Migration and Mobility during the Dust Bowl

Arun Ramesh

Senior Thesis

Abstract

The American Dust Bowl during the thirties conjured images of despair and penury, highlighted by the plight of newly-minted refugees fleeing storms and destroyed ecosystems. I utilize a census-linking approach to create a dataset linking people in the Great Plains region and their migratory responses from 1920 to 1940. By using a difference-in-differences as well as a difference-in-difference-in-differences approach, I look at how the Dust Bowl induced changes in migratory responses, selection and outcomes for those living in the Great Plains region. I find that migration rates were higher in the Dust Bowl regions, particularly in counties that experienced high levels of soil erosion. However, I only find modest changes in migrant selectivity and outcomes for Dust Bowl migrants. Interestingly, I find that migration to California increased, as did migrant selectivity and outcomes, providing some context to the westward migratory flows. Contrary to popular opinion, there was not a drastic mass exodus of poor farmers, and there were no drastic changes in migratory responses as a result of the Dust Bowl.

1 Introduction

The thirties saw a turbulent time that transformed the landscape of the United States, namely the Great Depression sparked from the stock market crash of 1929. However, one other significant event was the Dust Bowl - one that literally transformed the *landscape* of the Great Plains. The Dust Bowl was a period of dust storms and adverse climate conditions that ravaged the Great Plains region during the 1930s, causing heavy soil erosion in farm intensive areas, and sparking an exodus of refugees fleeing the storms. This exodus, one of the largest internal migrations ever in recorded history, has been immortalized through the arts, from John Steinbeck's *The Grapes of Wrath* to the iconic photographs of Dorothea Lange. In particular, the mass migration to the Pacific States, namely California, is noted as the legacy of the Dust Bowl.

The effects of the Dust Bowl did not end in the thirties, and continued beyond the end of the dust storms and the Great Depression. There were large declines in population in the Great Plains, along with declines in retail sales, land values and GDP (Hornbeck, 2012) (Fishback, Horrace, & Kantor, 2006). The economic adjustment was slow and continued through the 1950s, and agricultural recovery was just as slow. The Dust Bowl was not only an environmental shock, but an economic one as well, and so can be viewed as a 'push' factor that sparked mass out-migration, one that was able to overcome the negative effects of the Great Depression elsewhere in the country. The total degradation of farmland severely impacted local economies and agricultural production, and unlike usual temporary weather shocks, the effects of the Dust Bowl were systematic and permanent.

In this paper I look more into the migrants from this region, and attempt to distinguish how the Dust Bowl created refugees against the backdrop of the Great Depression. I utilize a census-linking approach to create a dataset to track the mobility of Americans from the decennial censuses in 1920, 1930 and 1940. This approach allows me to look at the different characteristics of the migrants, which can help identify what kind of migrants were the most able and willing to move. I will also look at how migration rates changed in the region, whether migrant selection changed during the Dust Bowl, as well as how outcomes were different for those who chose to migrate. By doing so, we can gain insight into the behaviours and incentives of Dust Bowl-induced migrants. I utilize a difference-in-differences approach to look at how migration rates changed in the Dust Bowl, and a difference-in-difference-in-differences approach to examine the causal effects of the Dust Bowl on migrant selection and outcomes.

I find that migration rates were higher in the Dust Bowl region in the 1930s, and counties in the Great Plains that experienced high and medium levels of soil erosion also saw more out-migration during the Dust Bowl, relative to counties that experienced relatively lower levels of soil erosion. Interestingly, out-migration rates remained relatively constant in counties affected by the Dust Bowl, while out-migration rates declined severely around the country, most likely due to effects of the Great Depression. While I find only modest changes in migrant selectivity and post-migration outcomes, I find that migrants to California (a notable destination for Dust Bowl refugees) had lower migrant selectivity and post-migration incomes than prior to the Dust Bowl.

The goal of my analysis is to get a comprehensive view of migration during the Dust Bowl, and changes in migratory responses. By looking at not only how migration rates changed, but also how migrant characteristics (selection) and migrant outcomes changed, I can provide context to validate the dark images of the Dust Bowl. Does empirical evidence provide basis for the depictions of the Dust Bowl migrant?

2 Background

The United States has a well-documented and storied relationship with migration (Abramitzky, Boustan, & Eriksson, 2014) (Abramitzky & Boustan, 2017), from the period of mass immigration from Europe (Abramitzky, Boustan, & Eriksson, 2012), to internal migration patterns from the Gold Rushes in the west (Clay & Jones, 2008) as well as the Great Migration which saw Black Americans move from the American South to northern states (Collins & Wanamaker, 2015) (Hornbeck & Naidu, 2014). After all, the Manifest Destiny was ingrained in the American folklore and policy, and the creation of the US Highway System and the Transcontinental railroad saw the creation of new destinations for Americans (Hudson, 1982) (Donaldson & Hornbeck, 2016). The Homestead Act of 1862 was passed to encourage new migrants into the Great Plains. Likewise, good weather and the general westward push attracted migrants to the Great Plains (Cunfer, 2011) (Galenson & Pope, 1989).

Likewise, the Dust Bowl was a shock that sparked a large internal migration immortalized in the annals of American history. Having started in 1930, and continuing through the decade with severe intensity, the Dust Bowl saw many dust storms in the Great Plains, a notable one occurring on April 14^{th}, 1935 (colloquially known as *Black Sunday*, during which dust travelled all the way to the East Coast). While dust storms were not a new phenomenon in the Great Plains, it was coupled with periods of drought and drastic changes in climate patterns and conditions (Cunfer, 2005). The cause of the Dust Bowl has been heavily scrutinized, with the debate of whether it was a man-made ecological problem brought forth by increased agriculture activity and mechanization, or whether it was indicative of broader changes in climatic conditions. However, its effects were very clear: the severe degradation of land, the decline of the economy, and the sparking of a large-scale exodus. In response to this, The US government introduced New Deal programs, which primarily targeted the effects of Great Depression, but also had programs to aid those in the Dust Bowl, such as the Agricultural Adjustment Act (Davis, 1936).

Donald Worster, in his *Dust Bowl: The Southern Plains in the 1930s*, one of the seminal works documenting the Dust Bowl, alludes to increased farming activity driven by profits as the main cause behind the degradation of the topsoil and the dust storms (Worster, 2004). Other accounts point to increased mechanization of agriculture and technological innovation in tandem to climatic conditions, along with the presence of many small farms with lack of incentive to adopt sustainable collective agricultural practices (Hurt, 1981) (Hansen & Libecap, 2004). However, Geoff Cunfer's *On the Great Plains* makes the claim that the Dust Bowl was an exogenous shock, accelerated by climatic conditions such as prolonged periods of drought (Cunfer, 2005). As there is still significant debate, one could note that if the Dust Bowl was a man-made phenomenon, and if residents were selected into the Dust Bowl to begin with, it could lead to problems with endogeneity. However, I follow the analysis from (Cunfer, 2005) and assume that it was indeed an exogenous shock (this is the same assumption taken in other works examining causal impacts of the Dust Bowl, such as (Arthi, 2018)). This is supported by analysis of weather and climate patterns and conditions, and through surveys undertaken in the Great Plains (Cunfer, 2005).

The boundaries of the Dust Bowl region have been disputed. From Geoff Cunfer's *On the Great Plains*, mapping software such as ArcGIS was used to digitize historical records to define the Dust

Bowl region, by looking at weather data as well as erosion data taken from the numerous surveys taken by local and federal agencies (Cunfer, 2005). Figure 1 is a map that shows the occurrence of major dust storms, along with the regions directly impacted. Henceforth, the counties in this region will be referred to the *Dust Bowl region*. This region is concentrated in the Southern Great Plains, most notably the panhandles of Oklahoma and Texas. While the panhandle region experiences significant wind erosion, the effects of the Dust Bowl were experienced across the great plains (Hansen & Libecap, 2004). For the sake of this paper, I will look at both regions and see whether the results from the cluster of counties in the *heart* of the Dust Bowl can be extrapolated to the entire Great Plains to test for internal validity.

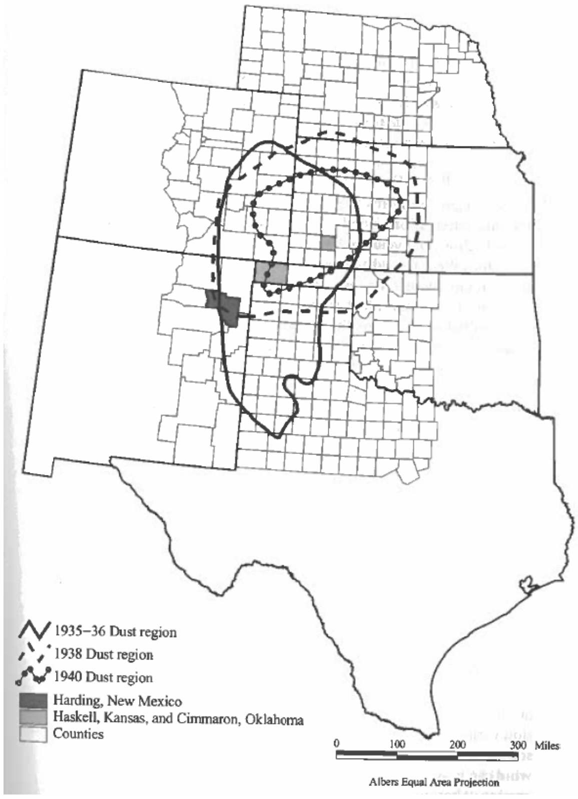

Figure 1: The Dust Bowl Region

There has been considerable interest in the American Dust Bowl, especially its implications on migration. There have been studies looking at migration during the period, using approaches such as county-wide migration flows (Hornbeck, 2020) and utilizing census linking algorithms to link individuals throughout the period (Long & Siu, 2018). (Fishback et al., 2006) looks at how New Deal Spending affected geographical mobility, which has significant overlap with the Dust Bowl. (Long & Siu, 2018) define the Dust Bowl region as a cluster of counties in the Southern Great Plains, based on the counties surveyed by the Soil Conservation Service in 1934 (Joel, 1937). On the other hand, (Hornbeck, 2020) focuses on the Great Plains, by using erosion intensity as a proxy for the severity of the Dust Bowl. While (Long & Siu, 2018) uses a census linking approach to look at migration decisions at the individual level, (Hornbeck, 2020) looks at county level migration rates between 1935 and 1940.

I will utilize a combination of both approaches to get a better understanding migration in the Dust Bowl region by looking at both definitions: the cluster of counties from (Cunfer, 2005) and the Great Plains region from (Hornbeck, 2020), in an attempt to isolate the effect of the Dust Bowl on migratory responses from the effect of the Great Depression as a whole. (Long & Siu, 2018) finds that migration rates remained the same in the 30s in the Dust Bowl region relative to the 20s, but actually fell in the rest of the country. (Hornbeck, 2020) takes a different approach and finds that migrants from highly eroded counties were negatively selected with respect to education. (Gregory, 1991) was one of the first papers to look at migration utilizing census data (using the 1940 1% sample) and finds that many migrants were actually white collar workers, contrary to the notion of refugees being comprised of farmers.

Studies on the Dust Bowl also look at other aspects than migration: (Arthi, 2018) finds long term negative effects on human capital of those born during the Dust Bowl, while (Moscona, 2020) looks at the long term impacts of the Dust Bowl on agricultural innovation, in terms of new varieties of crops. (Hornbeck, 2012) examines the short-term and long-term effects of the Dust Bowl on the economy, and finds that counties that experienced high soil erosion saw immediate, short-term and long-term decline in land values.

With respect to the broader topic of migration, there has been considerable literature on the selection of migrants (i.e. whether migrants are randomly selected, or whether certain types of migrants are induced to move, such as high-skilled professionals) (Roy, 1951) (Borjas, 1987) (Abramitzky et al., 2012), and the outcomes of migrants relative to natives in their migrant destinations and relative to those who chose not to migrate (Abramitzky, Boustan, Jacome, & Perez, 2021) (Cattaneo

& Peri, 2016). Migrant selectivity and impact on local labor markets have long been of interest for economists and policymakers (Boustan, Fishback, & Kantor, 2010) (Boustan, 2016).

There is significant debate as to how climate change will impact migration (Bohra-Mishra, Oppenheimer, & Hsiang, 2014) (Stern & Stern, 2007), and the economy as a whole (Desmet & Rossi-Hansberg, 2015), and by looking at the Dust Bowl, we can learn how environmental shocks generate and alter migratory responses and patterns. In this paper I will utilize empirical approaches to see how the Dust Bowl impacted migration, by looking at the concentration of counties surrounding the Texas and Oklahoma panhandles, and extrapolating my models to the entire Great Plains region to capture the effect of the Dust Bowl.

3 Data

There have been many initiatives to create multi-generational datasets that enable researchers to match individuals across years (Abramitzky, Mill, & Pérez, 2020) (Ferrie, 1996), and (Long & Siu, 2018) using this approach to look at migration in the Dust Bowl. New advancements in machine learning algorithms and the rise of genealogical organizations such as *Ancestry.com* have enabled researchers to scale up the linking while also improving accuracy. Rather than looking at county-level migration rates, I opt for this approach because it allows me to look at migration decisions at the individual level.

For this paper, I use the linking algorithm from the IPUMS Multigenerational Longitudinal Panel Project[1] to match individuals from the years 1920 to 1930, and from 1930 to 1940 (Ruggles et al., 2019). By combining these two datasets, I get a representative sample of around 55 million Americans. As the data being linked are the full count decennial census data, there is a plethora of information at the individual-level - particularly the county of residence in each year (1920, 1930 and 1940). The 1940 census was the first to ask where the individual was five years prior, so there is also information on where the individual was in 1935. This information allows me to track the geographic mobility of Americans before and during the Dust Bowl, as well as seeing nationwide

[1] This approach compares demographic variables which do not change for individuals, such as birth year, state of birth, given name, surname, and race across multiple censuses. Exact matches are corroborated with records to check for accuracy. This algorithm utilizes machine learning to automate and scale up the process, while managing the trade-off between accuracy and generating a representative sample of the population

trends in internal migration. While there were other linking algorithms available, this approach allowed me to match more individuals with great accuracy relative to other linking algorithms (Ruggles et al., 2019) (Abramitzky et al., 2020). I restrict my sample to males, as one problem with the census linking approach is that individuals are linked based on similarities of names- which poses a problem for women who change their last name after marriage. As of the time this paper was written, there is no way of weighting the sample that was chosen for linkage, and it could be the case that individuals who migrated between decennial years are under-represented (Ruggles et al., 2019).

Each row of the final dataset consists of an individual at the start of the decade, with the variables of interest, and the variables from ten years later (when the next decennial Census was taken). This link between the decennial censuses is done using the linking algorithm from IPUMS. I link the 1920 full census to the 1930 full census, and then the 1930 full census to the 1940 full census, thus creating a 'panel' dataset which allows me to compare outcomes in the 1930s (when the dust storms occurred), to the 1920s. As the Dust Bowl started in the 1930, and continued through the decade, this allows me to adopt a difference in difference approach (and in some cases a triple difference in difference in differences model), and compare how the dust storms in the 30s affected migration and migrant outcomes via a *quasi-experimental* approach. Additionally, the 1940 census was the first to ask the whereabouts of the respondent 5 years ago, so I have information on the county of residence in 1935, which can provide additional insight into migration patterns and decisions during the thirties (See Appendix A).

Next, I get the erosion data for the Great Plains (Hornbeck, 2012) - which utilized digitized records from the United States Department of Agriculture (USDA) and Soil Conservation Service (SCS, now the Natural Resources Conservation Service). The map below (Figure 2, which was adapted from (Hornbeck, 2012)) shows the counties affected, with three categories to represent erosion intensity. High erosion refers to more than 75 percent of the topsoil being removed, medium erosion refers to 25 to 75 percent of topsoil being removed, while low erosion refers to less than 25 percent of the topsoil being removed. This map shows the cumulative erosion during the Dust Bowl period, using data complied from the Soil and Conservation Service. This approach allows me to capture variation in the impact of the Dust Bowl in the Great Plains, by using soil erosion as a proxy[2]. I also utilize data of county-level New Deal spending (Fishback et al., 2006), as well as data on county-level demographic data to get control variables of county-level characteristics from the Census of Agriculture (Gibson & Jung, 2002).

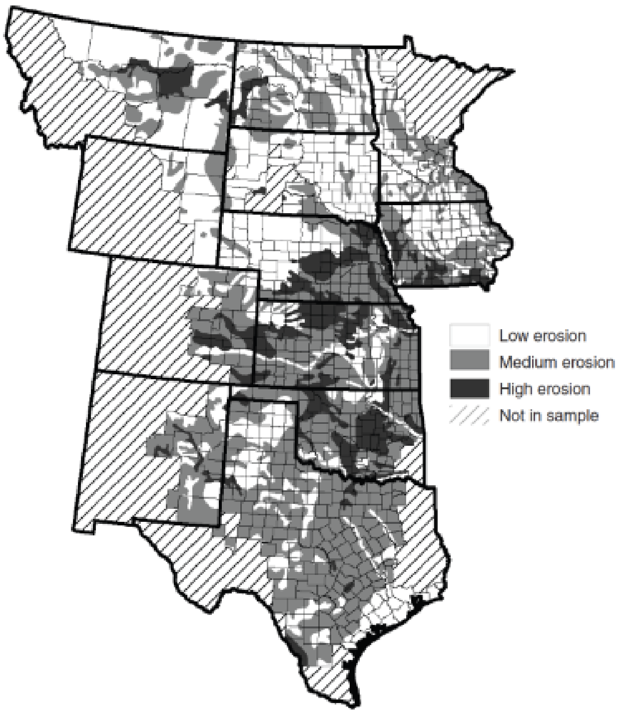

Figure 2: Erosion levels in the Great Plains (Hornbeck, 2012)

3.1 Summary Statistics

Let us look at the Dust Bowl region, identified by the counties within the boundaries in the figure (1) from (Cunfer, 2005). Figure 3 is a map showing the populations of the counties in 1930. We can see that many counties were largely agriculture-oriented, and so largely rural.

I also create heat maps showing the migrant destinations over the decade. For comparison, I look at migrant destinations from 1920 to 1930 (Figure 4), and then from 1930 to 1940 (Figure 5). We see that there is a westward push in migration, especially to Los Angeles County and the

[2]The erosion data is adjusted to 1940 county borders, accounting for changes in counties, following (Hornbeck, 2010)

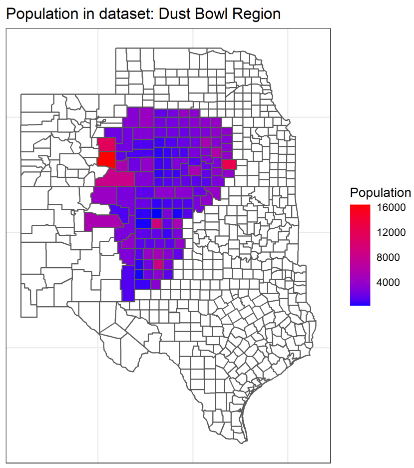

Figure 3: Population of counties in Dust Bowl region, 1930

San Joaquin Valley in Central California in both decades, although most migrants end up in nearby counties in the southern Great Plains.

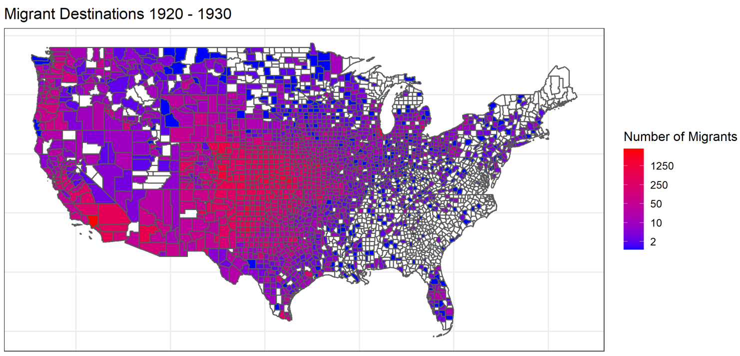

Figure 4: Migrant Destination Counties 1920-1930

This westward push is documented in American folklore, with vivid scenes of poor *Okies* (a pejorative name for poor Oklahoma migrants that arose from the Dust Bowl) leaving behind everything in search of opportunity in the Pacific States. John Steinback alludes to this in *The Grapes of Wrath*, with the protagonist documenting the hardships faced by refugees in California. In this paper, I will also conduct analysis to look at Dust Bowl migrants in California as well (Appendix C). Indeed,

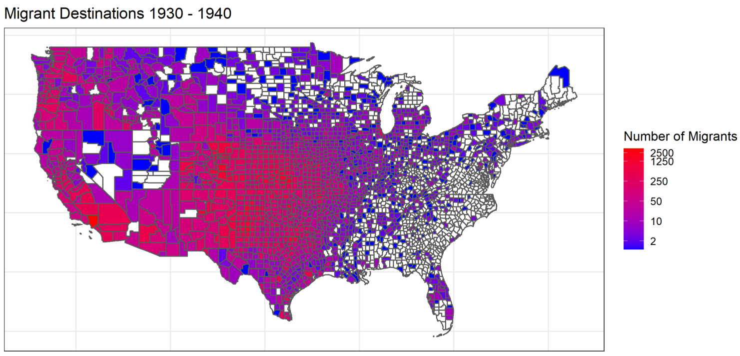

Figure 5: Migrant Destination Counties 1930-1940

we can see that the average distance migrated was higher in the 1930s than it was in the 1920s for migrants from the Dust Bowl regions: 322.92 miles compared to 303.41 miles in 1930. (Table 1). This could reflect a change in migratory patterns, with migrants moving outside the panhandle counties to newer migrant destinations.

Tables 2 and 3 show the transition matrix for the occupations in 1930 and 1940 for both non-migrants and migrants from the Dust Bowl region [3]. This preliminary analysis shows that migrant farmers seemed to have more upward occupational mobility relative to non-migrants, contrary to the vivid images of despair among migrants. However, there are no clear trends of occupational mobility for other occupation groups. In my empirical specification, I will look into both migrant selection and migrant outcomes for Dust Bowl migrants. One key thing to note is that my analysis only captures the earnings at the start and end of the decade, and one limitation is that I do not capture variation within the decade (due to limitations of the full census). My comparison with the decade prior (1920s) via a difference in differences model serves to mitigate some of these concerns.

[3] I divide this into four categories based on 1950 occupation delineations: high skilled (professionals and managers), semi-skilled (agents, sales workers, craftsmen, service workers etc.), laborers (farm laborers and general laborers), and farmers

Table 1: Distance Migrated (Miles): Dust Bowl Migrants

	Mean	Median	Std. Dev.	25th Percentile	75th Percentile
1920	303.41	161.29	336.15	62.63	396.58
1930	322.92	195.35	342.02	68.03	413.26

Table 2: Occupation Matrix: Dust Bowl Non-Migrants

	Occupation 1940				
Occupation 1930	High Skilled	Medium Skilled	Laborer	Farmer	Observations
High Skilled	**67.9**	22.0	3.4	6.4	10697
Medium Skilled	15.3	**71.5**	8.8	4.4	23828
Laborer	11.6	33.1	**29.5**	25.9	29206
Farmer	3.0	6.1	8.2	**82.6**	47449
Observations	15704	31956	15009	48511	111180

Table 3: Occupation Matrix: Dust Bowl Migrants

	Occupation 1940				
Occupation 1930	High Skilled	Medium Skilled	Laborer	Farmer	Observations
High Skilled	**55.1**	30.3	5.1	9.5	6310
Medium Skilled	16.5	**61.5**	11.3	10.7	13530
Laborer	12.7	38.3	**28.7**	20.3	17237
Farmer	7.4	21.2	22.6	**48.8**	18309
Observations	9245	20737	10924	14480	55386

4 Theoretical Framework

I aim to show how migrants during unexpected shocks differ from migrants in regular circumstances. While economic shocks may alter migration patterns, environmental shocks may actually induce different migration responses. Furthermore, permanent and long-lasting shocks such as the Dust Bowl may induce different responses and incentives over temporary environmental shocks. Understanding these difference can help us understand migrant decisions and how susceptible they are to different shocks.

Initial theory on migration revolved around aggregate migration flows, such as derivations of the Gravity model (i.e. focusing for net migration rates from areas with "repulsive" characteristics to "attractive" characteristics, and inversely proportional to distance). However, there has been a shift towards understanding micro-level decisions made by migrants and non-migrants. The Harris-

Todaro looks at how *expected* earnings (or utility) in both the initial location and destination plays a driving force for a person's decision to migrate, with the constraint of the cost of migrating (Harris & Todaro, 1970). This theory has been applied in an empirical context, with one study examining emigration of workers to South Africa to work in mines from bordering nations (Lucas, 1987), and another looking at the historic migration from Norway to the United States in the early 20^{th} century (Abramitzky, Boustan, & Eriksson, 2013). In my case, I wish to look at how an individuals decision to migrate changes when faced with an exogenous environmental and economic shock.

If we view the Dust Bowl as an exogenous shock, we can develop a simple migration model at the micro-level, by defining "push" or "pull" factors. I build off of Everret Lee's push and pull theory of migration (Lee, 1966), to construct a simple model that incorporates the cost of migration and the overall "push" effect of the Dust Bowl, either through the direct effects of the dust storms, or the indirect effects on the local economy and environment. Consider the following model:

$$M_i = f(Push_i^a, \quad Pull_i^b, X_i, \quad C_i^b, \quad \epsilon_i)$$

Here, M_i is the probability that individual i migrates, which is the function of a random noise term ϵ_i, the push and pull factors of migrating from county a to county b ($Push_i^a, Pull_i^b$), the characteristics of i on which i is selected (X_i), and the cost of migration incurred by i on migrating from county a to county b (C_i).

Let U_i be i's subjective net benefit or utility of moving, such that:

$$U_i = g(Push_i^a, \quad Pull_i^b, \quad X_i, \quad C_i^{a,b})$$

And so we get the following reduced form equation, where an individual's decision to migrate is naturally increasing on the utility incurred from migrating:

$$M_i = m(U_i, \quad \epsilon_i)$$

In the context of the Dust Bowl, we can see how migration decisions change as a result of the shock, through its effect of the individual's subjective utility of migrating. The Dust Bowl would not only affect push and pull factors, but also migration costs. This gives rise to a revised reduced form equation:

$$M_i^{DB} = n(U_i^{DB}, \quad \epsilon_i)$$

In the next section, I will talk about my empirical specification, using the reduced form models above. I aim to show how the Dust Bowl induced *changes* in migration decisions, by affecting the individual's net benefit of migration. If U_i^{DB} is different to U_i, we will see changes in migration patterns during the Dust Bowl (i.e. $\|M_i^{DB} - M_i\| \geq 0$).

5 Empirical Specification

The models I use to look at *changes* in migration rates, migrant selection, and migrant outcomes draw inspiration from (Long & Siu, 2018) and (Hornbeck, 2020). I look at both definitions of the Dust Bowl: I first run my models to compare outcomes between individuals the *Dust Bowl region* with the rest of the country country (following (Long & Siu, 2018)), as these counties represent the heart of the Dust Bowl storms. To test for internal validity, I add onto the specification from (Hornbeck, 2020) to limit my sample to the Great Plains States, as the effects of the Dust Bowl were felt throughout the region, and compare migrants from high erosion and medium erosion counties to low erosion counties within the Great Plains. By focusing on a specific region, I strengthen my assumption of parallel trends. This specification works as I assume the Dust Bowl to be exogenous. Suppose we were to have the opposite case, where individuals were selected into the Dust Bowl or if farming practices exacerbated the effects of the Dust Bowl, making it endogenous. In that case, it is not clear how this would affect my estimators in the migration model, and whether the Dust Bowl would affect residents differently than it would a random sample.

As the dataset I use has a large linked sample ((Long & Siu, 2018) did the linking themselves, and got a sample of around 4210 between 1920 and 1940), and makes use of multiple censuses ((Hornbeck, 2020) makes use of the 1940s census to only compare migration from 1935 to 1940 rather than opting to link multiple censuses), I hope to extrapolate from these models to investigate migration and mobility during the Dust Bowl. Appendix A delves deeper into migration in the 1930s: As I have the respondents' county of residence in 1935, I can compare early-decade migrants to late-decade migrants during the Dust Bowl period. Appendix B looks into more factors Dust Bowl migrants were selected on, while Appendix C looks more into migrants who moved to California.

I will assert causality through difference-in-differences and difference-in-difference-in-differences estimators, by looking at how the Dust Bowl impacted migration. Such methods have been used in the context of migration and its impacts (David Card examines the Muriel Boatlift to understand the effects of the influx of Cuban Migrants on the local labor market using such an approach (Card, 1990)). As the Dust Bowl disproportionately affected certain regions, I can employ such models to my design to estimate the causal impact of the Dust Bowl on migration. To get the effect of the Dust Bowl on migration, I estimate the difference-in-differences estimator comparing the Dust Bowl region before and during the Dust Bowl, to the rest of the country. For this empirical specification, I opt to use a linear probability model over other discrete dependent variable approach as it provides an intuitive interpretation of coefficients (in terms of its sign and magnitude) to estimate the impacts of the Dust Bowl.

Below is the estimator of the DD coefficient:

$$\widehat{\beta}_{DD} = (\overline{M}_{DB=1,post=1} - \overline{M}_{DB=1,post=0}) - (\overline{M}_{DB=0,post=1} - \overline{M}_{DB=0,post=0})$$

Here, M represents the migrant decision (i.e. 1 if the person migrated, otherwise 0). The subscripts represent the treatment and control groups, as well as for the pre and post-period, to illustrate the differences between them to obtain the estimator. In addition to this, for estimating migrant selection and migrant outcomes, consider the following model regressing dependent variable Y_i on the migrant dummy variable M_i. This model compares differences between migrants and non migrants.

$$Y_i = \delta \cdot M_i + \varepsilon_i$$

When I combine this to the DD estimator above, I obtain the following DDD estimator. This enables me to compare migrants from the Dust Bowl and non-migrants.

$$\widehat{\beta}_{DDD} = [(\overline{Y}_{DB=1,M=1,post=1} - \overline{Y}_{DB=1,M=1,post=0}) - (\overline{Y}_{DB=1,M=0,post=1} - \overline{Y}_{DB=1,M=0,post=0})] - [(\overline{Y}_{DB=0,M=1,post=1} - \overline{Y}_{DB=0,M=1,post=0}) - (\overline{Y}_{DB=0,M=0,post=1} - \overline{Y}_{DB=0,M=0,post=0})]$$

This estimator indicates the effects of the Dust Bowl on migrants from regions affected by the Dust Bowl. In the context of the migrant selection models, the dependent variable Y is the set of factors on which migrants are selected (e.g. pre-migration incomes, or farm status), and in the context of the migrant outcome models, Y represents the post-migration income to reflect how Dust Bowl migrants fared. I will expand on my specifications to reflect levels of intensity of soil erosion, to incorporate the variation in the effects of the Dust Bowl within the Great Plains Region.

5.1 Trends in Migration Rates

How likely were individuals in the Dust Bowl region likely to migrate? I employ difference-in-differences models on both definitions of the Dust Bowl: using the cluster of counties which felt the greatest effect of the dust storms, and using soil erosion as a proxy for Dust Bowl exposure in the Great Plains. The former allows me to look at the effects of a crisis (how strong the dust storms were in inducing migrants), while the latter allows me to compare migrants within the Great Plains and look at how soil erosion and degradation induced migrants, ceteris paribus.

5.1.1 Dust Bowl Region

Consider the simple difference in differences model in (1). The variable of interest is $(DB_i \cdot Post_t^{1930})$, for which the coefficient β represents the change in migration rate during the Dust Bowl period in the 1930's, relative to the period before the Dust Bowl (the 1920s). $X_{it}^{'}$ is a vector of control variables: including the $Post_t^{1930}$ and DB_i variables, which are dummy variables for the Dust Bowl period (the 1930s) and the Dust Bowl counties respectively. Other control variables include socioeconomic and demographic variables, including race [4], skill-level [5], whether the individual resides in an urban area, whether they live on a farm, the number of children they have, whether they own a house, and their age. I also include whether they reside in the same state as their birth-state and the same state as their parent's birth-state as a proxy for how strong their networks and connection with the Dust Bowl are.

$$mig_{it} = \beta_1 \cdot (DB_i \cdot Post_t^{1930}) + \Gamma X_{it}^{'} + \alpha_i + \varepsilon_{it} \tag{1}$$

where for t $\in$ {1920, 1930} , mig_{it} is 1 if the county of residence in the initial decennial census is different from the county of residence in the census taken ten years later, and 0 if the counties are the same - indicating that the respondent has not moved between the time the census was taken.

[4] I look at how outcomes varied for Black Americans and Native Americans, as the 1930s also saw the Great Migration: which saw Black Americans migrate from the South to the North and West. The Southern Great Plains also had significant Native American populations, due to the many reservations in the area.

[5] I use the same definitions as in table 2

Table 4: Migration Model: Dust Bowl Region

	(1) Baseline Model	(2) Controls Added	(3) County FE	(4) State FE
$(DB_i \cdot Post_t^{1930})$	0.0478***	0.0455***	0.0208**	0.0409
	(0.00121)	(0.00118)	(0.00714)	(0.0228)
DB_i	0.137***	0.0758***		0.0331
	(0.000897)	(0.000877)		(0.0166)
$Post_t^{1930}$	-0.0672***	-0.0634***	-0.0605***	-0.0642***
	(0.000108)	(0.000105)	(0.00309)	(0.00329)
Constant	0.229***	0.488***	0.446***	0.463***
	(0.0000831)	(0.000252)	(0.00334)	(0.00745)
N	55243751	54836954	54836954	54836954
adj. R^2	0.009	0.065	0.050	0.053

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results for this are in Table 4. For brevity, I include the interaction term, as well as the dummy variables for the Dust Bowl period and Dust Bowl region. Appendix C shows the coefficients of the control variables (see table 13). I run the baseline model without any fixed effects or control variables, then the model with the controls described above, then the model with controls and county fixed effects, and finally with controls and state fixed effects.[6]

We can see that it is the case that migration rates were higher in the Dust Bowl region. In the baseline model, we see that Dust Bowl residents were 4.78 percent more likely to migrate than the rest of the country. After adding controls for demographic variables I also find that the value on the coefficient does not change much (a 4.55 percent increase). Controlling for county fixed effects however reduces the coefficient to 2.1 percent, while when incorporating state fixed effects, the coefficient is not significant.

To put this into context with the overall trends in migration around the country, let us look at the baseline model in column (1). Prior to the Dust Bowl, residents in the Dust Bowl region were 13.7 percent more likely to migrate relative to the rest of the country, and this increased to 18.48 percent during the Dust Bowl. Overall, out-migration actually *fell* in both the Dust Bowl region

[6]For the fixed effects I use the initial county or state of residence. For all models, robust standard errors are reported in parenthesis, and are clustered at the county level.

and elsewhere during the Dust Bowl. However, migration rates in the rest of the country fell by 6.72 percent, whereas in the Dust Bowl region, it fell by only 1.94 percent during the Dust Bowl. This is not surprising due to the effects of the Great Depression which ravaged the entire country- although it seems the push factor from the Dust Bowl provided a great incentive to migrate despite the recession. These results are in line with (Long & Siu, 2018), who do not opt for a difference in difference approach, but rather running multiple probit models for each decade.

5.1.2 Great Plains Region

The specification above compares the cluster of counties in the heart of the Dust Bowl region with the rest of the country. To further check how the Dust Bowl induced migrants (rather than the effects of the Great Depression), I will now concentrate on the Great Plains Region, and utilize the soil erosion intensity data from (Hornbeck, 2012). By focusing on variation in erosion within the great plains, the difference in differences model can be adjusted to have a fairer comparison group (low erosion counties in the great plains), thus further strengthening the parallel trends claims for the difference in differences empirical specification. One drawback of the specification in 1 is that we are comparing those in a cluster of counties to the entire population (although results are significant even when controls and county fixed effects are added), and so this specification adds a layer of robustness to my findings.

Consider the baseline difference in differences model looking at migration rates in the Great Plains. This time, I account for differences in erosion intensity. by comparing counties with high erosion and medium erosion to counties with low erosion. H_i and M_i represent the percent of the county's land with high soil erosion and medium soil erosion respectively. The variables of interest are $(H_i \cdot Post_t^{1930})$ and $(M_i \cdot Post_t^{1930})$, where β_1 and β_2 estimate changes in migration rates in high erosion counties and medium erosion counties relative to low erosion counties during the Dust Bowl, *within* the Great Plains. $X_{it}^{'}$ is a vector of control variables: including the $Post_t^{1930}$, H_i and M_i variables, along with the control variables from equation 1.

$$mig_{it} = \beta_1 \cdot (H_i \cdot Post_t^{1930}) + \beta_2 \cdot (M_i \cdot Post_t^{1930}) + \Gamma X_{it}^{'} + \alpha_i + \varepsilon_{it} \tag{2}$$

The results are in Table 5, the model specifications are similar to Table 4. Once again, when I look at migration rates in the control group: residents in low erosion counties, I find that in all specifications, out-migration increased in high erosion counties relative to low erosion counties, while the claim is less strong for those in medium erosion counties. While out-migration was higher by 4 percent in high erosion counties than in low erosion county ones during the Dust Bowl (with the result being significant even with county/state fixed effects and added controls), Medium erosion county residents are only a bit more likely to migrate during the Dust Bowl than before (the coefficient is significant only in the model with control variables- and is only around 1 percent).

Table 5: Migration Model: Great Plains Region

	(1) Baseline Model	(2) Controls Added	(3) County FE	(4) State FE
$(H_i \cdot Post_t^{1930})$	0.0388***	0.0418***	0.0336***	0.0398*
	(0.00126)	(0.00121)	(0.00918)	(0.0141)
$(M_i \cdot Post_t^{1930})$	-0.000229	0.0114***	0.00830	0.0123
	(0.000932)	(0.000900)	(0.00739)	(0.0148)
H_i	0.0464***	0.00380***		-0.00807
	(0.000923)	(0.000896)		(0.0216)
M_i	0.0165***	0.00457***		0.00166
	(0.000697)	(0.000675)		(0.0240)
$Post_t^{1930}$	-0.0659***	-0.0735***	-0.0735***	-0.0756***
	(0.000608)	(0.000589)	(0.00450)	(0.0110)
Constant	0.295***	0.606***	0.560***	0.587***
	(0.000457)	(0.000742)	(0.00413)	(0.0168)
N	8680466	8620096	8620096	8620096
adj. R^2	0.006	0.079	0.065	0.068

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

To put this into context, and compare trends between counties with high and medium levels of erosion relative to low levels of erosion, consider column (2) from Table 5. Prior to the Dust Bowl, I find that relative to individuals from low erosion counties, individuals from high erosion counties were only 0.38 percent more likely to migrate, while individuals from medium erosion counties were 0.457 percent more likely to migrate. However, out-migration rates fell by 7.35 percent in low erosion counties during the Dust Bowl, while they only fell by 6.21 percent and 3.17 percent for medium and high erosion counties respectively. In fact, during the Dust Bowl, relative to low erosion counties, high erosion counties had a 4.56 percent larger out-migration rate, relative to only 1.6 percent for medium erosion counties. Similar to my findings in the Dust Bowl region, I find similar trends when accounting for erosion intensity: while there are larger disparities in out-migration for high erosion counties, it is much less so for medium erosion counties (it is important to note that in this case, the comparison group is not the rest of the country, but rather counties in the same Great Plains region with relativity low levels of erosion).

5.2 Migrant Selection

How were migrants selected in the Dust Bowl? Pop culture references to the Dust Bowl allude to images of poverty and the plight of the farmer refugees. However empirical analysis seems to contradict this, with claims that it was actually the wealthy, professional class that could move. While selection varies across migrants in regular circumstances based on the relative costs and benefits they incur from migrating, economic shocks, and *environmental* shocks, can certainly change a person's incentive to migrate, and so migrants during a shock may vary from migrants during regular circumstances. The following specifications aim to understand migrant selection during the Dust Bowl, and also how they changed relative to previous periods. Below, I look at migrant selectivity in terms of pre-migration income. In Appendix B, I will look at other factors that migrants were selected on during the Dust Bowl. Appendix C looks at changes in migrant selectivity for migrants who moved to California.

5.2.1 Dust Bowl Region

Consider the following difference-in-difference-in-differences model. The dependant variable is the log of the earnings at the start of the decade [7] (essentially the log of pre-migration income). My variable of interest is $(DB_i \cdot mig_i \cdot Post_t^{1930})$, which reflects the change in migrant selectivity (in

[7] As income is only asked in the 1940's census, I use the *OCCSCORE* variable, which is a proxy for the income, based on the median income for the person's occupation. One drawback of this approach is that while it captures differences between occupations, it does not account for variation within an occupation. This variable is used in other papers dealing with early 20^{th} century census (Abramitzky et al., 2012), (Collins & Wanamaker, 2015) and (Long & Siu, 2018)

terms of pre-migration income) from the Dust Bowl counties as a result of the Dust Bowl.

The triple interaction term ($DB_i \cdot mig_i \cdot Post_t^{1930}$) shows the *change* in migrant selection as a result of the Dust Bowl. I also include the double interaction terms ($DB_i \cdot mig_i$) , ($DB_i \cdot Post_t^{1930}$), and ($mig_i \cdot Post_t^{1930}$). $\Gamma X_{it}^{'}$ is a vector of control variables: including the $Post_t^{1930}$, mig_i and DB_i dummy variables (indicating the 1930s decade, migrant and the Dust Bowl region respectively), along with variables to control for skill: education level [8], age, and age squared.

$$\begin{aligned} log(Y_{it}) = \beta \cdot (DB_i \cdot mig_i \cdot Post_t^{1930}) + \gamma_1 \cdot (DB_i \cdot mig_i) + \gamma_2 \cdot (DB_i \cdot Post_t^{1930}) + \\ \gamma_3 \cdot (mig_i \cdot Post_t^{1930}) + \Gamma X_{it}^{'} + \alpha_i + \varepsilon_{it} \end{aligned} \tag{3}$$

The results are in Table 6. Column (1) is the baseline model, while Column (2) adds the control variables described above. Columns (3) and (4) follow the specification from (Hornbeck, 2020) to look at the out-selection and in-selection of migrants. This is done by varying the county fixed effects α_i: when fixed effects refer to the initial county at the start of the decade, I compare Dust Bowl migrants and Dust Bowl non-migrants in the initial county (out-selection). When the fixed effects refer to the final county at the end of the decade, I compare Dust Bowl migrants and the natives of their destination counties. My variable of interest is ($DB_i \cdot mig_i \cdot Post_t^{1930}$), which shows the change in selectivity as a result of the Dust Bowl.

From column (1) I find that migrant selectivity fell during the Dust Bowl, indicated by the negative coefficient for ($DB_i \cdot mig_i \cdot Post_t^{1930}$). In other words, the general fall in migrant selectivity seems to indicate the change in composition and characteristics of migrants as a result of the Dust Bowl: it seems that the Dust Bowl induced migrants with lower pre-migration incomes. However, column (2) this result is not significant when controlling for skill level.

However, the fixed effects models paint an interesting picture. From column (3) there is no compelling evidence of a change out-selection (comparing migrants and non-migrants in the Dust Bowl), however there is evidence that migrant selectivity fell when we account for in-selection (comparing differences between migrants and natives) - migrants from the Dust Bowl region in the 1930s

[8] As there is no proper education variable that is asked in all three (1920, 1930 and 1940) censuses, I use the proxy variable *EDSCOR50*, which is the proportion of having a bachelors degree for each occupation in 1950

Table 6: Migration Selection Model: Dust Bowl Region

	(1) Baseline Model	(2) Controls Added	(3) County FE: Out-selection	(4) County FE: In-selection
$(DB_i \cdot mig_i \cdot Post_t^{1930})$	-0.0128**	0.000120	0.00932	-0.0199**
	(0.00417)	(0.00344)	(0.00766)	(0.00702)
$(DB_i \cdot mig_i)$	0.0467***	0.0428***	-0.00396	-0.0130
	(0.00311)	(0.00254)	(0.00666)	(0.0158)
$(DB_i \cdot Post_t^{1930})$	0.0276***	0.0332***	0.0345***	0.0372***
	(0.00242)	(0.00193)	(0.00612)	(0.00642)
$(mig_i \cdot Post_t^{1930})$	0.0129***	-0.0102***	-0.00848***	0.0382***
	(0.000564)	(0.000477)	(0.00212)	(0.00398)
DB_i	-0.219***	-0.200***		-0.0468**
	(0.00181)	(0.00143)		(0.0158)
mig_i	-0.0238***	-0.0226***	0.0324***	-0.0252***
	(0.000396)	(0.000337)	(0.00239)	(0.00642)
$Post_t^{1930}$	0.0643***	0.0336***	0.0197***	0.0205***
	(0.000237)	(0.000198)	(0.00231)	(0.00239)
Constant	7.585***	6.433***	6.564***	6.545***
	(0.000181)	(0.000958)	(0.0281)	(0.0269)
N	25350341	25312598	25312598	25312598
adj. R^2	0.006	0.303	0.288	0.290

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

had lower pre-migration incomes relative to natives in their destination county than migrants in the decade prior. Essentially, we cannot conclude that the Dust Bowl saw changes in disparities in migrants and non-migrants from the Dust Bowl region, but we see that migrants had lower incomes relative to natives in their destination counties during the Dust Bowl. Overall, the results point to decreases in migrant selectivity in the Dust Bowl than before, although the results are not statistically significant in some specifications (Columns (2) and (3)).

To put this into context consider the baseline model in Column (1). If we look at the pre-period (i.e. $Post_t^{1930} = 0$), migrants in the Dust Bowl region tended to have a 4.8 percent higher income, and so were positively selected on income. During the Dust Bowl however (i.e. $Post_t^{1930} = 1$), migrants from the Dust Bowl region now only had 3.97 percent higher income. While migrants in the Dust Bowl were still positively selected on income, they had lower pre-migration incomes than migrants who migrated prior to the Dust Bowl. The Dust Bowl saw a 1.27 percent decrease in selectivity on income for migrants from the Dust Bowl region. When we look at in-selection, Dust Bowl migrants had lower selectivity with respect to pre-migration incomes relative to natives. However, these differences are rather small to conclude a paradigm shift in the composition of migrants.

5.2.2 Great Plains Region

I repeat the specification in equation 3, to now incorporate differences in erosion intensity within the Great Plains. I now have the following difference-in-difference-in-differences. $(H_i \cdot mig_i \cdot Post_t^{1930})$ and $(M_i \cdot mig_i \cdot Post_t^{1930})$ reflect changes in migrant selection during the Dust Bowl for migrants from high erosion and medium erosion counties respectively. $X_{it}^{'}$ is a vector of control variables: including the $Post_t^{1930}$, H_i and M_i variables, along with the demographic variables from above. This way I can compare migrant selectivity within the great plains, and see how disparities in soil erosion induced migration.

$$
\begin{aligned}
log(Y_{it}) = \beta_1 \cdot (H_i \cdot mig_i \cdot Post_t^{1930}) + \beta_2 \cdot (M_i \cdot mig_i \cdot Post_t^{1930}) + \\
\gamma_1 \cdot (H_i \cdot mig_i) + \gamma_2 \cdot (H_i \cdot Post_t^{1930}) + \gamma_3 \cdot (M_i \cdot mig_i) + \gamma_4 \cdot (M_i \cdot Post_t^{1930}) + \\
\gamma_5 \cdot (mig_i \cdot Post_t^{1930}) + \Gamma X_{it}^{'} + \alpha_i + \varepsilon_{it}
\end{aligned}
\tag{4}
$$

The results are in table 7. The structure of the table is similar to that in table 6. When controlling our sample to the Great Plains Region, I find interesting results. From columns (1) and (2), the positive coefficient on $(H_i \cdot mig_i \cdot Post_t^{1930})$ shows that for high erosion counties, migrant selectivity actually increased during the Dust Bowl in terms of income (even controlling for skill). Columns (3) and (4) indicate that migrant selectivity increased relative to non-migrants and natives respectively during the Dust Bowl highlighting an increase in out and in-selection. Essentially, migrants from high erosion counties had higher pre-migration incomes during the Dust Bowl compared to non-migrants in their home county *and* natives in their destination county, relative to migrants prior to the Dust Bowl.

Table 7: Migration Selection Model: Great Plains Region

	(1) Baseline Model	(2) Controls Added	(3) County FE: Out-selection	(4) County FE: In-selection
$(H_i \cdot mig_i \cdot Post_t^{1930})$	0.0616***	0.0597***	0.0532**	0.0364***
	(0.00491)	(0.00406)	(0.0189)	(0.0109)
$(M_i \cdot mig_i \cdot Post_t^{1930})$	0.0164***	0.00908**	0.0144	-0.00622
	(0.00386)	(0.00317)	(0.00758)	(0.00774)
$(H_i \cdot mig_i)$	-0.0444***	-0.0233***	-0.0312***	-0.00372
	(0.00355)	(0.00289)	(0.00946)	(0.0304)
$(H_i \cdot Post_t^{1930})$	0.0104***	0.00472*	0.00947	0.00935
	(0.00261)	(0.00205)	(0.00914)	(0.00936)
$(M_i \cdot mig_i)$	-0.0545***	-0.0267***	-0.0221**	-0.0189
	(0.00274)	(0.00223)	(0.00812)	(0.0245)
$(M_i \cdot Post_t^{1930})$	0.0194***	0.00423**	0.0136*	0.0130
	(0.00200)	(0.00157)	(0.00656)	(0.00679)
$(mig_i \cdot Post_t^{1930})$	-0.0161***	-0.0289***	-0.0285***	0.00879
	(0.00256)	(0.00211)	(0.00443)	(0.00534)
H_i	0.00714***	0.00422**		-0.0320
	(0.00194)	(0.00152)		(0.0277)
M_i	0.00423**	0.00669***		-0.00756
	(0.00151)	(0.00118)		(0.0269)
mig_i	0.0566***	0.0462***	0.0643***	-0.00787
	(0.00182)	(0.00149)	(0.00422)	(0.0158)
$Post_t^{1930}$	0.0585***	0.0368***	0.0167***	0.0183***
	(0.00131)	(0.00104)	(0.00355)	(0.00371)
Constant	7.406***	6.074***	6.122***	6.122***
	(0.000992)	(0.00246)	(0.0281)	(0.0255)
N	4079437	4072810	4072810	4072810
adj. R^2	0.004	0.363	0.355	0.356

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Applying the same for medium erosion counties, the positive coefficient on ($M_i \cdot mig_i \cdot Post_t^{1930}$) in columns (1) and (2) indicates migrants from medium erosion counties saw a smaller increase in migrant selectivity relative to those in low erosion counties during the Dust Bowl. However, there is no compelling evidence for in- or out selection when we account for county fixed effects. Essentially, while medium erosion migrants had higher incomes during the Dust Bowl than migrants from low erosion counties, there is no evidence that it was higher relative to non-migrants and relative to natives in their destination counties.

Putting this into context, similar to before, consider the baseline model in column (1). Looking at the pre-Dust Bowl period in the 1920s (i.e. $Post_t^{1930} = 0$), migrants from medium erosion counties were negatively selected, and had a 5.6 percent lower income. Migrants from high erosion counties were also negatively selected, but had only a 4.3 percent lower income. When we look at the Dust Bowl period (i.e. $Post_t^{1930} = 1$), we see that migrants were negatively selected from medium erosion counties (but now only had 3.7 percent lower incomes). On the other hand, migrants from high erosion counties were *positively* selected during the Dust Bowl, and had 1.7 percent higher incomes.

When we restrict our sample to the Great Plains region, the Dust Bowl seems to have induced an increase in migrant selectivity in counties that experienced high and medium erosion relative to counties with less erosion, which seems contrary to the stereotypes of Dust Bowl migrants. When we look at in- and out-selection, the results are statistically significant for High Erosion counties, with increases in selectivity, but less so for Medium Erosion counties. While the results for the Dust Bowl region in the southern Great Plains did not show particularly significant results, we see stronger results when we incorporate differences within the Great Plains.

5.3 Migrant Outcomes

After looking at whether the Dust Bowl spawned more migrants, and *how* the Dust Bowl spawned migrants, I look at what happened to the Dust Bowl migrants. The image of the Dust Bowl migrant is one of penury, with many migrants being forced into refugee camps or having to settle for lower paying jobs in their new homes. I examine changes in post-migration incomes for Dust Bowl migrants, and look at the divisions across different occupational groups. Appendix C looks into the outcomes of migrants who moved to California.

5.3.1 Dust Bowl Regions

I will now use the same empirical specification as in (3), but now Y_{it} is now the individual's income a the end of the decade. This way we can see the individual's outcome at the end of the decade for both migrants and non-migrants, or the *post-migration* outcome, and how it changed as a result of the Dust Bowl. $X_{it}^{'}$ reflects the pre-migration income as a control variable, to reflect the *change* in outcomes for Dust Bowl migrants. Column (1) refers to the base line model with this control, while column (2) adds county fixed effects. To control for differences across occupational skill groups, I repeat the specification in column (2), for high-skilled individuals, medium-skilled individuals, laborers, and farmers (Columns (3), (4), (5) and (6) respectively). My variable of interest is $(DB_i \cdot mig_i \cdot Post_t^{1930})$ which indicates how post-migration incomes changed for migrants from the Dust Bowl relative to the decade prior.

The results are in table 8. We can see that incomes fell for Dust Bowl migrants relative to those who migrated in the decade prior. The reasons of this could vary from the depressed wages as a result of the Great Depression, or migrants not finding work or being forced to work in lower paying jobs in their migrant destinations. The negative coefficient on $(DB_i \cdot Migrant_i)$ indicates that migrants from the Dust Bowl region saw an increase in income on moving, however this fell during the Dust Bowl period. Columns (3) to (6) paint an interesting story. While there is no evidence that migrant outcomes changed for high-skilled, medium-skilled or laborers, (interestingly high-skilled and medium-skilled migrants saw declines in income in general, but there is no evidence of changing for those who migrated as a result of the Dust Bowl). However, farmers actually saw modest increases in incomes on migrating during the Dust Bowl.

If we put this into context, if we compare migrants to non-migrants from the Dust Bowl region in the 1920s, we find that migrants had around a 5.1 percent increase in income upon migrating. However, during the Dust Bowl, migrants from the Dust Bowl only saw a 3.1 percent increase on incomes upon migrating. Overall, migrants during the Dust Bowl saw a 1.96 percent lower incomes than non-migrants, than those who migrated prior to the Dust Bowl. While the results point to modest declines in incomes for Dust Bowl migrants, it is not clear whether this is a result of discrimination (migrants were forced to settle for lower paying jobs), or a result of the Great Depression depressing wages in their destination counties. Interestingly enough, the rise in incomes

Table 8: Labor Market Outcomes: Dust Bowl Region

	(1) Baseline Model	(2) County FE	(3) High Skilled	(4) Medium Skilled	(5) Laborers	(6) Farmers
$(DB_i \cdot mig_i \cdot Post_t^{1930})$	-0.0198***	-0.0173**	-0.0146	-0.00857	0.00672	0.0142*
	(0.00354)	(0.00612)	(0.0122)	(0.0102)	(0.0117)	(0.00616)
$(DB_i \cdot mig_i)$	0.0500***	0.0366***	-0.0262**	-0.0333***	-0.0258	-0.00616
	(0.00266)	(0.0105)	(0.00965)	(0.00960)	(0.0140)	(0.00571)
$(DB_i \cdot Post_t^{1930})$	0.0113***	0.0172***	0.00641	0.00944	-0.000351	-0.000555
	(0.00176)	(0.00435)	(0.00739)	(0.00626)	(0.00909)	(0.00304)
$(mig_i \cdot Post_t^{1930})$	-0.0414***	-0.0399***	0.0297***	0.0101***	-0.0669***	-0.0600***
	(0.000456)	(0.00244)	(0.00180)	(0.00304)	(0.00251)	(0.00148)
DB_i	-0.0616***					
	(0.00132)					
mig_i	0.0569***	0.0718***	-0.0700***	-0.0299***	0.133***	0.170***
	(0.000323)	(0.00460)	(0.00324)	(0.00490)	(0.00467)	(0.00168)
$Post_t^{1930}$	0.0139***	0.00809***	-0.0423***	-0.0133***	0.0108***	0.00601***
	(0.000169)	(0.00175)	(0.00118)	(0.00101)	(0.00158)	(0.000713)
$\log(earnings)$	0.535***	0.466***	0.841***	0.404***	0.244***	0.142***
	(0.000205)	(0.00247)	(0.00264)	(0.00468)	(0.00155)	(0.00524)
Constant	3.598***	4.122***	1.167***	4.704***	5.720***	6.259***
	(0.00158)	(0.0176)	(0.0221)	(0.0370)	(0.0106)	(0.0380)
N	21891274	21891274	2583351	9493165	4320001	5494757
adj. R^2	0.389	0.276	0.270	0.099	0.084	0.032

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

for migrant farmers could potentially be explained by many farmers moving to other occupations during the Dust Bowl relative to the decade prior (See tables 2 and 3). This fall in migrant incomes may also be explained by the fall in migrant selectivity in the first place, as the migrants who moved during the Dust Bowl tended to be poorer, and coupled with the effects of the Great Depression on wages in their destination counties, seeing declines in incomes in understandable within that context.

5.3.2 Great Plains Region

As above, I use the same empirical specification as in (4), with Y_{it} representing the individual's income a the end of the decade. Similar to the specification for the Dust Bowl region, this model now reflects the *change* in outcomes for migrants from high-erosion counties and medium erosion counties relative to migrants from low erosion counties in the Great Plains. ($H_i \cdot mig_i \cdot Post_t^{1930}$) and ($M_i \cdot mig_i \cdot Post_t^{1930}$) reflect this change in migrant outcomes for high erosion and medium erosion counties during the Dust Bowl relative to low erosion counties. This way, by restricting the empirical specification to the Great Plains, I can look at how migrants were impacted disproportionately by the Dust Bowl.

The results are in table 9, and is structured similarly to table 8. Columns (1) and (2) do not point to strong evidence of changing outcomes for migrants from high erosion counties during the Dust Bowl. However, there is evidence of rising incomes for migrants from medium erosion counties during the Dust Bowl. Similarly, we see that farmers and medium skilled migrants from medium erosion counties saw a rise in wages during the Dust Bowl, whereas there is no evidence for changes in outcomes across different occupational groups for high erosion counties. However, the coefficients for both models are rather small, despite the statistical significance, I can only find modest economically significant results. While migrant selectivity in high erosion and medium erosion counties increased during the Dust Bowl with respect to pre-migration income, I find that only migrants from medium erosion counties saw modest increases in incomes during the Dust Bowl.

Table 9: Labor Market Outcomes: Great Plains Region

	(1) Baseline Model	(2) County FE	(3) High Skilled	(4) Medium Skilled	(5) Laborers	(6) Farmers
$(H_i \cdot mig_i \cdot Post_t^{1930})$	-0.00283	-0.00298	-0.00440	-0.0160	0.0174	0.00682
	(0.00400)	(0.0109)	(0.0123)	(0.0157)	(0.0156)	(0.00928)
$(M_i \cdot mig_i \cdot Post_t^{1930})$	0.0128***	0.0151*	-0.00137	0.0142*	0.0113	0.0252**
	(0.00311)	(0.00634)	(0.0102)	(0.00692)	(0.0121)	(0.00785)
$(H_i \cdot mig_i)$	-0.0210***	-0.0225	-0.0144	-0.00434	-0.0611**	-0.0362***
	(0.00288)	(0.0176)	(0.0107)	(0.0118)	(0.0213)	(0.00961)
$(H_i \cdot Post_t^{1930})$	-0.0116***	-0.00783	-0.0190*	-0.00142	-0.0173	-0.00947*
	(0.00185)	(0.00583)	(0.00839)	(0.00689)	(0.0100)	(0.00428)
$(M_i \cdot mig_i)$	-0.0226***	-0.0208	0.000993	-0.00603	-0.0254	-0.0442***
	(0.00219)	(0.0150)	(0.00909)	(0.0102)	(0.0186)	(0.00701)
$(M_i \cdot Post_t^{1930})$	-0.0214***	-0.0194***	-0.00821	-0.0151**	-0.0176*	-0.0157***
	(0.00140)	(0.00510)	(0.00769)	(0.00553)	(0.00890)	(0.00414)
$(mig_i \cdot Post_t^{1930})$	-0.0439***	-0.0430***	0.0422***	0.0108*	-0.0604***	-0.0501***
	(0.00205)	(0.00444)	(0.00593)	(0.00475)	(0.00776)	(0.00548)
H_i	-0.000622					
	(0.00135)					
M_i	0.0146***					
	(0.00103)					
mig_i	0.102***	0.109***	-0.101***	-0.0465***	0.162***	0.175***
	(0.00146)	(0.00911)	(0.00521)	(0.00590)	(0.0116)	(0.00468)
$Post_t^{1930}$	0.0292***	0.0218***	-0.0421***	-0.00606	0.0131*	0.00889***
	(0.000913)	(0.00360)	(0.00445)	(0.00372)	(0.00596)	(0.00257)
$\log(earnings)$	0.505***	0.462***	0.811***	0.371***	0.249***	0.136***
	(0.000484)	(0.00315)	(0.00490)	(0.00455)	(0.00273)	(0.0112)
Constant	3.770***	4.098***	1.419***	4.972***	5.636***	6.294***
	(0.00366)	(0.0221)	(0.0413)	(0.0360)	(0.0182)	(0.0811)
N	3585376	3585376	372143	964722	754472	1494039
adj. R^2	0.353	0.288	0.243	0.077	0.094	0.033

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6 Conclusion

The specifications above aim to illustrate how migration was affected as a result of the Dust Bowl. The migration models show conclusive evidence of migration rates being higher in the Dust Bowl region relative to the rest of the country. Within the great plains, we see increases in migration rate from counties with high and medium levels of erosion, relative to counties with low levels of erosion. Interestingly, migration rates fell significantly around the country, but only fell slightly during the Dust Bowl. Even the effects of an overall period of recession around the country could not overcome the push factor of the dust storms and land degradation. By controlling my sample to the Great Plains region, I get to capture varying effects of the Dust Bowl (by using soil erosion intensity as a proxy). I also look at the case of California migrants, in an attempt to examine trends for the westward push. Looking at when migrants moved during the Dust Bowl also illustrates disparities in the region. By comparing early-decade and late-decade migrants, I find that farmers were more likely to move after 1935. I also find that migration to California increased during the Dust Bowl, and was higher than the average migration during the Dust Bowl.

The migrant selectivity and migrant outcome models serve to answer the following: which kind of migrants were induced to migrate as a result of the Dust Bowl, relative to general trends in migration, and how did migrants fare during the Dust Bowl relative to general trends? While the promise of higher wages could drive migration, and the shock of a climate disaster could spawn different kinds of migrants who would not have otherwise migrated, did the Dust Bowl change the status quo on internal migration? If so, how much of that is due to the overall Great Depression, and how much is due to the Dust Bowl?

I find that migrants from the Dust Bowl region in the southern great plains were selected more negatively relative to natives during the Dust Bowl. However, when I control for migrants from the Great Plains region, migrants from high erosion and medium erosion counties had higher selection during the Dust Bowl than before. When I find that migrant selectivity increased for migrants from counties with high levels of erosion during the Dust Bowl period, relative to those in low erosion counties. In other words, migrants tended to have slightly higher incomes during the Dust Bowl than migrants before the Dust Bowl, contrary to stereotypes. I only find evidence that migrants from high erosion counties had increased in- and out-selection (i.e. they had higher incomes than non-migrants and natives during the Dust Bowl). When I look at California migrants, however, I find decreases in migrant selectivity during the Dust Bowl, from Dust Bowl counties and medium erosion counties (not so much for migrants from high erosion counties). When I look at other factors that migrants could potentially be selected on, I find modest changes in selectivity for Dust Bowl migrants.

When I look at migrant outcomes, migrants from the Dust Bowl region in the southern great plains saw decreases in incomes during the Dust Bowl. However, I find moderate increases in incomes for farmers who migrated. However, when I restrict my sample to the Great Plains region, I find modest increases in incomes for migrants from medium erosion counties during the Dust Bowl. On the other hand, when I look at migrants who moved to California during the Dust Bowl, I find decreases in post-migration incomes, and this trends holds across all occupations (high-skilled, medium-skilled, laborers and farmers). The decreases for California migrants are larger than for general Dust Bowl migrants.

There were some limitations to this approach however. As the full census only provides information of residence in the year the census was taken, I do not have information on temporary displaced migrants, and when during the decade the person migrated. The linking approach is currently being improved, and at the time of writing there has been no option to provide weights to the linked sample. This means that certain demographics may be under-represented in my sample (such as married women who change their surnames, and those who migrated), along with general errors in the census records. However, with new iterations of the linking algorithm, robustness checks can be performed to see if the results still hold with the new samples. Improvements in machine learning technologies can also help scale up the linking process, with better accuracy.

These results have interesting implications on future policy regarding climate change and climate induced migration. The parallels between the two are interesting (in that they disproportionately affect certain regions more than others). Although one could point that the Dust Bowl was a shock, while climate change and rising sea levels are more gradual, it is important to note that the Dust Bowl was a brewing storm long before the 1930s, although the severity of the Dust Bowl was unanticipated. In terms of the migration question, the Dust Bowl saw changes in migration rates, but only modest changes migrant selection and migrant outcomes, and its effects were felt much after the Dust Bowl as well. More research must be done to see how potential climate-induced disasters affect local populations at the micro-level, and how they induce changes in migration. I find that farmers who were arguably the most affected by the Dust Bowl, were actually less likely to migrate during the Dust Bowl, and the Dust Bowl did not induce changes in migration patterns and migrant characteristics.

Could the same be said for climate change? While the effects of climate change would be felt over a larger area than the Dust Bowl, affect a much larger population and significantly affect global economies (Nordhaus, 2019), would it induce a paradigm shift in migration flows? Future research should also look at the long term impacts of Dust Bowl migrants. How did migrants fare relative to non-migrants and natives in the long-run? I would also like to see more work on the impacts of migrants on local labor markets, and how justified reports of discrimination in California and other Pacific States were. The 1930s saw different shocks that could have potentially affect migration responses, from economic shocks such as the Great Depression, to environmental shocks such as the Dust Bowl.

There is however, limited literature on examining the difference between economic and environmental shocks on migration, and even less so when we think about temporary shocks compared to long-term systematic changes, and when we compare "anticipated" versus unanticipated shocks. The last point serves to highlight the challenges to the external validity of my findings. In what ways are the shocks from climate change different from Dust Bowl shocks? One could note that there was little to no preparation undertaken to prevent the Dust Bowl, but can the same be said for climate change? If not, how would that induce migratory responses? By understanding how migratory incentives and costs change with respect to different shocks, we can develop a better understanding about the driving forces of migration, not only in terms of flows, but also in terms of selection, adaptation and outcomes.

The Dust Bowl was a dark chapter in American history, set against the backdrop of one of the roughest periods in American history. Understanding its impacts on the local populations through empirical analysis can serve to better understand the effects of the Dust Bowl, and to provide context to migratory responses to one of the most extraordinary climatic shocks.

References

Abramitzky, R., & Boustan, L. (2017). Immigration in american economic history. *Journal of economic literature*, *55*(4), 1311–45.

Abramitzky, R., Boustan, L., Jacome, E., & Perez, S. (2021). Intergenerational mobility of immigrants in the united states over two centuries. *American Economic Review*, *111*(2), 580–608.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, *102*(5), 1832–56.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2013). Have the poor always been less likely to migrate? evidence from inheritance practices during the age of mass migration. *Journal of Development Economics*, *102*, 2–14.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy*, *122*(3), 467–506.

Abramitzky, R., Mill, R., & Pérez, S. (2020). Linking individuals across historical sources: a fully automated approach. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *53*(2), 94–111.

Arthi, V. (2018). " the dust was long in settling": Human capital and the lasting impact of the american dust bowl. *Journal of Economic History*, *78*(1), 196–230.

Bohra-Mishra, P., Oppenheimer, M., & Hsiang, S. M. (2014). Nonlinear permanent migration response to climatic variations but minimal response to disasters. *Proceedings of the National Academy of Sciences*, *111*(27), 9780–9785.

Borjas, G. J. (1987). *Self-selection and the earnings of immigrants* (Tech. Rep.). National Bureau of Economic Research.

Boustan, L. P. (2016). *Competition in the promised land: Black migrants in northern cities and labor markets*. Princeton University Press.

Boustan, L. P., Fishback, P. V., & Kantor, S. (2010). The effect of internal migration on local labor markets: American cities during the great depression. *Journal of Labor Economics*, *28*(4), 719–746.

Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *ILR Review*, *43*(2), 245–257.

Cattaneo, C., & Peri, G. (2016). The migration response to increasing temperatures. *Journal of Development Economics*, *122*, 127–146.

Clay, K., & Jones, R. (2008). Migrating to riches? evidence from the california gold rush. *The Journal of Economic History*, 997–1027.

Collins, W. J., & Wanamaker, M. H. (2015). The great migration in black and white: New evidence on the selection and sorting of southern migrants. *The journal of economic history*, *75*(4), 947–992.

Cunfer, G. (2005). *On the great plains: agriculture and environment* (No. 20). Texas A&M University Press.

Cunfer, G. (2011). The southern great plains wind erosion maps of 1936–1937. *Agricultural History*, *85*(4), 540–559.

Davis, C. C. (1936). The agricultural adjustment act and national recovery. *Journal of Farm Economics*, *18*(2), 229–241.

Desmet, K., & Rossi-Hansberg, E. (2015). On the spatial economic impact of global warming. *Journal of Urban Economics*, *88*, 16–37.

Donaldson, D., & Hornbeck, R. (2016). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, *131*(2), 799–858.

Duncan, O. D. (1961). A socioeconomic index for all occupations. *Class: Critical Concepts*, *1*, 388–426.

Ferrie, J. P. (1996). A new sample of males linked from the public use microdata sample of the 1850 us federal census of population to the 1860 us federal census manuscript schedules. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *29*(4), 141–156.

Fishback, P. V., Horrace, W. C., & Kantor, S. (2006). The impact of new deal expenditures on mobility during the great depression. *Explorations in Economic History*, *43*(2), 179–222.

Galenson, D. W., & Pope, C. L. (1989). Economic and geographic mobility on the farming frontier: Evidence from appanoose country, iowa, 1850-1870. *Journal of Economic History*, 635–655.

Gibson, C., & Jung, K. (2002). *Historical census statistics on population totals by race, 1790 to 1990, and by hispanic origin, 1790 to 1990, for the united states, regions, divisions, and states.* US Census Bureau Washington, DC.

Gregory, J. N. (1989). Dust bowl legacies: The okie impact on california, 1939-1989. *California History*, *68*(3), 74–85.

Gregory, J. N. (1991). *American exodus: the dust bowl migration and okie culture in california.* Oxford University Press, USA.

Hansen, Z. K., & Libecap, G. D. (2004). Small farms, externalities, and the dust bowl of the 1930s. *Journal of Political Economy*, *112*(3), 665–694.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review*, *60*(1), 126–142.

Hornbeck, R. (2010). Barbed wire: Property rights and agricultural development. *The Quarterly Journal of Economics*, *125*(2), 767–810.

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.

Hornbeck, R. (2020). Dust bowl migrants: Environmental refugees and economic adaptation.

Hornbeck, R., & Naidu, S. (2014). When the levee breaks: black migration and economic development in the american south. *American Economic Review*, *104*(3), 963–90.

Hudson, J. C. (1982). Towns of the western railroads. *Great Plains Quarterly*, *2*(1), 41–54.

Hurt, R. D. (1981). *The dust bowl: an agricultural and social history.* Taylor Trade Publications.

Joel, A. H. (1937). *Soil conservation reconnaissance survey of the southern great plains wind-erosion area* (No. 556). US Department of Agriculture.

Lee, E. S. (1966). A theory of migration. *Demography*, *3*(1), 47–57.

Long, J., & Siu, H. (2018). Refugees from dust and shrinking land: Tracking the dust bowl migrants. *The Journal of Economic History*, *78*(4), 1001–1033.

Lucas, R. E. (1987). Emigration to south africa's mines. *The American Economic Review*, 313–330.

Moscona, J. (2020). Environmental catastrophe and the direction of invention: Evidence from the american dust bowl.

Nordhaus, W. (2019). Climate change: the ultimate challenge for economics. *American Economic*

Review, *109*(6), 1991–2014.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, *3*(2), 135–146.

Ruggles, S., Fitch, C., Goeken, R., Hacker, J. D., Helgertz, J., Roberts, E., … Wellington, J. (2019). Ipums multigenerational longitudinal panel.

Stern, N., & Stern, N. H. (2007). *The economics of climate change: the stern review.* cambridge University press.

Worster, D. (2004). *Dust bowl: the southern plains in the 1930s.* Oxford University Press.

Ziebarth, N. L. (2013). The radio and bank distress in the great depression. *Manuscript, University of Iowa*.

Appendices

A Migration in the 1930s

As the dataset has information on the whereabouts of the individual in 1935, we can conduct a simple analysis on early migrants (those who moved between 1930 and 1935), and late migrants (those who moved between 1935 and 1940). Consider the simple linear probability model in 5. I reduce my sample to those residing in the Dust Bowl region in 1930. mig_{it} is now a dummy variable that equals 1 if the individual migrated post 1935, and 0 if the individual migrated between 1930 and 1935. This simple model follows from one of the specifications from (Long & Siu, 2018). This way I can compare late-decade migrants to early-decade migrants. A negative sign on the coefficient indicates an increased likelihood of migrating early in the decade, whereas a positive coefficient indicates higher likelihood of moving during the second half of the decade. Column (1) shows the baseline linear probability model, while column (2) includes county fixed effects.

I use the same controls above, but I also include log wage[9], the New Deal Expenditure per capita in the 1930 county of residence[10], as well as the percent of the population in the county with a radio. There have been studies showing the relationship between New Deal Spending on migration (Fishback et al., 2006), as well as looking at the relationship between radio ownership and access to information on dust storms, economic conditions, as well as information about emigration in surrounding areas (Ziebarth, 2013) (Long & Siu, 2018).

$$mig_{it} = \Gamma X_{it}^{'} + \varepsilon_{it} \tag{5}$$

From the results (in table 10), we can see that medium skilled migrants and laborers were more likely to migrate before 1935 relative to high skilled migrants, whereas farmers were more likely

[9] As income is only asked in the 1940's census, I use the *OCCSCORE* variable, which is a proxy for the income, based on the median income for the person's occupation.

[10] It should be noted that New Deal Spending could potentially be endogenous to conditions in the 1930s, so I will merely look at the sign of the coefficient, rather than the magnitude. In past papers, it has either been omitted or added as a robustness check.

to migrate later in the decade (as were those residing on farms), indicating that farmers only left after the large dust storms which directly affected their livelihood (such as *Black Sunday* in 1935). Similarly, those who reside in their birth-state were less likely to move early in the decade, as were homeowners and those with children. Those in urban areas were more likely to move early in the decade, as were Black Americans. One interesting observation is that migrants from counties with a large percent of radio ownership were more likely to move later in the decade, this could be due to making an informed decision about the severity of economic conditions in their home county and potential destination county.

B Other Selection models

In my migrant selection models, I look at selection based on pre-migration incomes. However, the Dust Bowl could have caused changes in migrant selectivity based on several other factors. In this section, I will look at other factors that could impact migration. Going back to the specification in equation 3, I can examine this by replacing my dependent variable with the other factors migrants are selected on.

I look more into the person's residence: whether they lived on a farm or in an urban area prior to migrating. I then look at occupations: how were farmers and highly skilled professionals selected? Finally, as I used the OCCSCORE proxy variable earlier, I look at other proxy scores for occupational prestige: the SEI[11] (occupational score) and EDSCOR50 (educational score) variables. These proxies provide robustness checks to my findings on income based selection.

Table 11 shows the results. For brevity, I only include the coefficient for the triple interaction term for each factor the migrants are selected on. In particular, I look at the baseline model, out-selection, and in-selection for the Dust Bowl region, high erosion counties, and medium erosion counties. These show the change in selectivity for each factor for migrants during the Dust Bowl.

For the Dust Bowl region, I find that migrants who lived on farms were more positively selected during the Dust Bowl. However, farmers were actually more negatively selected. On the other hand, migrants from urban areas were actually more negatively selected, as were higher skilled

[11] This is the Duncan Socioeconomic Index, which is a two digit weighted sum of occupational income and educational measures. For more, see (Duncan, 1961)

professionals. When I look at the proxies for educational attainment and occupational prestige, I also find that migrants were more negatively selected during the Dust Bowl. When we look at out-selection, I find that relative to non-migrants, migrants were less likely to be farmers and high-skilled workers. They were also negatively selected with respect to occupational prestige. Similarly for in-selection, migrants were negatively selected with respect to being farmers, living in urban areas, education and occupation compared to natives. On the other hand, they were more positively selected with respect to living on farms, relative to natives. However, these changes in in- and out-selection are modest.

When I restrict my analysis to the Great Plains region, I find that migrants from high erosion counties were more negatively selected with respect to living on farms as well as being farmers, and were positively selected with respect to living in urban areas. On the other hand, migrants from medium erosion counties were more positively selected with respect to living on farms and being farmers, while being more negatively selected with respect to living in urban areas. When I look at in- and out-selection, I only find modest changes in selectivity for migrants from counties with high and medium levels of erosion.

C California Migrants

One topic that always comes up when discussing the Dust Bowl is the plight of the migrants in California. Many Californians today have some *Okie* heritage, reflecting the mass migration from the Dust Bowl. I repeat the same analysis from my empirical specification, this time focusing on migrants who moved to counties in California. Table 12 shows the migration rate, as well as the migrant characteristics of all migrants and migrants to California. Table C shows the migrant selectivity and migrant outcome models, for migrants who moved to California. Rather than focus on the Great Plains region, I opt to focus on the cluster of counties in the Southern Great Plains, as many of the migrants who moved to California moved along highway routes such as Route 66 from the southern Great Plains (especially from the Oklahoma and Texas panhandle counties).

In table 12, column (1) refers to all migrants from the Dust Bowl region, while column (2) refers to migrants to California. The coefficient of the variable ($DB_i \cdot Post_t^{1930}$) shows that migration to

California was higher than the average migration rate from the Dust Bowl region during the Dust Bowl. It seems that most migrants were highly skilled, and there is strong evidence that farmers were less likely to migrate in the first pace. The control variables do not point to significant disparities between California migrants relative to overall migrants, and the model points to mainly high-skilled white migrants moving during the Dust Bowl.

Table 13 shows the migrant selection (in terms of pre-migration income) and outcome models for migrants to California. Columns (1), (2) and (3) show the migrant selection models ((1) is the baseline model, while (2) and (3) check for out-selection and in-selection), while columns (4) through (8) show the migrant outcome models (Column (4) shows the full sample, while columns (5) to (8) show the outcomes for each occupational group). The variable of interest is the variable ($DB_i \cdot mig_i \cdot Post_t^{1930}$), which shows the change in selectivity and outcomes for migrants from the Dust Bowl region to California during the Dust Bowl. I also look at the Great Plains Region, namely the ($H_i \cdot mig_i \cdot Post_t^{1930}$) and ($M_i \cdot mig_i \cdot Post_t^{1930}$) variables, which show the change in selectivity and outcomes for Dust Bowl migrants relative to low erosion counties during the Dust Bowl.

When we look at the migrant selection models, we see that migrant selectivity fell in the Dust Bowl region, but there is no evidence of a change in in- or out-selection during the Dust Bowl. However, when we look at migrant outcomes for California migrants, we see clear disparities throughout the board (from the negative and statistically significant coefficient on the triple interaction variable). The full sample of California migrants saw a large decrease in post-migrant incomes during the Dust Bowl. These decreases are consistent across all occupational groups, with laborers and high skilled migrants to California seeing larger relative decreases in incomes on migrating during the Dust Bowl. Despite there not being strong evidence for changes in migrant selectivity to California, I find clear decreases in incomes for those who moved to California relative to the average Dust Bowl migrant.

Restricting the sample to the Great Plains region, I do not find any evidence of a change in migrant selectivity. However, I find that migrant selectivity in medium erosion counties fell during the Dust Bowl, and so did in-selection (migrants had lower wages than natives). Looking at migrant outcomes, I find that incomes fell for migrants from both high erosion and medium erosion counties during the Dust Bowl. Farmers who migrated also saw declines in income relative to those who migrated prior to the Dust Bowl.

Dust Bowl migrants to California did not only have larger decreases in selectivity than migrants to other states, but also had larger decreases in post-migration incomes during the Dust Bowl than other migrants. Migrant farmers to California also fared much worse than elsewhere. Many depictions of Dust Bowl migrants focus on the migrant flows to California. My results seem to provide some context to these depictions. While, overall, Dust Bowl migrants showed modest differences to migrants prior to the Dust Bowl, migrants to California were not only poorer, they also fared worse upon migrating than those who migrated to California prior to the Dust Bowl.

These results differ from the results for general Dust Bowl migrants. Many stereotypes of Dust Bowl migrants arose from perceptions of migrants to California (Gregory, 1989), and I find that Dust Bowl migrants performed worse in California than elsewhere, which could have impacted popular perceptions about Dust Bowl migrants overall.

Table 10: Early- vs. Late-Decade Migrants

	(1) Baseline Model	(2) County FE
Occupational Group		
Semi skilled	-0.0254***	-0.0203**
	(0.00569)	(0.00701)
Farmers	0.0162**	0.0184**
	(0.00625)	(0.00688)
Laborers	-0.0588***	-0.0552***
	(0.00511)	(0.00574)
same state as birth	0.0577***	0.0597***
	(0.00326)	(0.00456)
same state as parents birth	-0.0116**	-0.00246
	(0.00398)	(0.00548)
Race		
Black	-0.0765***	-0.0827**
	(0.0193)	(0.0291)
Native American	-0.000841	0.0184
	(0.0793)	(0.0863)
lives on farm	0.0497***	0.0437***
	(0.00379)	(0.00543)
lives in urban area	-0.0137***	0.00634
	(0.00404)	(0.00860)
Nativity		
first generation	0.0337*	0.0314*
	(0.0135)	(0.0153)
second generation	0.0245***	0.0187*
	(0.00575)	(0.00761)
Owns House	0.133***	0.132***
	(0.00317)	(0.00591)
Number of Children	0.00419***	0.00427**
	(0.00111)	(0.00140)
New Deal Spending per capita	-0.0000168	
	(0.0000104)	
percent having radio	0.405***	
	(0.0134)	
age	0.000620	0.000556
	(0.000366)	(0.000382)
age squared	-0.000000308	0.000000803
	(0.00000557)	(0.00000511)
Constant	0.0921***	0.196***
	(0.00635)	(0.00654)
N	98649	98649
adj. R^2	0.043	0.030

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Migrant Selection Models

	$(DB_i \cdot mig_i \cdot Post_t^{1930})$			$(H_i \cdot mig_i \cdot Post_t^{1930})$			$(M_i \cdot mig_i \cdot Post_t^{1930})$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Baseline Model	Out Selection	In Selection	Baseline Model	Out Selection	In Selection	Baseline Model	Out Selection	In Selection
Lives on Farm	0.0058*	-0.0124	0.0279***	-0.0506***	-0.0442*	-0.0172	0.0062**	-0.0017	0.0273***
	(0.0026)	(0.0071)	(0.0079)	(0.0030)	(0.0201)	(0.0125)	(0.0023)	(0.0077)	(0.0084)
Lives in Urban Area	-0.0087*	0.0067	-0.0339***	0.0151***	0.0117	-0.0266*	-0.0058*	0.0058	-0.0310**
	(0.0021)	(0.0053)	(0.0106)	(0.0026)	(0.0091)	(0.0136)	(0.0020)	(0.0071)	(0.0109)
Farmer	-0.0062**	-0.0144***	0.0024	-0.0194***	-0.0155*	-0.0092	0.0045*	0.00001	-0.0113**
	(0.0022)	(0.0041)	(0.0039)	(0.0023)	(0.0077)	(0.0057)	(0.0018)	(0.0040)	(0.0037)
High Skilled	-0.0045***	-0.0035*	-0.0048**	-0.0018	-0.0022	-0.0041*	-0.0004	-0.0002	-0.0018
	(0.0011)	(0.0016)	(0.0017)	(0.0013)	(0.0027)	(0.0020)	(0.0010)	(0.0014)	(0.0017)
Educational Score	-0.3470**	-0.2088	-0.4724*	-0.0728	-0.0766	-0.3795	0.1247	0.1928	-0.0559
	(0.1216)	(0.1685)	(0.1935)	(0.1417)	(0.3514)	(0.2261)	(0.1095)	(0.2168)	(0.2010)
Occupational Score	-0.7360***	-0.5078***	-0.9085***	0.0124	-0.0841	-0.3947	0.1238	0.1587	-0.1263
	(0.0956)	(0.1451)	(0.2001)	(0.1109)	(0.2787)	(0.3311)	(0.0863)	(0.1687)	(0.2674)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Dust Bowl Migrant Characteristics

	(1) All Migrants	(2) California Migrants
$(DB_i \cdot Post_t^{1930})$	0.0208**	0.0315***
	(0.00714)	(0.00792)
$Post_t^{1930}$	-0.0605***	-0.0418***
	(0.00309)	(0.00884)
Occupational Group		
Semi skilled	-0.00525***	0.00208
	(0.00117)	(0.00145)
Farmers	-0.0344***	-0.0302***
	(0.00114)	(0.00382)
Laborers	-0.00355***	-0.00263
	(0.000699)	(0.00225)
same state as birth	-0.102***	-0.0240*
	(0.00209)	(0.0115)
same state as parents birth	-0.0323***	-0.0207**
	(0.00113)	(0.00646)
Race		
Black	-0.101***	-0.0666***
	(0.00432)	(0.00743)
Native American	-0.126***	-0.135***
	(0.00662)	(0.0169)
lives on farm	-0.0476***	-0.0321***
	(0.000999)	(0.00494)
lives in urban area	-0.0286***	-0.0296***
	(0.00106)	(0.00334)
Nativity		
first generation	-0.113***	-0.0618***
	(0.00475)	(0.0133)
second generation	-0.0565***	-0.0432***
	(0.00194)	(0.00687)
Owns House	-0.108***	-0.0838***
	(0.00325)	(0.0135)
age	0.000702***	0.000871*
	(0.000107)	(0.000340)
age squared	-0.0000342***	-0.0000253***
	(0.00000143)	(0.00000743)
Number of Children	0.00183***	-0.00122
	(0.000335)	(0.00121)
Constant	0.446***	0.450***
	(0.00334)	(0.0223)
N	54836954	2341465
adj. R^2	0.050	0.033

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Dust Bowl Migrant Selection and Outcomes: California

	Migrant Selection				Migrant Outcomes			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline Model	Out Selection	In Selection	Full Sample	High Skilled	Medium Skilled	Laborers	Farmers
$(DB_i \cdot mig_i \cdot Post_t^{1930})$	-0.0267**	-0.0103	-0.0149	-0.0717***	-0.0958**	-0.0563***	-0.110**	-0.0492*
	(0.0010)	(0.0155)	(0.0110)	(0.0131)	(0.0309)	(0.0152)	(0.0336)	(0.0214)
$(H_i \cdot mig_i \cdot Post_t^{1930})$	-0.00730	0.0226	0.0151	-0.0849***	-0.0763	-0.0788**	-0.0414	-0.0844**
	(0.0131)	(0.0245)	(0.0108)	(0.0209)	(0.0427)	(0.0284)	(0.0442)	(0.0316)
$(M_i \cdot mig_i \cdot Post_t^{1930})$	-0.0436***	0.00179	-0.0278**	-0.0673***	-0.0480	-0.0265	-0.109**	-0.0720**
	(0.0111)	(0.0145)	(0.0104)	(0.0154)	(0.0294)	(0.0204)	(0.0359)	(0.0275)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$