*Preliminary and incomplete: Comments welcome*

# Extreme Weather Events and Rural-to-Urban Migration

Trudy Ann Cameron*
Department of Economics, University of Oregon

Raisa Saif
Program in Environmental Studies, University of Oregon

Eric Duquette
Economic Research Service, U.S. Department of Agriculture

## ABSTRACT

In numerous regions around the globe, climate change can be expected to change the pattern of severe weather events. The nature of future changes in these patterns can be difficult to predict, but it is instructive to consider some of the potential consequences of extreme weather on household migration decisions based on past events. We examine county-to-county migration decisions in the U.S., treating various types of extreme weather events as random exogenous shocks to the affected communities and their economies. We are particularly interested in whether rural-to-urban migration flows are altered systematically in the wake of extreme weather events. We explore a variety of specifications for a panel of roughly half a million significant annual U.S. county-to-county flows. Our models demonstrate that the effects of a number of different types of extreme weather events (i.e. flooding, heat waves, and wildfires) in the origin county on county-to-county migration flows is statistically significantly greater when the destination county is relatively more urbanized. The effect of the number of fatalities from flooding and heat waves in the origin county on migration flows is also amplified when the destination county is more urbanized. Thus it appears that even in a developed country like the U.S., extreme weather events continue to exacerbate rural-to-urban migration flows.

JEL Classifications: Q54, R11, R23, O15,
Keywords: migration, rural-to-urban migration, natural hazards, extreme weather, urbanization, gravity model

* Corresponding author: Department of Economics, 435 PLC, 1285 University of Oregon, Eugene, OR 97403-1285; (541) 346-1242; cameron@uoregon.edu

Extreme Weather Events and Rural-to-Urban Migration

## 1. Introduction

Climate change has the potential to change the pattern of extreme weather events. It is difficult to predict the nature of future changes in the frequency, severity or geographical distribution of extreme weather events, but it is instructive to consider some of the potential consequences of extreme weather on household migration decisions, based on events in the recent past.

Of course, many developing countries are more vulnerable than the U.S. to physical and economic damage from these types of shocks. For example, observation in some developing countries (e.g. Bangladesh) suggests that increasing rates of weather-related disasters have the effect of driving increasing numbers of rural dwellers off the land and into urbanized areas. This type of displacement can put considerable strain on a society's resources. Unfortunately, detailed migration data at a high level of spatial and temporal resolution (and conformable weather data) for countries like Bangladesh are not readily available.

Instead, we examine six years of annual county-to-county migration decisions in the U.S. between 2005 and 2010, treating various types of extreme weather events as random exogenous shocks to the affected counties and their economies. We consider ten types of events: floods, droughts, hailstorms, heat waves, hurricanes, severe storms, tornadoes, wildfires, wind storms, and severe winter weather. We are particularly interested in whether rural-to-urban migration flows are altered systematically in the wake of extreme weather events of different types. Our data for the U.S. suggest that this can be the case, especially for some types of weather-related natural hazards.

We explore a variety of specifications for a panel of roughly half a million significant annual U.S. county-to-county flows. Our models demonstrate that the effects of a number of different types of extreme weather events in the origin county on county-to-county migration flows is statistically significantly greater when the destination county is relatively more urbanized. In many cases, severe weather events increase rural-to-urban flows and decrease urban-to-rural flows. In other cases, almost all flows are increased by severe weather events (i.e. wildfires), but rural-to-urban flows are increased by more than urban-to-rural flows. Even in a developed country like the U.S., therefore, some types of extreme weather events continue to exacerbate net rural-to-urban migration flows.

### a. Literature

An early survey of the literature on the determinants of gross migration between areas in the U.S. is provided in Greenwood (1975). Much of the research on migration in economics has had to do with the question of labor mobility. He notes that most such models are "gravity-type" models where migration is hypothesized to be directly related to the size of the relevant origin and destination populations and inversely related to distance (e.g. Schwartz (1973)). Other determinants of migration can include differential rates of regional income growth from the perspective of individual potential migrants, although income opportunities tend to provide better explanations of in-migration than out-migration.

Concerning rural-urban migration, Greenwood (1975) describes some evidence that the unemployment rate in the urban area can serve as an indicator of whether potential rural-urban migrants will find employment within a given time-period, but notes that some studies have found unanticipated signs or insignificant coefficients on an unemployment rate variable (attributed to simultaneity bias).[1] At the individual level, Nakosteen and Zimmer (1980) consider the selectivity-bias question in detail, as do Borjas et al. (1992).

The urban economics literature during the 1970's was also concerned with the determinants of migration, for example Graves and Linneman (1979). In the urban economics view of migration, market rents and wages are assumed to adjust continuously to leave utility levels essentially constant across space so that any observed wage differentials are compensated by differences in other amenities.[2] Climate is considered as a determinant of migration in Graves (1979) and Graves (1980), although the influence of climate in general as a factor affecting migration is different from the influence of specific extreme weather events addressed in the current paper.

Greenwood (1985) emphasizes the effects of growing quantities of micro-data that permit more refined models of the effects of life-cycle stages on migration for individual households (e.g. Davanzo (1978), Davanzo (1983)). However, detailed information about place is censored in most of these datasets (e.g. the PSID), precluding much study of the effects of locational characteristics on migration. However, he also notes the growing focus on time-series information about migration, although the decade-long census interval prevents much time-wise resolution, and the notion that lags in the migration decision process may be important. However, most of the cited data is inter-country, or at most between four regions of the U.S. Stark and Bloom (1985) emphasize the potential for micro-data and improved econometric

[1] Muth (1971), for example, questions whether differential rates of migration are induced by differential growth in job opportunities or employment, or whether differential changes in employment are induced by differential rates of in-migration.

[2] Greenwood et al. (1991) allow for both equilibrium and disequilibrium in their model of net migration, and find only minor departures from equilibrium, lending support to the literature where the value of amenities is estimated from compensating differentials in housing prices and wages.

methods to enhance economists' understanding of the determinants of individual migration decisions based on labor market opportunities.

In the regional science literature, Partridge et al. (2012) note that there has been a significant downward shift in U.S. gross migration rates after 2000 and question whether this is evidence of an approach to spatial equilibrium. Examining U.S. counties, they find "only slight ebbing of natural amenity-based migration after 2000 and little slowing of population redistribution from peripheral towards core urban areas."

### b. Climate, weather, and migration

The impact of climate and weather on migration has been considered for sub-Saharan Africa by Barrios et al. (2006) and by Marchiori et al. (2012), based on annual cross-country data. The latter paper estimates that temperature and rainfall anomalies caused a total net displacement of five million people during 1960 through 2000, and projects that future weather anomalies could lead to an addition 11.8 million displaced people by the end of the current century.

For Malawi, Lewin et al. (2012) examine whether rainfall conditions influence rural workers' decision to move to urban areas or other rural areas. They find that rainfall shocks are negatively associated with rural out-migration, that migrants choose communities where rainfall variability and the chances of a drought are lower.[3]

Concerning droughts, Gray and Mueller (2012) use a household panel dataset from Ethiopia (1,500 households over 15 years) in a multinomial logit model to examine how different types of mobility are influenced by proportion of households exposed to drought (where the effects are heterogeneous by household characteristics). The poor are more severely affected, and women's marriage-related mobility is constrained by income losses due to drought. One

[3] Still awaiting access to the full article.

important insight is that adverse conditions can reduce mobility by undermining the necessary resources for mobility.[4] We find some evidence in our own study to support reductions in mobility due to extreme weather events that damage assets for many households.

Boustan et al. (2012) examine net migration away from disaster-struck areas during the 1920's and 1930's in the U.S., using information on disasters from the American Red Cross and migration activity using two new panel dataset from Census sources, one tracking individuals from 1920 to 1930, and another between 1935 and 1940. Using 15,000 randomly selected men who could be unambiguously matched across census years, they find that, "on net, young men move away from areas hit by tornadoes but are attracted to areas experiencing floods." These authors do not emphasize their results for hurricanes or earthquakes, since there were so few such events. Other controls indicate that migrants sought warmer winters and cooler summers.[5] The main argument in their paper is that public efforts at disaster mitigation may counteract individual migration decisions, making current residents less likely to move out and prospective residents more likely to move in. This work reflects the issues examined in Charney (1993), who reviews the relationships between migration flows and public policy decisions (national policies such as defense spending, migration subsidies, and intergovernmental transfers, and regional policies such as welfare and unemployment benefits, taxation, education and other public goods).

Hornbeck (2009) confirms the expected out-migration from Dust Bowl states in the mid-1930's but Deryugina (2011) finds no net population changes for U.S. counties struck by hurricanes in the 1980s and 1990s. Transfer payments to victims of severe weather can cushion

[4] See Laczko and Aghazarm (2009).

[5] Boustan et al. (2012) also control for (but do not include coefficients for ) "the logarithm of total population and of land area; the black population share; a quadratic in latitude and in longitude; a dummy variable equal to one for SEAs with some coastal exposure; a proxy for employment growth in the SEA; total disaster count in the previous or subsequent decade; and a quartic in distance between the SEA of origin and the current SEA."

the impacts of these events, either directly, or indirectly through compensating for job losses and other displacements.

Rappaport (2007) observes that U.S. residents have been moving to places with nicer weather. They move towards places with warmer winters, but also to places with cooler and less-humid summer seasons, despite the role of air-conditioning. Migration patterns cannot be explained entirely by shifting industrial activities or by the residential location choices of the elderly. Instead, it is argued that rising per-capita income has increased demand for nice weather as an amenity. Earlier, however, Cragg and Kahn (1999) observed that between 1960 and 1990, consumption of warmer winter climates rose for both working families and especially for seniors. At the same time, the relative price of a desirable climate increased for seniors but fell for working families.

Deschenes and Moretti (2009) consider extreme heat and extreme cold events and their effects on mortality rates in the U.S. They note that mobility in the U.S. between the Northeast and the Southwest over the past thirty years may account for as much as 4% to 7% of total tains in life expectancy. However, their model does not explicitly track mobility as a consequence of extreme weather. The return-versus-migrate decisions of evacuees in the wake of a single extreme weather event—Hurricane Katrina—are considered by Landry et al. (2007).

Cragg and Kahn (1997) infer willingness to pay for a more moderate climate from migration decisions based on Census data.

### c. Interdisciplinary research

Piguet (2010) offers an interdisciplinary survey of approaches that have been used to explore the connections between migration and environmental factors, and Black et al. (2011)

conceptualize five broad “families of drivers,” that influence migration decisions—economic, political, social, demographic, and environmental.

Fielding (2011) considers the likely impact of climate change on internal migration in the U.K. and concludes that climate change can be accommodated without any major redistribution of its population. Minor exceptions are predicted to involve river and coastal flooding which will make some areas hazardous to settlements and/or costly to protect from flooding.

## 2. Model

We design our model to explain county-to-county migration based on a traditional “gravity” model for migration as described in the early survey by Greenwood (1975) and as used by many subsequent researchers. The standard equation for gravitational attraction is $A = g\, m_1 m_2 / d^2$, where the arguments include the mass of each object and the distance between them, and the parameter $g$ is the gravitational constant. We adapt this equation to explain migration as a function of the populations of the origin and destination counties, and this distance between the county centroids. We convert the gravitational “constant” into a systematic varying parameter that is allowed to depend upon the attributes of the origin and destination counties, as dictated by the data. If we include a random error term $\varepsilon$, our basic formula to explain migration between county $i$ and county $j$ in year $t$ is given by:

$$migration_{ijt} = g\left[X_{ijt}\right] \frac{\left(Pop_i\right)^{\beta_1} \left(Pop_j\right)^{\beta_2}}{\left(distance_{ij}\right)^{2\beta_3}} \exp\left(\varepsilon_{ijt}\right) \tag{1}$$

The parameters $\beta_1$, $\beta_2$ and $\beta_3$ are all equal to one in the standard gravitational formula, but we will allow them to take on whatever values the data imply and test whether $\beta_1 = \beta_2 = \beta_3 = 1$ can be rejected statistically.

Taking logarithms to produce an expression that is linear in the unknown parameters produces:

$$\begin{aligned}\log\left(migration_{ijt}\right) &= \log\left(g\left[X_{ijt}\right]\right) \\ &+\beta_1 \log\left(Pop_i\right) + \beta_2 \log\left(Pop_j\right) - 2\beta_3 \log\left(distance_{ij}\right) + \varepsilon_{ijt}\end{aligned} \tag{2}$$

We will model the first term, $\log(g[X_{ijt}])$, as a function of set of variables that measures different types of weather hazards in origin county $i$ in year $t$, $weather_{kit}$, where $k =$ floods, droughts, hailstorms, heat waves, hurricanes, severe storms, tornadoes, wildfires, wind storms and winter weather. We will also allow the derivatives of $\log\left(migration_{ijt}\right)$ with respect to each of these measures of weather hazards to depend upon the differences between destination and origin counties in proportion of the population in urbanized areas. Finally, to account for unobserved heterogeneity at the state level, we include fixed effects for each origin state and fixed effects for each destination state, as well as year fixed effects:

$$\begin{aligned}\log\left(migration_{ijt}\right) = \beta_0 &+ \textstyle\sum_{k=1}^{10} \beta_{4k} weather_{kit} \\ &+\beta_4\left[PropUrban_j - PropUrban_i\right] \\ &+\textstyle\sum_{k=1}^{10} \beta_{5k}\left(weather_{kit} \times \left[PropUrban_j - PropUrban_i\right]\right) \\ &+\beta_6 stateFE_i + \beta_7 stateFE_j + \beta_8 yearFE_t \\ &+\beta_1 \log\left(Pop_i\right) + \beta_2 \log\left(Pop_j\right) - 2\beta_3 \log\left(distance_{ij}\right) + \varepsilon_{ijt}\end{aligned} \tag{3}$$

For this paper, our main interest concerns the question of whether extreme weather events in the origin county affect county-to-county migration flows. But in particular, we are curious to know if these effects are greater for flows from more-rural to urbanized counties and less for flows from urbanized to more-rural counties. The model in equation (3) implies that the derivative of expected migration flows between county $i$ and county $j$, with respect to extreme weather events of type $k$ in the origin county is given by:

$$\frac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right], \qquad k = 1,\ldots,10 \tag{4}$$

We estimate several versions of the model in equation (3), for different measures of the impact of each type of extreme weather. Based on each version of the model, we consider the ten different derivatives in equation (4). A given type of extreme weather affects migration if the corresponding expression is statistically significantly different from zero for at least some value of $\left[PropUrban_j - PropUrban_i\right]$ in the data. This weather event favors rural-to-urban migration of the estimated coefficient $\beta_{5k}$ is positive and statistically significantly different from zero. In other words, a given type of extreme weather favors rural-to-urban migration if the derivatives in equation (4) are plotted as a function of $\left[PropUrban_j - PropUrban_i\right]$ and (a) their confidence bounds exclude zero for at least some portion of the range and (b) the slope of the function is positive.

## 3. Data

We combine two main sources of data for this study: (1) migration information from the U.S. Internal Revenue Service, and (2) information on weather-related natural hazards from the Spatial Hazard Events and Losses Database for the United States (SHELDUS) database. We also make use of several other minor sources that permit us to connect these two data sets and/or provide the necessary control variables.

### a. Migration data (IRS)

We measure county-to-county migration using the Statistics of Income (SOI) Tax Stats files provided by the U.S. Internal Revenue Service.[6] These data files include state and county codes for each origin and destination county, the state abbreviation and county name, plus the number of tax returns that were submitted by the same person from the origin county on their previous return and the destination county on their current return. Also provided are the number of exemptions claimed on that return (roughly the total number of household members associated with all of those tax return), and the "adjusted gross income" aggregated across these returns.

The first salient feature of these data is their sheer volume. There are roughly 3100 counties in the U.S. and the potential number of migration flows to account for in any one year is thus on the order of 10 million. We will build a panel of data for a span of six years, meaning that the entire dataset involves on the order of 60 million observations.

There are a number of limitations to these data, however, described in detail in Gross (unknown). For our purposes, the most significant limitation is that the flow of "households" (which we will refer to as "migrants") between any pair of counties is censored if the number of tax returns is less than ten. In cases like this, the flows are aggregated with a total for a larger origin or destination area. Censoring according to the absolute magnitude of the key dependent variable in a study presents statistical problems. With a smaller dataset, it would be appropriate to use maximum likelihood methods developed specifically for censored data, such as a truncated Poisson or truncated negative binomial estimator, where we could explicitly acknowledge the absence of information about pairs of counties characterized by low migration flows. For the present study, we merely acknowledge that our models cannot explain minor

[6] These files may be accessed via http://www.irs.gov/taxstats/article/0,,id=212695,00.html .

flows of migrants and we focus our attention on those county pairs that display “significant flows” in each year of our dataset. This reduces the size of our (unbalanced) pane of county pairs over six years to a more manageable sample on the order of 500,000 observations. Table 1 shows our data on annual county-to-county migration rates, per county pair (for flows>=10). The two available measures of migration flows are the number of tax returns (roughly the number of households) and the number of exemptions (roughly the number of individuals), keeping in mind that households who do not submit tax returns are not reflected by these measures.

The map in Figure XX [pending] shows those counties that appear as origins in at least one year of our six-year sample, and the map in Figure XX [pending] shows those counties that appear as destinations in at least one year of our six-year sample.

### b. Extreme Weather Events (SHELDUS)

The SHELDUS dataset is a county-level hazard data set for the United States.[7] The dataset contains information about eighteen different types of hazards, but we use just ten of these hazards, all related to weather. This includes information about events that involve flooding, droughts, hail, heat, hurricanes or tropical storms, severe storms, tornadoes, wildfires, wind, and winter weather.

The SHELDUS data record beginning dates and ending dates for each event. We use this information to calculate a duration variable for each event. We also use the beginning date to allocate each specific event to a particular tax year. We define a variable called taxdue_year. If a weather event occurs on or before April 15, it is assigned to the current taxdue year. If it occurs

[7] The SHELDUS files may be accessed via http://webra.cas.sc.edu/hvri/products/sheldus.aspx .

after April 15, it is assigned to the next taxdue year. This is necessary because the migration data hinge upon changes of address between tax filings.[8]

We then collapse the detailed data on individual weather events to produce counts or totals during the relevant April-to-April years for our study. We produce a count of the number of distinct events (events), a sum of the total number of days for each type of event during each year (tdays), a sum of the total number of injuries attributed to each event (tinju) and the total number of fatalities (tfata). The SHELDUS data also provides some information about property damages, so we calculate the total amount of crop damage (tcrop) and the total amount of property damage (tprop).

Table 2 lists data for each type of event, by year, for the years 2005 through 2010 inclusive, which are the years covered by our migration data.

### c. Populations of Origin and Destination Counties

County populations are based on U.S. Census data from the American Community Survey (ACS) provided as averages for the years 2006-2010.[9] Our dependent variable is the number of tax returns (if greater than ten) for which the tax-filer moved between counties, so we choose the household as the corresponding unit of measurement for population in the origin and destination counties. These population measures thus vary across counties but not across time.

[8] We acknowledge that some people will submit their taxes prior to April 15 and some weather events that begin before April 15 will continue past that date. However, we adopt this benchmark date as the most practical approximation.

[9] The data can be downloaded from the U.S. Census Bureau's State & County Quickfacts database http://quickfacts.census.gov/qfd/download_data.html. Other available county-level variables include area, population per square mile, population same house one year ago, foreign born, percent speaking language other than English at home, educational attainment, average travel time to work, owner-occupied housing units, multi-dwelling structures, median house values, average household size, per capita income, median household income, percent in poverty.

### d. Distances Between Counties

For the distances between each pair of counties in the U.S., we use spherical distance data prepared by Merryman (2005) and Pisati (2001). Great circle distances are measured in miles and are calculated using:

$$d\left(x_1, y_1; x_2, y_2\right) = 6378\left[ \operatorname{acos}\left\{ \begin{array}{l} \sin\left(x_1 \dfrac{\pi}{180}\right) \cdot \sin\left(x_2 \dfrac{\pi}{180}\right) \\ + \left[ \begin{array}{l} \cos\left(x_1 \dfrac{\pi}{180}\right) \cdot \cos\left(x_2 \dfrac{\pi}{180}\right) \\ \cdot \cos\left(\left(y_1 - y_2\right) \dfrac{\pi}{180}\right) \end{array} \right] \end{array} \right\} /1.609 \right] \tag{5}$$

The variables $y$ and $x$ denote latitude and longitude respectively.

### e. County Proportions Urbanized

Ideally, of course, one would like to know whether a specific state-to-state migrant moved from a non-urban portion of his or her origin county to an urban portion of his or her destination county, but this information is not publicly available. Presumably, the IRS knows this information, since the raw data are based on origin and destination addresses for individual tax filers. Unfortunately, confidentiality precludes access to this richer individual data.

Instead, we construct our crucial urbanization variable using U.S. Census data for the year 2000 census. By county, we use the counts of persons living in areas within the county that are designated as “urban” by the census. We calculate the share of the county population living in urbanized areas for each origin and destination county as of the year 2000.[10]

[10] Subsets of the urban population include counts of the population living in “urbanized areas” (densely settled suburban developments in the vicinity of large cities) and “urban clusters” (areas containing at least 2,500 and less than 50,000 people—the built-up territory around small towns and cities). In 2000, about 68% of the U.S. population lived in 452 urbanized areas, and about 11% lived in 3,158 urban clusters. The rural population (anything

To capture whether a migration flow between counties can be considered as "rural-to-urban" or "urban-to-rural," we use the difference between the destination county's urban population share and the origin county's urban population share. The resulting variable ranges between -1 and +1, with positive values signifying rural-to-urban flows and negative values signifying urban-to-rural flows. Zero values imply moves between counties with equal proportions of their populations in urbanized areas.

This constructed variable measures the degree to which county-to-county migration is rural-to-urban if we assume that individual migrant households are drawn randomly from their origin counties and distributed randomly across their destination counties. To the extent that this assumption is violated, there will be a degree of measurement error in this variable. However, this constructed variable seems to be the best available candidate to capture this crucial characteristic of each possible county-to-county migration flow.

### f. Other data sources

The weather events that we use in this paper to explain migration patterns are exogenous in the sense that migration patterns in the current year can be assumed to have no effect on the weather. In a reduced form equation, we can expect little in the way of omitted variables bias if we leave out other determinants of migration flows between counties. However, we can reduce the variance in the estimates if we include other logical explanatory variables, and we can also assess the fitted model for the plausibility of the estimated coefficients.

One factor to consider is that weather events may have both direct and indirect effects on migration. People may move because their households have been directly affected by severe weather, as with floods or hurricanes, for example. Alternatively, they may move because these

outside urban areas—open country and settlements with fewer than 2,500 residents) is simply the complement of the urban population for each county.

events affect the local economy and their livelihoods may be so adversely affected that they move to a new county to seek new employment. If job losses or decreased earnings or a reduced number of establishments is attributable to weather, then weather indirectly affects migration decisions. In one case, we may wish to know the "reduced form" impact of weather on migration, suggesting that county employment, county wages, or the number of establishments in the county should be left out of the model. In other cases, we may wish to control for job changes to learn how much of the migration between a pair of counties is not attributable to local economic conditions in the origin or destination counties.

For this paper, we concentrate on reduced form estimates of the effects of extreme weather events in origin counties on migration flows between counties.

## 4. Estimation and Results

Our generalized gravity-type model stipulated a log-log specification for the main variables, so we adhere to this restriction in the current paper. Alternative specifications are summarized in the section on sensitivity analyses.

The generic generalized gravity model displayed in equation (3) leaves open the definition of "weather" to be used to quantify the impacts of the ten types of extreme weather events considered in this study. We are interested in establishing the "weight of the evidence" across alternative measures of the incidence of extreme weather in each origin county. For some types of extreme weather, only a subset of possible measures of these events produce statistically significant evidence that this type of extreme weather induces rural-to-urban migration. For full disclosure, we have elected to present the key results for each type of extreme weather—i.e. the derivatives in equation (4)—from all seven alternative measures of incidence, and we will resort to graphical depictions of these results in the discussion to follow.

Preliminarily, however, it is relevant to consider whether the rest of the specification behaves as expected. In general, the basic "gravity" portion of the model, that is the three coefficient $\beta_1$, $\beta_2$, and $\beta_3$, are hugely statistically significant and bear the correct signs and *relative* magnitudes. We conduct tests of whether all three coefficients are equal to one, however, and this joint test is soundly rejected in virtually all specifications.

Another set of important hypothesis tests concerns whether the $\beta_{4k}$ coefficients on the incidence variable alone, not interacted with $\left[PropUrban_j - PropUrban_i\right]$, are statistically significantly different from zero. If these coefficients are no different from zero, then we cannot reject the hypothesis that a plot of the derivative of migration with respect to $\left[PropUrban_j - PropUrban_i\right]$ is zero when $\left[PropUrban_j - PropUrban_i\right]$ is zero—that is, when the origin and destination counties have the same proportions of their populations designated as urban. If $\beta_{4k} = 0$ and $\beta_{5k} > 0$ simultaneously, this constitutes evidence that extreme weather of type $k$, according to the current measure of incidence, induces rural-to-urban migration. In contrast, if $\beta_{4k} = 0$ and $\beta_{5k} < 0$ simultaneously, we infer that this type of extreme weather, according to the current measure of incidence, induces urban-to-rural migration.

Tables XX-XX [not yet included] give the numerical estimates produced by our models. We include (a) the $\beta_{5k}$ slopes on the ten key interaction terms, (b) the $\beta_{4k}$ "intercepts" of each derivative that determine whether the vertical intercept for the plot of the derivative is zero, (c) the single-subscript $\beta_4$ coefficient on the urbanization differential variable, (d) the results of two hypothesis tests concerning the three standard "gravity" terms, and (e) mutually exclusive counts of the numbers of fixed effects that are significant at the 5% and at the 10% levels, for origin states, destination states, and years.

Figures 3a through 3i provide succinct visual summaries of the key derivatives given by equation (4). Each component graph in each of these multi-graph figures table depicts our key derivative as a function of $\left[PropUrban_j - PropUrban_i\right]$. For each decile of the distribution of $\left[PropUrban_j - PropUrban_i\right]$ in our sample, we show the calculated confidence bounds for the derivative. The background shading for each graph provides a quick visual indicator of a preliminary type of hypothesis test: the background is *white* if the confidence bound for the key derivative calculated at any decile of the distribution of the urbanization differential in a particular graph excludes zero (i.e. does not overlap the horizontal line drawn at zero on the vertical axis). The background is *grey* if the estimated derivative is *not* statistically significantly different from zero at any of the deciles of the distribution of $\left[PropUrban_j - PropUrban_i\right]$.[11]

In each figure, there are seven graphs per type of extreme weather. Each of these seven graphs reflects the two key coefficients for the specified type of extreme weather, and for a given type of weather, each graph is implied by a separate regression.  Each regression uses the specification in equation (3) including *all ten* weather types, but employing a *different* measure of the $weather_{kit}$ variable:  (1) an indicator for the occurrence of any weather events of that type, (2) a count of the number of events, (3) total days, (4) crop damage, (5) property damage, (6) human injuries, and (7) human fatalities.

[11] The median of the distribution of the differential urbanization variable is zero, and we show a vertical line at zero in each graph. Despite the loss of county pairs for which migration was less than 10, if the flow in one direction is greater than 10, the flow in the other direction is likely to be at least 10 as well, so the distribution is approximately symmetric.

Perusal of each of the graphs for each type of extreme weather reveals whether there is evidence of this type of weather, occurring between April 16th of one year and April 15 of the following year in the origin county, leading to statistically significant changes in the migration patterns captured by the significant flows (>10 households) in the IRS data. For graphs with grey backgrounds, the derivative is not statistically significant at any decile of the distribution of the urbanization differential. These cases may suggest increasing (or sometimes decreasing) derivatives as the urbanization differential varies from -1 to +1, but we cannot reject a derivative of zero. In other cases, however, the derivatives in question are statistically significantly different from zero for at least some values of the urbanization differential.

In the color versions of these figures, a yellow-highlighted frame around a graph indicates a statistically significant change in migration as a result of a change in the weather measure in question, for *at least some* decile of the urbanization differential, and *positive* dependence of the derivative on the urbanization differential.[12] This implies a larger increase in migration flows when the destination is relatively more urbanized (and correspondingly a smaller increase, or even a decrease in flows when the destination is relatively less urbanized). Notice that the vertical intercept of the derivative function, where the urbanization differential is zero, conveys the size and confidence interval for the parameter $\beta_{4k}$ in equation (4). At this point in the domain of the plotted function, the destination and origin counties have the same proportion of their populations designated as urban.

The horizontal intercept in each graph also has a useful interpretation. For positively sloped derivative functions, increases in the specified type of extreme weather event in the origin county *increase* migration flows along pathways where the urbanization differential is greater

[12] In black and white versions, given our current color palette, the yellow borders show up only faintly, whereas the gold borders are visible on most printers. Statistical significance and a *positive* slope fo the derivative as a function of the urbanization differential are therefore conveyed by a white background to the plot and no dark frame.

than the horizontal intercept of the plot. This illustrates that severe weather events *can* increase urban-to-rural flows as well, but in this case, these effects will be smaller than the increases in rural-to-urban flows. When the parameter $\beta_{4k}$ is (statistically) no different from zero, we can state with greater certainty that “weather events of this type increase rural-to-urban migration and decrease urban-to-rual migration.”

A gold highlighted frame around a graph indicates, again, a statistically significant change in migration for at least some decile of the urbanization differential, but negative dependence of this derivative on the urbanization differential, suggesting a decrease in flows.

### a. Robustness tests

We have estimated the models described here with a number of variations, including each of the following:

1. The same basic specification, but without clustering of the errors on the origin state
2. The same basic specification, but with robust standard errors
3. The same specification, including the “gravity model” controls of origin population, destination population, and distance between counties (but without the fixed effects for origin state, destination state, and year)
4. The same specification, including fixed effects for origin state, destination state, and year, but without the “gravity model” controls
5. A model with none of the fixed effects and none of the standard “gravity model” controls. Given that weather events are purely exogenous over the six-year time horizon, there should be no omitted variables bias from failure to control for other factors, just a loss of precision.

6. The same specification featured in the paper, but with the log of the number of exemptions claimed on the tax return, which converts the dependent variable from the log of the number of migrating households to the log of the number of migrating persons (at least those covered in the sample of tax returns for each county)
7. The same specification featured in the paper, but augmented by four additional controls: the proportions of the origin and destination counties living in poverty and the proportions over age 65. This specification is intended to reveal whether migration is less than otherwise predicted when either county has a larger fraction of people who would potentially be missed by the IRS counts of tax returns.

The consequences of these robustness checks vary somewhat across these six variations. However, the following themes emerge:

a.) Failure to control for the "gravity model" variables and failure to include fixed effects tends to result in confidence bounds for the predicted derivatives that are more likely to include zero. Thus these two classes of regressors are very important to our ability to discern patterns of rural-to-urban migration in response to extreme weather events in origin counties.

b.) Clustering the errors dramatically expands the confidence bounds for the key derivative. Apparent estimator precision is much greater, either without clustering or with the use of robust standard errors as opposed to the clustering employed in the paper.

c.) The qualitative results are in general not sensitive to the choice of log migration (theoretically indicated) versus the level of migration as the dependent variable. This finding is crucial to our upcoming step of re-integrating the "flows less than 10 households" into the data as aggregates.

## 5. Discussion and Extensions

We show results for various different measures of the impact of extreme weather demonstrate the preponderance of the evidence across these alternative specifications. Overall, we find the following. First, there is substantial evidence that when the derivative of migration flows with respect to severe weather events are statistically significantly different from zero, severe weather events tend to increase rural-to-urban migration flows (and/or decrease urban-to-rural flows) in the cases of: hail storms, hurricanes, severe storms, wildfires (especially) and severe winter weather.

The evidence is mixed in the case of wind storms. For the most part, wind storms decrease migration between equally urbanized counties, but there is relatively little significant effect of windstorms at the extremes of the distribution of relative urbanization. There is some evidence that "any windstorms" decrease urban-to-rural migration, but no strong evidence that they increase rural-to-urban migration.

In four other cases, rural-to-urban migration is increased by severe weather (when significant) for all but one measure of weather impact: floods (except for property damage), droughts (except for injuries), heat waves (except for crop damage) and tornadoes (except for crop damage). We speculate that these exceptions may reflect that damage to major assets for numerous households that might otherwise prefer to move. This loss of assets may reduce mobility. For example, property damage from flooding decreases rural-to-urban flows but does not significantly increase urban-to-rural flows. Crop damage from heat waves reduces rural-to-urban flows a lot, and urban-to-rural flows a little. Crop damage from tornadoes increases flows between equally urbanized county pairs, or towards slightly more-urbanized counties, but does not have a statistically significant effect on flows between counties with extreme differentials in urbanization.

### a. Planned extension: Recovering minor flows (<10 households)

Among possible extensions, we note that it would be possible to recover the aggregated data for county-to-county flows when individual flows are less than ten households. However, practicality would necessitate using a model where the dependent variable is in levels form. If the specified relationship among the dependent and explanatory variables is valid, and if it remains linear and additively separable in its parameters, then it is possible to aggregate both sides of the equation across destinations (for censored other flows within the same state as the origin county, and censored other flows within the same region as the origin county). Each individual explanatory variable would have to be aggregated across the identical set of counties to which the aggregation of the migration flows applies. This is straightforward, but tedious.

We anticipate extending the analysis to include these other censored flows as our specifications stabilize. Only about 500,000 flows are modeled in the current paper. The remaining roughly 59,500,000 flows over the six years of our sample may contain interesting patterns as well. The fact that our qualitative results are robust to the use of levels of the dependent variable suggests that reversion to a specification using levels of migration would not greatly obscure any potential findings.

Use of the aggregated flows will not increase the size of the estimating sample to this extent, however. For each of the roughly 3100 counties, we will add one observation for “other flows within the same state” and four more observations for “other flows to the Northeast, Midwest, South and West Census regions” for a total of no more than about 75,000 additional observations.[13] A weighting scheme will certainly be appropriate during estimation, given the vastly different scale of the individual versus aggregated county observations. Weights for each observation (individual and two types of aggregates) should probably be based on population in some way.

### b. Planned extension: Derivatives allowed to be quadratic in urbanization difference

The models discussed in the current version of this paper limit the functional form to one where the derivatives of the expected value of the log of migration flows with respect to extreme weather are specified as linear in the urbanization differential between destination and origin counties. We are in the process of estimating more-general specifications where the key derivatives (as pictured in the multi-graph figures in this paper) are curvilinear.

As evidence from the positions of the deciles of the distribution of urbanization differentials, most of the information in our data corresponds to urbanization differentials between about -0.4 and +0.4. The apparent derivatives at the extrema of this distribution may be heavily influenced by the linearity that is assumed in the current set of results.

## 6. Conclusions

Many countries are concerned about the prospect of increased rural-to-urban migration that may be induced by changes in the pattern of extreme weather events associated with climate change. Data are scarce for many jurisdictions, however, making it difficult to assess recent patterns in rural-to-urban migration in response to extreme weather at a fine enough degree of geographic and temporal disaggregation. Thus we have focused our investigation on patterns in the United

[13] The process of identifying the sets of counties over which to aggregate all of the explanatory variables will be machine intensive, but the eventual estimation will involve comparable demands to the models described here.

States over a recent six-year period, utilizing detailed county-to-county migration information from the Internal Revenue Service.

Patterns in a developed country, of course, will reflect the resilience afforded by relatively high incomes, strict building standards, and access to insurance. The findings of this paper are thus not quantitatively transferable to developing countries.  However, the fact that we can detect considerable evidence of rural-to-urban migration, even when relatively little might be expected, is important because it suggests that the influence of extreme weather on migration will be even greater where populations are less resilient.

There is still considerable censoring in the data we have explored in the current paper (i.e. we use only flows greater than 10 households between any pair of counties). However, we find considerable evidence that most types of extreme weather events tend to increase rural-to-urban migration flows (and/or decrease urban-to-rural flows), except in cases where the loss of major assets for a substantial number of affected households may prevent much mobility. More-general models are still being explored, so the next revision of this paper may reveal even greater detail.

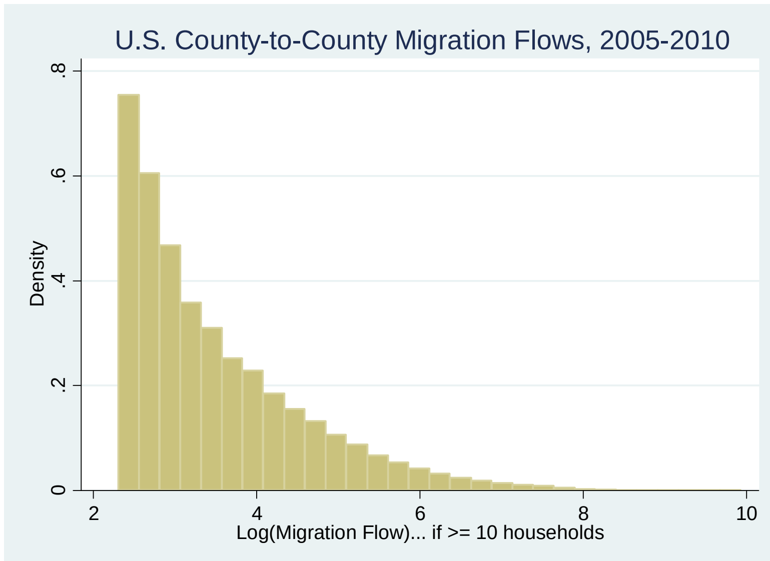


Figure 1 – Distribution of “significant” migration flows is highly skewed; note censoring at 10 households (tax returns)

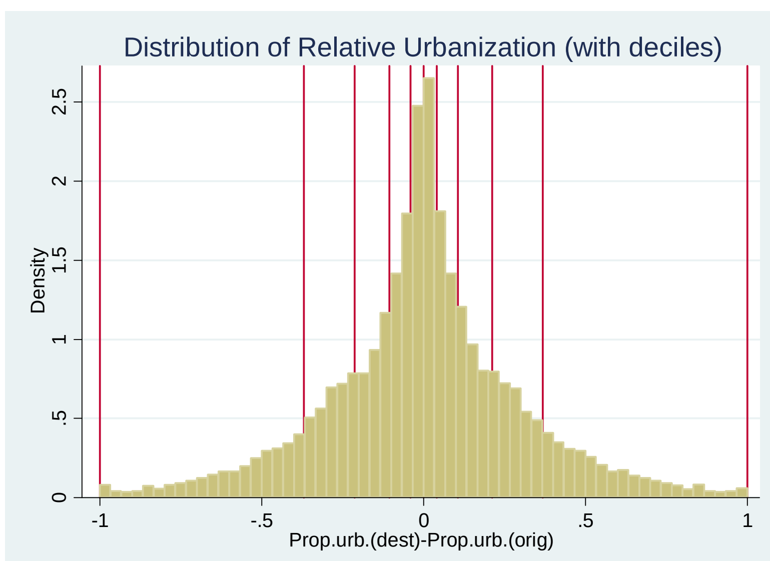


Figure 2 – Distribution of $\left[PropUrban_j - PropUrban_i\right]$

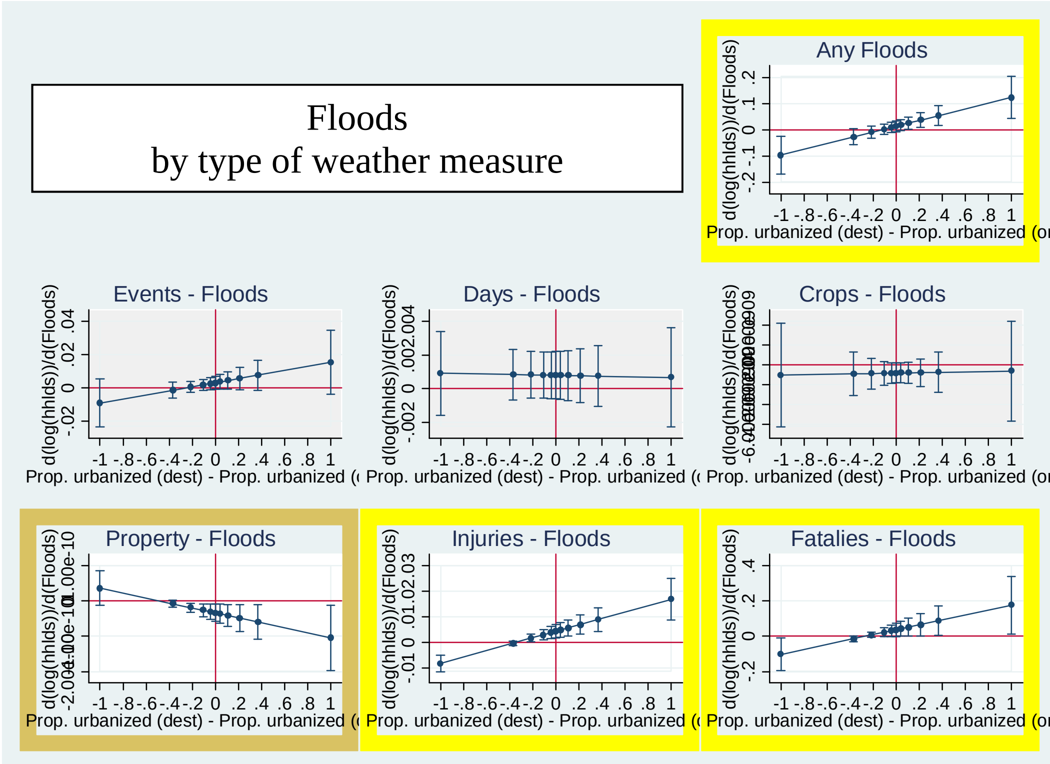


Figure 3a – Floods: $\dfrac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

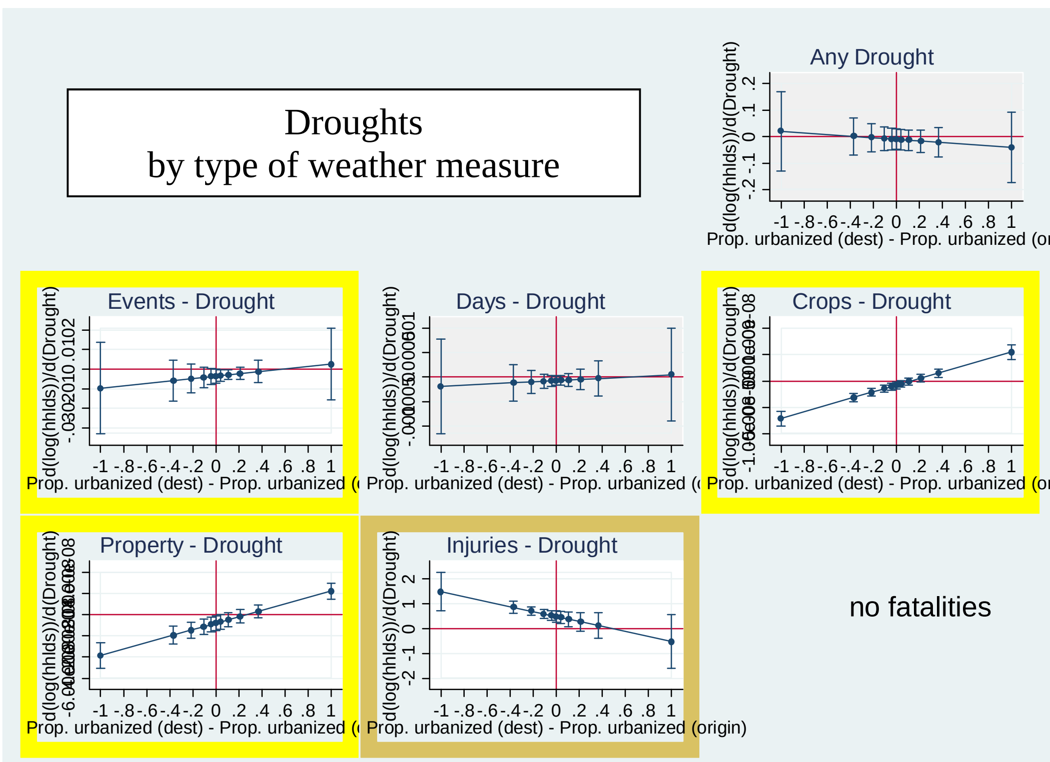


Figure 3b – Droughts: $\dfrac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

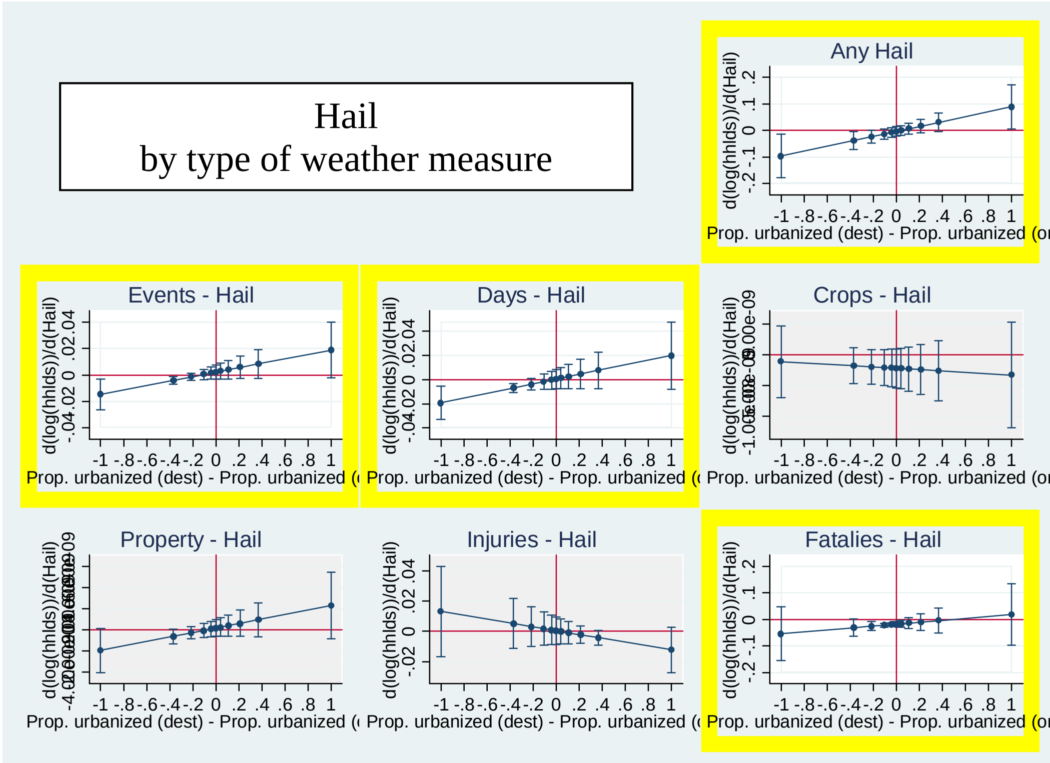


Figure 3c – Hail: $\frac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

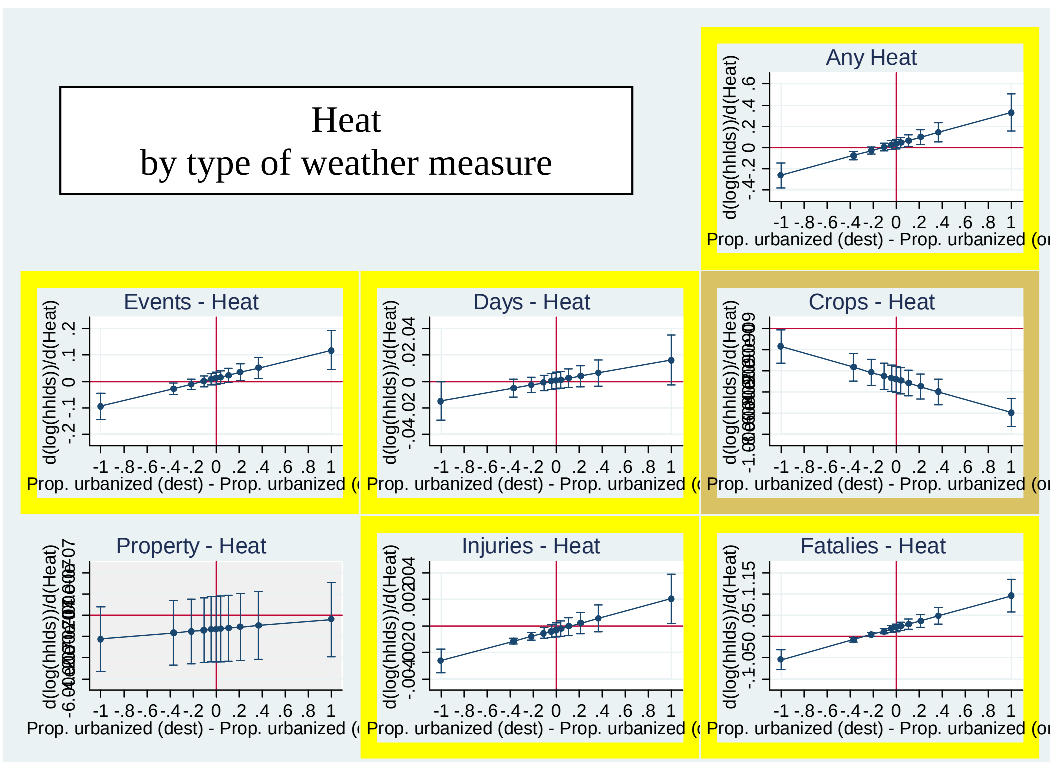


Figure 3d – Heat: $\frac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

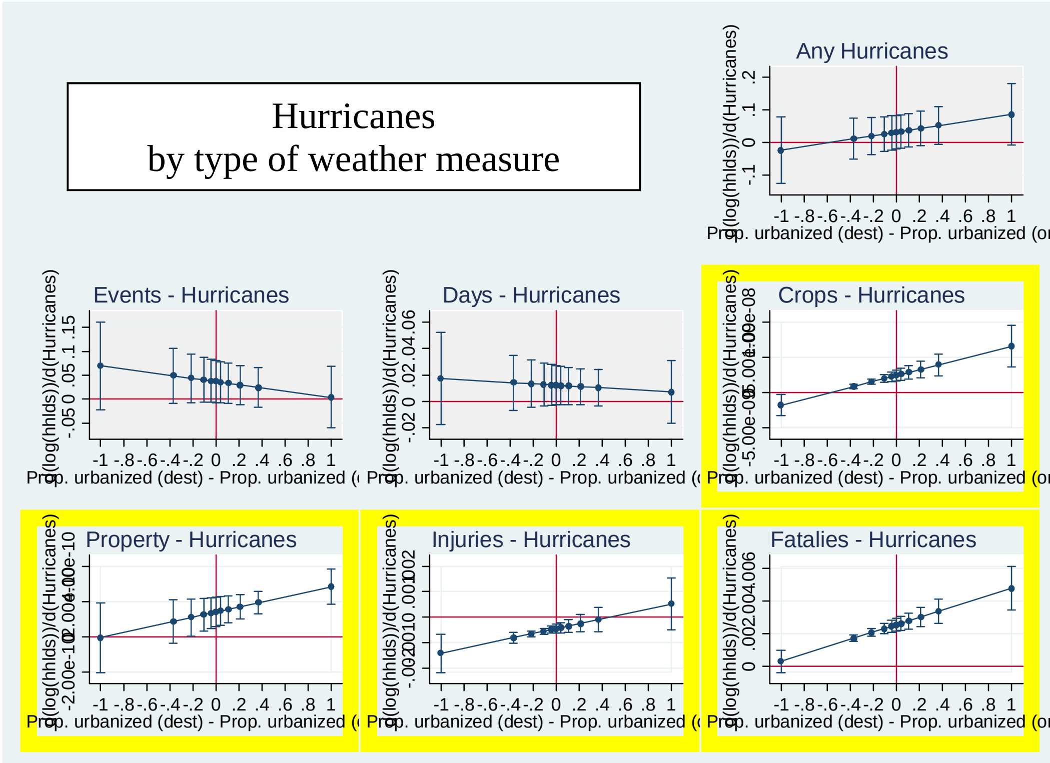


Figure 3e – Hurricanes: $\dfrac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

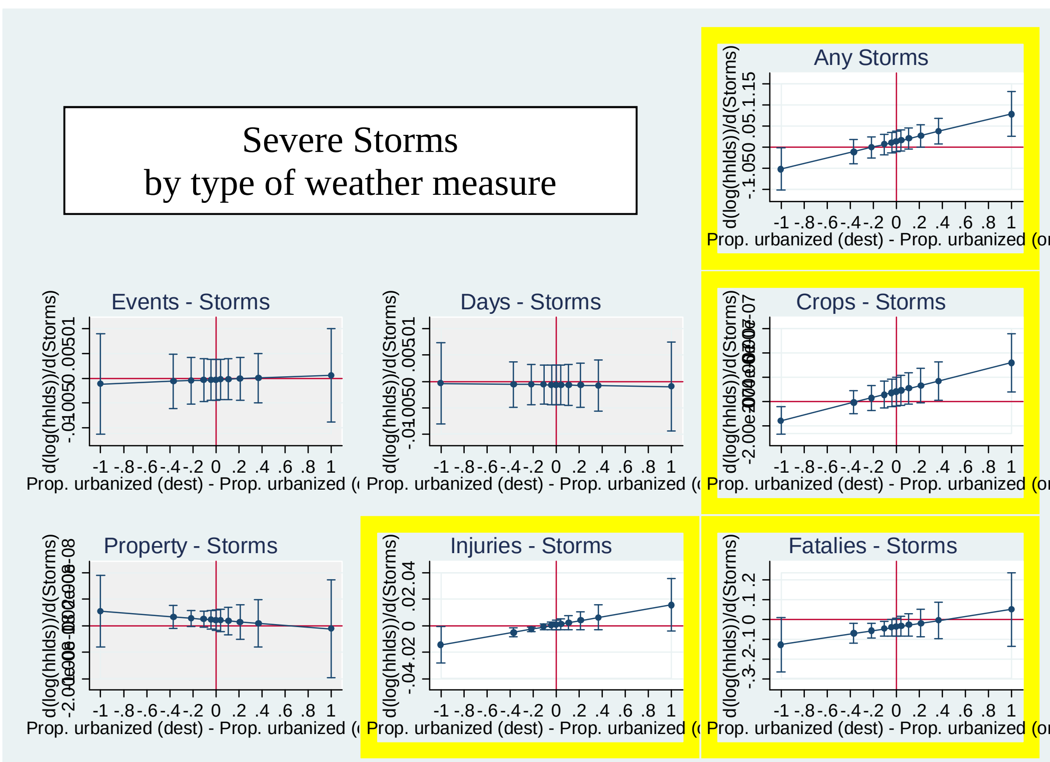


Figure 3f – Storms: $\dfrac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

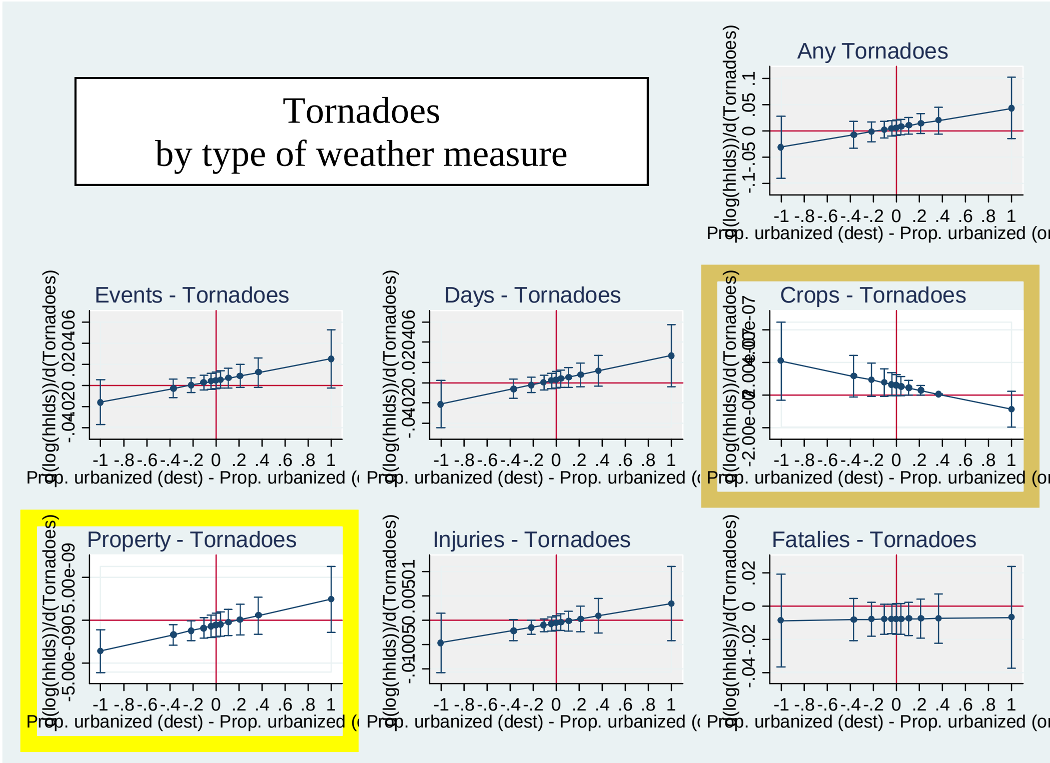


Figure 3g - Tornadoes: $\dfrac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

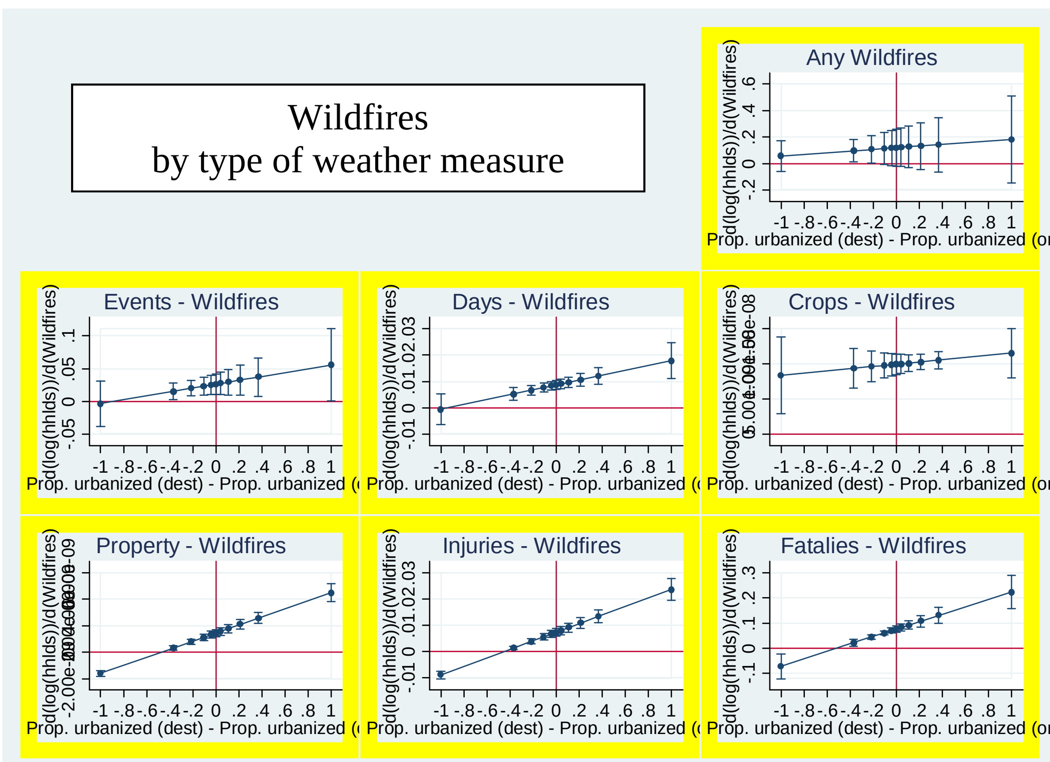


Figure 3h – Wildfires: $\dfrac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

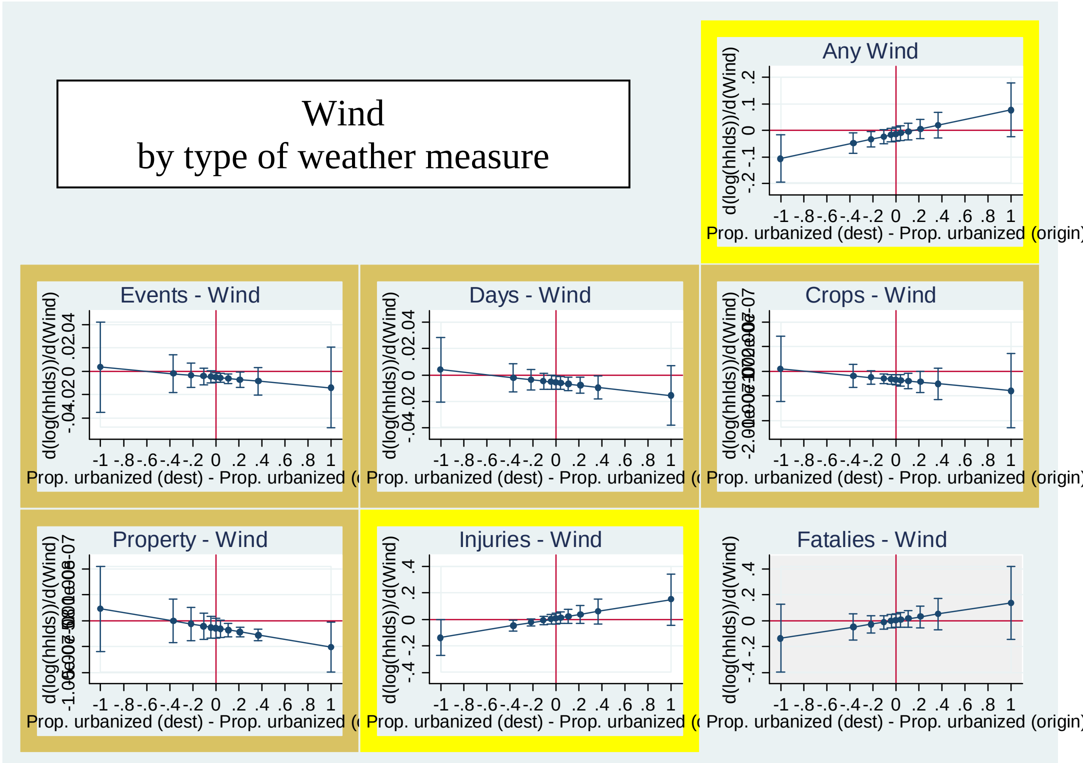


Figure 3i – Wind: $\dfrac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

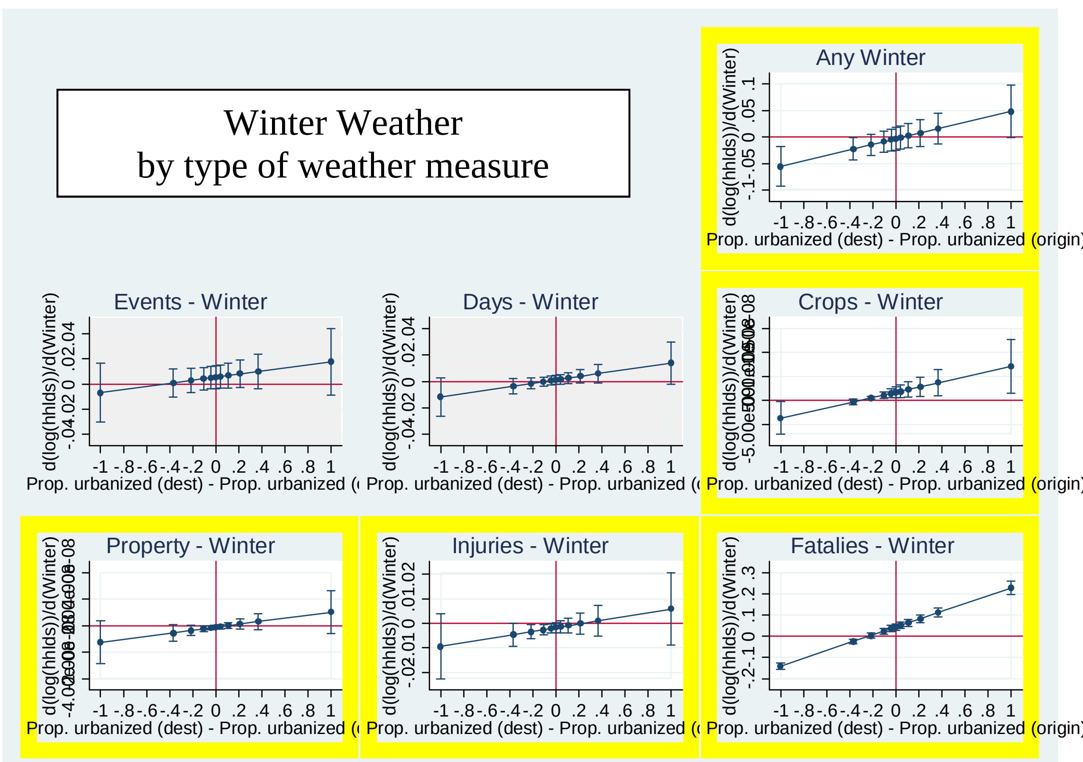


Figure 3j – Winter Weather: $\dfrac{\partial E\left[\log\left(migration_{ijt}\right)\right]}{\partial\left(weather_{kit}\right)} = \beta_{4k} + \beta_{5k}\left[PropUrban_j - PropUrban_i\right]$

Table 1: Annual county-to-county migration rates, per county pair (flows>=10)

| VARIABLES | (1)<br>2005 | (2)<br>2006 | (3)<br>2007 | (4)<br>2008 | (5)<br>2009 | (6)<br>2010 |
|---|---|---|---|---|---|---|
| Tax returns | 87.15 | 87.62 | 86.33 | 86.95 | 86.49 | 87.40 |
| (~households) | (323.5) | (327.8) | (314.9) | (316.4) | (316.4) | (327.4) |
| Total exemptions | 166.7 | 167.3 | 163.6 | 162.0 | 160.9 | 163.0 |
| (~individuals) | (636.9) | (647.5) | (611.4) | (595.7) | (597.4) | (621.2) |
| Observations | 94,814 | 98,401 | 96,759 | 98,861 | 96,426 | 91,419 |

Table 2: Measures of Extreme Weather Events (annual per origin county)

| VARIABLES | (1)<br>At least one | (2)<br># Events | (3)<br>Total days | (4)<br>Crop Damage (x $\$10^6$) | (5)<br>Property Damage (x $\$10^6$) | (6)<br>Injuries | (7)<br>Fatalities |
|---|---|---|---|---|---|---|---|
| Floods | 0.312 | 0.694 | 1.847 | 0.120 | 1.119 | 0.0166 | 0.0214 |
| | (0.463) | (1.685) | (7.739) | (2.859) | (56.56) | (0.732) | (0.207) |
| Droughts | 0.0268 | 0.103 | 3.270 | 0.186 | 0.030 | 0.000208 | n.a. |
| | (0.162) | (0.795) | (26.12) | (2.352) | (0.390) | (0.00408) | |
| Hail | 0.178 | 0.478 | 0.479 | 0.0486 | 0.262 | 0.0129 | 0.00103 |
| | (0.382) | (1.688) | (1.690) | (0.905) | (7.073) | (0.669) | (0.0849) |
| Heat | 0.0398 | 0.0533 | 0.269 | 0.0267 | 0.00039 | 0.103 | 0.0257 |
| | (0.196) | (0.296) | (1.946) | (1.370) | (0.014) | (4.054) | (0.458) |
| Hurricanes | 0.0493 | 0.0725 | 0.141 | 0.175 | 3.932 | 0.0477 | 0.0500 |
| | (0.216) | (0.356) | (0.735) | (2.894) | (60.28) | (2.571) | (3.905) |
| Storms | 0.226 | 0.696 | 0.708 | 0.00139 | 0.044 | 0.0262 | 0.00266 |
| | (0.418) | (1.961) | (2.001) | (0.0554) | (1.402) | (0.551) | (0.0608) |
| Tornadoes | 0.165 | 0.249 | 0.250 | 0.00745 | 0.284 | 0.203 | 0.0196 |
| | (0.371) | (0.682) | (0.685) | (0.526) | (4.930) | (3.006) | (0.377) |
| Wildfires | 0.0224 | 0.0312 | 0.132 | 0.0151 | 0.094 | 0.0257 | 0.00233 |
| | (0.148) | (0.261) | (1.776) | (0.961) | (4.563) | (1.196) | (0.110) |
| Wind | 0.0583 | 0.114 | 0.139 | 0.0022 | 0.0082 | 0.00399 | 0.00206 |
| | (0.234) | (0.684) | (0.869) | (0.099) | (0.338) | (0.100) | (0.0421) |
| Winter | 0.260 | 0.496 | 0.977 | 0.186 | 0.030 | 0.0553 | 0.0142 |
| | (0.438) | (1.408) | (2.738) | (2.352) | (0.390) | (1.077) | (0.209) |
| Observations | 18,445 | 18,445 | 18,445 | 18,445 | 18,445 | 18,445 | 18,445 |

## References


Barrios, S., Bertinelli, L., Strobl, E., 2006. Climatic change and rural-urban migration: The case of sub-Saharan Africa. *Journal of Urban Economics* 60, 357-371.

Black, R., Adger, W.N., Arnell, N.W., Dercon, S., Geddes, A., Thomas, D.S.G., 2011. The effect of environmental change on human migration. *Global Environmental Change-Human and Policy Dimensions* 21, S3-S11.

Borjas, G.J., Bronars, S.G., Trejo, S.J., 1992. Self-selection and internal migration in the United States. *Journal of Urban Economics* 32, 159-185.

Boustan, L.P., Kahn, M.E., Rhode, P.W., 2012. Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century. *American Economic Review* 102, 238-244.

Charney, A.H., 1993. Migration and the public sector - A survey. *Regional Studies* 27, 313-326.

Cragg, M., Kahn, M., 1997. New estimates of climate demand: Evidence from location choice. *Journal of Urban Economics* 42, 261-284.

Cragg, M.I., Kahn, M.E., 1999. Climate consumption and climate pricing from 1940 to 1990. *Regional Science and Urban Economics* 29, 519-539.

Davanzo, J., 1978. Does unemployment affect migration - Evidence from micro data. *Review of Economics and Statistics* 60, 504-514.

Davanzo, J., 1983. Repeat migration in the United States - Who moves back and who moves on. *Review of Economics and Statistics* 65, 552-559.

Deryugina, T., 2011. The dynamic effects of hurricanes in the U.S.: The role of non-disaster transfer payments, MIT, Job Market Paper.

Deschenes, O., Moretti, E., 2009. Extreme weather events, mortality, and migration. *Review of Economics and Statistics* 91, 659-681.

Fielding, A.J., 2011. The impacts of environmental change on UK internal migration. *Global Environmental Change-Human and Policy Dimensions* 21, S121-S130.

Graves, P.E., 1979. Life-cycle empirical analysis of migration and climate, by race. *Journal of Urban Economics* 6, 135-147.

Graves, P.E., 1980. Migration and climate. *Journal of Regional Science* 20, 227-237.

Graves, P.E., Linneman, P.D., 1979. Household migration - Theoretical and empirical results. *Journal of Urban Economics* 6, 383-404.

Gray, C., Mueller, V., 2012. Drought and Population Mobility in Rural Ethiopia. *World Development* 40, 134-145.

Greenwood, M.J., 1975. Research on internal migration in the United States - Survey. *Journal of Economic Literature* 13, 397-433.

Greenwood, M.J., 1985. Human migration - Theory, models, and empirical studies. *Journal of Regional Science* 25, 521-544.

Greenwood, M.J., Hunt, G.L., Rickman, D.S., Treyz, G.I., 1991. Migration, regional equilibrium, and the estimation of compensating differentials. *American Economic Review* 81, 1382-1390.

Gross, E., unknown. U.S. Population Migration Data: Strengths and Limitations. Statistics of Income Division, Internal Revenue Service.

Hornbeck, R., 2009. The enduring impact of the American Dust Bowl, Working paper, Harvard University.

Laczko, F., Aghazarm, C., 2009. *Migration, Environment and Climate Change: Assessing the Evidence*. International Organization for Migration, Geneva.

Landry, C.E., Bin, O., Hindsley, P., Whitehead, J.C., Wilson, K., 2007. Going home: Evacuation-migration decisions of Hurricane Katrina survivors. *Southern Economic Journal* 74, 326-343.

Lewin, P.A., Fisher, M., Weber, B., 2012. Do rainfall conditions push or pull rural migrants: evidence from Malawi. *Agricultural Economics* 43, 191-204.

Marchiori, L., Maystadt, J.F., Schumacher, I., 2012. The impact of weather anomalies on migration in sub-Saharan Africa. *Journal of Environmental Economics and Management* 63, 355-374.

Merryman, S., 2005. USSWM: Stata Module to Provide US State and County Spatial Weight (Contiguity) Matrices., *Statistical components S448405*. Boston College Department of Economics.

Muth, R.F., 1971. Migration - Chicken or egg. *Southern Economic Journal* 37, 295-306.

Nakosteen, R.A., Zimmer, M., 1980. Migration and income - Question of self-selection. *Southern Economic Journal* 46, 840-851.

Partridge, M.D., Rickman, D.S., Olfert, M.R., Ali, K., 2012. Dwindling US internal migration: Evidence of spatial equilibrium or structural shifts in local labor markets? *Regional Science and Urban Economics* 42, 375-388.

Piguet, E., 2010. Linking climate change, environmental degradation, and migration: a methodological overview. *Wiley Interdisciplinary Reviews-Climate Change* 1, 517-524.

Pisati, M., 2001. sg162: Tools for spatial data analysis. *Stata Technical Bulletin* 21-37.

Rappaport, J., 2007. Moving to nice weather. *Regional Science and Urban Economics* 37, 375-398.

Schwartz, A., 1973. Interpreting the effect of distance on migration. *Journal of Political Economy* 81, 1153-1169.

Stark, O., Bloom, D.E., 1985. The new economics of labor migration. *American Economic Review* 75, 173-178.