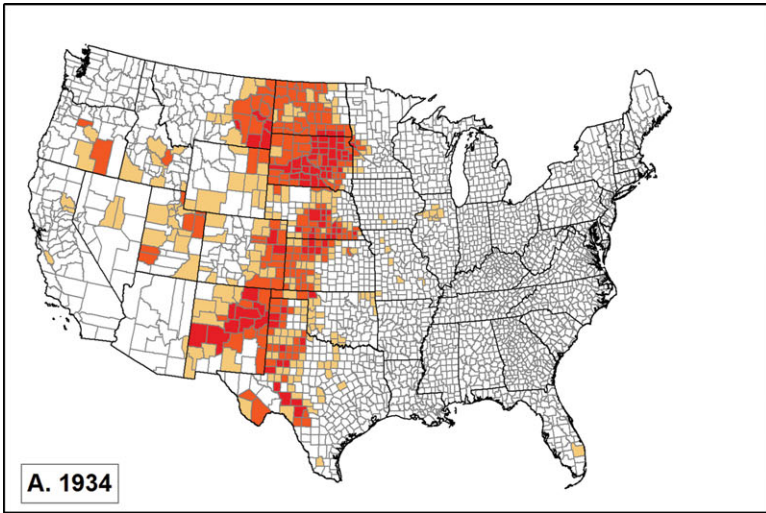
In order to attempt to find new ways to gauge the impact of the weather on agricultural production, we developed three other measures. The most ambitious of our measures estimates the percent change in crop production from 1929 to 1934 and from 1929 to 1939, for each county's three largest crops as indicated in the 1930 Census. We describe the methods we used to derive these estimates in Appendix A. This measure captures the overall falloff in production, including that due to farmers not having the resources to plant or believing that the crop would fail anyway. As we show in [figure 6](#), a very large portion of the United States experienced major production shortfalls in 1934, something that still had not been reversed in 1939. The spatial pattern changes between 1934 and 1939, with a greater falloff in production in the later period in Georgia, Mississippi, Alabama, and Kentucky, plus New England and New York, and less in the corn belt states (Illinois, Iowa, northern Missouri), and the upper Midwest (Wisconsin, Minnesota).

We also experimented with two other measures. In one, we made the same calculation we did for the three largest crops, and instead focused only on corn, wheat, and cotton. In a second, we attempted to gauge the impact of severe weather and poor agricultural conditions on livestock by estimating livestock inventories in 1935 and 1940 as a percentage of what they were in 1930 ([figure 7b](#)).³ The results shown in [figure 7](#) are interesting because they show that, in most parts of the United States, livestock had increased between 1930 and 1935 ([figure 7.a](#)), with some exceptions in areas of the central United States with the worst weather. The situation worsened between 1935 and 1940 ([figure 7.b](#)), but livestock nonetheless continued to increase in numbers in most of the country.

Although we hypothesize that changes in agriculture drove most of the out-migration in the 1930s, other economic factors played a role. One of these is unemployment, which was enumerated in a special census of "partial employment, unemployment, and occupations," in 1937 (Biggers and United States 1938).⁴ [Figure 8](#)

3. We chose cattle and swine for our analysis of livestock because these data are most suitable for comparison across census years. The data for other stock animals made comparison more difficult. Work animals such as horses, donkeys, and mules had been in steady decline for some time due to the mechanization of farms; thus, they are not reliable indicators of the health of farms. Livestock units are calculated using ratios developed by the FAO, which counts each cattle as one unit and each swine as one-quarter of a unit (United Nations Food and Agriculture Organization [FAO] 2011). Livestock numbers are inflated for the census year of 1935 in comparison to the other two years (1930 and 1940). Due to the collection of data earlier in the year and slightly different questions posed to farmers for 1935, this year's data does not include those animals that died or were slaughtered between January 1 and April 1 of that year. Thus, the top map shows a greater increase (or lesser decrease) in livestock than that actually experienced between 1930 and 1935.

4. The data are available in digital format in Haines and Inter-university Consortium for Political and Social Research (2010). The number of unemployed is the number of persons 15–74 years of age who reported being totally unemployed but wanting and able to work on the Unemployment Report Card. This does not include emergency workers. Farmers and farm laborers were included; cards were given to each



Crop Failure (%)

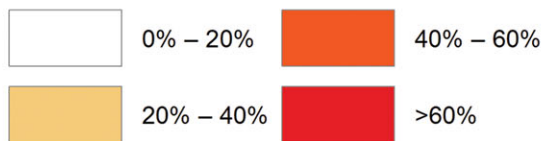
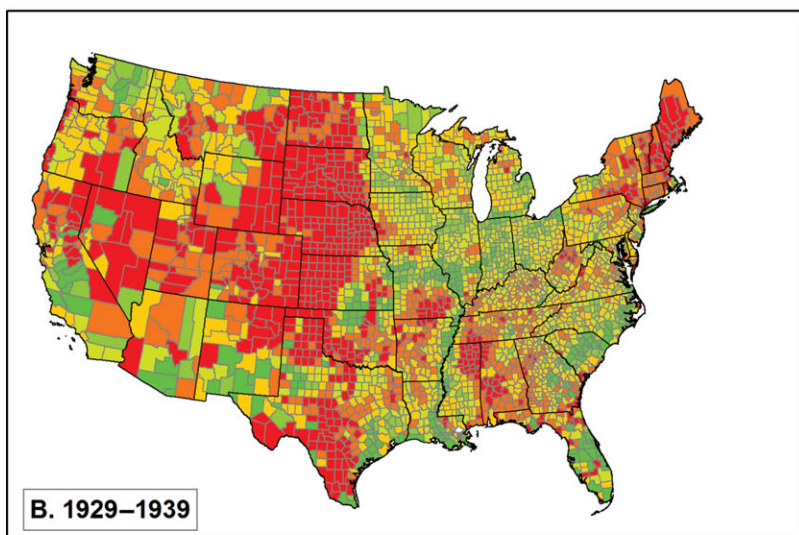
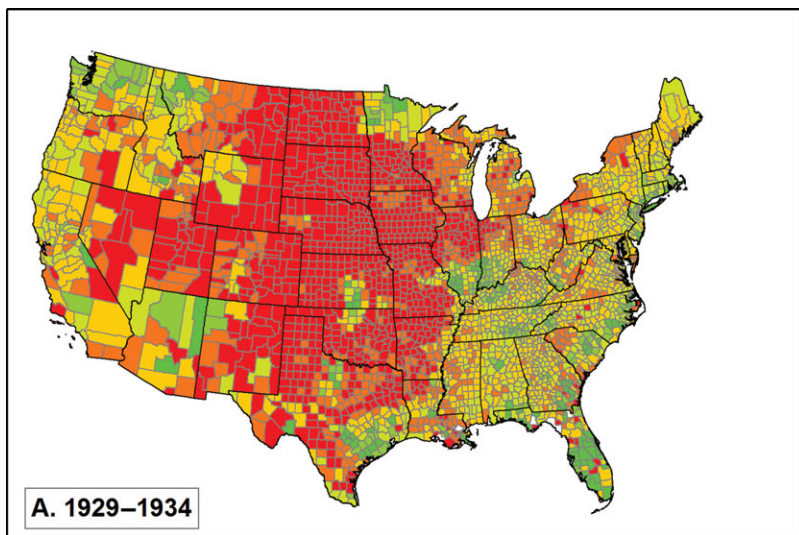


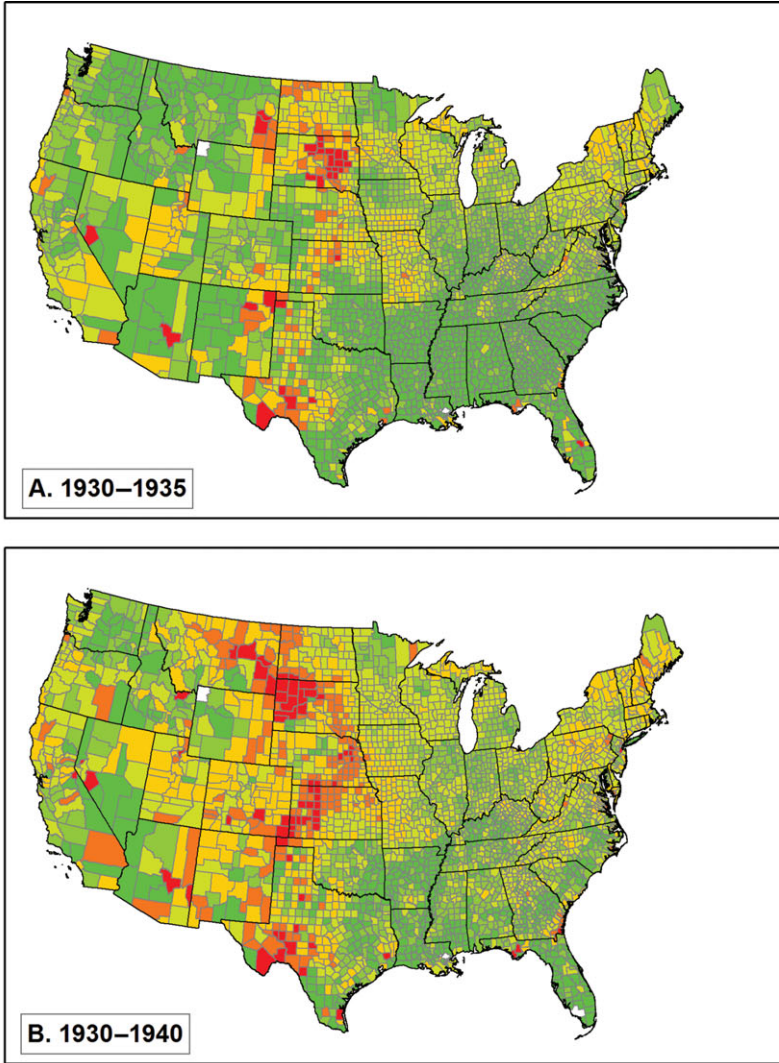
FIGURE 5. *Crop failure acreage as percentage of total crop acreage.*



Production of Top 3 Crops: Weighted Percent Change



FIGURE 6. *Weighted percent change in county's three top crops.*



Cattle and Swine: Percent Change in Livestock Units



FIGURE 7. *Percent change in cattle and swine by livestock unit (1930–35 and 1930–40).*

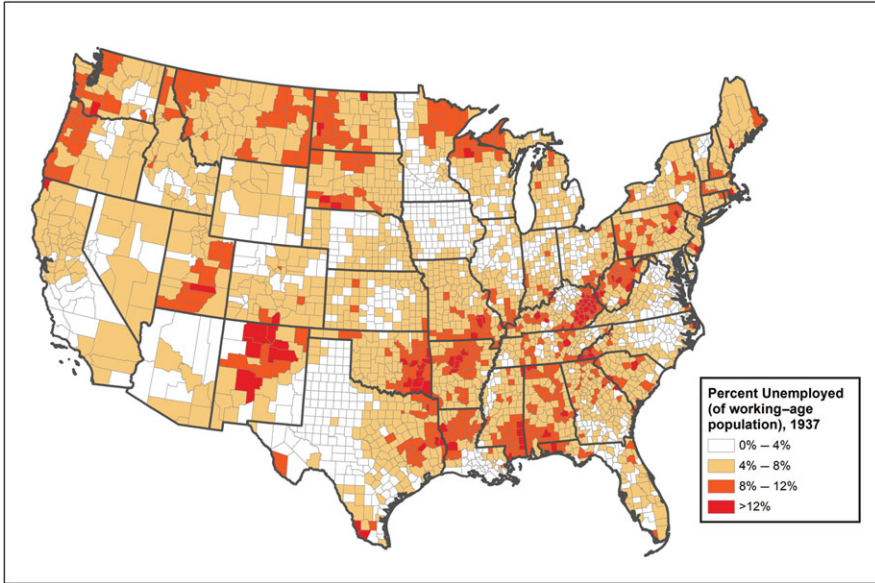


FIGURE 8. *Percent of working age population unemployed, 1937.*

displays the spatial distribution of unemployment, with the highest levels of unemployment in the Deep South, Appalachia, in the northeast, along the northern tier of the United States, and in Utah and New Mexico. As we might predict this does not appear to align easily with either weather or agriculture. We believe the economy played a broader role, so we have also examined employment in various industries as a potential indicator that the economy was capable of doing better or worse in different areas of the United States, with an impact on migration.

Our hypotheses are simple and straightforward. We expect that rural counties that experienced severe weather, poor agricultural results, and high unemployment should have had higher levels of out-migration than those that did not, and that urban counties and counties with higher levels of employment in manufacturing and retail sales should have had lower levels of out-migration, all other things being equal.

Data and Methods

The variety of data necessary for this analysis, originally gathered at different spatial and temporal scales, were transformed in order to produce a data set that may be

household, but there does not seem to have been any follow-up to ensure compliance. In order to estimate an unemployment rate in 1937, we estimated a total working-age population for each county by summing the population aged 15 and over in 1930 and 1940, and interpolating to a 1937 population (70 percent of the difference).

analyzed at the county level. Our dependent variable is the rate of out-migration from each county of the United States between 1935 and 1940, based on the data in the 1940 US Census of Population. The independent variables draw on environmental, agricultural, and economic data. The methods we use include both descriptive and multivariate approaches.

The list of counties of the United States has changed over time, even during as brief a period as the decade from 1930 to 1940. For this analysis, we began with the 1940 list of counties and their geography, and modified that geography to take into account changes in the list of counties and the ways that we and others have aggregated counties to optimize analysis. First, there are counties that existed in 1930 but not in 1940 (or vice versa). Campbell and Milton counties were merged into Fulton County, Georgia, between 1930 and 1940; we combined these three counties into one. In another case, we aggregated spatial units in Virginia in order to combine independent cities with their surrounding counties.⁵ In a third modification, we needed to merge counties from a single metropolitan area where confusing naming practices made it impossible to distinguish separate counties (New York City's five counties and Saint Louis City and Saint Louis County, Missouri). These modifications are summarized in Appendix B. The resulting data set contains 3,069 counties from the contiguous 48 states that existed in 1940. In the cases in which we combined counties and the data were counts, we summed the counts across all geographic units. In the case of data that were rates or averages, we calculated means, which were spatially weighted when appropriate.

Census Data

The 1940 US Census of Population full count data have a wealth of information about individuals, including demographic, social and cultural, economic, and location data (Ruggles et al. 2010), and constitute the primary data source for our analysis. In the 1940 Census, respondents were asked to provide their county of residence and that of all their household members in 1935. At the time we undertook this analysis, the full-count 1940 data were available in two forms, and we made use of both of them. The University of Minnesota's Integrated Public Use Microdata Series (IPUMS) project has released a preliminary version of the 1940 data in IPUMS coded format, which includes every person in the United States but not every variable from the census. The IPUMS version of the 1940 data set includes standardized variables for a person's 1935 state of residence and the scale of their residential movements since 1935, delineating those who moved within or between counties, states, and countries, but the coded data do not yet contain a variable for *which* county a person lived in in 1935. That information is available from a restricted release of a version of the 1940 data that does include the detailed response text for every person. We have merged

5. Following the example of Fishback et al. (2003), we combined independent cities with their surrounding counties to ensure consistent and useful data. In doing so, we assigned new FIPS codes to these aggregated units.

these data sets, and coded—to the extent possible—origin county from the 1935 textual residence variable, in addition to recoding 1940 places of residence that were erroneously coded in the IPUMS data (e.g., coding that conflated Brown and Boone counties in Indiana, and Richmond City and Richmond County in Virginia). That turned out to be a challenging task, which we describe in Appendix C.

Table 1 presents a subset of the data, tabulated by state of residence in 1935. This allows us to show the level of out-migration from counties within states, for migrants who stayed within their state of residence (“intercounty move”) and those who left their state of residence (“interstate move”). We exclude people under age 20 in 1940, people living in group quarters in 1940, those living outside the United States in 1935, and those whose residence in 1935 is unknown. For the contiguous United States as a whole, using our data set, 5.5 percent of the population moved across county boundaries (but stayed in state),⁶ and 4.1 percent moved across state lines, a total of 9.6 percent. Variation from state to state is substantial, as we would expect.

Weather and Climate Data

We derived our climatological variables from data sets developed by the PRISM Climate Group at Oregon State University. These PRISM data are 4km-grid rasters of temperature and precipitation, modeled at a monthly resolution, stretching back to the late nineteenth century, and covering the contiguous United States. Assuming that physiographic factors such as elevation and aspect have influenced local climate in a similar way in both the more recent and more distant past, the PRISM group incorporated a “climate fingerprint” from 30-year normals for 1971–2000 into their expert system to fill in the gaps between scattered weather monitoring stations for earlier years. Further details on PRISM’s method of topographically informed interpolation may be found in the group’s publications and web documentation (Daly et al. 2002, 2008). For our analysis, we aggregated the PRISM data to counties by calculating zonal statistics, using 1940 county boundaries from the National Historical Geographical Information System, and aligned with the PRISM rasters.⁷ In order to do this we averaged the grid cell values within each county to calculate the maximum and minimum temperatures and the total precipitation.⁸

6. We treat moves within New York City but across borough lines as being within the same county, despite the fact that the five New York City boroughs are each a separate county.

7. For counties that were too small to contain even a single 4km-by-4km raster cell and for which zonal statistics could not be calculated, temperature and precipitation values were obtained by creating a centroid for each county and intersecting that centroid with the raster data.

8. We employed a combination of ArcGIS’s *arcpy*, the statistical package R, and Beyer’s Geospatial Modeling Environment (GME; <http://www.spatial ecology.com/gme/index.htm>), software that acts as a go-between for the two programming languages. For instance, GME allowed us to process our data county by county, thus circumventing some of the known issues with the way ArcGIS performs its rasterization and statistical processes.

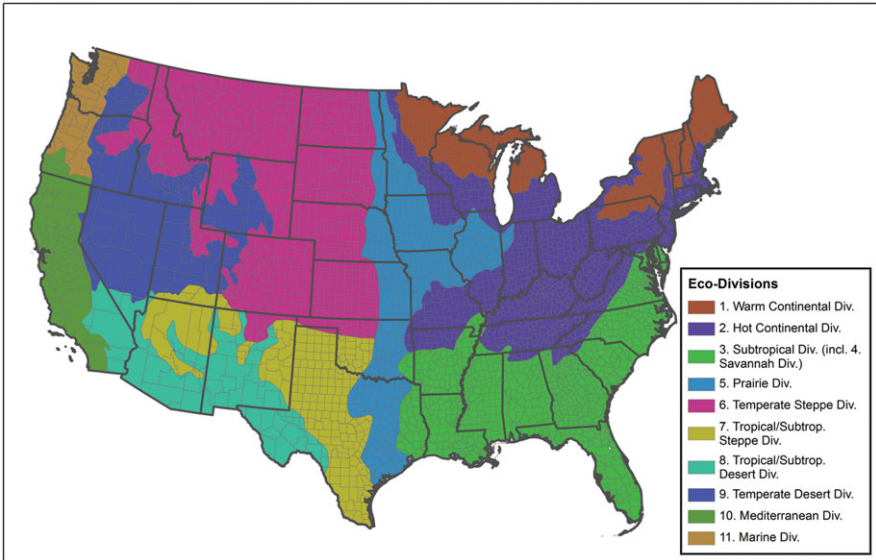


FIGURE 9. *Ecodivisions as defined by a United States Forest Service report in 1997.*

Urban and Regional Status

We recognize that there is regional variation in how demographic, economic, agricultural, and environmental characteristics affected the migration patterns of the 1930s. In order to understand this variation, we decided to classify counties into broader groups, using agro-ecological categories defined by the US Forest Service in 1997 (Bailey 1997).⁹ Bailey's report divides the United States into a hierarchical system of ecoregions. These levels of ecoregions include 4 climatic domains, 15 divisions, and 63 provinces. Each of these regional classifications was developed at an increasing level of climatic precision. The division level has the appropriate amount of resolution for our analysis, with 11 divisions for the continental United States, and an additional nine mountain regime subdivisions. To simplify our analysis, we have joined these mountain regimes with their climatic lowland counterpart (i.e., "Subtropical Regime Mountains" is joined with "Subtropical Division") (see figure 9). We also merged division 4 (Savannah) with division 3 (Subtropical), both of which represent parts of the far southeastern United States, because division 4 only contains four counties at the southern tip of the state of Florida.

Along with regional variation, we assume that the dominance of agricultural factors should be less in areas where agriculture plays a smaller role in the economy. One way to do that is to identify urban areas (figure 2). Our starting point for this was a set

9. This report is a revised and updated version of a map published in 1981 by the US Fish and Wildlife Service (Bailey and Cushwa 1981).

of historical classifications of metropolitan statistical areas available at the Minnesota Population Center's IPUMS website ("County Composition," in Ruggles et al. 2010). We created an urban scale variable from the metropolitan area data for 1930. This variable has three values: rural/nonmetro, urban/nonmetro (all nonmetro counties with an urban population greater than 50,000 in 1930), and urban/metro, which are counties classified as metropolitan in 1930, which we further aggregated to two values (rural vs. urban) for our analysis.

Methods

At this stage in our research, most of our data analysis has been descriptive and visual, using maps to show the spatial patterns of migration alongside other spatial patterns—weather, agricultural production and failure, and unemployment. The results, as we will show, are striking, within the usual constraints that it is very difficult to understand correlation or causation visually. What looks like a striking relationship might just be something that catches our eye.

In order to provide a more conventional statistical presentation, we have taken a variety of directions. We begin with ordinary least squares (OLS) regression models with county as the unit of analysis and the natural log of the number of adults (aged 15 or over in 1935) who were resident in the county in 1935 and moved outside the county by 1940 as the dependent variable. All models include the log of the estimated base 1935 population (those who moved plus those who did not) among the explanatory variables. Each model presented explores different combinations of potential explanatory variables. Because there is reason to believe that the relationships of our explanatory variables to our dependent variable vary in different parts of the United States, we also explore the role of ecological divisions in the models. We first consider ecological divisions as a set of binary independent variables. In these OLS models, we find significant heteroskedasticity and a handful of outliers, so we also estimate the models using robust regression methods. While the robust regression models we use might have addressed these issues, in our models it did not do so, as indicated by the Breusch-Pagan *p*-values in tables 3, 4, and 5.

When residuals from adjacent counties are correlated, that is, spatial autocorrelation, standard OLS estimates of standard errors are artificially small and the goodness of fit measures and chances of finding statistical significance are inflated. Spatial autocorrelation can be measured with Moran's *I*, and our results show that there is significant spatial autocorrelation in the standard OLS models. It is not possible to calculate Moran's *I* for the Robust Regression techniques we use, but we suspect that there are spatial effects present.

The third approach we use is to consider ecological divisions as characteristics that interact with all the other variables, and use a method that allows those interactions to uncover how the effects of the other independent variables may depend upon ecological context. Following Anselin (2007) we refer to the latter as "regime" models because each ecodivision constitutes a separate spatial regime with its own set of regression coefficients. Despite the advantages of the regime models, treating our

ecological divisions as different spatial regimes still does not account fully for the spatial effects in the system. This requires a spatial regression approach to reveal the nature and extent of the effects from neighboring counties. Anselin's (2007) decision tree is the accepted method for choosing between alternative specifications of spatial effects, leading us to choose the "spatial error" model (as opposed to a spatial lag or combined model), because it outperformed other spatial regression and OLS models in our tests. A spatial error model specifies that the unexplained out-migration in a focal county is directly affected by the residual out-migration in the surrounding counties. Regression analyses were largely performed in R, and work that explored patterns in the residuals for different models was conducted in GeoDa (Anselin et al. 2006).

Results

We begin our discussion of results with a visual presentation of adult out-migration by county of residence in 1935 (the results are tabulated by state in [table 1](#)). These results are presented in [figure 1](#). These are data where the numerator is the number of people known to live in a specific county in 1935 and who had left by 1940. The denominator for these computations is the sum of all adults (over age 15 in 1935) whose residence was known in 1935 and was within the United States. All individuals living in group quarters in 1940 or living outside the United States in 1935 are excluded.

People who lived in the western United States in 1935 appear more likely to have moved in the next five years than those in the east (and especially New England and the Middle Atlantic states), with the greatest likelihood of out-migration in two north-south bands, one from western Texas and New Mexico north to the Dakotas and Montana (roughly what we consider the Great Plains), and the second away from the coasts in states on the western edge of the United States, especially Arizona, Nevada, California, Idaho, Oregon, and Washington.

What causes these patterns? Certainly, looking back to [figure 3.a](#) (temperature in 1934) and [figure 4.a](#) (precipitation in 1934), we see possible connections. There is a great deal of out-migration from those places that were hottest and driest in the worst year of the drought. Is this a real relationship? How does it work? Is it a direct connection, or one that works through agriculture? What is the relationship between migration and other factors, such as unemployment, the extent to which the county's population is engaged in farming, or New Deal public programs?

It is possible to quantify the relationships we see with a multivariate statistical analysis. We estimated a series of multivariate OLS models, with various combinations of independent variables. We estimate multivariate models for two families of regressions, one ("Crop Failure") where the main independent variable is the level of crop failure, and the other ("Climate"), where the main independent variables are precipitation and temperature. We began by estimating univariate regressions between the amount of migration and each of the potential independent variables. For the crop failure model, we achieved stronger associations by dividing the values into four categories (as opposed to a single continuous variable), including 0 percent failure as

the reference category. For the climate models, the data for 1934 as a percentage of 1920–40 averages was most predictive of migration. We also include the weighted percent change in production of a county's three top crops as reported in 1930 and 1940 in these models. This measure was more predictive of migration than the other measures, such as the change from 1930 to 1935, or 1935 to 1940, or changes in livestock. In both the crop failure and climate models we include per capita retail and manufacturing employment in 1930, the percent of the working age population unemployed in 1937, and a binary category delineating whether a county is urban (having an urban population in excess of 50,000). For the ecodivisions, we chose the warm continental region in the northeastern United States as the reference category; ecodivisions were added before the other independent variables (except for the log of the total population) when incorporated into models as a main effect.

We hypothesize that there are attributes of counties that encourage or discourage migration, and that these are generally linked to their impact on livelihood. At the core of our analysis are agriculture and the climate forces that shape it, such as drought. We hypothesize that poor agricultural conditions lead to greater out-migration from a county, and that better agricultural conditions are associated with less out-migration. The same holds for unemployment, with greater unemployment associated with higher out-migration. On the other side, we hypothesize that economic activities outside of agriculture will be protective and associated with less out-migration. In that category, we have employment in retail sales and manufacturing, plus the urban status of the county.

The first and most basic of these results are included in [table 2](#). We begin with a model that just includes the natural log of the baseline population (the denominator in a migration rate) on the right side of the equation. We then add the ecodivisions to the model to show regional effects. All regions (except the Hot Continental division) are significantly different in their level of migration from the Warm Continental reference category. Looking back to [figures 3.a](#) and [4.a](#), we see that the largest coefficients are in areas with the hottest and driest weather in 1934, with the complication that heat and drought did not overlap perfectly—it was hot but not especially dry in the Pacific northwest, and dry but not especially hot in California.

In [table 3](#) we show three versions of a model in which the main independent variable is crop failure, divided into four categories (zero, 1 to 5 percent, 5 to 25 percent, and 25 percent and over). We present two versions of the crop failure OLS model, one with and one without the ecodivisions, as well as a model using robust regression. We present the results as odds ratios. Overall, the OLS model with the ecodivision regions shows a better fit than the one without the ecodivisions, and the robust regression model has an r-squared that is still larger. The results in [table 3](#) largely confirm our assumptions about how environmental and economic stress contributed to migration flows. As expected, higher levels of crop failure led to more migration, as did higher unemployment. By contrast, counties with more manufacturing employment and counties with large urban populations had slightly lower (and not always significantly different) levels of out-migration. The one result that does not necessarily confirm our assumptions is the percent of the population employed in retail, often suggested as an indicator of overall economic activity. Our results show that higher

TABLE 2. OLS models of county out-migration with population and ecodivisions only

		Population only	Population and ecodivisions
DIAGNOSTICS	R squared	0.7952	0.8540
	Adj. R squared	0.7952	0.8535
	F statistic	1.192e + 04 on 1 and 3069 DF	1790 on 10 and 3060 DF
	Shapiro test W value	0.8670	0.8052
	Breusch-Pagan p value	0.0900	0.0049
	VIF values > 3		
	Moran's I (if significant)	0.4847	0.3198
INDEPENDENT VARIABLES	Population	Odds (sig.) 2.2305***	Odds (sig.) 2.4131***
	Ecodivisions (<i>ref = warm continental</i>)		
	<i>Hot continental</i>		1.0534 [‡]
	<i>Subtropical</i>		1.1134 ***
	<i>Prairie</i>		1.5662***
	<i>Temperate steppe</i>		1.8560***
	<i>Subtropical and tropical steppe</i>		2.0443***
	<i>Subtropical and tropical desert</i>		1.8300***
	<i>Temperate desert</i>		1.7112***
	<i>Mediterranean</i>		1.9235***
	<i>Marine</i>		1.8707***

Significance codes: 0: ***; 0.001: **; 0.01: *; 0.05: [‡].

levels of employment in retail sales are associated with higher levels of out-migration from the county, perhaps because of the ubiquity of retail employment throughout the United States in 1930.

The US ecodivisions are consistently significant in their impact on out-migration, although there is little difference between the two continental divisions in the northeast and northern Midwest. What the different odds ratios show us is that there was more out-migration in some regions than others (generally in the West), even after we take into account the rest of the model. We will return to this issue later.

In [table 4](#) we replicate the analysis from [table 3](#), with the main independent variables reflecting a combination of climate and agriculture. These variables are precipitation and temperature in 1934 (compared with 1920–40), plus production of the three main crops in the county in 1939, as a percentage of the production of those crops in 1929. These Climate model results confirm what we saw with the Crop Failure models, with a slightly better overall model fit. In these models higher temperatures and lower precipitation, as well as lower agricultural production, are also associated with more out-migration. The other variables generally behave in the same way as they did in the Crop Failure models. One interesting finding is that the unemployment variable only becomes significant when ecodivision is included in the model.

We report the results of our spatial regime models in summary form in [table 5](#), which compares model diagnostics using the ecodivision regions in three ways: as categorical independent variables, a spatial regime model, and a spatial regime model with a

TABLE 3. OLS models of county out-migration—crop failure models

		<i>Crop failure</i>	<i>Crop failure with ecodivisions</i>	<i>Crop failure with ecodivisions (robust)</i>
DIAGNOSTICS	R squared	0.8431	0.8613	0.9185
	Adj. R squared	0.8427	0.8605	0.9180
	F statistic	2056 on 8 and 3062 DF	1115 on 17 and 3053 DF	
	Shapiro test W value	0.8166	0.7868	0.7756
	Breusch-Pagan p value	0.0273	0.0006	0.0006
	VIF values > 3		ecodiv	log of pop, ecodiv, crop failure
	Moran's I (if significant)	0.3385	0.2874	
INDEPENDENT VARIABLES		Odds (sig.)	Odds (sig.)	Odds (sig.)
	Population	2.3009***	2.4190***	2.3741***
	Crop failure (<i>ref</i> = 0%)			
	<i>Crop failure 1–5%</i>	1.1245***	1.0970***	1.0808***
	<i>Crop failure 5–25%</i>	1.3765***	1.1976***	1.1822***
	<i>Crop failure > 25%</i>	1.5937***	1.2571***	1.2270***
	% of population in retail employment, 1930	1.1076***	1.0498***	1.0647***
	% of population in manufacturing employment, 1930	0.9816***	0.9860***	0.9820***
	Est. % of working age (15–64) population unemployed, 1937	1.0033	1.0067**	1.0080***
	County is urban	0.9390 [‡]	0.9292*	0.9665
	Ecodivisions (<i>ref</i> = <i>warm continental</i>)			
	<i>Hot continental</i>		1.0589 [‡]	1.0788***
	<i>Subtropical</i>		1.1413***	1.2341***
	<i>Prairie</i>		1.3862***	1.3810***
	<i>Temperate steppe</i>		1.5525***	1.5249***
	<i>Subtropical and tropical steppe</i>		1.7413***	1.7568***
	<i>Subtropical and tropical desert</i>		1.5797***	1.6033***
<i>Temperate desert</i>		1.4833***	1.4369***	
<i>Mediterranean</i>		1.8149***	1.7567***	
<i>Marine</i>		1.8791***	1.8730***	

Significance codes: 0: ***; 0.001: **; 0.01: *; 0.05: †.

spatial error term. Because the spatial regime models involve interactions between the 10 ecodivisions and all the other variables, the coefficients are voluminous. We have chosen not to report them here, but rely in the following text on residual maps to show the spatial characteristics of fit. The diagnostics in table 5 show that by including all of the interactions between ecodivision and the other independent variables we are able to develop a model that explains virtually all of the variance in out-migration, based on an r-squared greater than 0.99. When we then add the spatial error term, improve the fit still further (as indicated by the significantly reduced AIC), and eliminate evidence of spatial autocorrelation (as indicated by an insignificant Moran's I).

TABLE 4. OLS models of county outmigration—climate models

		<i>Climate failure</i>	<i>Climate with ecodivisions</i>	<i>Climate with ecodivisions (robust)</i>
DIAGNOSTICS	R squared	0.8458	0.8655	0.9210
	Adj. R squared	0.8454	0.8647	0.9206
	F statistic	2100 on 8 and 3062 DF	1155 on 17 and 3053 DF	
	Shapiro test W value	0.8213	0.7852	0.7728
	Breusch-Pagan p value	<0.0001	<0.0001	<0.0001
	VIF values > 3		ecodiv, temperature	log of pop, ecodiv, max temperature
	Moran's I (if significant)	0.3569	0.2688	
INDEPENDENT VARIABLES		Odds (sig.)	Odds (sig.)	Odds (sig.)
	Population	2.3751***	2.4558***	2.4000***
	% change in 1934 average daily max temperature from 20-year normal	1.0175***	1.0069**	1.0074***
	% change in 1934 total precipitation from 20-year normal	0.9951***	0.9940***	0.9950***
	% of population in retail employment, 1930	1.0857***	1.0447***	1.0592***
	% of population in manufacturing employment, 1930	0.9815***	0.9866***	0.9825***
	Weighted % change in 3 top crops production, 1930–40	0.9991***	0.9993***	0.9995***
	Est. % of working age (15–64) population unemployed, 1937	1.0018	1.0067**	1.0082***
	County is urban	0.9253*	0.9266*	0.9717
	Ecodivisions (<i>ref = warm continental</i>)			
	<i>Hot continental</i>		1.0144	1.0306
	<i>Subtropical</i>		1.1941***	1.2674***
	<i>Prairie</i>		1.3636***	1.3479***
	<i>Temperate steppe</i>		1.3812***	1.3633***
	<i>Subtropical and tropical steppe</i>		1.7012***	1.7037***
	<i>Subtropical and tropical desert</i>		1.4150***	1.4617***
	<i>Temperate desert</i>		1.3686***	1.3221***
<i>Mediterranean</i>		1.7142***	1.6475***	
<i>Marine</i>		1.9351***	1.8888***	

Significance codes: 0: ***; 0.001: **; 0.01: *.

When we introduce the full set of interactions in the regimes model, we achieve a strong model fit, but reveal dramatic heterogeneity among the effects for each combination of independent variable values. We illustrate this heterogeneity in figures 10 and 11 by mapping model residuals. In figure 10 we map the residuals for four versions of our crop failure models, and in figure 11 we do this for the climate models: without ecodivisions, with ecodivisions included as categorical variables, with ecodivisions

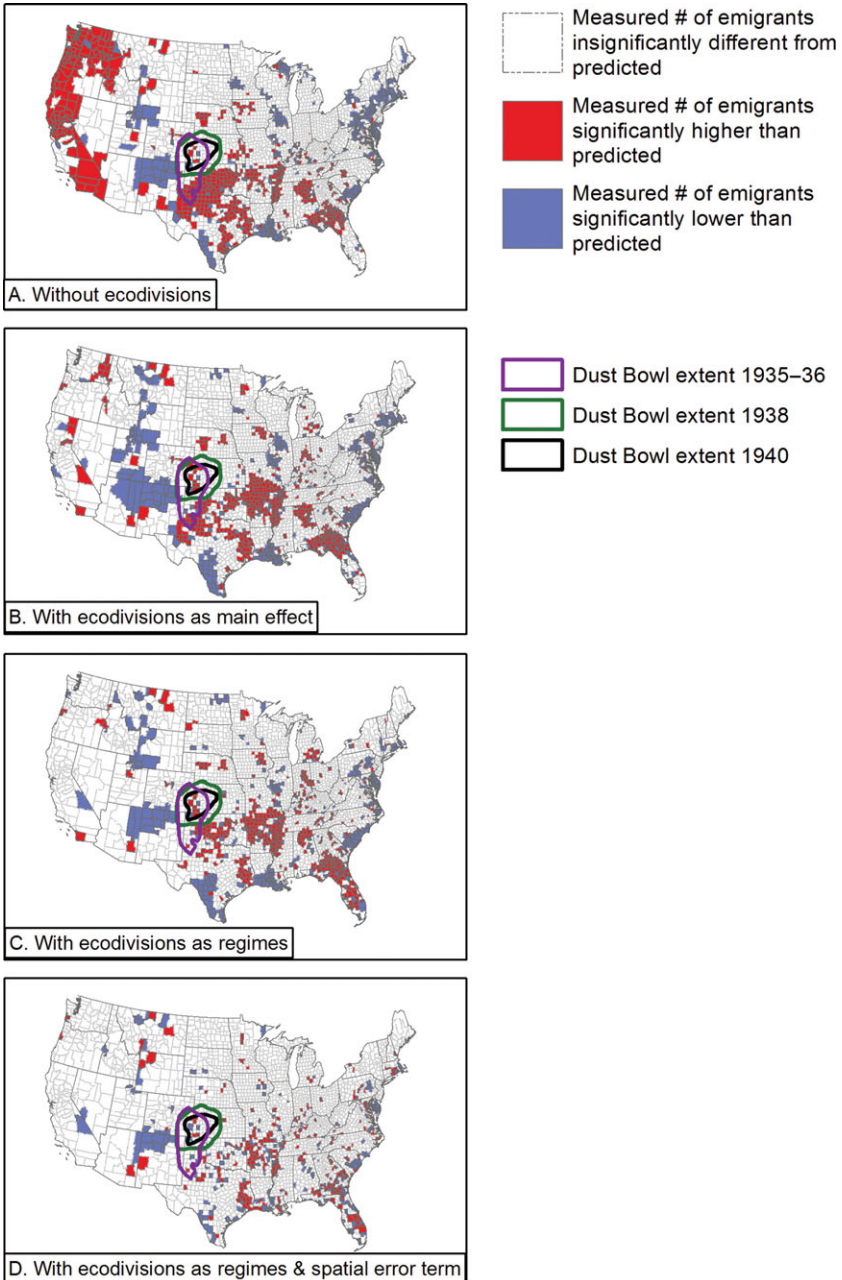


FIGURE 10. Mapped residuals of four different implementations of the crop failure model.

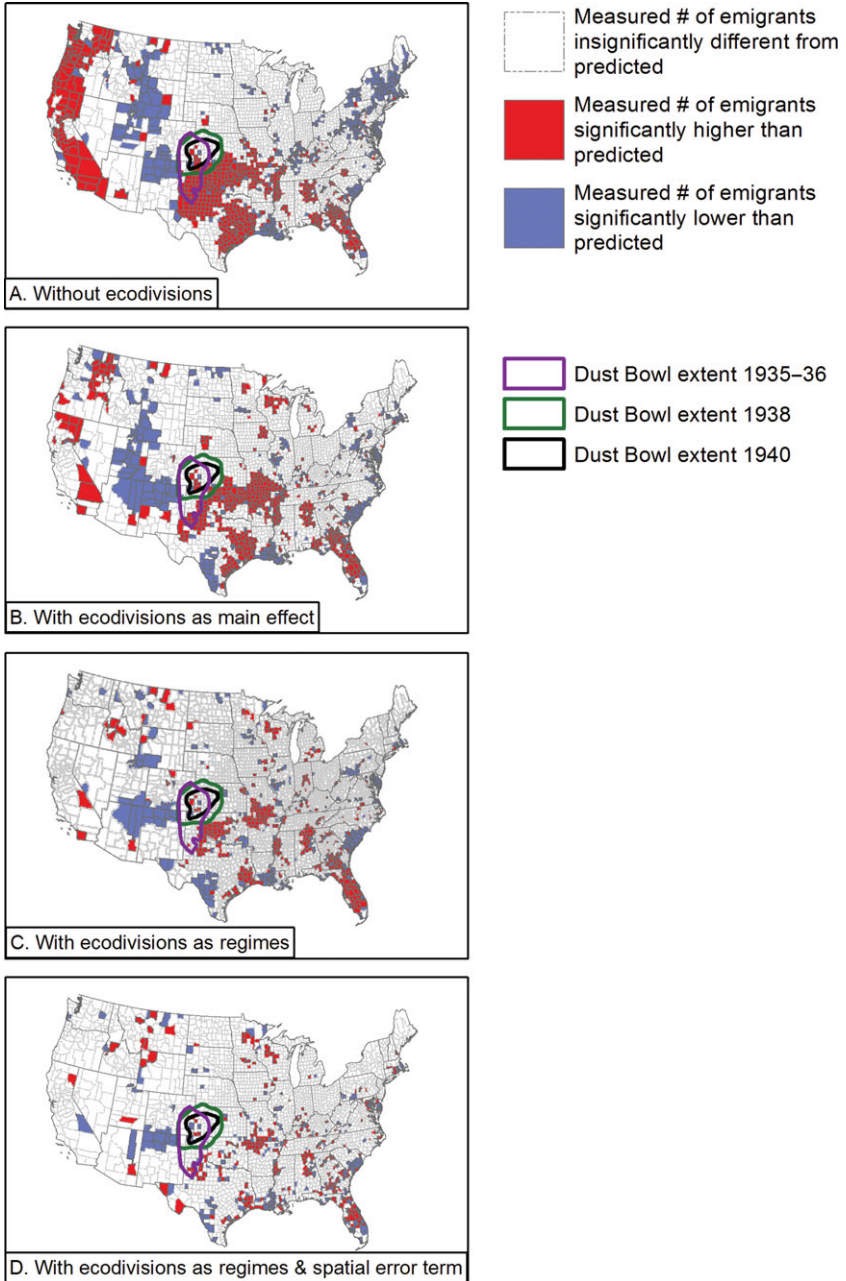


FIGURE 11. Mapped residuals of four different implementations of the climate model.

TABLE 5. *Comparison of models using ecodivision with varying techniques*

		<i>With ecodivision as categorical independent variable</i>	<i>With ecodivision as regime</i>	<i>With spatial error term and ecodivision as regime</i>
CROP FAILURE MODELS				
DIAGNOSTICS	R squared	0.8613	0.9976	
	Adj. R squared	0.8605	0.9976	
	AIC	2589.697	2480.67	2105.541
	Shapiro test W value	0.7868	0.7772	0.7238
	Breusch-Pagan p value	0.0006	0.1210	0.5830
	Moran's I (if significant)	0.2874	0.2309	insignificant
CLIMATE MODELS				
DIAGNOSTICS	R squared	0.8655	0.9978	
	Adj. R squared	0.8647	0.9977	
	AIC	2496.307	2265.556	2032.81
	Shapiro test W value	0.7852	0.7661	0.7278
	Breusch-Pagan p value	<0.0001	0.0911	0.3594
	Moran's I (if significant)	0.2683	0.1785	insignificant

incorporated as regimes, and with both the ecodivision regimes and a spatial error term to explicitly correct for spatial autocorrelation. We use LISA (local indicators of spatial association) statistics, calculated in GeoDa using contiguity to define neighbor relationships, to better convey the patterns of geographical dependence in our data and our model fits. These four steps allow us to use increasingly effective means for understanding the role of spatial dependence in our data and models. In the maps, white counties are those where the model fits well. Blue counties are areas where the measured migration is significantly lower than predicted by the model, while red counties are ones where the measured migration is significantly higher than predicted by the model.

As we progress toward more nuanced appreciations of spatial effects, our overall model fit improves; the number of counties with measured migration significantly different from that predicted by the models (the blue and red counties) decreases. This is a good outcome, confirming the statistical results in [table 5](#), where r-squared, AIC, and other diagnostics improve as we add spatial information. Some areas of poor model fit persist, however. An area in the Southwest remains blue, indicating measured migration is lower than expected and suggesting that an unidentified factor may be providing protection against migration drivers, while red patches scattered across the South indicate that our variables are not accounting for all the drivers of out-migration. These maps also begin to show multicounty patterns that are more localized than what is captured by the ecodivisions. Along the Gulf Coast of Texas, for instance, Harris County, the location of Houston, appears in blue in the final maps, indicating a place of lower than expected out-migration bordered by counties of higher than expected out-migration, perhaps suggestive of an especially strong

rural to urban pull due to the rapid growth of the oil industry in Harris County in the 1930s.

The last thing we note in discussing these results is the visual relationship between the residuals and the location of the various periods of dust activity in the 1930s and 1940. In our best models, those areas appear to fit the model relatively well, suggesting that their behavior is well explained by our models, and may not be unusual when compared with other areas of high temperature and low precipitation.

Conclusion

We began this article by asking how much we really knew about the causes of migration in the 1930s. There was a relationship between weather and migration in the 1930s, and it operated beyond the borders of the Dust Bowl. In most of the United States, people left places that were very hot or very dry, and stayed in places that were relatively cool and wet, although that was not true everywhere. Much of the migration-weather process worked through agriculture, but it did not always operate in ways that can be generalized across the whole United States. That is why the models that use temperature and precipitation appear to explain more of the variation in migration than most of the agricultural variables we could find or estimate.

The factors that determine levels of migration during an era of environmental and economic stress are both national and regional in scale. The United States is a large country with strong regional variations in climate, agriculture, and economy, which our analysis reveals. While temperature and precipitation had generalizable impacts on migration in the second half of the Depression, there were noticeable exceptions to the general pattern of how temperature and precipitation related to migration in various parts of the United States. Moreover, while spatial error models captured unobserved spatial processes and reduced the number of counties with unexplained migration outcomes, significant patterns remain. These patterns suggest a wage decline mechanism, rather than an environmental push mechanism, so that people left areas where wages declined. In the Great Plains, heat and drought reduced production, lowering wages and leading to migration. In Georgia, Florida, and along the Mississippi delta, higher production more than production failures drove down wages, which also led to out-migration.

In revealing both national and regional patterns, our work does not discredit the visceral conventional story of 1930s migration that is beautifully illustrated by Steinbeck, or Lange, or described historically by Worster and Gregory; rather, our analysis grounds the drama of what we know about specific cases within a wider context, the nuances of which may only be sketched with the refined and complex data we now have at our disposal. The coexistence of spatially dependent and general processes suggest that future work must continue to explore the nature of multiscale interactions, and to treat the responses of individual migrants as influenced by places of origin and destination. Even as the work presented here anticipates individual-level

analysis, it reminds us that the motivations migrants share, even during an era of widely shared environmental and economic suffering, remain influenced by history and the characteristics of place.

Acknowledgments

This research has been supported by the Institute of Behavioral Science, the Graduate School, the Office of the Vice Chancellor for Research, and the Provost of the University of Colorado. It has benefited from research, administrative, and computing support from the University of Colorado Population Center (CUPC), funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), Project 2P2CHD066613-06. The content is solely the responsibility of the authors and does not necessarily represent the official views of the CUPC or NICHD. An earlier version of this article was presented at the Annual Meeting of the Social Science History Association in Baltimore, November, 2015. We are grateful for the comments we received at that time, and for the additional comments provided by an anonymous reviewer for *Social Science History*. Tom Dickinson assisted with data preparation for the environmental variables.

APPENDIX A

Deriving Variables about Changes in Agricultural Production for Each County's Three Largest Crops

Using agricultural census data for the years 1930, 1935, and 1940 (which respectively represented farming in 1929, 1934, and 1939) (Haines et al. 2014), we have examined several different types of data as indicators of agricultural production during the period. These data include:

- Percent of total cropland that failed;
- Change in production of major market crops (corn, cotton, and wheat);
- Change in production of major livestock (cattle and pigs). We chose cattle and pigs for our analysis of livestock because these data are most suitable for comparison across census years. The data for other stock animals made comparison more difficult. Work animals such as horses, donkeys, and mules had been in steady decline for some time due to the mechanization of farms; thus, they are not reliable indicators of the health of farms; and
- An agricultural production composite index of the top three crops for each county.

The first three data sets are relatively self-explanatory. The fourth data set, however, requires more explanation. Identifying the most important crops for each of the approximately 3,069 counties in the contiguous United States in our modified data set and then calculating the percent change in production from 1929 to 1939 required several steps. First, we identified the

TABLE A.1. Frequency distribution of the top three crops for each county (only five most frequent 1st, 2nd, and 3rd crops are shown)

Largest Crop (by acreage)	%	2nd Largest Crop	%	3rd Largest Crop	%
Corn	34.61	Corn	29.76	Hay	22.57
Hay	31.49	Hay	20.74	Wheat	14.13
Cotton	19.21	Oats	17.55	Oats	13.94
Wheat	10.16	Wheat	11.85	Corn	8.73
Oats	0.72	Cotton	6.02	Barley	8.01
All Others	3.81	All Others	14.07	All Others	32.63

Note: E.g., corn is the predominant crop in 34.61 percent of all counties in 1930 (by acreage), the second-most in 29.76 percent of counties, and third in 8.73 percent. Thus, corn was a top three crop in 73.1 percent of all counties in 1930.

top three crops in each county by the acreage harvested in 1929 (table A.1). The Agricultural Census data for the period contains acreage for nearly all *significant* crops, meaning any crop that appears in the top three for at least one county. The one major exception is fruit trees: Census takers only recorded the number of trees, rather than the amount of acreage. However, 20 years later, the 1950 Census recorded both acreage and number of trees. This allowed us to estimate the number of acres devoted to fruit trees in each county, assuming that the average number of trees per acre was stable from 1929 to 1949.

Second, for each county, we identified the amount of production for each of the top three crops (whether recorded in bushels, bales, pounds, or tons), which we then used to calculate changes in production for the period. The data posed several problems in this effort. Census takers of the period did not record any production values for some crops, most notably vegetables, for which enumerators noted farms reporting, acreage, and dollar values for each vegetable, but not the quantity of production (e.g., bushels). Vegetables were only reported as one of the top three crops in 126 counties, making this gap less problematic. For other crops, the unit of measurement used to record production changed between 1930 and 1940. Most of the time, a simple conversion was sufficient to make the data comparable, an example being that one bushel of cherries weighs approximately 56 pounds, allowing us to convert production in bushels into production in pounds, or vice versa. In a few cases, the data required more complicated conversion calculations. Peanuts are a more complicated example, because they were recorded in bushels in 1930 and 1935, but in pounds in 1940. This posed a problem as different types of peanuts had significantly different rates of pounds/bushel. We worked around this by applying different conversion rates for each state depending on the dominant type of peanut found in that state.

Once we had accounted for all discrepancies in the data, we calculated the percent change in production for each crop across three time periods (1929–34, 1935–39, and 1929–39). In addition, we created a composite index of the percent change of the three crops combined. We created this composite measure by calculating a weighted percent change. For example, in Sutter County, California, barley, wheat, and hay represented 48.8 percent, 40.2 percent, and 11 percent, respectively, of the county's acreage devoted to the top three crops. We multiplied the percent change in production of each crop by the crop's relative size as a weight to create a composite weighted percent change figure.

Appendix B

Counties Combined in the Analysis

<i>Original Counties</i>	<i>New Units (following Fishback et al. 2003)</i>
Campbell (13041), Fulton (13121), and Milton (13203) counties, GA (for 1930 only; these counties were already merged by the 1940 Census)	Fulton (13121)
New York City: Bronx (36050), Kings (36470), New York (36610), Queens (36810), and Richmond (36850) counties	"New York City County" (36150)
St. Louis City (29510) and County (29189)	"St. Louis City and County" (29300)
<i>Virginia Independent Cities and Surrounding Counties:</i>	
Albemarle (51003) and Charlottesville City (51540)	"Albemarle (incl. Charlottesville)" (51200)
Allegheny (51005) and Clifton Forge City (51560)	"Allegheny (incl. Clifton Forge)" (51210)
Arlington (51013) and Alexandria City (51510)	"Arlington (incl. Alexandria City)" (51215)
Augusta (51015) and Staunton City (51790)	"Augusta (incl. Staunton)" (51220)
Campbell (51031) and Lynchburg City (51680)	"Campbell (incl. Lynchburg)" (51230)
Dinwiddie (51053) and Petersburg City (51730)	"Dinwiddie (incl. Petersburg City)" (51240)
Hampton (51630) and Elizabeth City (51055)	"Elizabeth City (incl. Hampton)" (51250)
Frederick (51069) and Winchester City (51840)	"Frederick (incl. Winchester City)" (51260)
Henrico County (51087) and Richmond City (51760)	"Henrico (incl. Richmond City)" (51270)
Martinsville (51690) and Henry (51089)	"Henry (incl. Martinsville)" (51280)
James City County (51095) and Williamsburg City (51830)	"James City County (incl. Williamsburg City)" (51290)
Montgomery County (51121) and Radford City (51750)	"Montgomery (incl. Radford City)" (51300)
Nansemond (51123) and Suffolk City (51800)	"Nansemond (incl. Suffolk City)" (51310)
Norfolk (51129), Norfolk City (51710), South Norfolk City (51785), and Portsmouth City (51740)	"Norfolk County (combined)" (51320)
Danville (51590) and Pittsylvania (51143)	"Pittsylvania (incl. Danville)" (51330)
Prince George (51149) and Hopewell City (51670)	"Prince George (incl. Hopewell City)" (51340)
<i>Original Counties</i>	
<i>New Units (following Fishback et al. 2003)</i>	
Roanoke (51161) and Roanoke City (51770)	"Roanoke (incl. Roanoke City)" (51350)
Rockbridge (51163) and Buena Vista City (51530)	"Rockbridge (incl. Buena Vista City)" (51360)
Rockingham (51165) and Harrisonburg City (51660)	"Rockingham (incl. Harrisonburg City)" (51370)
Fredericksburg (51630) and Spotsylvania (51177)	"Spotsylvania (incl. Fredericksburg)" (51380)
Warwick (51189) and Newport News City (51700)	"Warwick (incl. Newport News)" (51390)
Bristol (51520) and Washington (51191), VA	"Washington (incl. Bristol)" (51400)

Appendix C

Managing Places of Origin and Matching the Coded and Uncoded Versions of the 1940 Full-count Census

The process of matching the coded and uncoded (raw text) 1940 Census data appears simple, requiring that the data user match on references to the original microfilm reel and manuscript page and line number; all are reported in both data sets. What appears simple turns out to be extremely difficult because IPUMS coding rules (developed for the older sample-based data

sets) call for the page number to be “recoded” in order to ensure that all members of a household have the same page number, even when they span multiple pages of the original manuscript. Put another way, when a household spans multiple pages, the page number in the coded data set stays the same, but in the raw text data set page number changes as the page turns. This is a particularly troublesome characteristic if the “household” is an institution (e.g., group quarters, military base, hospital), spanning multiple pages or if intermediate pages are missing. We have developed a script in SAS that corrects the majority of these problems, resulting in a data set that contains 131,438,236 observations (not including Alaska and Hawaii, which have limited other data); the cases lost through our merging process are primarily residents of group quarters, who would be systematically excluded from our analysis in any event. With two exceptions, our data represent between 99 and 101 percent of the official population of each state—and almost 99.9 percent of the official population of the contiguous United States as a whole—as reported in the Historical Statistics of the United States (Carter et al. 2006).

Effectively merging the two versions did not produce useable information about the 1935 county of origin of migrants, however. We needed to assign a unique identifier (state and county Federal Information Processing Standard [FIPS] geographical codes) to the 1935 place of residence for everyone enumerated in 1940. For most people, enumerated as being in the “same house” or “same place” and coded as having stayed in their origin county by IPUMS’s *migrate5* variable, this task was easy: We coded their 1935 county of residence FIPS to the 1940 county of residence FIPS coded by IPUMS. Individuals classified by IPUMS as moving between US counties ($21 < \text{migrate5} < 40$) were more challenging. We coded the 1935 county of individuals who IPUMS reported as making an intercounty move but who remained within our modified county boundaries (e.g., moves between any of the five counties encompassed by New York City that we have combined into a single “county”) directly from the 1940 county. For individuals who made intercounty or interstate moves by our definition we began by creating a dictionary of unique state and county text strings (from official lists), and used those to assign county IDs for individuals for whom the enumerator had written down a 1935 county of residence. This worked reasonably well, giving us an exact 1935 county of residence for roughly half of all migrants. Two sorts of problems remained: a combination of clerical mistakes by the enumerator and misspelled or incorrectly identified counties, plus respondents reporting only the city of previous residence, and not the county. We resolved this issue by attempting to match the enumerated city-state combination in the uncoded data with a dictionary of unique city-county-state combinations drawn from the coded IPUMS version of the data.

We composed this city-county-state dictionary by extracting all the unique 1940 city-county-state combinations from the coded IPUMS data set using the data set’s *MIGCITY* variable, and removing any entries for cities that spanned more than one county, thus preventing the ambiguous assignment of a county name. Residents with an unknown county but a known city that spanned two or more counties were assigned to the county with the largest area within the city limits (e.g., Amarillo residents with an unknown 1935 county were assigned to Potter County). These changes are shown in Appendix D. Changes made to align the individual-level data with county-level data, described in the following text, also resolved county assignment issues for New York City, St. Louis, and Virginia’s independent cities. We have also fixed the miscoding in IPUMS’s coded data that coded Richmond County, Virginia, as Richmond City, Virginia.

At the end of these processes, among our population of interest (nongroup quarter adults who remained within the contiguous United States between 1935 and 1940), less than 5 percent had an unknown county of origin.

Appendix D

Cities in Multiple Counties

<i>1935 City</i>	<i>Possible 1935 Counties</i>	<i>1935 County Assigned If Unknown</i>
Amarillo, TX (160)	Potter (48375), Randall (48381)	Potter (48375)
Atlanta, GA (350)	Fulton (13121), DeKalb (13089)	Fulton (13121)
Bethlehem, PA (730)	Northampton (42095), Lehigh (42077)	Northampton (42095)
Centralia, IL (1021)	Marion (17121), Clinton (17027)	Marion (17121)
Elgin, IL (2030)	Kane (17089), Cook (17031)	Kane (1789)
Elwood City, PA (2061)	Lawrence (42073), Beaver (42007)	Lawrence (42073)
Fostoria, OH (2351)	Seneca (39147), Hancock (39063)	Seneca (39147)
Huntington, WV (2910)	Cabell (54011), Wayne (29099)	Cabell (54011)
Joplin, MO (3210)	Jasper (29097), Newton (29145)	Jasper (29097)
Rocky Mount, NC (5974)	Nash (37127), Edgecombe (37065)	Nash (37127)
St Cloud, MN (6693)	Stearns (27145), Sherburne (27141), Benton (27009)	Stearns (27145)
Watertown, WI (7311)	Jefferson (55055), Dodge (55027)	Jefferson (55055)

References

- Adger, W. N. (2006) "Vulnerability." *Global Environmental Change* 16 (3): 268–81.
- Agee, J., and W. Evans (1941) *Let Us Now Praise Famous Men*. Boston: Houghton Mifflin.
- Anselin, L. (2007) *Spatial Regression Analysis in R: A Workbook*. Santa Barbara, CA: Center for Spatially Integrated Social Science.
- Anselin, L., I. Syabri, and Y. Kho (2006) "GeoDa: An introduction to spatial data analysis." *Geographical Analysis* 38 (1): 5–22.
- Bailey, R. G. (1997) *Map: Eco-regions of North America* (rev.). Washington, DC: USDA Forest Service in cooperation with The Nature Conservancy and the US Geological Survey; <https://www.sciencebase.gov/catalog/item/54244abde4b037b608f9e23d> (accessed September 16, 2016).
- Bailey, R. G., and C. T. Cushwa (1981) *Map: Eco-regions of North America*. FWS/OBS-81–29. Washington, DC: US Fish and Wildlife Service.
- Bakewell, O. (2013) "Relaunching migration systems." *Migration Studies* 2 (3): 300–18.
- Berkes, F., F. Carl, and C. Johan (1998) *Linking Social and Ecological Systems Management Practices and Social Mechanisms for Building Resilience*. New York: Cambridge University Press.
- Biggers, J. D., and United States (1938) *Census of Partial Employment, Unemployment, and Occupations: 1937*. Washington, DC: US Government Printing Office.
- Black, R., N. W. Arnell, W. N. Adger, D. Thomas, and A. Geddes (2013) "Migration, immobility and displacement outcomes following extreme events." *Environmental Science and Policy* 27 (S1): S32–43.
- Black, R., W. N. Adger, N. W. Arnell, S. Dercon, A. Geddes, and D. Thomas (2011) "The effect of environmental change on human migration." *Global Environmental Change* 21, Supplement 1 (0): S3–S11.
- Bogue, D. J., and M. J. Hagood (1953) *Subregional Migration in the United States, 1935–40*. Oxford, OH: Scripps Foundation, Miami University.
- Bogue, D. J., H. S. Shryock Jr., and S. A. Hoermann (1957) *Subregional Migration in the United States, 1935–40*. Oxford, OH: Scripps Foundation, Miami University.
- Borjas, G. J. (1987) "Self-selection and the earnings of immigrants." *The American Economic Review* 77 (4): 531–53.
- Boustan, L. P., M. E. Kahn, and P. W. Rhode (2012) "Moving to higher ground: Migration response to natural disasters in the early twentieth century." *The American Economic Review* 102 (3): 238–44.

- Boustan, L. P., Price, V. Fishback, and S. Kantor (2010) "The effect of internal migration on local labor markets: American cities during the Great Depression." *Journal of Labor Economics* 28 (4): 719–46.
- Carter, S. B., S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright (2006) *Historical Statistics of the United States Millennium Edition Online*. New York: Cambridge University Press.
- Cook, B. I., R. Seager, and J. E. Smerdon (2014) "The worst North American drought year of the last millennium: 1934." *Geophysical Research Letters* 41 (20): 7298–7305.
- Cunfer, Geoff (2005) *On the Great Plains: Agriculture and Environment*. College Station: Texas A&M University Press.
- Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. P. Pasteris (2002) "A knowledge-based approach to the statistical mapping of climate." *Climate Research* 22 (2): 99–113.
- Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris (2008) "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States." *International Journal of Climatology* 28 (15): 2031–64.
- DeWaard, J., K. J. Curtis, and E. Fussell (2014) *Demographic Signatures of Migration Systems: Population Recovery after Hurricanes Katrina and Rita*. Minneapolis: Minnesota Population Center.
- Egan, T. (2006) *The Worst Hard Time: The Untold Story of Those Who Survived the Great American Dust Bowl*. Boston: Houghton Mifflin.
- FAO (2011) *Guidelines for the Preparation of Livestock Sector Reviews*. Rome: UN Food and Agriculture Organization.
- Ferrie, J. P. (2006) "Internal migration," in S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright (eds.) *Historical Statistics of the United States*. Millennium Edition Online. New York: Cambridge University Press.
- Fishback, Price V., S. Kantor, and J. Wallis (2003) "Can the New Deal's Three R's Be Rehabilitated? A Program-by-Program, County-by-County Analysis." *Explorations in Economic History* 40: 278–307.
- Fishback, P. V., W. C. Horrace, and S. Kantor (2006) "The impact of New Deal expenditures on mobility during the Great Depression." *Explorations in Economic History* 43 (2): 179–222.
- Fussell, E., K. Curtis, and J. DeWaard (2014) "Recovery migration to the City of New Orleans after Hurricane Katrina: A migration systems approach." *Population and Environment* 35 (3): 305–22.
- Giesen, J. C. (2011) *Boll Weevil Blues: Cotton, Myth, and Power in the American South*. Chicago: University of Chicago Press.
- Greenwood, M. J. (1997) "Internal migration in developed countries," in R. R. Mark and S. Oded (eds.) *Handbook of Population and Family Economics*, Vol. 1. The Netherlands: Elsevier: 647–720.
- Gregory, J. N. (1989) *American exodus: The Dust Bowl migration and Okie culture in California*. New York: Oxford University Press.
- (2005). *The Southern Diaspora: How the Great Migrations of Black and White Southerners Transformed America*. Chapel Hill: University of North Carolina Press.
- Haines, M. R., and Inter-university Consortium for Political and Social Research (2010) *Historical, Demographic, Economic, and Social Data: The United States, 1790–2002*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.
- Haines, M. R., Price, V. Fishback, and P. W. Rhode (2014) *United States Agriculture Data, 1840–2010*. Ann Arbor: Inter-university Consortium for Political and Social Research.
- Hall, P. K., and S. Ruggles (2004) "'Restless in the midst of their prosperity': New evidence on the internal migration of Americans, 1850–2000." *The Journal of American History* 91 (3): 829–46.
- Hornbeck, R. (2012) "The enduring impact of the American Dust Bowl: Short- and long-run adjustments to environmental catastrophe." *American Economic Review* 102 (4): 1477–1507.
- Kanbur, R., and H. Rapoport (2005) "Migration selectivity and the evolution of spatial inequality." *Journal of Economic Geography* 5 (1): 43–57.
- Lee, E. S. (1966) "A theory of migration." *Demography* 3 (1): 47–57.
- Lively, C. E., and C. Taeuber (1939) *Rural Migration in the United States*. Washington, DC: US Government Printing Office.
- Long, J., and H. Siu (2013) "Refugees from dust and shrinking land: Tracking the Dust Bowl migrants." Wheaton College: unpublished manuscript.

- Massey, D. S. (1990) "Social structure, household strategies, and the cumulative causation of migration." *Population Index* 56 (1): 3–26.
- Massey, D. S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, and J. E. Taylor (1993) "Theories of international migration: A review and appraisal." *Population and Development Review* 19 (3): 431–66.
- (1998) *Worlds in Motion: Understanding International Migration at the End of the Millennium*. New York: Oxford University Press.
- McEwan, R. W., N. Pederson, A. Cooper, J. Taylor, R. Watts, and A. Hruska (2014) "Fire and gap dynamics over 300 years in an old-growth temperate forest." *Applied Vegetation Science* 17 (2): 312–22.
- McLeman, R. A. (2013) *Climate and Human Migration: Past Experiences, Future Challenges*. New York: Cambridge University Press.
- McLeman, R. A., and B. Smit (2006) "Migration as an adaptation to climate change." *Climatic Change* 76 (1–2): 31–53.
- Olmstead, A. L., and P. W. Rhode (2008) *Creating Abundance: Biological Innovation and American Agricultural Development*. New York: Cambridge University Press.
- Ravenstein, E. G. (1885) "The laws of migration." *Journal of the Statistical Society of London* 48 (2): 167–235.
- (1889) "The laws of migration." *Journal of the Royal Statistical Society* 52 (2): 241–305.
- Roy, A. D. (1951) "Some thoughts on the distribution of earnings." *Oxford Economic Papers* 3 (2): 135–46.
- Ruggles, S., J. T. Alexander, K. Genadek, R. Goeken, M. B. Schroeder, and M. Sobek (2010) *Integrated Public Use Microdata Series: Version 5.0*. Minneapolis: University of Minnesota.
- Steinbeck, J. (1939) *The Grapes of Wrath*. New York: Modern Library.
- Tolnay, S. E., K. J. C. White, K. D. Crowder, and R. M. Adelman (2005) "Distances traveled during the Great Migration: An analysis of racial differences among male migrants." *Social Science History* 29 (4): 523–48.
- US Bureau of the Census (1946) *Sixteenth Census of the United States: 1940. Population. Internal migration, 1935 to 1940. Economic Characteristics of Migrants*. Washington, DC: US Government Printing Office.
- (2002) *Measuring America: The Decennial Censuses from 1790 to 2000*. Washington, DC: US Department of Commerce, Economics and Statistics Administration, US Census Bureau.
- (2012) *United States Summary, 2010 Population and Housing Unit Counts*. Washington, DC: US Dept. of Commerce, Economics and Statistics Administration, US Census Bureau.
- White, K. J. C. (2005) "Women in the Great Migration: Economic activity of black and white southern-born female migrants in 1920, 1940, and 1970." *Social Science History* 29 (3): 413–55.
- White, K. J. C., K. Crowder, S. E. Tolnay, and R. M. Adelman (2005) "Race, gender, and marriage: Destination selection during the great migration." *Demography* 42 (2): 215–41.