How California's wildfires spark migration

Rima Parekh

March 17, 2021

Advisor: Alisa Tazhitdinova

Abstract

This paper explores the relationship between California wildfires and human migration, and whether it can be reasonably assumed that California counties with a higher frequency and/or severity of wildfires experience greater out-migration than counties that experience a lower fire risk. Using county-to-county migration data from 2010 to 2018 and wildfire data from 2009 to 2017, I run regressions with two different models: the multiple regression and fixed-effect model. Source counties, i.e. counties where people are migrating from, observed in this study are only in California, but destination counties, i.e. counties where people are migrating to, include all counties in the U.S. In the case where destination counties are out of state, I aggregate counties by state so that I have county-to-county flows within California and county-to-state flows for the other 49 states. While it is possible to find literature that explores the effects of extreme climate events on human migration, little research exists on climate-induced migration in California, specifically with respect to wildfires.

Introduction

Existing research shows that climate disasters affect human migration, especially for climate disasters such as extreme heat, drought, and flooding, which can have a severe impact on countries whose economies rely heavily on resources and agriculture. However, most papers focus on developing countries, where climate-induced disaster could displace millions of people who have nowhere else to go. In contrast, fewer research exists on the effect that climate disasters will have on migration patterns within developed countries. While limited papers explore climate migration in the U.S., some research does exist on the relationship between wildfires and migration. The existing literature on climate migration in the U.S. shows that increased presence of wildfires near residential areas is the result of anthropogenic climate change, which has posed an existential threat to a larger and larger population with every passing year as extreme wildfire events become more potent and frequent.

In the U.S., the West Coast is most prone to facing these severe wildfire events. I want to explore the interaction between wildfires and migration: As climate-induced wildfires become more frequent and pervasive, specifically in the state of California, how will this affect migration patterns? I hypothesize that counties that experience greater frequency and severity of wildfire events will also experience higher rates of out-migration, and that people in these counties making migration decisions based on the wildfires will migrate to states or other counties in California that have a relatively lower fire-risk.

I look at wildfires and migration patterns by county in California to see if there is an association between wildfires in the previous year and out-migration flows. I use historical wildfire data to gather fire statistics for each county in California and state-wide fire statistics for the other 49 states from 2009 to 2017. For migration data, I gather county-to-county flows within California and county-to-state out-flows for counties in California and the 49 remaining states. The migration data is aggregated and separated into two time periods: 2010 to 2014, and 2014 to 2018. My model operates under the assumption that climate-induced migration follows a one-year lag from extreme weather events—the migration data is collected for eight years following 2009, with 2009 being the first year of observations for fire data.

I use two types of models when measuring these effects: the multiple regression model and the fixed-effect model. For each model, I run multiple regressions on different fire measures to see how migration responds, if at all, to the frequency or severity of fires. Frequency measures include the number of fires and number of fires per capita; severity measures include the number of acres burned per capita and proportion of acres burned. Using the multiple regression model, most of my regressors are statistically significant, indicating a relationship between wildfires events and migration. However, given that this paper uses time series data to answer the question, I also apply fixed effects for time and region. When applying the same regressions to the fixed-effect model, my findings are no longer statistically significant, indicating that the variation in migration for different counties and time periods skews my data, causing the model to incorrectly attribute changes in migration to wildfire events. Given these findings, I am unable to claim that a relationship exists between migration and California wildfire events.

Literature Review

Current literature relating to climate-induced migration is fairly limited for Western countries—most papers that focus on climate-migration explore this effect in developing countries with widespread poverty where the economy is highly based on agriculture or natural resources. Bangladesh is one country that has been studied more extensively to explore the impacts of climate change on migration within the country. In 2012, Hassani-Mahmooel and Parris modeled the impacts of extreme weather events on migration and predicted that there could be anywhere between 3 and 10 million internal migrants over the following 40 years. Given that Bangladesh is a coastal country with a tropical monsoon climate, the country is prone to hurricanes, flooding, and sea-level rise. Hassani-Mahmooel and Parris (2012) studied the net out-migration from the Bangladesh Bureau of Labor Statistics, and found a dramatic increase from 1970 to 1990 to 2010. They found that differences in out-migration were significant across gender. While other countries that are often studied to measure climate-migration bring mixed results, climate-migration appears to have a more robust link in Bangladesh—approximately two-thirds of households in the South-Asian country were displaced at least once in their lifetimes.

Bangladesh is divided into seven districts; climate and migration statistics are measured by these districts accordingly. The migration decision is modeled with push factors, e.g. climate-change scenarios and socioeconomic measures of the district, pull factors, e.g. socioeconomic conditions of the possible destination districts, and intervening factors, e.g. employment statistics and land or home ownership. The model also includes climate scenario data that follows a Poisson distribution and takes values between 0 and 1. This data is treated as a time series with 600 points, an observation for all 12 months over 50 years. Each data point represents the intensity of a climate event for a given district at a given time and accounts for the vulnerability of each of the districts to climate shocks. The paper finds, however, that the level of vulnerability across the districts is relatively heterogeneous. Hassani-Mahmooel and Parris' model (2012) predicts that over the next few decades, Bangladeshis would primarily live in the central and eastern part of the country, since people residing in the western and southern regions would become increasingly more vulnerable to drought and flooding, respectively. In contrast, the central and eastern part of the country are less susceptible to climate shocks, such as rising sea levels, flooding, and drought.

Hassani-Mahmooel and Parris' paper (2012) was useful in informing how I should construct a model for my own research question, but I also looked at other research papers that focused on climate-induced migration in the U.S., specifically with a focus wildfires events, in order to gain a better understanding of how to create an even more effective model. "Amenities or disamenities? Estimating the impacts of extreme heat and wildfire on domestic U.S. migration" by Winkler and Rouleau (2020) explored how heat waves and wildfires impacted net migration in the West and Southwest, since these regions are particularly susceptible to these sorts of climate events. Winkler and Rouleau (2020) explored how net migration was associated with extreme heat and wildfires by looking at its effects on in- and out-migration. They also explored how metropolitan areas and places with environmental amenities are impacted by these disasters harder than other locations since many people migrate for the amenities, which are then threatened by these disasters. Since migration choices experience a lag after natural disasters, Winkler and Rouleau (2020) looked at migration after one year of a natural disaster in order to provide a one-year lag. The paper put forth three hypotheses: (1) extreme heat in the prior year is associated with reduced net migration rates, (2) wildfire in the prior year is associated with reduced net migration rates, and (3) non-metropolitan counties and counties with more environmental amenities show a stronger response to heat and fire migration.

To observe these effects, Winkler and Rouleau (2020) looked at in-migration and out-migration by county and measured environmental amenities by using a scale from the U.S.

Department of Agriculture (USDA). This scale measures the climate conditions of a county that are considered desirable to live in, such as warm and sunny winters, temperate summers, topographic variation, and water area, to name a few. This scale is associated with population growth, making it relevant to the regression—it accounts for the influence that the attractiveness of a region as a place to live has on people's migrations decisions. Heat waves were measured by county using data from the PRISM (Parameter elevation Regression on Independent Slopes Model) Climate Group, which includes the maximum and minimum temperatures in a given county over a certain time period. Wildfire data was gathered from the Federal Emergency Management Agency (FEMA), which provides county-level data for FEMA-declared disasters and indicates the year and type of disaster. Winkler and Rouleau (2020) aggregated the number of wildfires every year for each county in order to create a dummy variable that took 1 for a given county that experienced at least one FEMA-declared wildfire in a given year. Migration data was collected from the Internal Revenue Service (IRS) Statistics of Income (SOI) data sets, and was gathered based on the year-to-year change of address on individual income tax returns.

In order to measure the impacts of disaster on migration, Winkler and Rouleau (2020) used fixed effects and random effects models to absorb variation in migration rates that was caused by factors other than the measures for extreme heat and wildfires. The fixed effects model included time lags and controlled for economic changes and changes in the number of non-migrants; the random effects model controlled for natural amenities, outdoor recreation, and metropolitan status by using time constant variables for these three regressors. The results from the fixed effects model supported the paper's first two hypotheses: heat waves and wildfires were both associated with reducing in-migration and increasing out-migration, thus reducing net-migration rates. However, the effect of heat waves on migration was relatively smaller than that of wildfires. The models also supported their third hypothesis: that in-migration to high-amenity counties is likely more susceptible to extreme heat waves and wildfires. Winkler and Rouleau (2020) ultimately concluded that heat waves and wildfires resulted in reduced net migration rates, and that counties with greater environmental amenities were more strongly impacted than lower-rated counties on the amenities scale.

Other relevant papers I looked at were focused on the effects of climate change on the West, but measured impacts in different ways. "The effects of wildfire and environmental amenities on property values in northwest Montana, USA" focused on observing the impacts on property values as mentioned in the title, while other papers attempted to model impacts of climate change on wildfire risk. While these papers were helpful in informing me how I wanted to establish my research question and empirical process, they were less relevant to my paper than "Amenities or disamenities? Estimating the impacts of extreme heat and wildfire on domestic U.S. migration," seeing as the latter paper's research question more similarly reflected mine.

Instead of measuring the effect of wildfires on migration, Stetler and Calkin (2010) take a more direct approach of measuring the economics of climate events by observing the impact of wildfires on property values in northwest Montana. "The effects of wildfire and environmental amenities on property values in northwestern Montana, USA" attempts to examine the effects of wildfires on human welfare by evaluating how the changes in environmental amenities and perceived wildfire risk due to wildfires are reflected in property values. Similar to Winkler and Rouleau (2020), Stetler and Calkin (2010) incorporate environmental amenities in their model to account for the desirability of a location due to its natural amenities, and they attempt to examine whether a relationship exists between changes in natural amenities due to wildfires and changes in property values. Northwest Montana is a desirable region for people interested in outdoor recreation activities, given the abundant open land, proximity to Glacier National Park, and other national forests. Stetler and Calkin (2010) use a hedonic price model that is a function of vectors of structural characteristics, neighborhood attributes, and environmental attributes, in order to measure the house sale price. Stetler and Calkin (2010) assume that prospective homeowners purchase a home that maximizes their utility of the hedonic price model, subject to their budget constraint.

Data for house sale prices, structural characteristics, and neighborhood characteristics from 1996 to 2007 were collected from the Northwest Montana Association of Realtors (NMAR). The data included relevant information for these parameters of the hedonic price model, such as the number of rooms, square footage, and "housing zones" (i.e. neighborhoods). Using geospatial data, Stetler and Calkin (2010) estimated the straight-line distances of homes to major natural amenities, such as lakes, rivers, and wilderness areas, and to wildfire burned areas. The model incorporated a wildfire variable that indicated whether wildfire burned area was visible from the home, and 17,963 home sale transactions were included in the model, where each observation contained information of these varying characteristics that Stetler and Calkin (2010) predicted would have an effect on the property values.

The model indicated that proximity to natural amenities had a large positive effect on property values. In contrast, proximity to a wildfire burned area had a large negative effect. On average, properties that had views of wildfire burned areas had lower property values than properties that did not have these views; however, this measure seemed to have a far smaller effect on properties than just the proximity measure. Given the results from their model, Stetler and Calkin (2010) concluded that wildfires have large negative effects on property values in northwest Montana, which can be explained by the impact wildfires have on environmental amenities and perceived wildfire risk. Stetler and Calkin's (2012) focus was not on climate-migration, but rather on the economic impacts of wildfires, specifically as it relates to the housing market. This type of literature is still useful in informing models for climate-induced migration, since migration is driven by several different factors, such as house prices, and this paper highlights how these other migration-inducing factors can be linked with climate change as well. If other regressors, such as house prices and economic opportunity, are included in the model, it is important to separate the effects of climate change on these other parameters as well, even if a regressor that measures for climate events is already included in the model.

Theory

This paper attempts to answer whether a relationship exists between migration and wildfires in California, and if climate-migration is a phenomenon that occurs within and/or outside of California. This question is one of the environment and the implications it has on our societies, but it is also indirectly a question of economics. A greater understanding of the association between wildfire events and migration, or lack thereof, can allow for more empirical-based discussion on the impacts of economic and environmental policy. More effective strategies for wildfire suppression and management, along with prescribed burning, have become increasingly imperative in California, given the alarming frequency and severity of California fires over the last few years (Temple, 2020) and an empirical understanding of the effects of wildfires on different aspects of society and the economy can allow for more targeted and productive discussion on how to address the issue.

Migration has always been closely linked to the economy of a region—the migration patterns of a region inform a great deal about the economy (*Effects of Immigration*, 2018). If it were an autonomous country, California would have the fifth largest economy in the world (*California,* 2019), making it the strongest economy of any state in the U.S. California also has the highest proportion of immigrants of any state in the country—as of 2019, 24% of immigrants in the U.S. live in California (Budiman, 2020), which has allowed California to have a robust labor force. As of 2018, immigrants made up 27 % of California's population, while 33% of California's labor force was comprised of immigrants (*Immigrants in California,* 2020); the share of immigrants in the workforce is 6 points higher than their share in the total population, indicating that immigrants in the workforce make up a higher proportion of the immigrant population than for non-immigrants in California. By testing for an association between wildfires and migration, my goal is to provide a framework that will facilitate predictions of the displacement or arrival of migrants in different counties in

California. This could ultimately allow for informed policy and legislative discussion as it relates to immigration, housing, industries that rely heavily on immigrants, etc.

California counties with strong economies, such as Los Angeles County, San Francisco County, and Santa Clara County, could potentially see loss of talent if these regions become more prone to fire risk and if residents thus chose to migrate to other relatively safer counties or states. The reverse could also be true; more rural regions that are closer to natural amenities may be at higher fire risk due to their proximity to natural preserves and open lands, causing people to leave these regions and migrate to counties with lower-fire risks. This could cause even greater socioeconomic disparity between urban and rural counties, and would be an important factor to consider for policymakers interested in mitigating the concentration of wealth in certain pockets of the state.

Empirical Strategy

To go about observing the effect that wildfire has on human migration, specifically by county in California, I look at two different migration flows: county-to-county within California and California county-to-state level for all other 49 states. I first collect California county level fire data from CalFire and county-to-county migration flows from the Current Population Survey (CPS) from the U.S. Census Bureau.

County-to-County Migration & Fire Data

The county-to-county migration data is aggregated over four years. My model includes data from 2010 to 2018—I merge migration flows from 2010 to 2014 with the migration flows from 2014 to 2018. When creating my model, I operate under the assumption that there is a one-year lag between a major fire event and human migration decisions. That is, I expect that people who migrate primarily due to fire-risk concerns take a year to move to a different area. Due to this assumption, I group fire data into two different time groups and then aggregate them by county. Thus, I group my fire data 2009 to 2013 and aggregate it by county, since I assume that the migration decisions made due to fire events are reflected in the 2010 to 2014 aggregated migration flow. I repeat this process with fire data from 2013 to 2017 and the 2014 to 2018 migration data.

I then merge the matching fire and migration periods by county so that I have two resulting data sets for the two different time periods that I am measuring. After combining this data, I assign a dummy variable that takes 0 for the first time period (2010-2014 migration flow) and 1 for the second time period (2014-2018). Using the dummy variable to measure the two different time periods, I can treat this as time series data.

For the county-to-county data, I run three simple ordinary least squares regressions on the migration flow from the source county (county B) to the destination county (county A), where B and A are two different counties in California. The flow is calculated by taking the number of people migrating from county B to county A, and then dividing by the total population of county B. My regression models take the following form:

$$\boldsymbol{Y_{ijt} = \beta_0 + \beta_1 FireMeasure_{it} + \beta_2 FireMeasure_{jt} + \epsilon_{it}} \quad (1)$$

In this model, i is each observation for county A and j is each observation for county B, meaning that $FireMeasure_i$ is a given fire measure for county A and $FireMeasure_j$ is a given fire measure for county B. I run these three regressions separately for each time period: once for the observations that took 0 for the time period dummy variable, and a second time for observations that took 1. The regression results for each of the three regressions on each time period are given in Tables 1-3.

I run regressions using another model that includes county-fixed effects for county A (δ_i) and county B (δ_j), and time fixed-effects (γ_t) in order to hold constant the average effects of county A and county B, and of the time periods, respectively. Because this model accounts for time by incorporating fixed effects to absorb variation by time, I run the model on the entire dataset that includes observations from both time periods, rather than splitting the observations from different time periods and running the regression separately.

$$\boldsymbol{Y_{ijt} = \beta_0 + \beta_1 FireMeasure_{it} + \beta_2 FireMeasure_{jt} + \delta_i + \delta_j + \gamma_t + \epsilon_{it}} \quad (2)$$

For both models, the first regression is of migration flow on the number of fires in county A and county B. In the second regression, I regress on the number of fires per capita in both counties, and in the third regression, I regress on the number of acres burned per capita. I also calculate the averages for each of these statistics—number of fires, number of fires per capita, and number of acres burned per capita—for each county in California, in order to explain the magnitude of effects. The results of the fixed-effect regressions are given in Tables 1-3 for each of the three regressions.

County-to-State Migration & Fire Data

I employ a similar strategy to measure the effect of wildfires on migration flows between California counties and other states. I use the same migration data as my county-to-county model, but include flows with other states, rather than just flows within California. I aggregate the county data by state for the other 49 states to ensure that I am looking at state level data for all states other than California. I use the Insurance Information Institute to gather state level fire data for all 49 states, and group this data into two different periods: 2009-2013 for the earlier migration period, and 2013-2017 for the later migration period.

Like with the county-to-county model, I aggregate the fire data by state and then merge it with the state-level migration data. I merge my California data that I cleaned using CalFire with the California migration flows by state. Once again, I do this for the two different time periods, so that I can treat the data as a time series, with the first period taking a value of 0 (2010-2014 migration flow) and the second period taking a value of 1 (2014-2018).

For the county-to-state data, I run three ordinary least square regressions on the migration flow from a given county in California to another state. This flow is calculated by taking the number of people from a given California county to a given state and dividing by the total population of the California county in question. My regression model takes the same form as regression (1), except in this case i is each California county observation and j is each state observation. Similarly, $FireMeasure_i$ is a given fire measure for the counties and $FireMeasure_j$ is a given fire measure for the states. As with the county-to-county model, I apply these three regressions on both time periods; the results are displayed in Tables 4-6.

I also run regressions using a model with county and state fixed-effects, as well as time fixed-effects. This model takes the same form as (2), where δ_{i} was the parameter for county fixed-effects, δ_{j} for state fixed-effects, and γ_t for time fixed-effects. Since time fixed-effects absorb variation in the model due to time, it is not necessary for us to run the regressions two times for each time period, as it was with the ordinary least squares model.

For both models, the first regression contains the number of fires in the county and state as the regressors, while the second regression included the proportion of acres burned. The third regression contains the number of fires per capita. As with the county-to-county regressions, I calculated the averages for these three statistics—the number of fires, the proportion of acres burned, and the number of fires per capita—for each California county and each state (excluding California) in order to explain magnitude effects.

Data

Migration

To measure whether a causal link exists between wildfires and migration, I use 2010 to 2018 county-to-county migration data from the Current Population Survey (CPS), which is jointly sponsored by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics (BLS). The CPS data records in, out, net, and gross county-to-county migration for every county in the U.S., aggregated in four year intervals. I gather data that only includes migration flows for which the source county was in California. I use two different migration data sets: one that aggregates county-to-county migration from 2010 to 2014, and one from 2014 to 2018. I clean the data so that I have two separate data sets for the county-to-county out-migration flows within California and county-to-county out-migration flows where the destination county is outside California.

Instead of simply looking at the out-migration flow of the source county in isolation with respect to fire risk, I also account for the county which people are migrating to in my model in order to get a better understanding of whether these migration flows are driven by fire-risk or some other confounding factors. If the data showed that people in high-risk counties are migrating to equally risky or higher risk counties, this indicates that migration decisions aren't being driven by wildfire risk, but rather by some other variables outside of the model. Because California has 58 counties and my migration flows account for the destination county, I expected to have 3,306 observations for each time period (6,612 total), since each of the 58 counties are measuring the out-migration to the other 57 counties. However, each source county didn't necessarily have migrants leaving for all 57 counties over the time periods, which resulted in fewer than 3,306 observations for each time period. Additionally, for both time periods, migration patterns weren't necessarily the same, meaning that in the second time period, source counties may have seen migration to new counties that weren't included in the first time period, resulting in an unequal number of observations for the first and second time periods. For the migration data aggregated from 2010 to 2014, 2,606 flows are included, while 2014 to 2018 aggregated migration data includes 2,604 flows, accounting for 5,210 observations in total over the two time periods.

For the county-to-county migration data for source counties inside California and destination counties outside of California, I use a similar strategy. However, instead of looking at migration flows at the county level for both the source and destination of migrants, I aggregate the destination counties by state, and measure the relationship between migration out of California counties and the fire-risk of a different state. The flows are then representative of county-to-state out-migration for all 58 California counties to the other 49 states. Thus, I expect to have 2,842 observations for each time period (and 5,684 observations total), since this is the product of 58 and 49. Similar to the county-to-county data, the county-to-state data has fewer observations than expected, since not all counties have migrants leaving for all 49 states. For migration flows aggregated for the first time period, there are 2,104 observations, and flows from the second time period have 2,097 observations, coming out to a total of 4,201 observations.

Wildfires

I gathered my California wildfire data from CalFire, which has recorded fire events in California since 2013 up to date, with the exception of a smattering of entries that have been included from years prior to 2013. For each fire event recorded, the data set includes the number of acres burned, the county of occurrence, and the exact date. Some wildfires spanned across more than one county; in this case, I duplicate the entry for each county the fire event occurred in. That is, if a fire spreads across three different counties, I create two new observations, in addition to the one that already exists, that takes the same value for the number of acres burned. When incorporating wildfire statistics into my datasets, I operate under the assumption that migration decisions based on wildfire events follow a one-year lag; if a family experienced an extreme wildfire event in proximity to their home and decided to leave the county as a result, I expect that their migration out of the county would be reflected in the following year. For this reason, I gather fire data from 2009 to 2017, since the migration data spans from 2010 to 2018. However, the fire data from 2009 to 2013 is extremely limited compared to 2013 to 2017. Thus, a causal relationship between wildfires and migration in the first time period may be harder to observe since fewer observations for fire events are recorded. I then aggregate the fire data by county so that I have 58 observations for all the California counties that measures the total number of fires and total number of acres burned for each county.

I employ a similar strategy when collecting state-wide wildfire data for the county-to-state migration flows. The state fire data is sourced from the National Interagency Coordination Center (NIIC), which publishes tables of the fire statistics each year for all 50 states. Similar to the CalFire data, the statistics included are the number of fires and the number of acres burned by state. While the tables include statistics for both wildland fires and prescribed burns, I only include wildland fires since I assume that prescribed burns won't affect residential areas, and thus won't have an impact on people's migration decisions. When including the state fire data in my migration data sets, I exclude California fire statistics, since my county-to-state migration flows only include destination states other than California. Rather, all California fire statistics used in the migration datasets are at the county-level.

Other Data

For both the California county and state fire statistics, I create my own fire measures to better account for the population and size of each county and state, respectively. For California populations by county, I use data from the Census Bureau. I extract the population totals for each year that I was measuring fire data, i.e. 2009 to 2013. For the first time period, 2009 to 2013, corresponding to the migration data from 2010 to 2014, I calculate the population totals by county for each year from 2009 to 2013, and the average population over the four years. I repeat this process for the second time period, with population data from 2013 to 2017. For state populations, I again use data from the Census Bureau, and use the same strategy as with the county population data. I merge the county and state population data with the county and state wildfire data, respectively, and calculate my own fire statistics accordingly.

I also collect county land area measurements from the U.S. Census Bureau to obtain the size of each California county by filtering out all counties outside of California. Land area is measured in square miles, which I convert to acres, and merge with the fire data by county. I repeat this procedure with land area by states, and merge it with the state fire data. In addition to the total number of fires and acres burned, I include the number of fires per capita, number of acres burned per capita, and proportion of acres burned. In order to calculate these additional statistics, I gather population data for all 58 California counties and for the other 49 states. For the number of fires per capita and number of acres burned per capita, I simply divide the total number of fires and total number of acres burned, respectively, by the population of the county, or state, depending on which fire data set I am using. I calculate the proportion of acres burned per capita by dividing the total number of acres burned by the total land area of each county, or state, depending on the data set.

Merging

Once the migration data is gathered and the fire data is merged with other relevant data to create new fire statistics, I merge the county-to-county migration from the first time period with the county fire data from the first time period twice: first by the source county, and then by the destination county. I repeat this with county-to-county migration and county fire data for the second time period. I follow the same procedure for the two different time periods for the county-to-state migration and fire data. I then end up with four merged data sets: county-to-county and county-to-state data for two different time periods. In order to treat this as time series data, I appended the county-to-county data from the first time period to the second, and did the same for the county-to-state data. I end up with two different data sets, one for county-to-county, and the other for county-to-state. Once all data sets have been merged and appended, I divide the out-flow of migrants for each observation by the total population of the source county for both the county-to-county and county-to-state data sets. This is to account for the variation in population of each California county, since larger counties such as Los Angeles County will obviously have high raw numbers of out-migrants by virtue of the fact that it is the largest county by population in California, and thus will seem to experience greater out-migration relative to smaller, more rural counties with populations counts in the thousands.

Results

Regressions on County-to-County Migration Flow

I run three different regressions on three fire statistics using both the multiple regression and fixed effect models for the county-to-county migration data. The migration data included in the regression is out-migration expressed as a proportion of the population of the source county. Table 1 shows the results of the first regression in which migration is regressed on the number of fires in both the source county (County A) and the destination county (County B). The output in column (1) is from the model without fixed-effects, i.e., the multiple regression model, for the first time period (2010-2014 aggregated out-migration), and the second column displays the output from the multiple regression model for the second time period (2014-2018 aggregated out-migration). The third column contains output from the fixed-effects model, in which fixed-effects are applied to county A, county B, and the time period. In the multiple regression models for both time periods, both regressors are statistically significant at the 0.01 level; however, once time and county fixed-effects are applied, the results are no longer statistically significant. This indicates that the statistical significance achieved in the first model was likely due to variation in counties and/or the time periods. The coefficients for these regressors are extremely small due to the fact that the out-migration flow is measured in proportion of the source county, making the values negligible. In the multiple regression model applied to the first time period (Period 0), the coefficient for County A Number of Fires is positive, suggesting that a one-point increase in the number of fires in County A (the source county) leads to a 0.00005 unit increase in the proportion of out-migrants to County B. The coefficient for County B Number of Fires, however, is negative, indicating that a one-point increase in the number of fires in County B (the destination county) leads to a 0.00001 decrease in the out-flow proportion of migrants from County A to County B. The coefficients for the multiple regression applied to the second time period (Period 1) are also extremely small, positive for County

A, and negative for County B, with statistical significance at 0.001 level. These results align with my hypothesis that people in higher-risk counties make migration decisions based on the risk level of their origin county and their destination county. That is, on average, I expect that people will choose to migrate to safer counties with respect to fire risk. However, In the fixed effects model, the coefficients for both regressors are positive, but effectively 0, suggesting that the number of fires in both counties has no bearing on migration out of County A or into County B. Since both models include time-series data, the overall conclusion is that there is no evidence of a causal relationship between the number of fires in both counties and the out-flow of migrants between both those counties, since it is standard practice to include fixed-effects for time-series data.

The second regression is of migration flow from County A to County B on the number of fires per capita. Table 2 shows that, as with the first regression, the regressors are statistically significant when applied to the multiple regression model without fixed-effects for both time periods, but when fixed-effects are applied, there is no evidence of a causal relationship between out-migration and number of fires per capita. Despite the fact that the coefficients are statistically significant for the multiple regression model, the direction of the effect is opposite from what my hypothesis proposes. Table 2 implies that, in Period 0, a 1 point increase in the number of fires per capita in County A leads to a 2.571 decrease in the proportion of out-migrants from County A to County B, while a 1 point increase in the number of fires per capita in County B leads to a 11.520 unit increase in the proportion of out-migrants. This is contrary to the results from the multiple regression model in Table 1, which suggested that higher fire-risk in County A and lower-risk in County B both lead to higher proportions of flow of out-migrants from County A to County B.

Table 3 shows a similar trend for the regression of migration flow on the number of acres burned per capita for both counties. In the multiple regression on both time periods, the coefficients are close to 0 and negative for both counties in Period 0. However, in Period 1, only County A takes a negative coefficient. All coefficients, with the exception of County B's fire statistic measure in Period 0, are statistically significant. However, when the fixed-effects model is applied to the time-series data including observations from both time periods, the coefficients fir the fire regressors are no longer statistically significant, and are positive for both County A and County B measures.

Regressions on County-to-State Migration Flow

For the county-to-state migration data, I use the same method for evaluating effects of wildfires on migration in the county-to-county data; I run three multiple regressions on observations for both time periods separately, and then another three regressions using the same regressors as the multiple-regressions, with fixed effects applied to county, state, and time period. In Table 4, the coefficients of the multiple regression on the number of fires for both time periods are positive for the county and negative for another state. Similar to the county-to-county regression on number of fires, the coefficients take extremely small values that are close to 0; this is natural given that the out-migration flow is measured as a proportion of the population of the source county. Since the response variable in the regression takes on values less than 1, it follows that the coefficient for the regressor should be small. The coefficients for both counties in both time periods are statistically significant; however, when the fixed-effect model is applied, the coefficients for both county and state are positive, indicating that out-migration to another state increases as the number of fires in the state increases, which is incongruous with my hypothesis. However, the magnitude of the coefficient for number of fires in the state under the fixed-effect model is effectively 0, making its effect negligible. When fixed-effects are applied, the regressors are no longer statistically significant, indicating that the significance of the regressors in the multiple-regression models applied on both time periods are absorbed by the county, state, and/or time fixed effects.

Table 5 shows the output for the regression on the proportion of acres burned. For the multiple regression in both time periods, the coefficients for another state are greater than 0, but are less than 0 for the county measure, which goes against the hypothesis that increased frequency of fires in a county negatively impacts migration to that county. The results are statistically significant for all regressors, except for the county fire measure in Period 1. When the fixed-effects model is applied, the direction of effect flips for the county and state fire measures, which matches with the hypothesis, unlike the output of the multiple-regression model. However, with the change in direction of magnitude came the loss of statistical significance. Thus, I cannot confidently attribute the change in migration to the regressors in the model, which in this case are the proportion of acres burned in the source county and the destination state.

The multiple regression model on number of fires per capita is not statistically significant in Period 0, but the regressors have a much larger magnitude compared to the other fire measures used in the regressions from Table 5 and 6. The county measure in Period 1 is the only statistically significant result, and is also the only measure in all six regressions that remains statistically after the fixed-effects are applied. The model indicates that, on average, a one unit increase in the number of fires per capita in the source county leads to a 1.701 unit decrease in out-migration from the county to another state. Even though the result is statistically significant, the direction of this effect is contrary to what I expected; I would have expected that an increase in the number of fires per capita in the county would lead to an increase in out-migration from that county, rather than a decrease, since people might feel at higher-risk and would want to move to a safer location.

Conclusion

Contrary to my hypothesis, my model indicates that no association exists between California wildfires and migration. While certain regressors appear to be statistically significant under the multiple regression models, the effect is often in the opposite direction than I had expected. Under the fixed effects model, only the County Number of Fires is statistically significant, but the direction of this regressor on migration is negative, which is contrary to my hypothesis. Given these contradictions, I cannot make a conclusive statement about the effect of wildfires on migration. Despite the incompatible findings with my hypothesis and existing research, as well as the lack of statistical significance for my regressors, I believe it would be inappropriate to conclude that a relationship does not exist between migration and wildfires in California for a number of reasons.

While wildfires have always been a common natural phenomenon in California, the record-breaking nature of California wildfires has been more recent, occurring over the last five years. The severity and frequency of these more recent extreme fire events are likely attributable to rising temperatures, drought, and less rainfall, all partly due to anthropogenic climate change. As a result, fire events from the years included in my model probably had little to no bearing on the migration decisions of Californians in proximity to these events, since fires back then were less frequent and rarely life-threatening. Given that the fire data in my model only includes fire events from 2009 to 2017, my model doesn't reflect the harsh, damaging nature of the fire seasons over the last few years. Thus, my model might produce more clarity of whether a relationship exists between wildfires and migration when historical fire data includes more recent data with years during which there were extreme fire seasons.

My model also fails to include regressors other than fire measures that account for frequency and severity of wildfires. As mentioned in the literature review, many papers on climate and migration account for amenities in their model, given that people tend to migrate to areas with higher environmental and natural amenities. Given that California has strong migration pulls—a robust economy, moderate Mediterranean-climate, coastal location, topographical variation, and a high number of environmental amenities—these positive effects on migration should be controlled in the model to obtain more accurate results. To accomplish this, I should include similar regressors for the other states in my county-to-state model to account for other variation in migration that may have been mistakenly attributed to wildfires in my current model, when in it should have instead been attributed to an omitted variable. Moving forward, I would include regressors in the model that account for these other variables that would likely have an effect on migration, in order to remove any confounding effects or omitted variable bias that may exist in my current model. If incorporating more recent data and additional non-fire regressors still fail to produce significant results, I could then say more conclusively whether or not California wildfires and human migration are associated.

Tables

Table 1: Regression of Migration Flow on Number of Fires

	Dependent variable:		
	Migration Flow Proportion from County A to County B		
	No FE Period 0	No FE Period 1	FE
	(1)	(2)	(3)
County A Number of Fires	0.00005*** (0.00002, 0.0001)	0.00001*** (0.00001, 0.00002)	0.00000(-0.00001,0.00001)
County B Number of Fires	-0.0001*** (−0.0001, −0.00003)	-0.00001*** (−0.00002, −0.00001)	0.00000(-0.00001,0.00001)
Constant	0.001*** (0.001, 0.001)	0.001*** (0.001, 0.001)	
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2: Regression of Migration Flow on Number of Fires Per Capita

	Dependent variable:		
	Migration Flow Proportion from County A to County B		
	No FE Period 0	No FE Period 1	FE
	(1)	(2)	(3)
County A Number of Fires Per Capita	-2.571*** (−3.282, −1.860)	-0.662*** (−0.786, −0.539)	0.043(-0.339,0.425)
County B Number of Fires Per Capita	11.520* (1.469, 21.571)	1.259*** (0.641, 1.878)	0.134(-0.248,0.516)
Constant	0.001*** (0.0005, 0.001)	0.001*** (0.001, 0.001)	
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3: Regression of Migration Flow on Number of Acres Burned Per Capita

	Dependent variable:		
	Migration Flow Proportion from County A to County B		
	No FE Period 0	No FE Period 1	FE
	(1)	(2)	(3)
County A Number of Acres Burned Per Capita	-0.0001*** (−0.0002, −0.0001)	-0.0002*** (−0.0002, −0.0001)	0.00002(-0.0001,0.0001)
County B Number of Acres Burned Per Capita	-0.00001(-0.0001,0.0001)	0.0001*** (0.00004, 0.0002)	0.00003(-0.00005,0.0001)
Constant	0.001*** (0.001, 0.001)	0.001*** (0.001, 0.001)	
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 4: Regression of Migration Flow on Number of Fires

	Dependent variable:		
	Migration Flow Proportion from County to State		
	No FE Period 0	No FE Period 1	FE
	(1)	(2)	(3)
State Number of Fires	0.00000*** (0.000, 0.00000)	0.00000*** (0.00000, 0.00000)	0.000(-0.00000,0.00000)
County Number of Fires	-0.00005** (−0.0001, −0.00001)	-0.00000** (−0.00001, −0.00000)	0.00001(-0.00000,0.00002)
Constant	0.001*** (0.0003, 0.001)	0.0004*** (0.0003, 0.0005)	
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 5: Regression of Migration Flow on Proportion of Acres Burned

	Dependent variable:		
	Migration Flow Proportion from County to State		
	No FE Period 0	No FE Period 1	FE
	(1)	(2)	(3)
State Proportion Acres Burned	0.029*** (0.018, 0.040)	0.029*** (0.023, 0.034)	-0.001(-0.016,0.014)
County Proportion Acres Burned	-0.003** (−0.005, −0.001)	-0.0004(-0.001,0.0001)	0.001(-0.001,0.003)
Constant	0.0003*** (0.0002, 0.0004)	0.0002*** (0.0002, 0.0003)	
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 6: Regression of Migration Flow on Number of Fires Per Capita

	Dependent variable:		
	Migration Flow Proportion from County to State		
	No FE Period 0	No FE Period 1	FE
	(1)	(2)	(3)
State Number Fires Per Capita	0.052(-0.095,0.199)	-0.001(-0.017,0.015)	-0.088(-0.384,0.209)
County Number Fires Per Capita	32.620(-1.147,66.386)	1.095*** (0.598, 1.592)	-1.701*** (−2.372, −1.029)
Constant	-0.0002(-0.001,0.0004)	0.0004*** (0.0003, 0.0004)	
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 7: County-to-County Regressor Averages

Regressor	**County Average**
Number of Fires	9.93167
Number of Fires Per Capita	0.00009
Number of Acres Burned Per Capita	0.34965

Table 8: CA County-to-State Regressor Averages

Regressor	**State Average**	**County Average**
Number of Fires	6178.06713	10.13925
Propotion of Acres Burned	0.01000	0.02436
Number of Fires Per Capita	0.00159	0.00007

References

Budiman, A. (2020, September 22). *Key Findings about U.S. Immigrants.* Pew Research Center..

California. (2019, December). Forbes.

California Department of Forestry and Fire Protection. *Incidents Overview*. Cal Fire Department of Forestry and Fire Protection.

The Effects of Immigration on the United States' Economy. (2018, December 19). Penn Wharton Budget Model.

Facts + Statistics: Wildfires. Insurance Information Institute.

Hassani-Mahmooel, B., & Parris, B. W. (2012, December). Climate change and internal migration patterns in Bangladesh: an agent-based model. *Environment and Development Economics.*

Immigrants in California. (2020, August 11). American Immigration Council.

Stetler, K. M., Venn, T. J. & Calkin, D. E. (2010). The effects of wildfire and environmental amenities on property values in northwest Montana, USA. *Ecological Economics.*

Temple, J. (2020, October 13). *This Is What California Needs to Do about Its Fires.* MIT Technology Review.

US Census Bureau. (2020, June 22). *County Population Totals: 2010-2019*. The United States Census Bureau.

US Census Bureau. (2020, August). *County-to-County Migration Flows.* The United States Census Bureau.

US Census Bureau. (2020, August 6). *USA Counties: 2011*. The United States Census Bureau.

Winkler, R. L., & Rouleau, M. D. (2020). Amenities or disamenities? Estimating the impacts of extreme heat and wildfire on domestic US migration. *Springer Nature B.V.*